The Canadian Journal of Statistics
Vol. 33, No. ?, 2005, Pages ???–???
La revue canadienne de statistique

Regression analysis of interval-censored failure time data with linear transformation models

Zhigang ZHANG, Liuquan SUN, Xingqiu ZHAO and Jianguo SUN

Key words and phrases: Covariate effects; interval-censored failure time data; linear transformation models; semiparametric regression.

MSC 2000: Primary 62G05; secondary 62F12.

Abstract: The authors consider the estimation of regression parameters in the context of a class of generalized proportional hazards models, termed linear transformation models, in the presence of interval-censored data. They present an estimating equation approach whose good performance is demonstrated through simulations and which they illustrate in a few concrete cases.

Analyse par régression de temps de décès censurés par intervalle au moyen de modèles à transformation linéaire

Résumé : Les auteurs s'intéressent à l'estimation des paramètres de régression dans le cadre d'une classe de modèles à risques proportionnels généralisés, dits modèles à transformation linéaire, en présence de données censurées par intervalle. Ils proposent une approche par équations d'estimation, dont ils montrent la bonne performance au moyen de simulations et qu'ils illustrent dans quelques cas concrets.

1. INTRODUCTION

This paper discusses regression analysis of interval-censored failure time data, which arises when the time of occurrence of the event of interest is known only to lie in an interval. Such data occur naturally when a failure time arises from a clinical trial or a longitudinal study that entails periodic follow-ups (Finkelstein 1986; Sun 2004). An individual due for weekly or monthly observation for a clinically observed change in disease status ("response") may miss a few weeks' or months' observations and may return with a changed status, thus contributing an interval-censored time of the occurrence of the change.

An example of interval-censored data, that motivated this study, is given in Finkelstein (1986) and arose from a retrospective study on early breast cancer patients with a periodic follow-up. The study involves two treatments: radiation therapy alone and radiation therapy with adjuvant chemotherapy. During the study, each patient was supposed to be checked every 4 or 6 months by physicians for the appearance of breast retraction, a response that has a negative impact on overall cosmetic appearance. However, actual examination times differ from patient to patient and only interval-censored data are available for the appearance.

Several methods have been proposed in the literature for regression analysis of interval-censored failure time data. For example, Betensky, Lindsey, Ryan & Wand (2002), Cai & Betensky (2003), Finkelstein (1986) and Pan (2000) investigated the fitting of the Cox model to the data mentioned above. Huang & Rossini (1997) proposed to use the proportional odds model and Sun (1997)

considered a logistic regression model. More references on this can be found in Sun (2004). Note that most of the existing methods were developed for specific models. Corresponding to this, we will consider a class of generalized Cox models, termed linear transformation models (Chen, Jin & Ying 2002), for interval-censored data.

The remainder of the paper is organized as follows. We will begin in Section 2 by introducing linear transformation models along with some notation and assumptions. The linear transformation models provide flexibility in specifying the effects of covariates on survival times and include as special cases the Cox model and the proportional odds model. Section 3 presents an estimating equation approach for estimating regression parameters with interval-censored data and the asymptotic properties of the proposed estimates are discussed. Section 4 reports some results from simulation studies investigating the finite sample properties of the estimates and in Section 5, the methodology is applied to two well-known examples including the breast cancer data discussed above. Section 6 concludes with some remarks.

2. LINEAR TRANSFORMATION MODELS

Consider a survival study and let T and Z denote the survival time of interest and a vector of covariates, respectively, and u be an unknown strictly increasing function. A linear transformation model assumes that

$$u(T) = Z^\top \beta + \epsilon, \tag{1}$$

where β is the vector of unknown regression parameters and ϵ has a completely known distribution function F. An advantage of the linear transformation model is its generality as it includes some commonly used models as special cases. For example, (1) gives the Cox model (Kalbfleisch & Prentice 2002) if $F(t) = 1 - \exp\{-\exp(t)\}$, an extreme value distribution. If F is the standard logistic distribution, (1) becomes the proportional odds model (Huang & Rossini 1997). Let S_Z denote the survival function of T given Z. Then (1) is equivalent to

$$g\{S_Z(t)\} = u(t) - Z^\top \beta,$$

where $g^{-1}(s) = 1 - F(s)$.

In this paper, we will discuss the fitting of model (1) to interval-censored failure time data (Finkelstein 1986; Sun 2004). By interval-censored data, we mean that instead of observing T, we only observe two random variables $L \leq R$ such that $L \leq T < R$. That is, we have $\mathrm{P}(L \leq T < R) = 1$. Note that the censoring here is different from what is commonly referred to as interval truncation, meaning that a subject is observed if and only if T is in $[L, R)$. By interval-censoring, we mean that a subject is always observed, but its true failure time may not be exactly known and instead is known only to belong to an interval bracketed by L and R as in the example discussed above. If $L = R$ or $R = \infty$, we then have an exact or right-censored failure time, respectively, and in this case, several methods have been proposed for inference about the regression parameter β in model (1). For example, Cheng, Wei & Ying (1995) proposed a class of estimating functions and Chen, Jin & Ying (2002) generalized the partial likelihood estimator of β under the Cox model. However, there seems to be no existing method available for interval-censored data.

In the following, we will confine our attention to the situation where Z is a categorical variable, which frequently occurs in survival studies. In the breast cancer study discussed above, for example, Z denotes treatment indicators. As do most authors (e.g., Gómez, Espinal & Lagakos 2003), we will assume that the mechanism generating censoring intervals for T is independent of T given Z. In other words, for the development of the presented methodology, we need

$$\mathrm{P}(T \leq t | L = \ell, R = r, L \leq T < R, Z) = \mathrm{P}(T \leq t | \ell \leq T < r, Z)\ . \tag{2}$$

3. INFERENCE PROCEDURES

This section considers inference about β when only interval-censored failure time data are observed. Suppose that we have n i.i.d. replicates $\{(L_i, R_i, T_i, Z_i) : i = 1, \ldots, n\}$ of (L, R, T, Z) and that observed data consist of $\{([L_i, R_i), Z_i) : i = 1, \ldots, n\}$. Let H_Z denote the distribution function of T given Z and $\hat{H}_Z$ the maximum likelihood estimator of H_Z based on the observed data on subjects with $Z_i = Z$. Some comments about $\hat{H}_Z$ will be given below.

To estimate β, we will focus on the quantities $I(T_i \geq T_j)$ or the ranks of the T_i's. This is motivated by the fact that they are invariant under model (1) and provide efficient inference through the partial likelihood for β for right-censored data under the Cox model (Kalbfleisch & Prentice, 2002). Since $I(T_i \geq T_j)$ is not observed, we have to consider its expectation given Z_i and Z_j and under model (1), we have

$$\mathrm{E}\{I(T_i \geq T_j)|Z_i, Z_j\} = \mathrm{E}\{I(\epsilon_i - \epsilon_j \geq Z_{ij}^\top \beta)|Z_i, Z_j\} = \tau(Z_{ij}^\top \beta), \tag{3}$$

where $\epsilon_i = u(T_i) - Z_i^\top \beta$ and $Z_{ij} = Z_j - Z_i$, $i, j = 1, \ldots, n$. The function $\tau(t)$ can be expressed as

$$\tau(t) = \int_{-\infty}^{\infty} \{1 - F(s + t)\} dF(s) \, .$$

On the other hand, we show in Appendix A that

$$\mathrm{E}\left\{(a_i a_j)^{-1} \int_{L_i}^{R_i} \int_{L_j}^{R_j} I(t_i \geq t_j) dH_{Z_i}(t_i) dH_{Z_j}(t_j)|Z_i, Z_j\right\} = \tau(Z_{ij}^\top \beta) \tag{4}$$

according to the noninformative censoring mechanism (2), where

$$a_i = \int_{L_i}^{R_i} dH_{Z_i}(t_i), \quad i = 1, \ldots, n.$$

Motivated by the idea behind the generalized estimating equation and equations (3) and (4), we propose to use the estimating equation

$$U(\beta) = \sum_{i=1}^{n} \sum_{j=1}^{n} \tau'(Z_{ij}^\top \beta) \left\{(\hat{a}_i \hat{a}_j)^{-1} \int_{L_i}^{R_i} \int_{L_j}^{R_j} I(t_i \geq t_j) d\hat{H}_{Z_i}(t_i) d\hat{H}_{Z_j}(t_j) - \tau(Z_{ij}^\top \beta)\right\} Z_{ij} = 0$$

for estimation of β. In the above,

$$\hat{a}_i = \int_{L_i}^{R_i} d\hat{H}_{Z_i}(t_i)$$

and $\tau'(t)$ is the first derivative of $\tau(t)$ and has the form

$$\tau'(t) = -\int_{-\infty}^{\infty} f(s + t) dF(s),$$

where f is the density function of ϵ.

Let $\hat{\beta}$ denote the solution to $U(\beta) = 0$. Then it is shown in Appendix B that $\hat{\beta}$ is unique for large n and consistent. Furthermore, motivated by the weighted partial likelihood estimator, we can generalize $U(\beta)$ to

$$U_w(\beta) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(\hat{\beta}) \tau'(Z_{ij}^\top \beta) \left\{(\hat{a}_i \hat{a}_j)^{-1} \int_{L_i}^{R_i} \int_{L_j}^{R_j} I(t_i \geq t_j) d\hat{H}_{Z_i}(t_i) d\hat{H}_{Z_j}(t_j) - \tau(Z_{ij}^\top \beta)\right\} Z_{ij} \, ,$$

where w_{ij} is a positive bounded weight function. Let $\hat{\beta}_w$ denote the solution to the equation $U_w(\beta) = 0$ and β_0 the true value of β. Then $\hat{\beta}_w$ is a consistent estimator of β_0 and heuristically the distribution of $n^{1/2}(\hat{\beta}_w - \beta_0)$ can be approximated by a normal distribution with mean zero and covariance matrix $\hat{\Sigma} = \hat{D}\hat{\Gamma}\hat{D}^\top$ (see Appendix C), where

$$\hat{D}^{-1} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(\hat{\beta}_w)\{\tau'(Z_{ij}^\top\hat{\beta}_w)\}^2 Z_{ij}Z_{ij}^\top$$

and

$$\hat{\Gamma} = \frac{1}{n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k\neq j}^{n}(\hat{e}_{ij} - \hat{e}_{ji})(\hat{e}_{ik} - \hat{e}_{ki})Z_{ij}Z_{ik}^\top,$$

where $\hat{e}_{ij}$ is equal to

$$e_{ij}(\beta) = w_{ij}(\beta)\tau'(Z_{ij}^\top\beta)\left\{(a_i a_j)^{-1}\int_{L_i}^{R_i}\int_{L_j}^{R_j} I(t_i \geq t_j)dH_{Z_i}(t_i)dH_{Z_j}(t_j) - \tau(Z_{ij}^\top\beta)\right\}$$

with β and H_Z replaced by $\hat{\beta}_w$ and $\hat{H}_Z$, respectively. As pointed out by a referee, the above estimate $\hat{\Sigma}$ may in general underestimate the variance. However, our simulation results (see Section 4) suggest that the approximation to both the variance and the distribution work well for practical sample sizes.

To implement the above inference procedure, we need to determine the maximum likelihood estimator of H_Z based on interval-censored data. A common method for this, which is used in the following numerical studies and examples, is to use the self-consistency algorithm given in Turnbull (1976). A summary of some other algorithms for the determination of the maximum likelihood estimator can be found in Sun (2004).

4. NUMERICAL STUDIES

In this section, we report some results from simulation studies conducted for evaluating the proposed methodology. For the results reported below, we considered the two sample comparison problem with Z generated from a Bernoulli distribution with success probability 1/2. Furthermore, we assumed that the survival time of interest follows either the proportional hazards model or the proportional odds model. For the former case, we let $u(t) = \log(t)$ and rounded off generated survival times to their first decimal places. The censoring intervals were generated by adding and subtracting from the generated survival times random numbers from the uniform distribution $U\{0, 0.1, 0.2\}$, respectively. Note that this does not give completely independent observation times, but is more practical since it was motivated by and is equivalent to the usual set-up in follow-up studies. In these situations, each subject is observed at a sequence of time points and the censoring interval is given by the two observation time points that are immediately before and after the survival time, respectively.

For the case of the proportional odds model, we let $u(t) = \log(0.08t)$ and rounded off generated survival times to the nearest integers. The censoring intervals were generated in the same way as for the Cox model, but from $U\{0, 1, 2\}$. For both situations, a constant right-censoring time was used and chosen to give required percentages of right-censored survival times. The weight function in $U_w(\beta)$ was unity. The following results are based on $n = 200$ and 1000 replications for each set-up.

Table 1 presents the estimated average biases of the proposed estimators of the regression parameter β and the empirical 95% coverage probabilities for β based on the normal approximation given in Section 3 for different true values of β and different percentages of right-censoring. To assess the variance estimate, the table also includes the means of the standard error estimates given

in Section 3 and the sample standard deviations of the point estimates. Part (a) of the table is for the case of the proportional hazards model and part (b) corresponds to the proportional odds model. It is seen that the proposed method provides both reasonably accurate point estimates and confidence intervals and that the variance estimate seems reasonable for most cases. The bias for $\beta = -1$ seems larger than for other situations and the results that are not shown here suggest that it becomes smaller as the sample size increases.

We also assessed the appropriateness of the normal approximation to the distribution of $\hat{\beta}$ by studying the quantile plot of the standardized point estimates against the standard normal distribution. Figures 1–2 display such plots corresponding to $\beta = 1$ for the proportional hazards model with 10% right-censoring and the proportional odds model with 20% right-censoring, respectively. They suggest that the normal approximation to the distribution of $(\hat{\beta} - \beta)/\mathrm{SE}(\hat{\beta})$ is good. Similar quantile plots were obtained for other situations.

Table 1. Simulation results for estimation of β.

		a. Proportional hazards model				b. Proportional odds model			
Percentages of right censoring	β	Bias	SE	SEE	CP	Bias	SE	SEE	CP
10%	−1.0	0.001	0.188	0.168	0.928	0.064	0.177	0.208	0.963
	−0.5	−0.003	0.176	0.161	0.932	0.005	0.226	0.242	0.961
	0	−0.001	0.156	0.150	0.956	−0.001	0.237	0.242	0.957
	0.5	0.001	0.167	0.164	0.953	0.003	0.247	0.244	0.953
	1.0	0.001	0.180	0.177	0.948	−0.002	0.234	0.241	0.934
20%	−1.0	0.007	0.190	0.168	0.921	0.074	0.178	0.223	0.966
	−0.5	0.004	0.178	0.160	0.922	−0.006	0.236	0.242	0.958
	0	−0.002	0.156	0.158	0.950	−0.002	0.244	0.241	0.953
	0.5	−0.001	0.176	0.164	0.934	0.004	0.249	0.243	0.938
	1.0	−0.006	0.183	0.177	0.942	−0.004	0.251	0.241	0.942
30%	−1.0	0.041	0.185	0.167	0.911	0.080	0.219	0.241	0.968
	−0.5	0.031	0.173	0.159	0.929	0.021	0.239	0.240	0.952
	0	0.004	0.169	0.160	0.968	−0.004	0.246	0.234	0.944
	0.5	−0.025	0.168	0.162	0.932	0.025	0.253	0.241	0.929
	1.0	0.059	0.183	0.174	0.934	−0.018	0.263	0.248	0.924

Note: Bias represents the bias of the mean of the point estimates; SE represents the sample standard error of the estimates; SEE represents the mean of the standard error estimates; CP represents the empirical 95% coverage probability.

5. EXAMPLES

To illustrate the proposed methodology, we apply it to two sets of interval-censored failure time data. The first one arose from the breast cancer study discussed before (Finkelstein 1986) and the other is from a study of HIV-1 infection on patients with hemophilia (Kroner, Rosenberg, Adedort, Alvord & Goedert 1994).

5.1. Breast cancer study.

This study involves 94 early breast cancer patients with two treatments, radiotherapy alone and radiation therapy with adjuvant chemotherapy. Among them, 46 patients were given radiotherapy alone and 48 patients received radiation therapy with adjuvant chemotherapy. A main objective of the study was to compare the two treatments in terms of the time until the appearance of breast retraction, the survival time of interest. As mentioned before, the patients were examined

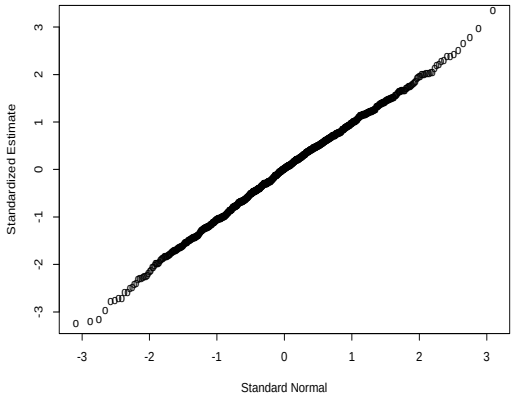

Figure 1. Quantile plot of the estimates for the proportional hazards model with $\beta = 1$.

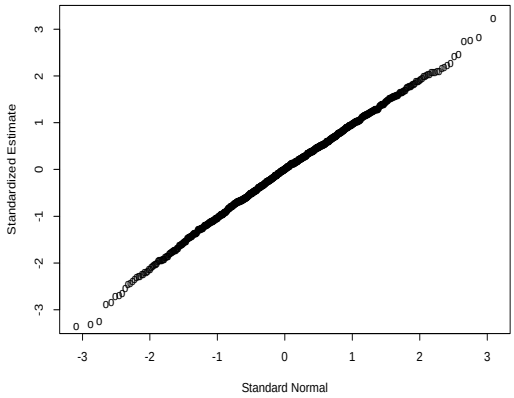

Figure 2. Quantile plot of the estimates for the proportional odds model with $\beta = 1$.

periodically and actual examination times differ from patient to patient since they often missed their regular visits as their recovery progressed. Thus only interval-censored data on the survival time were observed.

For the comparison of the two treatments, we define $Z_i = 0$ if the patient was given radiotherapy alone and $Z_i = 1$ otherwise and assume that the survival time can be described by model (1). Assuming the proportional hazards model, we obtained $\hat{\beta} = -0.697$ with the estimated standard error of 0.251. This yields a P-value of 0.006 based on the Wald test statistic and the standard normal distribution for testing $\beta = 0$, no treatment difference. By assuming the proportional odds model, the proposed method gave $\hat{\beta} = -1.041$ with the estimated standard error being 0.372. In this situation, for the test of the hypothesis $\beta = 0$, the Wald test statistic gave a P-value of 0.005. Both results suggest that the patients given adjuvant chemotherapy had significantly shorter survival times than the patients not having adjuvant chemotherapy. In other words, the adjuvant chemotherapy significantly increased the hazard of breast retraction.

We also tried other models within the class of linear transformation models and obtained similar P-values. Cai & Betensky (2003) and Finkelstein (1986) gave similar results by assuming the Cox model. In summary, the results obtained here confirmed that there were significant differences between the two groups of patients in terms of their hazards of developing breast retraction.

5.2. HIV-I Infection study.

A multi-center prospective study was conducted in 1980's to investigate HIV-1 infection rate among people with hemophilia (Kroner, Rosenberg, Adedort, Alvord & Goedert 1994). They were at risk of HIV-1 infection from blood products received for their treatment such as factor VIII and factor IX concentrate, which were made from the plasma of thousands of donors. In the study, only interval-censored observations were obtained for patients' HIV-1 infection times and the patients were grouped into different groups according to the average annual dose of the blood products they received. For illustration, here we will focus on 368 patients from five centers where patients were enrolled into the study without regard to HIV-1 antibody status and in the groups with low dose or no factor VIII concentrate. The goal is to compare the HIV-1 infection rates between the two groups. More details about the study can be found in Goedert et al. (1989).

To apply the proposed methodology, we let $Z_i = 0$ for patients receiving no factor VIII concentrate and $Z_i = 1$ otherwise. First we assumed that the HIV-1 infection time could be described

by the Cox model and in this case the method gave $\hat{\beta} = -1.016$ with the estimated standard error of 0.124. By assuming the proportional odds model, the method yielded $\hat{\beta} = -1.511$ with the estimated standard error being 0.182. In both cases, we have P-values less than 0.00001 for testing no treatment effect or $\beta = 0$. Similar P-values were obtained under other models from the linear transformation model family. In summary, the results suggest that the patients receiving low dose factor VIII concentrate had a significantly higher hazard rate of HIV-1 infection than the patients receiving no factor VIII concentrate.

6. CONCLUDING REMARKS

The proportional hazards model is the most commonly used regression model in survival analysis. However, as pointed out by some authors (Chen, Jin & Ying 2002), sometimes it may not provide a good fit to observed failure time data and thus a different or more general model is needed. This paper investigated the fitting of linear transformation models, a class of generalized Cox models, to interval-censored failure time data, which has not been discussed before. For inference about regression parameters, we presented an estimating equation approach based on rank information as the partial likelihood approach for right-censored data. Also like the partial likelihood approach, the presented method has the advantage that it does not involve estimation of the nonparametric part of the model. The simulation studies showed that the methodology works quite well for practical situations.

More research remains to be done. In this paper, we considered only the discrete covariate situation. Although this covers many common situations like the examples discussed in the previous section, it would be useful to generalize the proposed method to or develop similar methods for continuous covariate situations. Also sometimes it may be interesting to estimate the function u or the baseline cumulative hazard function. For this, under the Cox model, Betensky, Lindsey, Ryan & Wand (2002) and Cai & Betensky (2003) proposed to use the local likelihood method and the penalized spline approach, respectively. However, it does not seem to be straightforward to generalize these methods to model (1). Another topic for future research is the rigorous investigation of the asymptotic properties of the presented methodology. In the above, we only gave a heuristic derivation of the normality of the regression parameter estimator and an *ad hoc* estimator of the asymptotic covariance. Although the simulation result indicates that they seem reasonable, it would be helpful to provide a rigorous proof.

APPENDIX A: THE DERIVATION OF EQUATION (4)

For $s \le t$ and $i = 1, \ldots, n$, define $b_i(s,t) = \int_s^t dH_{Z_i}(v)$. It can be easily shown from (2) that

$$\mathrm{P}(T_i \le t | L_i = \ell_i, R_i = r_i, Z_i) = I(\ell_i \le t < r_i) b_i^{-1}(\ell_i, r_i) \int_{\ell_i}^{t} dH_{Z_i}(v) + I(t \ge r_i)$$

and thus

$$\begin{aligned} &\mathrm{E}\{I(T_i \ge T_j) | L_i = \ell_i, R_i = r_i, Z_i, L_j = \ell_j, R_j = r_j, Z_j\} \\ &= \int_{t_i, t_j} I(t_i \ge t_j) d\mathrm{P}(T_i \le t_i | L_i = \ell_i, R_i = r_i, Z_i) d\mathrm{P}(T_j \le t_j | L_j = \ell_j, R_j = r_j, Z_j) \\ &= \{b_i(\ell_i, r_i) b_j(\ell_j, r_j)\}^{-1} \int_{\ell_j}^{r_j} \int_{\ell_i}^{r_i} I(t_i \ge t_j) dH_{Z_i}(t_i) dH_{Z_j}(t_j) \, . \end{aligned}$$

Let $Q_{L,R}(\ell, r | Z)$ denote the joint distribution of (L, R) given Z. Then it follows from the above equation

$$\mathrm{E}\{I(T_i \ge T_j) | Z_i, Z_j\} \quad = \quad \mathrm{E}\left[\mathrm{E}\left\{I(T_i \ge T_j) | L_i, R_i Z_i, L_j, R_j, Z_j\right\} | Z_i, Z_j\right]$$

$$
\begin{aligned}
&= \int_{\ell_i \le r_i, \ell_j \le r_j} \mathrm{E}\{I(T_i \ge T_j) | L_i = \ell_i, R_i = r_i, Z_i, L_j = \ell_j, R_j = r_j, Z_j\} \\
&\quad \times dQ_{L_i,R_i}(\ell_i, r_i | Z_i) dQ_{L_j,R_j}(\ell_j, r_j | Z_j) \\
&= \int_{\ell_i \le r_i, \ell_j \le r_j} \{b_i(\ell_i, r_i) b_j(\ell_j, r_j)\}^{-1} \int_{\ell_j}^{r_j} \int_{\ell_i}^{r_i} I(t_i \ge t_j) dH_{Z_i}(t_i) dH_{Z_j}(t_j) \\
&\quad \times dQ_{L_i,R_i}(\ell_i, r_i | Z_i) dQ_{L_j,R_j}(l_j, r_j | Z_j) \\
&= \mathrm{E}\left[\{b_i(L_i, R_i) b_j(L_j, R_j)\}^{-1} \int_{L_j}^{R_j} \int_{L_i}^{R_i} I(t_i \ge t_j) dH_{Z_i}(t_i) dH_{Z_j}(t_j) \middle| Z_i, Z_j \right] .
\end{aligned}
$$

Hence by noting that $a_i = b_i(L_i, R_i)$ $(i = 1, \ldots, n)$, (4) follows from the above equation and equation (3).

APPENDIX B: THE DERIVATION OF CONSISTENCY AND UNIQUENESS

For the consistency and uniqueness of $\hat{\beta}$, following the arguments in Appendix 1 of Cheng, Wei & Ying (1995), we consider the quantity $n^{-2}U(\beta)(\beta_0 - \beta)$. Let G denote the distribution function of Z. It follows from the uniform consistency of $H_Z(t)$ (Groeneboom & Wellner 1992) and the strong law of large numbers for U-statistics that with probability one, $n^{-2}U(\beta)(\beta_0 - \beta)$ converges to

$$
\int_{z_1, z_2} \tau'(z_{12}^\top \beta)(z_{12}^\top \beta_0 - z_{12}^\top \beta) \left\{ \tau(z_{12}^\top \beta_0) - \tau(z_{12}^\top \beta) \right\} dG(z_1) dG(z_2)
$$

uniformly on any compact set of β, where $z_{12} = z_1 - z_2$. Since τ is a strictly decreasing function and $\tau' < 0$, the above limit is thus nonnegative and is zero only when $\beta = \beta_0$. This implies that $\hat{\beta}$ is unique and consistent.

APPENDIX C: A HEURISTIC DERIVATION OF THE ASYMPTOTIC NORMALITY of $\hat{\beta}_w$

To see the asymptotic distribution of $\hat{\beta}_w$, first note that the consistency of $\hat{H}_Z$ (Groeneboom & Wellner 1992) suggests that $n^{-3/2}U_w(\beta_0)$ can be approximated by

$$
n^{-3/2} \sum_{i=1}^n \sum_{j=1}^n w_{ij}(\beta_0) \tau'(Z_{ij}^\top \beta_0) Z_{ij} \times \left\{ (a_i a_j)^{-1} \int_{L_i}^{R_i} \int_{L_j}^{R_j} I(t_i \ge t_j) dH_{Z_i}(t_i) dH_{Z_j}(t_j) - \tau(Z_{ij}^\top \beta_0) \right\} ,
$$

which is asymptotically equivalent to a multivariate U-statistic. Thus it follows from the asymptotic theory of multivariate U-statistics (Wei & Johnson 1985) that the distribution of $n^{-3/2}U_w(\beta_0)$ can be asymptotically approximated by the normal distribution with mean 0 and covariance matrix

$$
\Gamma = \lim_{n \to \infty} \mathrm{E} \left\{ \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{k \ne j}^n (e_{ij} - e_{ji})(e_{ik} - e_{ki}) Z_{ij} Z_{ik}^\top \right\} ,
$$

which can be estimated by $\hat{\Gamma}$ given in Section 3. Now using the Taylor series expansion of $U_w(\hat{\beta}_w)$ around β_0, we have, asymptotically,

$$
n^{-3/2} U_w(\beta_0) = \left\{ -n^{-2} \frac{\partial U_w(\beta^*)}{\partial \beta} \right\} \left\{ n^{1/2}(\hat{\beta}_w - \beta_0) \right\} ,
$$

where β^* is on the segment between β_0 and $\hat{\beta}_w$. This plus the convergence of $-n^{-2}\partial U_w(\beta^*)/\partial \beta$ in probability suggests that the distribution of $n^{1/2}(\hat{\beta}_w - \beta_0)$ can be asymptotically approximated

by the normal distribution with mean 0 and covariance matrix $\hat{\Sigma}$ given in Section 3. Note that here the convergence of $-n^{-2}\partial U_w(\beta^*)/\partial\beta$ follows from the fact that asymptotically,

$$-n^{-2}\frac{\partial U_w(\beta^*)}{\partial\beta} = n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(\beta_0)\{\tau'(Z_{ij}^\top\beta_0)\}^2 Z_{ij}Z_{ij}^\top \ ,$$

which holds due to the consistency of $\hat{\beta}_w$ and the uniform strongly convergence of $\hat{H}_Z$.

ACKNOWLEDGEMENTS

The authors wish to thank the Editor, Professor Douglas P. Wiens, the Associate Editor and two referees for their many helpful comments and suggestions that greatly improved the paper. This work was partly supported by the National Natural Science Foundation of China grant 10471140 (to the second author) and a grant from the National Health of Institutes (to the last author).

REFERENCES

R. A. Betensky, J. C. Lindsey, L. M. Ryan & M. P. Wand (2002). A local likelihood proportional hazards model for interval censored data. *Statistics in Medicine*, 21, 263–275.

T. Cai & R. A. Betensky (2003). Hazard regression for interval-censored data with penalized spline. *Biometrics*, 59, 570–579.

K. Chen, Z. Jin & Z. Ying (2002). Semiparametric analysis of transformation models with censored data. *Biometrika*, 89, 659–668.

S. C. Cheng, L. J. Wei & Z. Ying (1995). Analysis of transformation models with censored data. *Biometrika*, 83, 835–846.

D. M. Finkelstein (1986). A proportional hazards model for interval-censored failure time data. *Biometrics*, 42, 845–854.

J. Goedert, C. Kessler, L. Adedort et al. (1989). A prospective-study of human immunodeficiency virus type-1 infection and the development of aids in subjects with hemophilia. *New England Journal of Medicine*, 321, 1141–1148.

G. Gómez, A. Espinal & S. W. Lagakos (2003). Inference for a linear regression model with an interval-censored covariate. *Statistics in Medicine*, 22, 409–425.

P. Groeneboom & J. A. Wellner (1992). *Information bounds and nonparametric maximum likelihood estimation*. Birkhäuser Verlag, Basel, Switzerland.

J. Huang & A. J. Rossini (1997). Sieve estimation for the proportional-odds failure time regression with interval censoring. *Journal of the American Statistical Association*, 92, 960–967.

J. D. Kalbfleisch & R. L. Prentice (2002). *The Statistical Analysis of Failure Time Data*. Wiley, New York.

B. Kroner, P. Rosenberg, L. Adedort, W. Alvord & J. Goedert (1994). HIV-1 infection incidence among people with hemophilia in the United States and Western Europe, 1978-1990. *Journal of Acquired Immune Deficiency Syndromes*, 7, 279–286.

W. Pan (2000). A two-sample test with interval censored data via multiple imputation. *Statistics in Medicine*, 19, 1–11.

J. Sun (1997). Regression analysis of interval-censored failure time data. *Statistics in Medicine*, 17, 497–504.

J. Sun (2004). Interval censoring. *Encyclopedia of Biostatistics*, Second Edition. Wiley, New York.

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

L. J. Wei & W. E. Johnson (1985). Combining dependent tests with incomplete repeated measurements. *Biometrika*, 72, 359–364.

Received 12 June 2004
Accepted 12 November 2004

Zhigang ZHANG: zhigang.zhang@okstate.edu
Department of Statistics, Oklahoma State University
Stillwater, OK 74078 USA

Liuquan SUN: slq@amt.ac.cn
Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100080, P.R.China

Xingqiu ZHAO: zhaox7@math.mcmaster.ca
Department of Mathematics and Statistics
McMaster University, Hamilton
Ontario, Canada L8S 4K1

Jianguo SUN: sunj@missouri.edu
Department of Statistics, University of Missouri
146 Middlebush Hall
Columbia, MO 65201 USA