Statistica Sinica **12**(2002), 1073-1083

NONPARAMETRIC SURVIVAL COMPARISONS FOR INTERVAL-CENSORED CONTINUOUS DATA

Hong-Bin Fang, Jianguo Sun and Mei-Ling Ting Lee

St. Jude Children's Research Hospital, University of Missouri-Columbia and Harvard Medical School

Abstract: The problem of two-sample survival comparisons has been investigated by several authors. Pepe and Fleming (1989) introduced a test for right-censored survival data. Petroni and Wolfe (1994) considered a similar test where the survival times can only take on a finite number of values. We extend their tests to general types of interval-censored data and introduce a class of test statistics based on integrated weighted differences between the two estimated survival functions.We derive the asymptotic distributions of the generalized test statistics, present a bootstrap test procedure and apply the proposed method to a set of interval-censored data from a breast cancer study.

Key words and phrases: Breast cancer, bootstrap method, continuous survival time, interval-censored data, two-sample comparison.

1. Introduction

Interval-censored survival data frequently arise in clinical trials and follow-up studies such as AIDS and cancer studies (DeGruttola and Lagakos (1989)). Let T denote the survival time for a subject. Instead of observing the value of T, we only know that $T \in (A, B]$ for some interval $(A, B]$. The analysis of such data has attracted a great deal of attention in the literature. For recent advances and applications, see Petroni and Wolfe (1994), Klein and Moeschberger (1997), Li, Watkins and Yu (1997), Sun (1998), Sun, Liao and Pagano (1999) and Yu, Li and Wang (2000).

Three special cases of interval-censored data have been well studied: (a) if $A = 0$, we have left-censored survival data; (b) if $B = \infty$, we have right-censored survival data (Kalbfleisch and Prentice (1980)); (c) if either $A = 0$ or $B = \infty$, the data are usually referred to as current status data. For right censored data, there exist many nonparametric test procedures in the literature for two-sample survival comparisons. Common classes of test statistics are weighted log-rank statistics and weighted Kaplan-Meier statistics (Gill (1980), Pepe and Fleming (1989)). The first class contains the well-known log-rank test (Kalbfleisch and Prentice (1980)). These two classes are based on integrated weighted differences in estimated hazard functions and in estimated survival functions and are sensitive to stochastically ordered hazard functions and survival functions, respectively. Among others, Sun and Kalbfleisch (1993) and Keiding, Begtrup, Scheike and Hasibeder (1996) consider statistical methods for current status data.

In this paper, we extend the two-sample tests considered by Pepe and Fleming (1989) and Petroni and Wolfe (1994) to situations where the intervals are not limited to the special cases listed above. Assume that the survival time of interest is a continuous random variable but only interval-censored continuous data $T \in (A, B]$ can be observed, where there is no restriction on the values of A and B. We derive a generalized class of test statistics based on integrated weighted differences in estimated survival functions to test the hypothesis $H_0 : S_1(t) = S_2(t)$, where $S_1(t)$ and $S_2(t)$ denote two survival functions. Pepe and Fleming (1989) introduced a test procedure for right censored data. We extend their procedure to general interval censored data. Petroni and Wolfe (1994) and Sun (1996) studied similar test statistics for the situation where the survival time can only take a finite number of values.

The remainder of the paper is organized as follows. In Section 2, we describe the data structure discussed here and introduce some notation and assumptions. In Section 3, we investigate the properties of the proposed test statistics along with their use for testing H_0. In particular we show that, under H_0 and some mild regularity conditions, the test statistic has an asymptotic normal distribution. Unlike the case of right-censored data, however, the asymptotic variance is very complicated and cannot be expressed in closed form. To accommodate this problem, in addition to giving a consistent estimate of the asymptotic variance, we also introduce a simple bootstrap procedure to determine the p-value of the hypothesis H_0. In Section 4, we demonstrate the proposed test procedure by applying it to a set of interval-censored data obtained from a breast cancer study. We conclude in Section 5 with some discussion.

2. Data, Notation and Assumptions

Let T denote the survival time of interest. Assume that it is not observable except for knowing that it belongs to an interval. Specifically, suppose that we observe two random variables U and V with $U \leq V$ and the indicator variables $\Delta_1 = I(T \leq U)$, $\Delta_2 = I(U < T \leq V)$ and $\Delta_3 = 1 - \Delta_1 - \Delta_2$, where I is the indicator function. Here U and V can be regarded as two examination times that bracket the survival time T assuming that each subject is observed at a sequence of examination times. The variables Δ_1, Δ_2 and Δ_3 indicate whether the survival event of interest has occurred before U, during the examination interval $(U, V]$, or after V, respectively. Throughout this paper, we assume that survival time is independent of the examination times.

As in the previous section, assume there exist two populations and let S_1 and S_2 denote their survival functions, respectively. Also let $F_1(t) = 1 - S_1(t)$ and $F_2(t) = 1 - S_2(t)$ denote the corresponding cumulative distribution functions. Suppose that n independent subjects are involved in a survival study and the observations are $\{(u_i, v_i, \delta_{1i}, \delta_{2i}, \delta_{3i}) : i = 1, \ldots, n\}$. Among the n subjects, we assume that n_1 individuals come from the population with survival function S_1 and n_2 individuals come from the population with survival function S_2. Let $H(u, v)$ denote the joint cumulative distribution function of U and V, and let H_1 and H_2 denote the marginal cumulative distribution functions of U and V, respectively. Also let $h(u, v)$, $h_1(u)$ and $h_2(v)$ denote the corresponding density functions; let S_0 and F_0 denote the common survival and cumulative distribution functions of interest under the hypothesis H_0 described in the previous section. Without loss of generality, we assume that the support of the survival functions of interest is $[0, M]$. By using the above notation, we can write the conditional likelihood function given the u_i's and v_i's as

$$L(S_0(t)) = \prod_{i=1}^{n} \{1 - S_0(u_i+)\}^{\delta_{1i}} \{ S_0(u_i+) - S_0(v_i+)\}^{\delta_{2i}} S_0^{\delta_{3i}}(v_i+) \qquad (1)$$

under H_0.

For the asymptotic properties of the test statistics U_n, we need the following regularity conditions.

(A) h_1 and h_2 are continuous with $h_1(t) + h_2(t) > 0$ for all $t \in [0, M]$;

(B) $h(u, v)$ is continuous, with uniformly bounded partial derivatives except at a finite number of points, where left and right partial derivatives exist;

(C) $P(V - U < \varepsilon_0) = 0$ for some ε_0 with $0 < \varepsilon_0 < M/2$, so h does not have mass close to the diagonal;

(D) $F_0 \ll H_1 + H_2$, and for $t \in (0, M)$, F_0 has a derivative f which is continuous at t and satisfies $f(t) \geq c$ for a constant $c > 0$ independent of t.

3. Nonparametric Test Procedures

Let $\hat{S}_{1n_1}(t)$ and $\hat{S}_{2n_2}(t)$ denote the maximum likelihood estimators of $S_1(t)$ and $S_2(t)$, respectively. To test the hypothesis H_0, we consider the following class of statistics

$$U_n = \sqrt{\frac{n_1 n_2}{n}} \int_0^M w(t)[\hat{S}_{1n_1}(t) - \hat{S}_{2n_2}(t)]dt,$$

where $w(s)$ is a weight function and M can be infinity or the longest follow-up time (it is formally defined below). Note that U_n is the integrated weighted difference in estimated survival functions and it reduces to the two-sample test statistics given in Pepe and Fleming (1989) if right-censored data are observed.

If $w(t) = 1$, U_n is the difference between the estimated mean survival times of the two populations, which is a natural choice as the test statistic for H_0.

Let $\hat{F}_{1n_1}(t) = 1 - \hat{S}_{1n_1}(t)$ and $\hat{F}_{2n_2}(t) = 1 - \hat{S}_{2n_2}(t)$. Then ML estimators $\hat{S}_{1n_1}$ and $\hat{S}_{2n_2}$ can be obtained using the likelihood function at (1), based on the observations associated only with S_1 and S_2, respectively. Turnbull (1976), Gentlemen and Geyer (1994) and Groeneboom (1996), among others, have proposed algorithms for computing the maximum likelihood estimator for interval-censored data.

We establish the asymptotic null distribution of U_n for testing the hypothesis H_0. For this purpose, let $Q_0(u, v, \delta_1, \delta_2)$ denote the distribution function of $(U, V, \Delta_1, \Delta_2)$ under H_0, and ϕ_{w,F_0} the solution to the Fredholm integral equation

$$\phi_{w,F_0}(t) = d_{F_0}(t) \left\{ w(t) - \int_0^M \frac{\phi_{w,F_0}(t) - \phi_{w,F_0}(t')}{|F_0(t) - F_0(t')|} h^*(t', t) dt' \right\}, \qquad (2)$$

where $d_{F_0}(t) = [F_0(t)\{1 - F_0(t)\}]/[h_1(t)\{1 - F_0(t)\} + h_2(t)F_0(t)]$ and $h^*(t', t) = h(t', t) + h(t, t')$. Define

$$\tilde{\theta}_{w,F_0}(u, v, \delta_1, \delta_2) = -\delta_1 \frac{\phi_{w,F_0}(u)}{F_0(u)} - \delta_2 \frac{\phi_{w,F_0}(v) - \phi_{w,F_0}(u)}{F_0(v) - F_0(u)} + (1 - \delta_1 - \delta_2) \frac{\phi_{w,F_0}(v)}{1 - F_0(v)}, \qquad (3)$$

which is the canonical gradient in the Hilbert space $\overline{T(Q_0)}$ defined as an extension of the tangent space at Q_0 (see Groeneboom (1996)). The proof of the following theorem is given in the Appendix.

Theorem. *Assume that regularity conditions* (A)–(D) *hold and that* $n_1/n \to a_1$, $n_2/n \to a_2$ *as* $n \to \infty$, *where* $0 < a_1, a_2 < 1$ *and* $a_1 + a_2 = 1$. *Assume that the weight function* $w(t)$ *has a bounded derivative on* $[0, M]$. *Then under* H_0, *as* $n \to \infty$, U_n *has an asymptotic normal distribution with mean 0 and variance* $\|\tilde{\theta}_{w,F_0}\|^2 = \int \tilde{\theta}^2_{w,F_0} \, dQ_0$.

To test the hypothesis H_0 using the above theorem, we offer two procedures. The first one is a simple bootstrap procedure, and the second is to derive a consistent estimate of the asymptotic variance, $\|\tilde{\theta}_{w,F_0}\|^2$, of the test statistic U_n.

3.1. A bootstrap procedure

The bootstrap test procedure can be described as follows. Let K be the number of repetitions, U^* the observed value of the statistics U_n based on the sampled data, and $U_1^*, \ldots, U_K^*$ values of the statistics U_n based on independent bootstrap samples with replacement. It follows from the theorem that, under H_0 and when n is large, the true bootstrap samples $U_1^*, \ldots, U_K^*$ follow a normal distribution. The p-value of the two-sided test of the hypothesis $H_0 : S_1(t) = S_2(t)$ can then be calculated as the proportion of $U_1^*, \ldots, U_K^*$ whose absolute values are greater than or equal to the absolute value of U^*, the observed test statistic.

3.2. A procedure based on a consistent estimate of the asymptotic variance

To derive a consistent estimate of the asymptotic variance of U_n, let $\hat{F}_n$ denote the maximum likelihood estimator of the common cumulative distribution function F_0 under H_0 and $0 < t_1 < \cdots < t_m < M$ the points at which $\hat{F}_n$ has jumps. Also let $z_i = \hat{F}_n(t_j)$, $j = 1, \ldots, m$, and let $\phi_{w,\hat{F}_n}$ denote the solution to equation (2) after replacing F_0 by $\hat{F}_n$. According to Theorem 3.5 of Groeneboom (1996), $\phi_{w,\hat{F}_n}$ is absolutely continuous with respect to $\hat{F}_n$ and a step function with jumps at the t_j's. The empirical distributions of (U, V), U and V are denoted by $\hat{H}_n$, $\hat{H}_{1n}$ and $\hat{H}_{2n}$, respectively. Let $y_j = \phi_{w,\hat{F}_n}(t_j)$ and define

$$\begin{aligned} d_j &= \frac{z_j(1-z_j)}{\Delta_j(h_1)(1-z_j) + \Delta_j(h_2)z_j}, \\ \Delta_j(h_r) &= \int_{t_j}^{t_{j+1}} h_r(t)dt \approx \int_{t_j}^{t_{j+1}} d\hat{H}_{rn}(t), \\ \Delta_{jl}(h) &= \int_{u=t_j}^{t_{j+1}} \int_{v=t_l}^{t_{l+1}} h(u,v)dvdu \approx \int_{u=t_j}^{t_{j+1}} \int_{v=t_l}^{t_{l+1}} d\hat{H}_n(u,v), \end{aligned}$$

$j, l = 1, \ldots, m$, $r = 1, 2$. Then it can be shown that the vector $\boldsymbol{y} = (y_1, \ldots, y_m)'$ is the unique solution to the following set of linear equations

$$y_j \left\{ d_j^{-1} + \sum_{l<j} \frac{\Delta_{lj}(h)}{z_j - z_l} + \sum_{l>j} \frac{\Delta_{jl}(h)}{z_l - z_j} \right\} = \Delta_j(w) + \sum_{l<j} \frac{\Delta_{lj}(h)}{z_j - z_l} y_l + \sum_{l>j} \frac{\Delta_{jl}(h)}{z_l - z_j} y_l,$$

for $j = 1, \ldots, m$. Define

$$\tilde{\theta}_{w,\hat{F}_n}(u, v, \delta_1, \delta_2) = -\delta_1 \frac{\phi_{w,\hat{F}_n}(u)}{\hat{F}_n(u)} - \delta_2 \frac{\phi_{w,\hat{F}_n}(v) - \phi_{w,\hat{F}_n}(u)}{\hat{F}_n(v) - \hat{F}_n(u)} + \delta_3 \frac{\phi_{w,\hat{F}_n}(v)}{1 - \hat{F}_n(v)}.$$

Also define

$$Q_n(u, v, \delta_1, \delta_2) = \sum \Delta\hat{H}_n(u_i, v_i) \hat{F}_n^{\delta_{1i}}(u_i) \{\hat{F}_n(v_i) - \hat{F}_n(u_i)\}^{\delta_{2i}} \{1 - \hat{F}_n(v_i)\}^{1-\delta_{1i}-\delta_{2i}}$$

for $(\delta_1, \delta_2) = (0, 0), (0, 1), (1, 0)$, the empirical estimate of Q_0, where the summation is over $\{i, u_i \le u, v_i \le v, \delta_{1i} = \delta_1, \delta_{2i} = \delta_2\}$. Then it can be shown that $\|\tilde{\theta}_{w,\hat{F}_n}\|^2 = \int \tilde{\theta}^2_{w,\hat{F}_n}(u, v, \delta_1, \delta_2) dQ_n(u, v, \delta_1, \delta_2)$ is a consistent estimate of

$\|\tilde{\theta}_{w,F_0}\|^2$. That is, $\|\tilde{\theta}_{w,\hat{F}_n}\|^2 \to \|\tilde{\theta}_{w,F_0}\|^2$ as $n \to \infty$ (see the proof in the Appendix). Hence the hypothesis H_0 can be tested using the statistic $U_n/\|\tilde{\theta}_{w,\hat{F}_n}\|$ based on the standard normal distribution, when n is large.

In the above, it is assumed that the weight function $w(t)$ is given. A natural and simple choice is $w(t) = 1$. In this case, as noted earlier, $\sqrt{n/n_1n_2}U_n$ is an estimate of the mean survival difference between the two populations over the study period. Another common choice is the decreasing function $w(t) = 1/(1+t)$ or the increasing function $w(t) = 1 - 1/(1+t)$, depending on whether one wants to emphasize early or later survival differences. More comments and discussions on the selection of the weight function can be foundin Pepe and Fleming (1989) and Petroni and Wolfe (1994).

4. An Example

In this section, we apply our nonparametric test procedure to the breast cancer study described in Finkelstein (1986) and Klein and Moeschberger (1997). The objective of the study was to compare early breast cancer patients who had been treated with radiotherapy alone (Population 1, 46 patients) to those treated with primary radiation therapy and adjuvant chemotherapy (Population 2, 48 patients). The survival time of interest is the time until the appearance of breast retraction. In this study, interval-censored data were observed due to irregular observation times of the patients. These are given in Table 1.

Table 1. Observed Intervals for Appearances of Breast Retraction

Radiotherapy					Radiotherapy and Chemotherapy				
[46,]	[26, 37]	[38,]	[5, 11]	[18, 25]	[9, 12]	[1, 5]	[31, 34]	[17, 20]	[14,]
[7, 10]	[47,]	[1, 5]	[34,]	[16,]	[1, 22]	[6, 8]	[14,]	[31, 36]	[19, 25]
[1, 7]	[27, 40]	[19,]	[47,]	[20, 26]	[25, 31]	[13, 20]	[11, 17]	[17, 24]	[19, 24]
[47,]	[47,]	[25,]	[12, 15]	[12, 18]	[18, 27]	[12,]	[9, 21]	[18, 26]	[36,]
[47,]	[28, 34]	[37,]	[38,]	[23,]	[18, 23]	[34, 40]	[5, 9]	[17, 60]	[33,]
[8, 16]	[37, 44]	[6, 12]	[39,]	[35,]	[25, 30]	[32,]	[12,]	[16, 22]	[36, 39]
[18,]	[47,]	[20, 35]	[47,]	[6, 12]	[17, 24]	[14, 39]	[15, 19]	[24,]	[12, 17]
[8, 14]	[37, 48]	[18, 25]	[37,]	[47,]	[14,]	[20, 32]	[5, 8]	[22,]	[45, 48]
[38, 44]	[38,]	[25,]	[1, 8]	[41,]	[12, 13]	[35,]	[35,]	[23, 32]	[12, 20]
[33,]					[15, 17]	[11, 35]	[49,]		

To compare the two treatment groups using the proposed method, we calculated the statistic U_n and obtained $U_n = 42.7130$ with the estimated standard deviation $\|\tilde{\theta}_{w,\hat{F}_n}\| = 12.4062$ for $w(t) = 1$. This yielded a p-value of 0.0006. By taking $w(t) = 1 - 1/(1+t)$, we obtained $U_n = 41.5610$ with $\|\tilde{\theta}_{w,\hat{F}_n}\| = 11.9093$, which gives a p-value of 0.0005. The results suggest that the patients treated with radiotherapy alone had a lower breast retraction rate than those treated with chemotherapy and radiotherapy together. In other words, the adjuvant chemotherapy increased the breast retraction rate. Finkelstein (1986) also analyzed the data set and reported a p-value of 0.004 using the score test derived under the Cox proportional hazards model. The difference between the p-values indicates that the difference between the survival functions of the patients with the two treatments is more significant than their corresponding hazard difference.

To investigate the approximation of the normal distribution to U_n under H_0, a simulation study was conducted with the variance of U_n estimated using the sample variance of bootstrap samples. Figure 1 presents quantile plots of the standardized U_n obtained against the standard normal distribution. For the figure, we have $n = 100$, 50 in each group, $K = 1000$, and 1000 replications. Figure 1 indicates that the approximation is satisfactory. The details of the simulation study will be reported elsewhere.

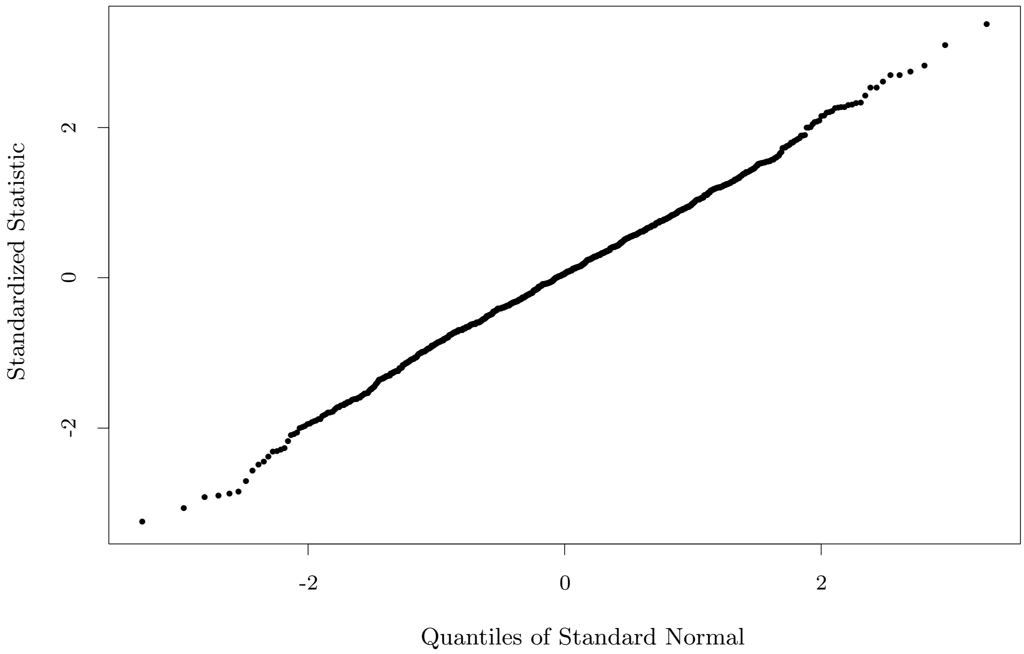

Figure 1. Quantile plot of the standardized test statistic.

5. Concluding Remarks

This paper considers the comparison of two continuous survival functions when interval-censored survival data are observed. The problem occurs in many survival studies such as clinical trials and longitudinal studies in which continuous observations on study subjects are not possible. To address the problem a class of test statistics, which are constructed based on integrated weighted survival difference, is proposed. For implementation a simple bootstrap procedure is suggested, although a consistent estimate of the asymptotic variance of the test statistic is derived. It has been our experience that the bootstrap procedure works well, while the asymptotic variance estimate has a very complex form. The test statistics are generalizations of the test statistics proposed by Pepe and Fleming (1989) for right-censored survival data.

It is worth noting that, as that given in Pepe and Fleming (1989), the test procedure proposed here is sensitive to stochastically ordered survival functions, often the situation in medical studies. If the difference between the two survival functions is nonproportional, however, our method may not be powerful enough to detect the difference. In this case, a different test procedure is preferred. This has been shown in an extensive simulation study, which will be reported elsewhere, conducted to study the finite sample properties of the test procedure and compare it with other test procedures for interval-censored data. The study suggests that the present method works well for stochastically ordered survival functions. In the simulation study, we considered the sample sizes of 100 or larger and in all situations, the normal approximation seems satisfactory. In terms of the number of bootstrap samples, it seems that $K = 1000$ works well and that larger values of K bring no significant differences.

Our focus has been on continuous survival time. If the time points at which the survival event can occur are finite, the method given in Petroni and Wolfe (1994) can be used. Then, unlike the continuous survival time considered here, the asymptotic variance of the test statistic U_n has a closed form and a simple consistent estimate of the asymptotic variance can be easily derived. If the hazard functions, instead of survival functions, are ordered and observed data are discrete interval-censored data, the test procedure presented in Sun (1996) can be used.

Acknowledgement

The authors wish to thank the Editor and two referees for their constructive comments and suggestions. Dr. Fang's research was supported in part by the Cancer Center Support Grant (CA21765) and ALSAC (American Lebanese Syrian Associated Charities). Dr. Sun's research was supported by a grant from the National Institutes of Health (NIH) and Dr. Lee's research was supported in part by NIH grants HL40619-09, EY12269-02 and GM55326-02.

Appendix. Proof of the theorem

Let $F_0(t)$, Q_0, $w(t)$, h_1, h_2, $h^*(t',t)$, $\tilde{\theta}_{w,F_0}$, $\|\tilde{\theta}_{w,F_0}\|^2$, $\hat{F}_n$, Q_n, $\tilde{\theta}_{w,\hat{F}_n}$, and $t_1 < \cdots < t_m$ be defined as in the previous sections. To prove the theorem, it is sufficient to prove that

$$\sqrt{n}\int_0^M w(t)[\hat{F}_n(t) - F_0(t)]dt \xrightarrow{\mathcal{D}} N(0, \|\tilde{\theta}_{w,F_0}\|^2). \tag{A.1}$$

First note that the left side of the above equation can be rewritten as

$$\sqrt{n}\int_0^M w(t)[\hat{F}_n(t) - F_0(t)]dt = \sqrt{n}\int \tilde{\theta}_{w,\hat{F}_n} dQ_0. \tag{A.2}$$

Define $\phi_{w,\hat{F}_n}(t)$ as the right-continuous solution to the equation

$$\phi_{w,\hat{F}_n}(t) = d_{\hat{F}_n}(t)\left\{w(t) - \int_0^M \frac{\phi_{w,\hat{F}_n}(t) - \phi_{w,\hat{F}_n}(t')}{|\hat{F}_n(t) - \hat{F}_n(t')|} h^*(t',t)dt'\right\},$$

which is unique when n is large according to Corollary 4.2 of Groeneboom (1996), where

$$d_{\hat{F}_n}(t) = \frac{\hat{F}_n(t)[1 - \hat{F}_n(t)]}{h_1(t)[1 - \hat{F}_n(t)] + h_2(t)\hat{F}_n(t)}.$$

Also define $\bar{\phi}_{w,\hat{F}_n}(t)$ as

$$\bar{\phi}_{w,\hat{F}_n}(t) = \begin{cases} \phi_{w,\hat{F}_n}(0), & t \in [0, t_1) \\ \phi_{w,\hat{F}_n}(t_i)\,, & t \in [t_i, t_{i+1}) \quad 1 \le i < m \\ \phi_{w,\hat{F}_n}(t_{i+1})\,, & t \in [t_{i+1}, M), \end{cases}$$

and let $\bar{\theta}_{w,\hat{F}_n}(u, v, \delta_1, \delta_2)$ be given by (3) with F_0 and ϕ_{w,F_0} replaced by $\hat{F}_n$ and $\bar{\phi}_{w,\hat{F}_n}$, respectively. Then we have

$$\int \bar{\theta}_{w,\hat{F}_n} dQ_n = 0. \tag{A.3}$$

Also, following the proof of Lemma 2.2 of Geskus and Groeneboom (1997), it can be shown that

$$\left|\int \left\{\bar{\theta}_{w,\hat{F}_n} - \tilde{\theta}_{w,\hat{F}_n}\right\} dQ_0\right| \le C\left\{\|\hat{F}_n - F_0\|_{H_1}^2 + \|\hat{F}_n - F_0\|_{H_2}^2\right\} = \mathcal{O}_p(n^{-2/3})\,, \tag{A.4}$$

where C is a constant.

It thus follows from (A.2), (A.3) and (A.4) that

$$\sqrt{n}\int_0^M w(t)[\hat{F}_n(t) - F_0(t)]dt = -\sqrt{n}\int \bar{\theta}_{w,\hat{F}_n} d(Q_n - Q_0) + \mathcal{O}_p(n^{-1/6}). \tag{A.5}$$

Note that we can rewrite the first term of the right hand of the above equation as

$$\sqrt{n}\int \bar{\theta}_{w,\hat{F}_n} d(Q_n - Q_0) = \sqrt{n}\int \tilde{\theta}_{w,F_0} d(Q_n - Q_0) + \sqrt{n}\int \left(\bar{\theta}_{w,\hat{F}_n} - \tilde{\theta}_{w,F_0}\right) d(Q_n - Q_0). \tag{A.6}$$

The first term of the right hand of (A.6) converges in distribution to a normal variable. According to Lemma 6.4 of Geskus and Groeneboom (1999), the second term of the right hand of (A.6) converges to zero in probability as $n \to \infty$. This together with equation (A.5) proves (A.1) and thus completes the proof.

References

DeGruttola, V. G. and Lagakos, S. W. (1989). Analysis of doubly-censored survival data, with application to AIDS. *Biometrics* **45**, 1-11.

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

Gentleman, R. and Geyer, C. J. (1994). Maximum likelihood for interval censored data: Consistency and computation. *Biometrika* **81**, 618-623.

Geskus, R. B. and Groeneboom, P. (1997). Asymptotically optimal estimation of smooth functionals for interval censoring, Part 2. *Statist. Neerlandica* **51**, 201-219.

Geskus, R. B. and Groeneboom, P. (1999). Asymptotically optimal estimation of smooth functionals for interval censoring, Case 2. *Ann. Statist.* **27**, 627-674.

Gill, R. D. (1980). *Censoring and Stochastic Integrals.* Mathematical Centre Tracts 124. Mathematical Centrum, Amsterdam.

Groeneboom, P. (1996). Lectures on inverse problems. In *Lecture Notes in Mathe.*, 1648. Springer-Verlag, Berlin.

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

Keiding, N., Begtrup, K., Scheike, T. H. and Hasibeder, G. (1996). Estimation from current status data in continuous time. *Lifetime Data Anal.* **2**, 119-129.

Klein, J. P. and Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data.* Springer, New York.

Li, L., Watkins, T. and Yu, Q. (1997). An EM Algorithm for Smoothing the Self-consistent Estimator of Survival Functions with Interval-censored Data. *Scand. J. Statist.* **24**, 531 - 542.

Pepe, M. S. and Fleming, T. R. (1989). Weighted Kaplan-Meier statistics: A class of distance tests for censored survival data. *Biometrics* **45**, 497-507.

Petroni, G. R. and Wolfe, R. A. (1994). A two-sample test for stochastic ordering with interval-censored data. *Biometrics* **50**, 77-87.

Sun, J. (1996). A non-parametric test for interval-censored failure time data with application to AIDS studies. *Statist. Medicine* **15**, 1387-1395.

Sun, J. (1998). *Interval Censoring.* Encyclopedia of Biostatistics 2090-2095. John Wiley, New York.

Sun, J. and Kalbfleisch, J. D. (1996). Nonparametric tests of tumor prevalence data. *Biometrics* **52**, 726-731.

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 with arbitrarily grouped censored and truncated data. *J. Roy. Statist. Soc. Ser. B* **38**, 290-295.

Yu, Q., Li, L. and Wong, G. Y. C. (2000). On consistency of the self-consistent estimator of survival functions with interval-censored data. *Scand. J. Statist.* **27**, 35-44.

Department of Biostatistics, St. Jude Children's Research Hospital, 332 North Lauderdale Street, Memphis, TN 38105, U.S.A.

E-mail: hfang@umm.edu

Department of Statistics, University of Missouri-Columbia, Columbia, MO 65211, U.S.A.

E-mail: tsun@stat.missouri.edu

Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical school, 181 Longwood Avenue, Boston, MA 02115, U.S.A.

E-mail: stmei@channing.harvard.edu

(Received February 2001; accepted March 2002)