Harvard University
Harvard University Biostatistics Working Paper Series

Year 2004 *Paper* 12

One- and Two-Sample Nonparametric Inference Procedures in the Presence of Dependent Censoring

Yuhyun Park* Lu Tian†
L. J. Wei‡

*Harvard University, ypark@hsph.harvard.edu
†Harvard University, ltian@hsph.harvard.edu
‡Harvard University, wei@hsph.harvard.edu

ONE- AND TWO-SAMPLE NONPARAMETRIC INFERENCE PROCEDURES IN THE PRESENCE OF DEPENDENT CENSORING

by Y. Park, L. Tian and L.J. Wei

Department of Biostatistics

Harvard University

677 Huntington Ave.

Boston, MA 02115

SUMMARY

In survival analysis, the event time T is often subject to dependent censorship. Without assuming a parametric model between the failure and censoring times, the parameter Θ of interest, for example, the survival function of T, is generally not identifiable. On the other hand, the collection Ω of all attainable values for Θ may be well-defined. In this article, we present non-parametric inference procedures for Ω in the presence of a mixture of dependent and independent censoring variables. By varying the criteria of classifying censoring to the dependent or independent category, our proposals can be quite useful for the so-called sensitivity analysis of censored failure times. The case that the failure time is subject to dependent interval censorship is also discussed in this article. The new proposals are illustrated with data from two clinical studies on HIV-related diseases.

Keywords: Competing risks; Martingale; Simultaneous confidence interval; Sensitivity analysis; Survival analysis.

1. INTRODUCTION

In survival analysis, the time to the event of interest is often subject to dependent right censorship. For example, in a recent double-blind clinical trial, ACTG 175, conducted by the AIDS Clinical Trials Group, two thousand four hundred sixty-seven patients were randomly assigned to one of four daily regimens (Hammer et al., 1996). The primary end point was the time T from randomization to one of the following events, a ≥ 50 percent decline in the CD4 cell count, development of the AIDS, and death. One thousand eight hundred ninety-nine event times were censored. Although the majority of these event times were censored administratively, six hundred fifty-six patients were off the treatments during the study due to, for example, toxicity or request from the patient or investigator, which was likely related to the primary endpoint. As indicated by Tsiatis (1975) and Peterson (1976) for handling the general dependent competing risks problem, serious errors can be made in estimating the survival function of the primary end point for a study such as ACTG 175 using the standard inference procedures in survival analysis.

With various parametric assumptions on the dependence structure between the event and censoring times, novel inference procedures and sensitivity analyses were proposed, for example, by Fisher & Kanarek (1974), Slud & Rubinstein (1983), Klein & Moeschberger (1988), Klein et al. (1992), Moeschberger & Klein (1995), Zheng & Klein (1995), Lin et al. (1996), DiRienzo & Lagakos (2001), and DiRienzo (2003). When auxiliary variables are available, innovative research has been done, for example, by Robins & Rotnitzky (1992), Robins (1993), Robins & Finkelstein (2000), Satten et al. (2001), and Scharfstein & Robins (2002).

In this article, we consider the case that the failure time T may be censored by either a dependent or an independent censoring variable under a purely nonparametric setting. Although the parameter Θ of interest for this case may not be identifiable, the collection Ω of all possible values of Θ is often well-defined. For example, if Θ is the survival function of T, Ω is the collection of non-increasing functions which are bounded by Peterson's bounds (Peterson, 1976). Here, we propose inference procedures for Ω under various

one- and two-sample settings. Specifically, we present a consistent estimate $\hat{\Omega}$ and a $(1-\alpha)$ confidence set $\hat{\Omega}_U$ for Ω such that $\text{pr}(\Omega \subset \hat{\Omega}_U) = 1-\alpha$, where $0 < \alpha < 1$. Such confidence *interval* estimation provides more informative than the single *point* estimation. Moreover, by varying the criteria of classifying censoring to the dependent or independent category, our proposal can be quite useful for sensitivity analysis of censored failure time observations. The new proposals are illustrated with the data from the aforementioned ACTG 175 study.

Lastly, we discuss the case that the failure time is subject to dependent interval censorship and present certain one- and two-sample inference procedures. The procedures are illustrated with the data from a well-known study on the HIV-1 infection incidence among hemophilia patients.

2. INFERENCES WITH RIGHT CENSORED OBSERVATIONS

2.1. ONE SAMPLE PROBLEMS

Let T be the continuous failure time of interest, D be the continuous, dependent censoring variable, and C be the independent censoring variable. Also, let $\{(T_i, D_i, C_i), i = 1, \cdots, n\}$ be n independent copies of (T, D, C). For the ith subject, one can only observe (X_i, η_i), where $X_i = \min(T_i, D_i, C_i)$, and

$$\eta_i = \begin{cases} 0 \text{ if } X_i = C_i, \\ 1 \text{ if } X_i = T_i, \\ 2 \text{ if } X_i = D_i. \end{cases}$$

First, suppose that we are interested in making inferences about the survival function $S(t)$ of T. In the presence of censoring, generally $S(t)$ cannot be estimated well nonparametrically for small or large t, here we let the parameter Θ be the function $S(\cdot)$ defined in a pre-determined, finite interval $\mathcal{I} = [\tau_1, \tau_2]$, where τ_1 and τ_2 are known constants such that $\text{pr}(X \le \tau_1, T < D \wedge C) > 0$ and $\text{pr}(X \ge \tau_2) > 0$. Without assuming a parametric dependence structure between T and D, $S(\cdot)$ is not identifiable. On the other hand, the

set Ω of all attainable values of $\{S(t), t \in \mathcal{I}\}$ is the collection of non-increasing functions which are bounded below by $S_L(t)$ and above by $S_U(t)$, where $S_L(t) = \text{pr}(T \wedge D \geq t)$ and

$$1 - S_U(t) = \text{pr}(T \leq t, T < D), \tag{2.1}$$

$t \in \mathcal{I}$ (Peterson, 1976). In the competing risks literature, the right hand side of (2.1) is the so-called cumulative incidence function.

Note that

$$S_U(t) = 1 - \int_0^t \exp(-\Lambda_T(s) - \Lambda_D(s)) d\Lambda_T(s),$$

and (2.2)

$$S_L(t) = \exp(-\Lambda_T(t) - \Lambda_D(t)),$$

where $\Lambda_T(s) = \int_0^s \lambda_T(u)du$, $\Lambda_D(s) = \int_0^s \lambda_D(u)du$, and $\lambda_T(t)$ and $\lambda_D(t)$ are the cause-specific hazard functions with respect to T and D, respectively (Aalen, 1978; Kalbfleisch & Prentice, 2002, p.251). To obtain a consistent estimate $\hat{\Omega}$ for Ω, one needs to estimate $S_L(t)$ and $S_U(t)$. To this end, let

$$\hat{\Lambda}_m(t) = \sum_{i=1}^n \int_0^t \frac{dI(X_i \leq s, \eta_i = I(m = T) + 2 \times I(m = D))}{\sum_{j=1}^n I(X_j \geq s)},$$

where $I(\cdot)$ is the indicator function, and $m = T, D$. Then, the Aalen-Nelson estimator $\hat{\Lambda}_m(t)$ is a consistent estimator for $\Lambda_m(t)$, and a consistent estimator $\hat{\Omega}$ for Ω is a set of non-increasing functions which are bounded above by $\hat{S}_U(t)$ and below by $\hat{S}_L(t)$, where $\hat{S}_U(t)$ and $\hat{S}_L(t)$ are obtained by replacing $\Lambda_m(t)$ in (2.2) with $\hat{\Lambda}_m(t), m = T, D$.

To obtain a $(1 - \alpha)$ confidence set $\hat{\Omega}_U$ of Ω, one needs the joint distribution of the process $\{(\hat{S}_L(s), \hat{S}_U(t)), s, t \in \mathcal{I}\}$. To this end, note that

$$\hat{\Lambda}_m(t) - \Lambda_m(t) = \sum_{i=1}^n \int_0^t \frac{dM_{mi}(s)}{\sum_{j=1}^n I(X_j \geq s)},$$

where $M_{mi}(s) = I(X_i \leq s, \eta_i = I(m = T) + 2 \times I(m = D)) - \int_0^s I(X_i \geq u) d\Lambda_m(u), m = T, D$. Since $M_{Ti}(\cdot)$ and $M_{Di}(\cdot)$ are orthogonal martingales (Fleming & Harrington, 1991, p.42), it follows that the processes $n^{1/2}(\hat{\Lambda}_T(s) - \Lambda_T(s))$ and $n^{1/2}(\hat{\Lambda}_D(t) - \Lambda_D(t))$ converge

jointly to a two-dimensional Gaussian process, for $s, t \in \mathcal{I}$, as $n \to \infty$. To relax the constraint that the cumulative hazard function is non-negative, one usually re-parametrizes this function by considering its log-transformation. By the functional δ-method, for large n, the distribution of the process, indexed by (s, t),

$$\left\{ \begin{pmatrix} \log \hat{\Lambda}_T(s) \\ \log \hat{\Lambda}_D(t) \end{pmatrix} - \begin{pmatrix} \log \Lambda_T(s) \\ \log \Lambda_D(t) \end{pmatrix} \right\}$$

can be well approximated by that of the process

$$\sum_{i=1}^{n} \begin{pmatrix} \hat{\Lambda}_T^{-1}(s) \int_0^s \frac{dM_{Ti}(u)}{\sum_j I(X_j \geq u)} \\ \hat{\Lambda}_D^{-1}(t) \int_0^t \frac{dM_{Di}(u)}{\sum_j I(X_j \geq u)} \end{pmatrix}. \tag{2.3}$$

Generally the distribution of a function of (2.3) may be rather difficult to obtain analytically. On the other hand, we may approximate the distribution of (2.3) utilizing a simple perturbation technique proposed by Lin, Wei & Ying (1993). To this end, let $\{G_1, \cdots, G_n\}$ be a random sample from the standard normal, which is independent of the data. Consider a process which is obtained by replacing $M_{mi}(t)$ in (2.3) by $G_i \times I(X_i \leq t, \eta_i = I(m = T) + 2 \times I(m = D)), m = T, D$. Then, for large n, conditional on the data, the distribution of the resulting process

$$\sum_{i=1}^{n} G_i \begin{pmatrix} \hat{\Lambda}_T^{-1}(s) \int_0^s \frac{dI(X_i \leq u, \eta_i = 1)}{\sum_j I(X_j \geq u)} \\ \hat{\Lambda}_D^{-1}(t) \int_0^t \frac{dI(X_i \leq u, \eta_i = 2)}{\sum_j I(X_j \geq u)} \end{pmatrix} \tag{2.4}$$

gives a good approximate to the unconditional distribution of (2.3). Note that the only random quantities in (2.4) are $G_i, i = 1, \cdots, n$. To obtain an approximation to the distribution of (2.4), one may generate a large number, say, N, of independent random samples $\{G_i, i = 1, \cdots, n\}$ to obtain N realizations of (2.4).

For convenience, define two random processes $\{(\Lambda_T^*(s), \Lambda_D^*(t))', s, t \in \mathcal{I}\}$ such that

$$\left\{ \begin{pmatrix} \log \Lambda_T^*(s) \\ \log \Lambda_D^*(t) \end{pmatrix} - \begin{pmatrix} \log \hat{\Lambda}_T(s) \\ \log \hat{\Lambda}_D(t) \end{pmatrix} \right\} = (2.4).$$

It follows that the distribution of the process $n^{1/2}(\hat{W}_L(s), \hat{W}_U(t))' = n^{1/2}\big(\log(-\log \hat{S}_L(s)) - \log(-\log S_L(s)), \log(-\log \hat{S}_U(t)) - \log(-\log S_U(t))\big)'$ is asymptotically Gaussian and it

can be approximated by the conditional distribution of the process $n^{1/2}(W_L^*(s), W_U^*(t))' = n^{1/2}\big(\log(-\log S_L^*(s)) - \log(-\log \hat{S}_L(s)), \log(-\log S_U^*(t)) - \log(-\log \hat{S}_U(t))\big)'$, $s, t \in \mathcal{I}$, where $S_L^*(t)$ and $S_U^*(t)$ are obtained by replacing $\Lambda_m(t)$ in (2.2) with $\Lambda_m^*(t), m = T, D$. Let $\sigma_L(s)$ and $\sigma_U(t)$ be the estimated standard errors for $W_L^*(s)$ and $W_U^*(t)$, respectively. These two standard errors can be obtained via the sample variances based on the above N realizations of $(W_L^*(\cdot), W_U^*(\cdot))'$.

A $(1-\alpha)$ confidence set $\hat{\Omega}_U$ for Ω is the collection of non-increasing functions $S(\cdot)$ which satisfy

$$\log(-\log \hat{S}_U(t)) - c\sigma_U(t) < \log(-\log S(t)) < \log(-\log \hat{S}_L(t)) + c\sigma_L(t), \tag{2.5}$$

where $t \in \mathcal{I}$ and c is chosen to satisfy

$$\text{pr}(\inf_{t\in\mathcal{I}} \frac{W_L^*(t)}{\sigma_L(t)} > -c, \quad \sup_{t\in\mathcal{I}} \frac{W_U^*(t)}{\sigma_U(t)} < c \mid \text{data}) \approx 1-\alpha. \tag{2.6}$$

Note that the probability measure (2.6) is generated by $\{G_i, i = 1, \cdots, n\}$, but conditional on the data.

Now, we use the data from ACTG 175 study to illustrate the above inference procedures. Although there were four treatment groups in the study, for illustration, we only compare the AZT monotherapy with the other three treatments combined. Six hundred nineteen out of 2467 patients were randomly assigned to the AZT monotherapy. There were 423 and 1479 failure times were censored in the AZT and combined groups, respectively. Here, we take a rather conservative way to define independent censoring. That is, we let D be the time that the patient was off-treatment due to toxicity or the request of the investigator or patient. There were 155 and 501 such dependent censored events for the AZT and combined groups, respectively. In Figure 1, we present point estimate $\hat{\Omega}$ and 0.95 confidence set $\hat{\Omega}_U$ based on (2.5) and (2.6) for the combined and the AZT groups. Here, for each group, τ_1 and τ_2 are chosen such that they cut off approximately the lower and upper 5% of the observed failure times, respectively. For the AZT group, $\tau_1 = 140$ (days) and $\tau_2 = 950$ (days), and for the combined group, $\tau_1 = 170$ and $\tau_2 = 995$. The standard error estimates $\sigma_L(t)$ and $\sigma_U(t)$ and the cutoff point c is obtained with $N = 1000$

realizations of $S_U^*(t)$ and $S_L^*(t)$. The collection of non-increasing functions, whose upper and lower bounds are denoted by the solid lines, is the point estimate $\hat{\Omega}$, and the region bounded by the dotted lines is the 0.95 confidence set $\hat{\Omega}_U$.

Now, suppose that we are interested in making inferences about the quantile process of the survival function, for example, the median or upper and lower quartiles of T. To this end, let t_p be the pth quantile of the survival function $S(\cdot)$, that is, $1 - S(t_p) = p$. Here, the parameter Θ is a function t_p of $p \in \mathcal{J} = [p_1, p_2]$, a pre-determined interval such that $[t_{p_1}, t_{p_2}] = \mathcal{I}$. Let t_{lp} and t_{up} be the "pth" quantiles for $S_L(\cdot)$ and $S_U(\cdot)$, respectively. Then, the set Ω of all possible values of Θ consists of non-decreasing functions $t_p, p \in \mathcal{J}$, which are bounded below by the function t_{lp} and above by t_{up}. A consistent estimator $\hat{\Omega}$ can be obtained easily via estimators $\hat{t}_{lp}$ and $\hat{t}_{up}$ for t_{lp} and t_{up} by solving the equations: $\hat{S}_L(t) = 1 - p$ and $\hat{S}_U(t) = 1 - p$.

One may use the aforementioned perturbation technique to obtain a $(1-\alpha)$ confidence set $\hat{\Omega}_U$ for Ω. Since the process $W_L^*(\cdot)$ and $W_U^*(\cdot)$ are tight, it follows that the asymptotic distribution of the process $n^{1/2}(\hat{W}_L(t_{lp}), \hat{W}_U(t_{ur}))'$, indexed by (p, r), is the same as the conditional distribution of $n^{1/2}(W_L^*(\hat{t}_{lp}), W_U^*(\hat{t}_{ur}))'$. Conditional on the data, let t_{lp}^* and t_{up}^* be the random variables such that

$$\hat{W}_L(t_{lp}^*) = W_L^*(\hat{t}_{lp}), \quad \hat{W}_U(t_{up}^*) = W_U^*(\hat{t}_{up}).$$

Then, using the results from Goldwasser et al. (2004), for large n, the distribution of the process $(\log \hat{t}_{lp} - \log t_{lp}, \log \hat{t}_{ur} - \log t_{ur})'$, indexed by (p, r), can be approximated well by that of the process $(\log t_{lp}^* - \log \hat{t}_{lp}, \log t_{ur}^* - \log \hat{t}_{ur})'$, where $p, r \in \mathcal{J}$. Let ϕ_{lp} and ϕ_{up} be the estimated standard errors of $\log t_{lp}^*$ and $\log t_{up}^*$, respectively. Then, $\hat{\Omega}_U$ consists of all non-decreasing functions t_p such that

$$\log \hat{t}_{lp} - c\phi_{lp} \le \log t_p \le \log \hat{t}_{up} + c\phi_{up}, \tag{2.7}$$

$p \in \mathcal{J}$. Here, c is chosen to satisfy

$$\text{pr}(\inf_{p \in \mathcal{J}} \frac{\log t_{up}^* - \log \hat{t}_{up}}{\phi_{up}} > -c, \quad \sup_{p \in \mathcal{J}} \frac{\log t_{lp}^* - \log \hat{t}_{lp}}{\phi_{lp}} < c \mid \text{data}) \approx 1 - \alpha. \tag{2.8}$$

To illustrate the above estimation procedure, we again use the data from Study ACTG 175. In Figure 2, we present the point estimate $\hat{\Omega}$ and a 0.95 *interval* estimate $\hat{\Omega}_U$ based on (2.7) and (2.8) with $\mathcal{J} = [0.04, 0.32]$ for the AZT group, and $= [0.03, 0.21]$ for the combined group. Here, $\hat{\Omega}$ is the region bounded by the solid lines, and $\hat{\Omega}_U$ is bounded by the dotted lines.

2.2. TWO SAMPLE PROBLEMS

In this section, we present non-parametric and semi-parametric inference procedures for various parameters which quantify the relative merit between two independent groups of failure times in the presence of dependent censoring. To this end, all the aforementioned theoretical and empirical quantities in Section 2.1 are sub-indexed by their group membership $k, k = 1, 2$. For example, the data now consist of $\{(X_{ki}, \eta_{ki}), i = 1, \cdots, n_k; k = 1, 2\}$.

First, suppose that we are interested in $\Theta = \{S_2(t) - S_1(t), t \in \mathcal{L}\}$, the difference of two underlying survival functions, where $\mathcal{L}$ is a pre-determined interval $[\tau_1, \tau_2]$ such that $\mathrm{pr}(X_{k1} \leq \tau_1, T_{k1} < D_{k1} \wedge C_{k1}, k = 1, 2) > 0$, and $\mathrm{pr}(X_{k1} \geq \tau_2, k = 1, 2) > 0$. Note that Ω consists of functions $S_2(\cdot) - S_1(\cdot)$, which satisfy

$$S_{2L}(t) - S_{1U}(t) \leq S_2(t) - S_1(t) \leq S_{2U}(t) - S_{1L}(t), \tag{2.9}$$

$t \in \mathcal{L}$. A consistent estimator $\hat{\Omega}$ of Ω can be obtained by replacing $S_{kL}(t)$ and $S_{kU}(t)$ in the lower and upper bounds of (2.9) with their empirical counterparts, $k = 1, 2$. A $(1 - \alpha)$ confidence set $\hat{\Omega}_U$ is the collection of functions of t, which belong to the intervals

$$(\hat{S}_{2L}(t) - \hat{S}_{1U}(t) - c(\xi^2_{2L}(t) + \xi^2_{1U}(t))^{1/2},\ \hat{S}_{2U}(t) - \hat{S}_{1L}(t) + c(\xi^2_{2U}(t) + \xi^2_{1L}(t))^{1/2}),$$

where $\xi_{kL}(t)$ and $\xi_{kU}(t)$ are the estimated standard errors of $S^*_{kL}(t)$ and $S^*_{kU}(t),\ k = 1, 2$, and c is chosen such that

$$\mathrm{pr}(\inf_{t \in \mathcal{L}} \frac{S^*_{2U}(t) - S^*_{1L}(t) - \hat{S}_{2U}(t) + \hat{S}_{1L}(t)}{(\xi^2_{2U}(t) + \xi^2_{1L}(t))^{1/2}} > -c,$$

$$\sup_{t \in \mathcal{L}} \frac{S^*_{2L}(t) - S^*_{1U}(t) - \hat{S}_{2L}(t) + \hat{S}_{1U}(t)}{(\xi^2_{2L}(t) + \xi^2_{1U}(t))^{1/2}} < c \ |\text{data}) \approx 1 - \alpha.$$

Again, an approximation to the above cutoff point c can be obtained via the perturbation technique discussed in Section 2.1.

For Study ACTG 175, we let $S_2(t)$ and $S_1(t)$ be the survival functions for the combined and AZT groups, respectively. First, we assume that the dependent censoring is due to toxicity or the request from the patient or investigator. In Figure 3(a), we present a point estimate $\hat{\Omega}$ and a 0.95 interval estimate $\hat{\Omega}_U$ for Ω with $\mathcal{L} = [170, 950]$. With $N = 1000$ sets of realizations from $\{S^*_{kL}(\cdot), S^*_{kU}(\cdot), k = 1, 2\}$, $\hat{\Omega}$ is composed of functions bounded by the solid lines, and $\hat{\Omega}_U$ is the set of functions bounded by the dotted lines. In Figure 3(b), we present a similar plot, but assume that the dependent censoring event is only due to the request from the patient or investigator. There are 88 and 288 such events in the AZT and combined groups, respectively. Lastly, we assume that all the censoring variables are independent of T, and in Figure 3(c) we provide the Kaplan-Meier estimate denoted by the solid line, and a 0.95 confidence set $\hat{\Omega}_U$ whose boundaries are the dotted lines. These plots provide valuable information regarding sensitivity of the censoring assumptions.

Now, suppose that there exists an unknown constant Θ such that

$$\log S_2(t) = e^{\Theta} \log S_1(t), \tag{2.10}$$

a two-sample proportional hazards model. We are interested in making inferences about Θ. Note that for $t \in \mathcal{L}$,

$$\Theta_L(t) = \log(-\log S_{2U}(t)) - \log(-\log S_{1L}(t)) \leq \Theta \leq$$

$$\log(-\log S_{2L}(t)) - \log(-\log S_{1U}(t)) = \Theta_U(t).$$

Let $\Theta_L = \sup_{t\in\mathcal{L}} \Theta_L(t)$ and $\Theta_U = \inf_{t\in\mathcal{L}} \Theta_U(t)$. It is not difficult to show that any member of the interval $\Omega = [\Theta_L, \Theta_U]$ is an attainable value for Θ in Model (2.10). Let $\hat{\Theta}_L(t)$ and $\hat{\Theta}_U(t)$ be the estimators obtained by replacing $S(t)$ with $\hat{S}(t)$ in $\Theta_L(t)$ and $\Theta_U(t)$, respectively. Similarly, $\Theta^*_L(t)$ and $\Theta^*_U(t)$ are obtained with $S(t)$ replaced by $S^*(t)$. A consistent point estimator for Ω is $\hat{\Omega} = [\hat{\Theta}_L, \hat{\Theta}_U]$, where $\hat{\Theta}_L = \sup_{t\in\mathcal{L}} \hat{\Theta}_L(t)$ and $\hat{\Theta}_U = \inf_{t\in\mathcal{L}} \hat{\Theta}_U(t)$.

To derive a $(1-\alpha)$ confidence set $\hat{\Omega}_U$, unfortunately, it is rather difficult, if not impossible, to obtain the joint distribution of $\hat{\Theta}_L$ and $\hat{\Theta}_U$ analytically or numerically. Now, consider the following class of interval estimates for Ω, indexed by time $t \in \mathcal{L}$,

$$(\hat{\Theta}_L(t) - c\psi_L(t), \quad \hat{\Theta}_U(t) + c\psi_U(t)), \tag{2.11}$$

where $\psi_L(t)$ and $\psi_U(t)$ are the estimated standard errors for $\Theta_L^*(t)$ and $\Theta_U^*(t)$. Note that for any *pre-determined* $t \in \mathcal{L}$, an interval (2.11) with $c \approx 1.96$ is a valid 0.95 confidence set for Ω. However, such an interval for Ω can be quite large. To obtain a robust interval estimate, first, we let the cutoff point c in (2.11) be chosen such that

$$\text{pr}(\inf_{t\in\mathcal{L}} \frac{\Theta_U^*(t) - \hat{\Theta}_U(t)}{\psi_U(t)} > -c, \quad \sup_{t\in\mathcal{L}} \frac{\Theta_L^*(t) - \hat{\Theta}_L(t)}{\psi_L(t)} < c \,|\text{data}) \approx 1-\alpha. \tag{2.12}$$

With this relatively larger threshold value than 1.96, the set of intervals (2.11) is a $(1-\alpha)$ simultaneous confidence band for Ω across $t \in \mathcal{L}$. Thus, the "narrowest" interval from this band is a valid $(1-\alpha)$ confidence set for Ω. For example, a possible choice for $\hat{\Omega}_U$ is the interval

$$(\sup_{t\in\mathcal{L}}\{\hat{\Theta}_L(t) - c\psi_L(t)\}, \quad \inf_{t\in\mathcal{L}}\{\hat{\Theta}_U(t) + c\psi_U(t)\}). \tag{2.13}$$

For Study ACTG 175, first, let us assume that the dependent censoring event is due to toxicity or the request from the patient or investigator. For this case, with $\mathcal{L} = [170, 950]$ the cutoff point c based on (2.12) is 2.48. In Figure 4, we present a 0.95 simultaneous confidence band (2.11) with $c = 2.48$. The minimizer for $\{\hat{\Theta}_U(t) + 2.48\psi_U(t)\}$ is $t = 844$, and the maximizer for $\{\hat{\Theta}_L(t) - 2.48\psi_L(t)\}$ is $t = 812$. It follows from (2.13) that a 0.95 confidence interval for Θ is $(-1.21, 0.03)$, indicating that even without assuming a parametric model between the failure and dependent censoring times, patients in the combined group were doing better than those in the AZT group. It is interesting to note that for any predetermined $t \in [600, 800]$, the corresponding 0.95 confidence interval $(\hat{\Theta}_L(t) - 1.96\psi_L(t), \ \hat{\Theta}_U(t) + 1.96\psi_U(t))$ is almost identical to our interval $(-1.21, 0.03)$. On the other hand, if one chooses $t < 500$, the resulting interval for Θ is quite wide.

For example, when $t = 200$, the pointwise interval is $(-2.21, 0.50)$, which is much larger than ours. Lastly, if one assumes that all censorings are independent of T, the maximum partial likelihood estimate for Θ is -0.60 and the corresponding 0.95 interval for $\Theta = \Omega$ is $(-0.78, -0.43)$.

Now, suppose that there exists an unknown Θ such that $S_2(e^{\Theta}t) = S_1(t), t > 0$, the so-called two-sample accelerated failure time model (Kalbfleisch & Prentice, 2002, p. 217-46). We are interested in making inferences about Θ. Note that

$$\Theta_L(p) = \log t_{2lp} - \log t_{1up} \leq \Theta \leq \log t_{2up} - \log t_{1lp} = \Theta_U(p),$$

for $p \in \mathcal{M} = [p_1, p_2]$, where t_{p_1} and t_{p_2} are the lower and upper boundaries of $\mathcal{L}$. Let $\Theta_L = \sup_{p\in\mathcal{M}} \Theta_L(p)$ and $\Theta_U = \inf_{p\in\mathcal{M}} \Theta_U(p)$. Then, $\Omega = [\Theta_L, \Theta_U]$. Let $\hat{\Theta}_L(p) = \log \hat{t}_{2lp} - \log \hat{t}_{1up}$ and $\hat{\Theta}_U(p) = \log \hat{t}_{2up} - \log \hat{t}_{1lp}$. Also, let $\Theta^*_L(p) = \log t^*_{2lp} - \log t^*_{1up}$ and $\Theta^*_U(p) = \log t^*_{2up} - \log t^*_{1lp}$. The point estimate $\hat{\Omega} = [\hat{\Theta}_L, \hat{\Theta}_U]$, where $\hat{\Theta}_L$ and $\hat{\Theta}_L$ are the empirical counterparts of Θ_L and Θ_U, respectively. Moreover, it follows from a similar argument in Section 2.1 that the distribution of $(\hat{\Theta}_L(p) - \Theta_L(p),\ \hat{\Theta}_U(r) - \Theta_U(r))'$ can be approximated well by that of $(\Theta^*_L(p) - \hat{\Theta}_L(p),\ \Theta^*_U(r) - \hat{\Theta}_U(r))'$, where $p, r \in \mathcal{M}$. Similar to the case of the proportional hazards model discussed above, a $(1-\alpha)$ confidence interval $\hat{\Omega}_U$ of Ω is

$$(\sup_{p\in\mathcal{M}}\{\hat{\Theta}_L(p) - c(\phi^2_{1up} + \phi^2_{2lp})^{1/2}\},\quad \inf_{p\in\mathcal{M}}\{\hat{\Theta}_U(p) + c(\phi^2_{2up} + \phi^2_{1lp})^{1/2}\}), \tag{2.14}$$

where the cutoff point c is chosen such that

$$\text{pr}(\inf_{p\in\mathcal{M}} \frac{\Theta^*_U(p) - \hat{\Theta}_U(p)}{(\phi^2_{2up} + \phi^2_{1lp})^{1/2}} > -c, \sup_{p\in\mathcal{M}} \frac{\Theta^*_L(p) - \hat{\Theta}_L(p)}{(\phi^2_{1up} + \phi^2_{2lp})^{1/2}} < c \mid \text{data}) \approx 1 - \alpha.$$

To illustrate the above estimation procedures with Study ACTG 175, we only consider the most conservative case that the dependent censoring is due to toxicity or the request from the patient or investigator. With $\mathcal{M} = [0.04, 0.21]$, a 0.95 confidence interval (2.14) for Θ is $(-0.20, 1.07)$.

3. INFERENCES WITH INTERVAL CENSORED DATA

Suppose that for each T_i, one cannot observe T_i directly, but only observe an interval (E_{Li}, E_{Ui}) which contains $T_i, i = 1, \cdots, n$. When E_{Li} and E_{Ui} are *independent* of T_i, nonparametric estimation procedures for $S(t)$ were proposed, for example, by Peto (1973), Turnbull (1976) and Gentleman & Geyer (1994). Regression methods have been studied, for example, by Bacchetti (1990), Rabinowitz et al. (1995), Rosenberg (1995), Huang (1996, 1999), Huang & Wellner (1997), Kooperberg & Clarkson (1997), Joly et al. (1998), Betensky et al. (2001, 2002), and Cai & Betensky (2003).

Unlike the case with the dependent right censorship, for the interval censored data, even if one can identify which interval censorings are informative and which are not, it is not clear how to utilize this valuable information to obtain *sharp* theoretical bounds such as the Peterson bounds for $S(t)$. Here, we propose inference procedures which are valid even when all interval censorings are informative. To this end, let $S_L(t)$ and $S_U(t)$ be the survival functions of E_{Li} and E_{Ui}, respectively. The parameter Θ is $\{S(t), t \in \mathcal{N}\}$, where $\mathcal{N}$ is the pre-determined interval $[\tau_1, \tau_2]$ such that $\mathrm{pr}(E_{Ui} \leq \tau_1) > 0$, and $\mathrm{pr}(E_{Li} \geq \tau_2) > 0$. The Ω consists of non-increasing functions $S(t)$ such that $S_L(t) \leq S(t) \leq S_U(t), t \in \mathcal{N}$. The $S_L(t)$ and $S_U(t)$ can be estimated well by $\hat{S}_L(t) = n^{-1} \sum_{i=1}^n I(E_{Li} \geq t)$ and $\hat{S}_U(t) = n^{-1} \sum_{i=1}^n I(E_{Ui} \geq t)$, and a consistent estimator $\hat{\Omega}$ for Ω can be obtained accordingly.

To obtain a $(1-\alpha)$ confidence set $\hat{\Omega}_U$ of Ω, note that for large n, the distribution of the process

$$\begin{pmatrix} \hat{W}_L(s) \\ \hat{W}_U(t) \end{pmatrix} = \begin{pmatrix} \log(-\log \hat{S}_L(s)) - \log(-\log S_L(s)) \\ \log(-\log \hat{S}_U(t)) - \log(-\log S_U(t)) \end{pmatrix}$$

can be approximated well by the conditional distribution of

$$\begin{pmatrix} W_L^*(s) \\ W_U^*(t) \end{pmatrix} = \sum_{i=1}^n G_i \begin{pmatrix} \{\log(\hat{S}_L(s))\hat{S}_L(s)\}^{-1}(I(E_{Li} \geq s) - \hat{S}_L(s)) \\ \{\log(\hat{S}_U(t))\hat{S}_U(t)\}^{-1}(I(E_{Ui} \geq t) - \hat{S}_U(t)) \end{pmatrix}, \tag{3.1}$$

where $s, t \in \mathcal{N}$. Now, let $S_L^*(s)$ and $S_U^*(t)$ be the random processes such that

$$\begin{pmatrix} \log(-\log S_L^*(s)) - \log(-\log \hat{S}_L(s)) \\ \log(-\log S_U^*(t)) - \log(-\log \hat{S}_U(t)) \end{pmatrix} = \begin{pmatrix} W_L^*(s) \\ W_U^*(t) \end{pmatrix}.$$

A $(1-\alpha)$ confidence set of Ω is exactly like (2.5), where $\sigma_L(t)$ and $\sigma_U(t)$ are the estimated standard errors for the $W_L^*(t)$ and $W_U^*(t)$ via (3.1), and the cutoff point c is obtained via (2.6) with the current $W_L^*(t)$ and $W_U^*(t)$.

Now, we use the so-called "five-center cohort" data set from a well-known, multi-center study on the HIV-1 infection incidence among hemophilia patients to illustrate the above interval estimation procedures for Ω (Kroner et al., 1994; Betensky et al., 2002). During the 1980s, persons with hemophilia had high risk of infection with HIV due to their need for infusion of factor VIII or factor IX concentration, products manufactured from the donor's plasma. For this five-center cohort, patients were enrolled without regard to their HIV antibody status. For each patient, repeated serum samples were taken between early 1978 and early 1987, and HIV seroconverters were individuals with both a last negative and first positive serum sample. Thus each infected subject was assigned a "window" of time in which he/she seroconverted. It is not clear from the literature if the sampling times for the patient were independent of the underlying T. In Figure 5, the solid lines are the upper and lower boundaries of the point estimate $\hat{\Omega}$ and a 0.95 interval estimate $\hat{\Omega}_U$ is the region bounded by the dotted lines. Here, we let $\tau_1 = 1000$ (days) and $\tau_2 = 5000$ (days).

Now, for comparing two independent groups of failure times $\{T_{ki}, i = 1, \cdots, n_k; k = 1, 2\}$ with the interval censored data $\{(E_{kLi}, E_{kUi})\}$, let us assume that the two failure times follow a proportional hazards model with parameter e^{Θ}. Using the arguments via (2.11)-(2.13) with the current $\hat{S}_L(t), \hat{S}_U(t), S_L^*(t)$ and $S_U^*(t)$, one can obtain a $(1-\alpha)$ confidence interval (2.13) for Θ.

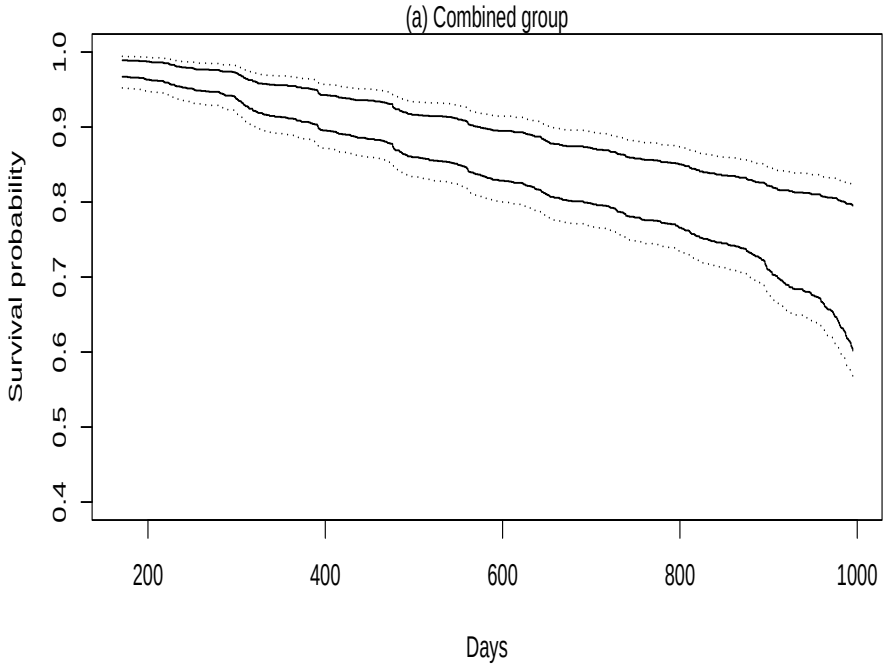
Let us illustrate this interval estimation procedure for Θ with the five-center cohort data from the above study for hemophilia patients. One of the goals of the study was to examine if the patient's average annual dose of non-heat-treated factor VIII concentrate used from 1978 (or birth) to 1984 was related to the time of seroconvertion. For all the analyses done for this study in the literature, the dose level was classified as high ($> 20,000$ U), low ($1 - 20,000$ U) or none. Let us assume that the failure time T_{2i}

for the high dose and T_{1i} for the group without using factor VIII concentration have a proportional hazards structure. First, we obtain the two bounds corresponding to (2.11) for $2500 \leq t \leq 4500$ and $c = 2.67$. Then, it follows from (2.13) that a 0.95 confidence interval $\hat{\Omega}_U$ is $(2.1, 3.6)$, indicating that even without any parametric assumption between the failure and interval censoring times, the high dose group of patients tended to have a much higher HIV incidence rate than the group of patients who did not use this particular concentration.

4. REMARKS

To the best of our knowledge, under the present nonparametric setting there are no confidence *interval* estimation procedures available for the set of all attainable values of the parameter of interest in the presence of a mixture of dependent and independent censoring. Our proposals for one- and two-sample problems are derived based on the *sharp* theoretical bounds for the underlying, non-identifiable survival function. Extending our procedures to the general regression problems seems quite challenging due to the difficulty of identifying possible values of the regression parameters with dependent censorship.

Figure 1: $\hat{\Omega}$ and 0.95 confidence set $\hat{\Omega}_U$ for possible values of $S(\cdot)$ with data from ACTG 175 (solid lines are the boundaries for $\hat{\Omega}$; dotted lines are boundaries for $\hat{\Omega}_U$)



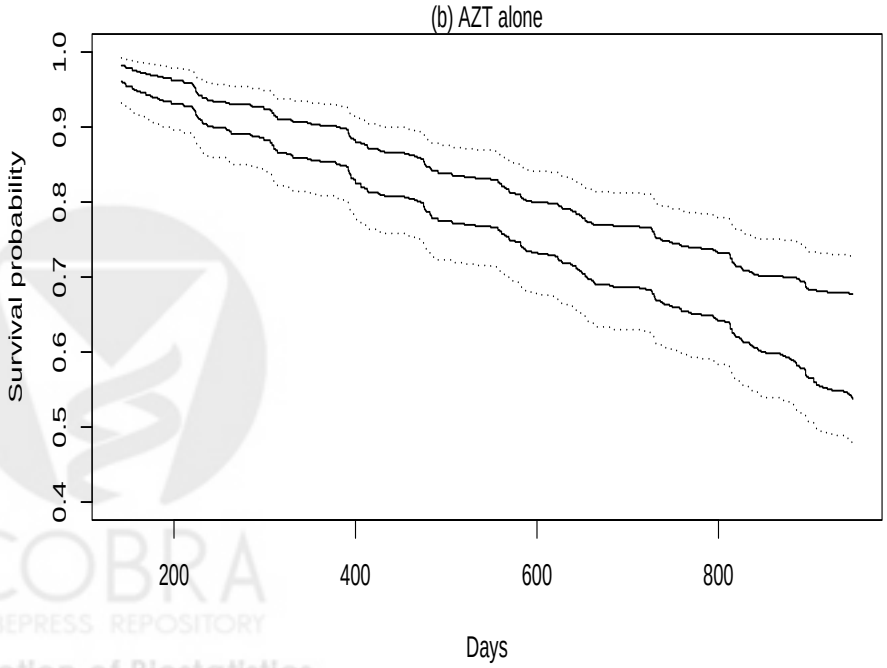
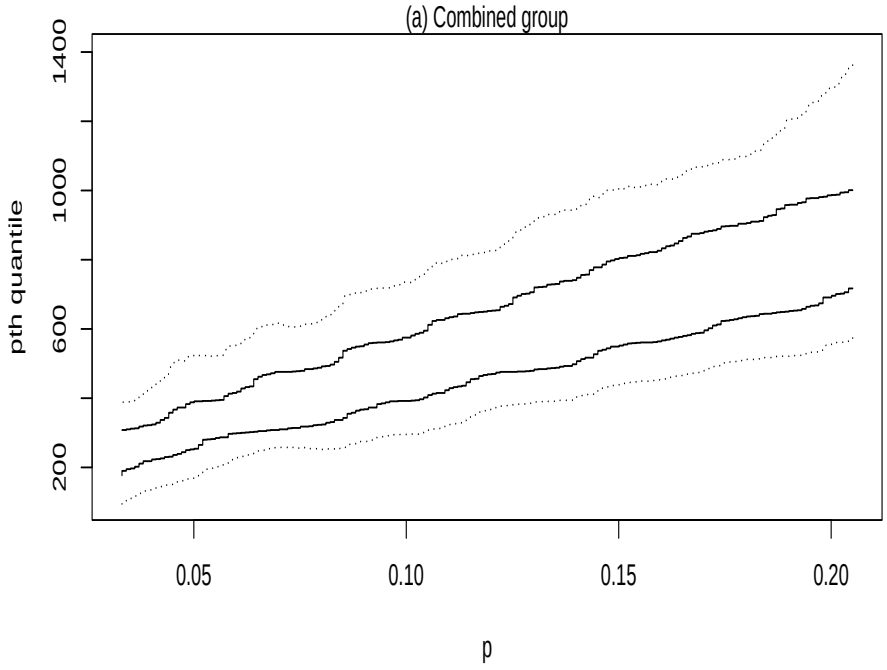


Figure 2: $\hat{\Omega}$ and 0.95 confidence set $\hat{\Omega}_U$ for possible values of pth quantiles with data from ACTG 175 (solid lines are the boundaries for $\hat{\Omega}$; dotted lines are boundaries for $\hat{\Omega}_U$)



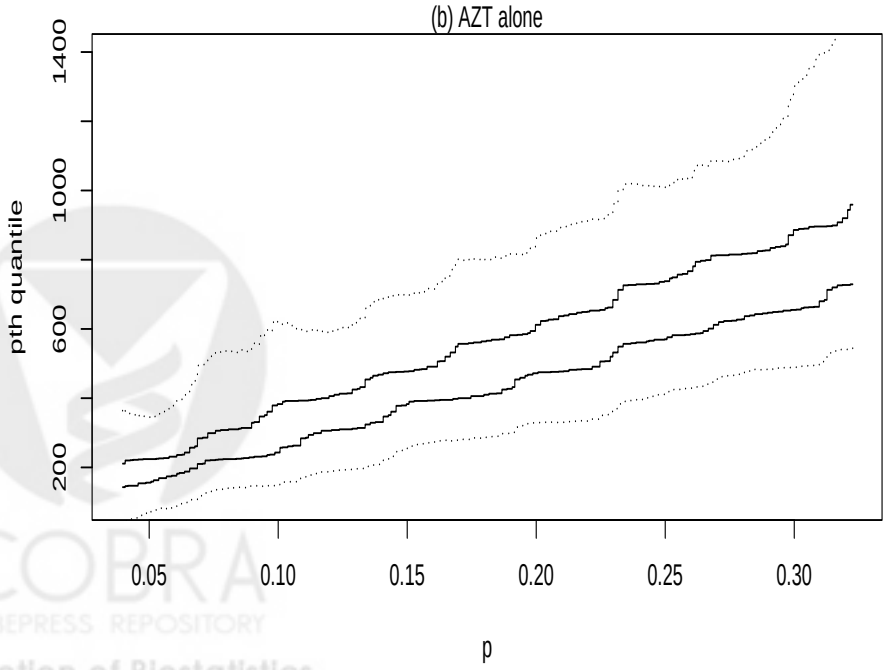


Figure 3: $\hat{\Omega}$ and 0.95 confidence set $\hat{\Omega}_U$ for possible values of $S_2(\cdot) - S_1(\cdot)$ with data from ACTG 175 under various independent censoring assumptions (solid lines are the boundaries for $\hat{\Omega}$; dotted lines are boundaries for $\hat{\Omega}_U$)

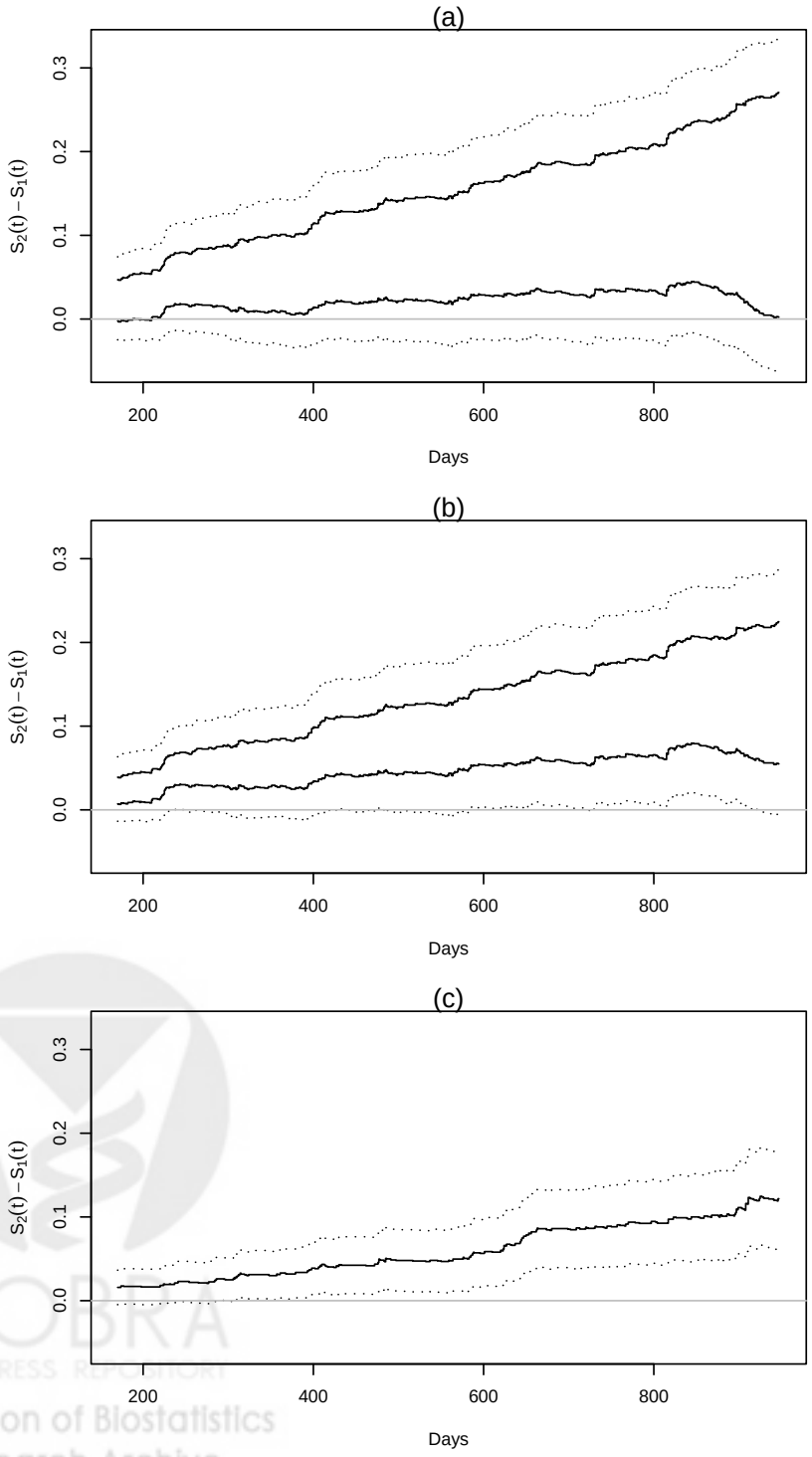


Figure 4: 0.95 simultaneous confidence band for $[\Theta_L(t), \Theta_U(t)]$ under the proportional hazards model with data from ACTG 175

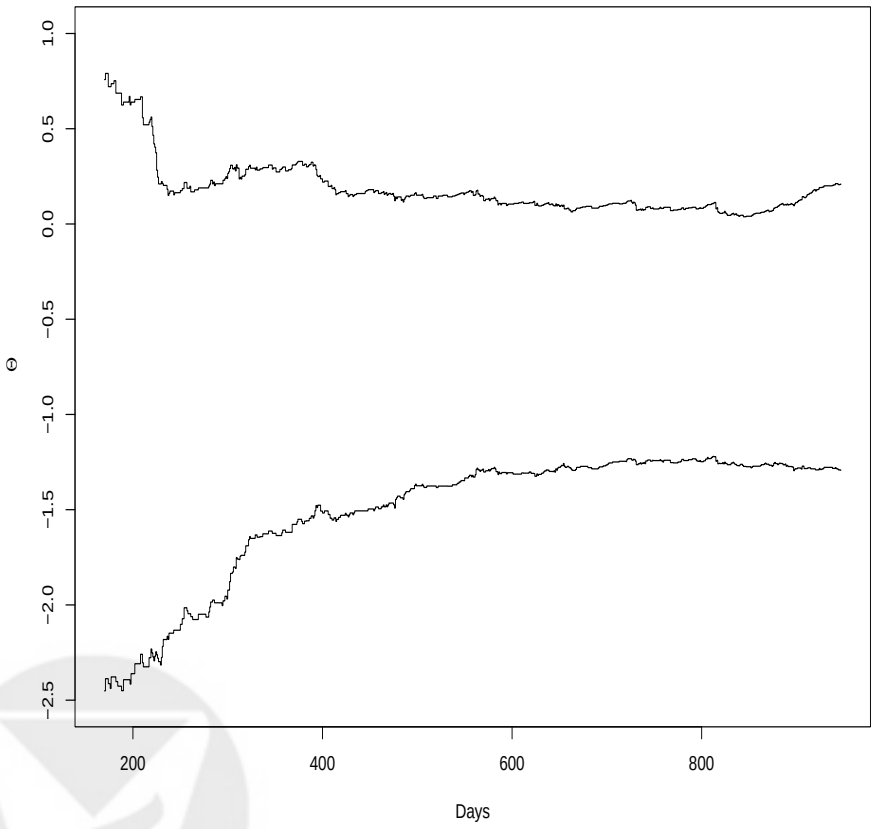
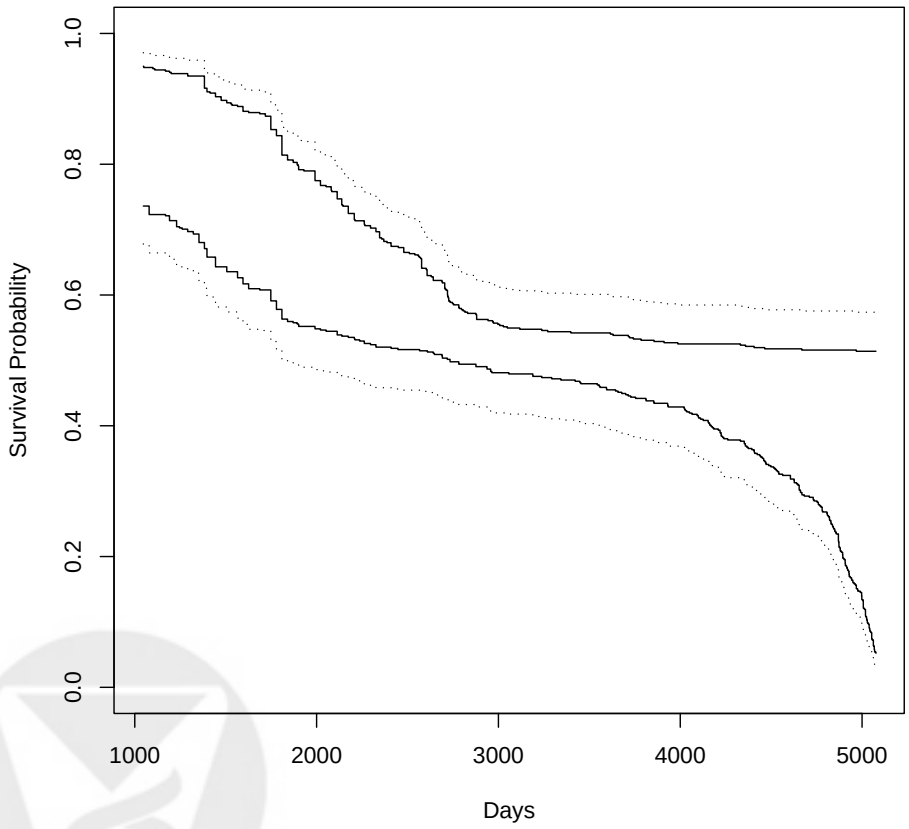


Figure 5: $\hat{\Omega}$ and 0.95 confidence set $\hat{\Omega}_U$ for possible values of $S(\cdot)$ with data from hemophilia study (solid lines are the boundaries for $\hat{\Omega}$; dotted lines are boundaries for $\hat{\Omega}_U$)



REFERENCES

Aalen. O. (1978). Nonparametric estimation of partial transition probabilities in multiple decrement models. *Annals of Statistics* **6**, 534-45.

Bacchetti, P. (1990). Estimating the incubation period of AIDS by comparing population infection and diagnosis patterns. *Journal of the American Statistical Association* **85**, 1002-8.

Betensky, R. A., Lindsey, J. C., Ryan, L. M., and Wand, M. P. (2002). A local likelihood proportional hazards model for interval censored data. *Statistics in Medicine* **21**, 263-75.

Betensky, R. A., Rabinowitz, D. and Tsiatis, A. A. (2001). Computationally simple accelerated failure time regression for interval censored data. *Biometrika* **88**, 703-11.

Cai, T. and Betensky, R. A. (2003). Hazard Regression for Interval Censored Data with Penalized Spline. *Biometics* **59**, 570-9.

DiRienzo, A. G. and Lagakos, S. W. (2001). Bias correction for score tests arising from misspecified proportional hazards regression models. *Biometrika* **88**, 421-434.

DiRienzo, A. G. (2003). Nonparametric Comparison of Two Survival-Time Distributions in the Presence of Dependent Censoring. *Biometrics* **59**, 497-504.

Fisher, L. and Kanarek, P. (1974). Presenting censored survival data when censoring and survival times may not be independent. In *Reliability and Biometry: Statistical Analysis of Lifelength*, F. Proschan and R. Serfling (eds), 303-326. Philadelphia: SIAM

Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* New York: Wiley.

Gentleman, R. and Geyer, C. J. (1994). Maximum likelihood for interval censored data: Consistency and computation. *Biometrika* **81**, 618-23.

Goldwasser, M. A., Tian, L. and Wei, L. J. (2004). Statistical Inference for Infinite Dimensional Parameters Via Asymptotically Pivotal Estimating Functions. *Biometrika* **91**, 81-94.

Hammer, S. M., Katzenstein, D. A., Hughes, M. D., et.al. (1996). A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine* **335**, 1081-90

Huang, J. (1996). Efficient estimation for the proportional hazards model with interval censoring. *Annals of Statistics* **24**, 540-68.

Huang, J. (1999). Asymptotic Properties of Nonparametric Estimation Based on Partly Interval-Censored Data. *Statistica Sinica* **9**, 501-19.

Huang, J. and Wellner, J.A. (1997). Interval censored survival data: a review of recent progress. *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*, Springer, New York.

Joly, P., Commenges, D. and Letenneur, L. (1998). A penalized likelihood approach for arbitrarily censored and truncated data: Application to age-specific incidence of dementia. *Biometrics* **54**, 185-194.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data.* 2nd ed. New York: Wiley.

Klein, J. P. and Moeschberger, M. L. (1988). Bounds on net survival probabilities for dependent competing risks. *Biometrics* **44**, 529-38

Klein, J. P., Moeschberger, M. L., Li, Y. H., and Wang, S. T. (1992). Estimating random effects in the Framingham Heart Study (with discussion). In *Survival Analysis: State of the Art*, J.Klein and P.Goel (eds), 99-120, Dordrecht: Kluwer.

Kooperberg, C. and Clarkson, D. B. (1997). Hazard regression with interval-censored data. *Biometrics* **53**, 1485-94.

Kroner, B. L., Rosenberg, P. S., Aledort, L. M., Alvord, W. G. and Goedert, J. J. (1994). HIV-1 infection incidence among persons with hemophilia in the United States and Western Europe, 1978-1990. *Journal of Acquired Immune Deficiency Syndromes* **7**, 279-86.

Lin, D. Y., Robins, J. M. and Wei, L. J. (1996). Comparing two failure time distributions in the presence of dependent competing risks. *Biometrika* **83**, 381-93.

Lin, D. Y., Wei, L. J. and Ying, Z. (1993). Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. *Biometrika* **80**, 557-72.

Moeschberger, M. L. and Klein, J. P. (1995). Statistical methods for dependent competing risks. *Lifetime Data analysis* **1**, 195-204.

Peterson, A. V. (1976). Bounds for a Joint Distribution Function with Fixed Sub-Distribution Functions: Application to Competing Risks. *Proceedings National Academy of Sciences of the United States of America*, **73**, 11-3.

Peto, R. (1973). Experimental survival curves for interval-censored data. *Applied Statistics* **22**, 86-91.

Rabinowitz, D., Tsiatis, A. A. and Aragon, J. (1995). Regression with interval-censored data. *Biometrika* **82**, 501-13.

Robins, J. M. and Rotnizky, A. (1992). Recoverry of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology: Methodological Issues*, N. Jewell and K.Dietz (eds), 297-331. Boston:Birkhauser.

Robins, J. M. (1993). Information recovery and bios adjustment in proporationl hazards regression analysis of randomized trials using surrogate markers. In *Proceedings of the biopharmaceutical Section, American statistical Association*, 24-33. Alexandria,Virginia: American Statistical Association.

Robins, J. M. and Finkelstein, D. H. (2000). Correcting for non-compliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank test. *Biometrics* **56**, 779-88.

Rosenberg, P. S. (1995). Hazard function estimation using B-splines. *Biometrics* **51**, 874-87

Satten, G. A., Datta, S. and Robins, J. M. (2001). An estimator for the survival function when data are subject to dependent censoring. *Statistics and Probability Letters* **54**, 397-403.

Scharfstein, D. O. and Robins, J. M. (2002). Estimation of the failure time distribution in the presence of informative censoring. *Biometrika* **89**, 617-634.

Slud, E. V. and Rubinstein, L. V. (1983). Dependent competing risks and summary survival curves. *Biometrika* **70**, 643-9.

Tsiatis, A. A. (1975). A nonidentifiability aspect of the problem of competing risk. *Proceedings National Academy of Sciences of the United States of America* **72**, 20-22.

Turnbull, B. W. (1976). The empirical distribution function with arbitrary grouped, censored and truncated data. *J. R. Statist. Soc.* B **38**, 290-5.

Zheng, M. and Klein, J. P. (1995). Estimates of Marginal Survival for Dependent Competing Risks Based on an Assumed Copula *Biometrika* **82**, 127-38.