Computational Statistics and Data Analysis 60 (2013) 123–131

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

A new class of generalized log rank tests for interval-censored failure time data

Xingqiu Zhao [a], Ran Duan [b], Qiang Zhao [c], Jianguo Sun [b,*]

[a] The Hong Kong Polytechnic University, Hong Kong

[b] University of Missouri, United States

[c] Texas State University at San Marcos, United States

ARTICLE INFO

Article history:
Received 26 January 2012
Received in revised form 8 August 2012
Accepted 1 November 2012
Available online 12 November 2012

Keywords:
Asymptotic distribution
Clinical trials
Interval-censoring
Survival comparison

ABSTRACT

This paper discusses nonparametric comparison of survival functions when one observes only interval-censored failure time data (Peto and Peto, 1972; Sun, 2006; Zhao et al., 2008). For the problem, a few procedures have been proposed in the literature. However, most of the existing test procedures determine the test results or *p*-values based on ad hoc methods or the permutation approach. Furthermore for the test procedures whose asymptotic distributions have been derived, the results are only for the null hypothesis. In other words, no nonparametric test procedure exists that has a known asymptotic distribution under the alternative hypothesis and thus can be employed to carry out the power and test size calculation. In this paper, a new class of generalized log-rank tests is proposed and their asymptotic distributions are derived under both null and alternative hypotheses. A simulation study is conducted to assess their performance for finite sample situations and an illustrative example is provided.

1. Introduction

This paper discusses nonparametric comparison of survival functions when one observes only interval-censored failure time data (Peto and Peto, 1972; Sun, 2006; Zhao et al., 2008). By interval-censored data, we mean that the failure times of interest are observed only to belong to some windows or intervals, instead of being observed or known exactly. This would occur if, for example, a survival study involves periodic follow-ups such as clinical trials. One would get an interval-censored observation for a survival event of interest if a subject has not experienced the event at one follow-up time but it is found at the next follow-up time that the event has already occurred. Interval-censored data include right-censored data (Kalbfleisch and Prentice, 2002) as a special case.

A well-known set of interval-censored failure time data that has been discussed by many authors arose from a breast cancer study (Finkelstein, 1986; Sun, 2006). The data consist of 94 early breast cancer patients treated at the Joint Center for Radiation Therapy in Boston between 1976 and 1980. For their treatments, the patients were given either radiation therapy alone or radiation therapy plus adjuvant chemotherapy. Each patient was supposed to have clinic visits every 4–6 months to be examined for cosmetic appearance such as breast retraction. However, actual visit times differ from patient to patient and as a consequence, with respect to the breast retraction time, only interval-censored data were observed. Specifically, some patients did not actually experience the breast retraction during the study and thus gave right-censored observations for the time. For all other patients, the observation was interval-censored with the intervals given by the last clinical visit time at which the breast retraction had not occurred and the first clinical visit time at which it was detected. In particular,

* Correspondence to: Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211, United States.
E-mail address: sunj@missouri.edu (J. Sun).

0167-9473/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2012.11.002

there are some patients for whom the breast retraction was detected at their first clinical visits, meaning that the censored intervals include zero. Another example of interval-censored data from an AIDS clinical trial will be discussed below.

Survival comparison is usually one of the main goals in survival studies. For the case of right-censored failure time data, there exist a number of well-established procedures such as the weighted log-rank tests and the weighted Kaplain–Meier tests (Fleming and Harrington, 1991; Kalbfleisch and Prentice, 2002). For the case of interval-censored failure time data, a few nonparametric test procedures have also been actually developed. For example, Finkelstein (1986) suggested a score test procedure, and Sun (1996) and Zhao and Sun (2004) generalized the log-rank test for right-censored data. However, most of the existing approaches for interval-censored data are ad hoc generalizations of those for right-censored data and have unknown asymptotic properties (Sun, 2006). Some exceptions are the procedures proposed by Fang et al. (2002), Sun et al. (2005) and Zhao et al. (2008), in which the null asymptotic distribution of the test statistics was established. It is clear that all of these test procedures cannot be used if one intends to perform some power or test size calculation as their asymptotic distributions under the alternative hypothesis are still unknown. In this paper, we propose a new class of test procedures whose asymptotic distributions are established under both null and alternative hypotheses.

The remainder of the paper is organized as follows. We will begin in Section 2 with introducing some notation and assumptions that will be used throughout the paper and then present the new test statistics. The asymptotic distributions of the test statistics will be established in Section 3. In Section 4, we will present some numerical results obtained from a simulation study for assessing the finite sample performance of the proposed test procedure. An illustrative example is also given in Section 4. Section 5 contains some concluding remarks.

2. Generalized log-rank test statistics

Consider a survival study that involves n independent subjects. Let T_i denote the survival time of interest for subject i, $i = 1, \ldots, n$. Suppose that for subject i, we only observe $\{U_i, V_i, \Delta_i = I(T_i \le U_i), \Gamma_i = I(U_i < T_i \le V_i)\}$, where U_i and V_i are non-negative random variables independent of T_i such that $U_i < V_i$ with probability one, $i = 1, \ldots, n$. This means that one only knows if T_i is smaller than U_i, between U_i and V_i, or larger than V_i. In other words, we only have interval-censored data on the T_i's. Assume that the study involves two groups, control (group 1) and treatment (group 2) groups. Let $F_1(t)$ and $F_2(t)$ denote the cumulative distribution functions of the T_i's for the subjects in the control and treatment groups, respectively. Suppose that the main goal is to compare the two groups or to test the hypothesis $H_0 : F_1(t) = F_2(t)$.

To construct the proposed test statistics, we first look at the test statistics given in Sun et al. (2005). For this, let $F(t)$ denote the common survival function under the null hypothesis H_0 and define

$$K_F(u, v, \delta, \gamma) = \delta \frac{\eta\{F(u)\} - c_0}{F(u)} + \gamma \frac{\eta\{F(v)\} - \eta\{F(u)\}}{F(v) - F(u)} + (1 - \delta - \gamma) \frac{c_0 - \eta\{F(v)\}}{1 - F(v)}.$$

Here η is a known function over $(0, 1)$ such that $\lim_{x\to 0} \eta(x) = \lim_{x\to 1} \eta(x) = c_0$, where c_0 is a constant. Also let $\hat{F}_n(t)$ denote the nonparametric maximum likelihood estimate of F based on all samples and S_l the set of indices for the subjects in group l, $l = 1, 2$. To test H_0, Sun et al. (2005) proposed the following test statistic

$$U_{SZZ} = \left(\sum_{i\in S_1} K_{\hat{F}_n}(U_i, V_i, \Delta_i, \Gamma_i), \sum_{i\in S_2} K_{\hat{F}_n}(U_i, V_i, \Delta_i, \Gamma_i) \right)^T$$

and derived its null asymptotic distribution.

On the other hand, it is easy to see that it would be difficult or impossible to derive the asymptotic distribution of U_{SZZ} under the alternative hypothesis partly because $\hat{F}_n$ is not well-defined if $F_1 \ne F_2$. To modify the test statistic U_{SZZ}, let n_1 and n_2 ($n_1 + n_2 = n$) denote the numbers of subjects in the control and treatment groups, respectively, and $\hat{F}_{n_1}$ and $\hat{F}_{n_2}$ the nonparametric maximum likelihood estimates of F_1 and F_2 based on the samples from the control and treatment groups, respectively. Naturally, by noting that

$$\sum_{i\in S_1} K_{\hat{F}_{n_1}}(U_i, V_i, \Delta_i, \Gamma_i) = 0$$

and

$$\sum_{i\in S_2} K_{\hat{F}_{n_2}}(U_i, V_i, \Delta_i, \Gamma_i) = 0,$$

one could define a new statistic as

$$\left(\sum_{i\in S_1} K_{\hat{F}_{n_2}}(U_i, V_i, \Delta_i, \Gamma_i), \sum_{i\in S_2} K_{\hat{F}_{n_1}}(U_i, V_i, \Delta_i, \Gamma_i) \right)^T$$

by replacing $\hat{F}_n(t)$ with $\hat{F}_{n_1}$ or $\hat{F}_{n_2}$ in U_{SZZ}. However, it would still be difficult to derive the asymptotic distribution of the statistic given above.

To construct a workable test statistic, define

$$K_{F_1,F_2}(u, v, \delta, \gamma) = \delta \frac{\eta\{F_2(u)\} - c_0}{F_1(u)} + \gamma \frac{\eta\{F_2(v)\} - \eta\{F_2(u)\}}{F_1(v) - F_1(u)} + (1 - \delta - \gamma) \frac{c_0 - \eta\{F_2(v)\}}{1 - F_1(v)}.$$

For testing the hypothesis H_0, we propose to use the statistic

$$\bar{U}_n = (\bar{U}_{n_1}, \bar{U}_{n_2})^T = \left(\frac{1}{\sqrt{n_1}} \sum_{i \in S_1} K_{\hat{F}_{n_1}, \hat{F}_{n_2}}(U_i, V_i, \Delta_i, \Gamma_i), \frac{1}{\sqrt{n_2}} \sum_{i \in S_2} K_{\hat{F}_{n_2}, \hat{F}_{n_1}}(U_i, V_i, \Delta_i, \Gamma_i) \right)^T.$$

In the next section, we will establish the asymptotic properties of $\bar{U}_{n_1}$ and $\bar{U}_{n_2}$ and hence present the resulting test procedure for H_0. Some comments will be given below on the determination of $\hat{F}_{n_1}(t)$ and $\hat{F}_{n_2}(t)$ as well as the selection of function η.

3. Asymptotic distributions and test procedures

In this section, we will first establish the asymptotic distributions of $\bar{U}_{n_1}$ and $\bar{U}_{n_2}$ and then present the test procedure. For this, let H and h denote the distribution and density functions of (U_i, V_i), respectively, and λ_2 and ν_2 denote the Lebesgue measure on R^2 and counting measure on the set $\{(0, 1), (1, 0), (0, 0)\}$, respectively. Define

$$q_F(u, v, \delta, \gamma) = h(u, v)\{F(u)\}^{\delta}\{F(v) - F(u)\}^{\gamma}\{1 - F(v)\}^{1-\delta-\gamma}$$

and similarly $q_{F_1}(u, v, \delta, \gamma)$ and $q_{F_2}(u, v, \delta, \gamma)$ with respect to $\lambda_2 \otimes \nu_2$. It is easy to see that $q_{F_l}(u, v, \delta, \gamma)$ is the density function of $(U_i, V_i, \Delta_i, \Gamma_i)$ for $i \in S_l$, $l = 1, 2$. Also for $l = 1, 2$, define $dQ_l = q_{F_l}\, d(\lambda_2 \otimes \nu_2)$ and the empirical measure

$$Q_{n_l}(u, v, \delta, \gamma) = \frac{1}{n_l} \sum_{i \in S_l} 1_{\{(U_i, V_i) \le (u, v), (\Delta_i, \Gamma_i) = (\delta, \gamma)\}}$$

with $Q_l f = \int f dQ_l$ and $Q_{n_l} f = \int f dQ_{n_l} = \frac{1}{n_l} \sum_{i \in S_l} f(U_i, V_i, \Delta_i, \Gamma_i)$ for any function $f(u, v, \delta, \gamma)$. Then we have

$$\bar{U}_{n_1} = \sqrt{n_1}\, Q_{n_1}(K_{\hat{F}_{n_1}, \hat{F}_{n_2}}), \qquad \bar{U}_{n_2} = \sqrt{n_2}\, Q_{n_2}(K_{\hat{F}_{n_2}, \hat{F}_{n_1}}).$$

For the result below, we will assume that the regularity conditions given in Groeneboom and Wellner (1992) for the strong consistency of $\hat{F}_{n_1}$ and $\hat{F}_{n_2}$ hold. Also following Sun et al. (2005), we will assume that $F_1(t)$ and $F_2(t)$ have their support in $[0, M]$ with continuous density functions, and that there exist $0 < \delta_0, \varepsilon_0 < M/2$ and $M_0 < M$ such that $Pr(U_i < \delta_0) = 0$, $Pr(U_i + \varepsilon_0 \le V_i \le M_0) = 1$, $0 < F_l(\delta_0) < F_l(M_0) < 1$ and $\min_{\delta_0 \le t \le M_0 - \varepsilon_0}[F_l(t + \varepsilon_0) - F_l(t)] \ne 0$, where M is a positive constant. These conditions usually hold for periodic follow-up studies such as clinical trials. The following theorem gives the asymptotic behavior of $\bar{U}_{n_1}$ and $\bar{U}_{n_2}$.

Theorem 1. *Suppose that the assumptions described above hold and η is a bounded Lipschitz function on $[a, 1]$ for any finite positive number $a < 1$. Also suppose that as $n \to \infty$, $n_k/n \to p_k$, where $0 < p_k < 1$ and $p_1 + p_2 = 1$. Then we have, asymptotically,*

$$\bar{U}_{n_1} = Z_{n1} + o_p(1)$$

and

$$\bar{U}_{n_2} = Z_{n2} + o_p(1),$$

where

$$Z_{n1} = \sqrt{n_1}(Q_{n_1} - Q_1)\left\{K_{F_1,F_2} - \tilde{\theta}_{g_1,F_1}\right\}$$

and

$$Z_{n2} = \sqrt{n_2}(Q_{n_2} - Q_2)\left\{K_{F_2,F_1} - \tilde{\theta}_{g_2,F_2}\right\}$$

with g_l and $\tilde{\theta}_{g_l,F_l}$ given in the Appendix.

The proof of the above theorem is sketched in the Appendix. Note that Z_{n1} and Z_{n2} are independent. Then it follows from the theorem and the central limit theorem that $\bar{U}_{n_1}$ and $\bar{U}_{n_2}$ converge in distribution to two independent normal random variables Z_1 and Z_2, where $Z_1 \sim N(0, \sigma_1^2)$ and $Z_2 \sim N(0, \sigma_2^2)$ with

$$\sigma_1^2 = Q_1\left[\left\{K_{F_1,F_2} - \tilde{\theta}_{g_1,F_1}\right\} - Q_1\left\{K_{F_1,F_2} - \tilde{\theta}_{g_1,F_1}\right\}\right]^2$$

and

$$\sigma_2^2 = Q_2\left[\left\{K_{F_2,F_1} - \tilde{\theta}_{g_2,F_2}\right\} - Q_2\left\{K_{F_2,F_1} - \tilde{\theta}_{g_2,F_2}\right\}\right]^2.$$

Define

$$S = \frac{\bar{U}_{n_1}^2/\sigma_1^2}{\bar{U}_{n_2}^2/\sigma_2^2}.$$

Then it follows from the theorem above that S has an asymptotic $F(1, 1)$ distribution and furthermore, under the hypothesis H_0 and as $n \to \infty$, the distribution of $S_0 = \bar{U}_{n_1}^2/\bar{U}_{n_2}^2$ can be approximated by the $F(1, 1)$ distribution. This suggests that one can carry out the test of the hypothesis H_0 by using the statistic S_0 based on the $F(1, 1)$ distribution.

To implement the test procedure proposed above, one needs to determine $\hat{F}_{n_1}$ and $\hat{F}_{n_2}$ and select the function η. For the former, the simplest method is to apply the self-consistency algorithm given in Turnbull (1976). Some alternatives can be found in Sun (2006). For the latter, a common choice, which will be used below for the numerical study, is $\eta(x) = 1 - (1 - x)\log(1 - x)(1 - x)^b x^c$, where b and c are some numbers between [0, 1]. More comments on this can be found in Sun et al. (2005).

As discussed above, in practice, one may be often interested in performing power calculations. For this based on the test procedure given above, for the given significance level α, let Z denote the random variable following the $F(1, 1)$ distribution and F_L and F_U be defined such that

$$P(Z < F_L) = \alpha/2 \quad \text{and} \quad P(Z > F_U) = \alpha/2.$$

Then the asymptotic power is given by

$$F_{1,1}\left(\frac{\sigma_2^2}{\sigma_1^2}F_L\right) + 1 - F_{1,1}\left(\frac{\sigma_2^2}{\sigma_1^2}F_U\right)$$

if F_1 and F_2 are known.

4. Numerical studies

Now we report some results obtained from a simulation study conducted to assess the finite sample performance of the class of test procedures proposed in the previous sections and its application to a real set of interval-censored data. For the simulation study, we assumed that a half of the subjects are from the control group and the other half from the treatment group. To generate the survival times of interest, we considered two set-ups. One is to assume that T_i follows the exponential distribution with the mean $\exp(\alpha + \beta z_i)$, where z_i is the treatment indicator, being equal to 0 for the subjects in the control group and 1 otherwise. The other is to generate T_i from the gamma distribution with the shape parameter equal to 2 and the scale parameter $1/(\alpha + \beta z_i)$.

To generate the censoring interval for subject i, we first generated U_{i1} and U_{i2} independently from the uniform distribution over $(1, \theta_1)$ and $(1, \theta_2)$, respectively. Here θ_1 and θ_2 are some positive constants chosen to give the desired percentages of left-censored, interval-censored and right-censored observations. Given U_{i1} and U_{i2}, we defined U_i to be the nearest integer to U_{i1} and V_i the nearest integer to the maximum of $U_{i1}+1$ and $U_{i1}+U_{i2}$. Also we assumed that the study ended at $t = 10$ and thus defined V_i to be 10 if the V_i generated above is larger than 10. The results given below are based on 1000 replications.

Table 1 presents the empirical or estimated size and power of the proposed test procedure based on the simulated data generated from the exponential distribution with $\alpha = 2$, $\beta = -3, -2, -1.5, 0, 1.5, 2$ or 3. Here we used the η function given in Section 3 with different values of b and c and the self-consistency algorithm for the determination of the maximum likelihood estimates $\hat{F}_{n_1}$ and $\hat{F}_{n_2}$. In the table, the first column gives the percentages of left-censored, interval-censored and right-censored observations in the generated data, which are roughly (20%, 20%, 60%) and (17%, 16%, 67%) for the two situations considered here. The results obtained under the gamma distribution are given in Table 2 and here we took $\alpha = 1$ and the same values for β as in Table 1. One can see from both Tables 1 and 2 that the proposed test procedure seems to give the right size and have good power for the situations considered here.

To illustrate the proposed approach, we apply it to the set of interval-censored data discussed in Goggins and Finkelstein (2000) and Sun (2006) among others. The data arose from an AIDS clinical trial concerning the opportunistic infection cytomegalovirus (CMV). During the study, among other activities, blood and urine samples were collected from the patients at their clinical visits and tested for the presence of CMV, which is also commonly referred to as shedding of the virus. These samples and tests provide observed information on the two variables, the times to CMV shedding in blood and urine, respectively. The study consists of 204 patients who provided at least one urine and one blood sample during the study. For some patients, their shedding had already occurred at their first clinical visits or they had not yet started shedding by the end of the study, giving either left- or right-censored observations on their shedding times. For the other patients, their shedding times were observed to belong to some intervals given by the last negative and first positive blood or urine test, respectively.

Table 1
Estimated size and power based on simulated data from exponential distribution.

Censoring percentages	b	c	β						
			3	2	1.5	0	−1.5	−2	−3
(20%, 20%, 60%)	0	0	0.478	0.158	0.085	0.057	0.260	0.838	1
	0	0.5	0.397	0.118	0.075	0.059	0.365	0.921	1
	0.5	0	0.708	0.242	0.103	0.043	0.212	0.804	1
	0.5	0.5	0.628	0.214	0.095	0.044	0.320	0.893	1
(17%, 16%, 67%)	0	0	0.489	0.171	0.118	0.047	0.262	0.844	1
	0	1	0.335	0.122	0.084	0.059	0.530	0.962	1
	1	0	0.855	0.397	0.172	0.040	0.199	0.793	1
	1	1	0.770	0.291	0.140	0.046	0.399	0.923	1

Table 2
Estimated size and power based on simulated data from gamma distribution.

Censoring percentages	b	c	β						
			3	2	1.5	0	−1.5	−2	−3
(12%, 12%, 76%)	0	0	0.997	0.946	0.692	0.047	0.854	0.996	1
	0	0.5	0.993	0.923	0.686	0.053	0.927	1	1
	0.5	0	1.000	0.966	0.719	0.041	0.852	1	1
	0.5	0.5	1.000	0.960	0.708	0.042	0.919	1	1
(10%, 15%, 75%)	0	0	0.997	0.930	0.704	0.041	0.853	0.999	1
	0	1	0.989	0.907	0.695	0.053	0.946	1	1
	1	0	1.000	0.958	0.714	0.043	0.817	0.998	1
	1	1	0.999	0.948	0.705	0.040	0.936	1	1

Table 3
Results on the analysis of AIDS clinical trial.

(b, c)	On blood shedding time						
	(0, 0)	(0, 0.5)	(0.5, 0)	(0.5, 0.5)	(0, 1)	(1, 0)	(1, 1)
Test statistics	0.00022	0.00029	0.00011	0.00036	0.00031	0.00023	0.00055
p-value	0.019	0.022	0.013	0.024	0.022	0.019	0.030
Exact p-value	0.039	0.037	0.032	0.047	0.046	0.036	0.044
	On urine shedding time						
Test statistics	0.0239	1.1376	0.0221	0.0002	0.0099	0.0073	0.4043
p-value	0.019	0.108	0.721	0.126	0.188	0.958	0.195
Exact p-value	0.021	0.063	0.408	0.134	0.097	0.551	0.169

In addition to the observed information about CMV shedding times in blood and in urine, the study also provided the range of each patient's baseline CD4 cell count. In particular, the patients were classified into two groups: those with their baseline CD4 cell counts less than 75 (cells/μl) and the others. Note that the CD4 cell count indicates the status of a person's immune system and is commonly used to measure the stage of HIV infection. For this data set, one problem of interest is to compare the two groups of patients with respect to their CMV shedding times. For this, we applied the test procedure developed in the previous sections to the data on the times to CMV shedding in blood and urine separately and the obtained results are presented in Table 3. They indicate that the CMV shedding times in blood were significantly different between the two groups of patients. However, it seems that there was no significant difference in CMV shedding times in urine. Suggested by a reviewer, we also applied the Monte Carlo exact test to the data and include the obtained p-values in Table 3. It can be seen that they gave similar conclusions and furthermore the results are also similar to those given by others (Sun, 2006).

5. Concluding remarks

This paper discussed the nonparametric comparison of survival functions when only interval-censored failure time data are available. For the problem, a class of nonparametric tests was proposed and both finite sample and asymptotic properties of the presented approach were established. One major advantage of the proposed test procedure is that its asymptotic distribution is known under both null and alternative hypotheses, which makes both power and test size calculation possible. In contrast, for all existing nonparametric test procedures, their asymptotic distribution is either unknown or known only under the null hypothesis. Note that another shortcoming for some existing test procedures is that the estimation or determination of the variance of the test statistics involve the dealing of high dimension matrices, which makes them unstable. It is easy to see that the proposed test procedure does not have the same problem.

In the previous sections, we only considered the two sample test and a natural question of practical interest would be if the proposed test procedure can be extended to the k-sample comparison problem. For this, consider a survival study that

involves k-groups with n_l subjects from group l and $n_1 + \cdots n_k = n$. Let $F_l(t)$ denote the cumulative distribution function of the survival time for group l. Suppose that one is interested in testing if the null hypothesis is $H_0 : F_1(t) = \cdots = F_k(t)$ and define

$$\bar{U}^*_{n_1} = \sqrt{n_1}\, Q_{n_1}(K_{\hat{F}_{n_1},\hat{F}_{n_2}}), \qquad \bar{U}^*_{n_l} = \sqrt{n_l}\, Q_{n_l}(K_{\hat{F}_{n_l},\hat{F}_{n_1}}), \quad l = 2, \ldots, k, k \geq 2,$$

and

$$S^*_0, = \frac{(\bar{U}^{*2}_{n_2} + \cdots + \bar{U}^{*2}_{n_k})/(k-1)}{\bar{U}^{*2}_{n_1}}.$$

In the above, Q_{n_l} denotes the empirical measure of sample l and $\hat{F}_{n_l}$ the nonparametric maximum likelihood estimate of F_l based on sample l from group l, $l = 1, \ldots, k$. One can expect to show that under H_0, S^*_0 asymptotically follows the $F(k-1, 1)$ distribution and thus can construct a test procedure based on it.

It should be noted that there exist some limitations about the proposed nonparametric test procedure. One is that so far we have assumed that no exact observation on the survival time of interest is observed. Although this may not be true in general, it holds in many situations such as studies with periodic follow-ups. Also we have assumed that the distributions of interest are continuous. But actually the procedure presented is valid if the distributions of interest have only finite support points. Note that for this latter case, the problem is much simpler as the standard maximum likelihood theory for parametric models could be applied. Of course, it would be useful to generalize the proposed approach to situations where the observed data include both exact and interval-censored observations on the survival time of interest.

Another limitation of the proposed approach is that we only considered the situation where the distributions generating censoring intervals are identical for the subjects in different treatment groups. Sometimes this may not be true as, for example, the subjects in different treatment groups may have different follow-up patterns in a periodic follow-up study. One specific example of this is given by a clinical trial in which patients receiving placebo treatment may feel worse compared to other patients and thus visit doctors more often. Among others, Sun (1999) discussed this problem for current status data, a special case of interval-censored data. However, there does not seem to exist a nonparametric test procedure similar to the one proposed here.

Acknowledgments

The authors wish to thank the co-Editor, Dr. Stanley Azen, the Associate Editor and two referees for their effort on the manuscript and many useful comments and suggestions. The work was partly supported by the Research Grant Council of Hong Kong (PolyU 5032/09P) and The Hong Kong Polytechnic University grants to the first author and a NIH grant to the last author.

Appendix. Proof of Theorem 1

To prove the theorem, we will only need to prove the first part on $\bar{U}_{n_1}$ as the proof for the second part is similar. For this, note that we can rewrite $\bar{U}_{n_1}$ as

$$\bar{U}_{n_1} = \sqrt{n_1}(Q_{n_1} - Q_1)\left(K_{\hat{F}_{n_1},\hat{F}_{2,n_2}} - K_{F_1,F_2}\right) + \sqrt{n_1}Q_1\left(K_{\hat{F}_{n_1},\hat{F}_{n_2}}\right) + \sqrt{n_1}(Q_{n_1} - Q_1)\left(K_{F_1,F_2}\right). \tag{1}$$

For the second term at the right side of the above equation, we have

$$\begin{aligned}\sqrt{n_1}Q_1\left(K_{\hat{F}_{n_1},\hat{F}_{n_2}}\right) &= \sqrt{n_1}Q_1\left[\left\{K_{\hat{F}_{n_1},\hat{F}_{n_2}} - K_{\hat{F}_{n_1},F_2}\right\} - \left\{K_{F_1,\hat{F}_{n_2}} - K_{F_1,F_2}\right\}\right] \\ &\quad + \sqrt{n_1}Q_1\left\{K_{F_1,\hat{F}_{n_2}} - K_{F_1,F_2}\right\} + \sqrt{n_1}Q_1\left(K_{\hat{F}_{n_1},F_2}\right)\end{aligned} \tag{2}$$

and

$$\begin{aligned}\sqrt{n_1}Q_1\left(K_{\hat{F}_{n_1},F_2}\right) &= \sqrt{n_1}\int\left\{K_{\hat{F}_{n_1},F_2}(u,v,\delta,\gamma)q_{F_1}(u,v,\delta,\gamma)\right\}d(\lambda_2\otimes\nu_2) \\ &= \sqrt{n_1}\int\left\{K_{\hat{F}_{n_1},F_2}(u,v,\delta,\gamma) - K_{F_1,F_2}(u,v,\delta,\gamma)\right\} \\ &\quad\times\left\{q_{F_1}(u,v,\delta,\gamma) - q_{\hat{F}_{n_1}}(u,v,\delta,\gamma)\right\}d(\lambda_2\otimes\nu_2) \\ &\quad + \sqrt{n_1}\int K_{\hat{F}_{n_1},F_2}(u,v,\delta,\gamma)q_{\hat{F}_{n_1}}(u,v,\delta,\gamma)\,d(\lambda_2\otimes\nu_2) \\ &\quad + \sqrt{n_1}\int K_{F_1,F_2}(u,v,\delta,\gamma)\left\{q_{F_1}(u,v,\delta,\gamma) - q_{\hat{F}_{n_1}}(u,v,\delta,\gamma)\right\}d(\lambda_2\otimes\nu_2).\end{aligned} \tag{3}$$

It can be easily shown that

$$\sqrt{n_1}Q_1\left[\left\{K_{\hat{F}_{n_1},\hat{F}_{n_2}}-K_{\hat{F}_{n_1},F_2}\right\}-\left\{K_{F_1,\hat{F}_{n_2}}-K_{F_1,F_2}\right\}\right]=o_p(1),$$

$$\sqrt{n_1}Q_1\left\{K_{F_1,\hat{F}_{n_2}}-K_{F_1,F_2}\right\}=0,$$

$$\sqrt{n_1}\int\left\{K_{\hat{F}_{n_1},F_2}(u,v,\delta,\gamma)-K_{F_1,F_2}(u,v,\delta,\gamma)\right\}\left\{q_{F_1}(u,v,\delta,\gamma)-q_{\hat{F}_{n_1}}(u,v,\delta,\gamma)\right\}d(\lambda_2\otimes\nu_2)=o_p(1),$$

and

$$\int K_{\hat{F}_{n_1},F_2}(u,v,\delta,\gamma)q_{\hat{F}_{n_1}}(u,v,\delta,\gamma)d(\lambda_2\otimes\nu_2)=0.$$

Thus it follows from (2) and (3) that

$$\sqrt{n_1}Q_1\left(K_{\hat{F}_{n_1},\hat{F}_{n_2}}\right)=-I_n+o_p(1), \tag{4}$$

where

$$I_n=\sqrt{n_1}\int K_{F_1,F_2}(u,v,\delta,\gamma)\left\{q_{\hat{F}_{n_1}}(u,v,\delta,\gamma)-q_{F_1}(u,v,\delta,\gamma)\right\}d(\lambda_2\otimes\nu_2).$$

Now we will show that

$$\sqrt{n_1}\,(Q_{n_1}-Q_1)(K_{\hat{F}_{n_1},\hat{F}_{n_2}}-K_{F_1,F_2})\to 0 \tag{5}$$

in probability as $n\to\infty$. For this, define

$$\mathcal{F}=\{F:F \text{ is a distribution function defined on } [0,M]\},$$

$$\mathcal{G}=\left\{F:F\in\mathcal{F},0<F(\delta_0)<F(M_0)<1,\min_{\delta_0\le t\le M_0-\varepsilon_0}[F(t+\varepsilon_0)-F(t)]\neq 0\right\}$$

and

$$\mathcal{H}=\left\{K_{F_3,F_4}(u,v,\delta,\gamma)-K_{F_1,F_2}(u,v,\delta,\gamma):(u,v)\in D,F_3,F_4\in\mathcal{G}\right\},$$

where $D=\{(u,v):u\ge\delta_0,u+\varepsilon_0\le v\le M_0\}$. Because $\mathcal{F}$ is a *P-Donsker* from the proof of Corollary 5.1 of Huang and Wellner (1995), $\mathcal{G}$ is a *P-Donsker* by Theorem 2.10.1 of van der Vaart and Wellner (1996). Note that for any $F_3,F_4,F_5,F_6\in\mathcal{G}$, $(u,v)\in D$, we have

$$\begin{aligned}&\left|\delta\frac{\eta(F_4(u))-c_0}{F_3(u)}+\gamma\frac{\eta(F_4(v))-\eta(F_4(u))}{F_3(v)-F_3(u)}+(1-\delta-\gamma)\frac{c_0-\eta(F_4(v))}{1-F_3(v)}\right.\\&\quad\left.-\delta\frac{\eta(F_6(u))-c_0}{F_5(u)}-\gamma\frac{\eta(F_6(v))-\eta(F_6(u))}{F_5(v)-F_5(u)}-(1-\delta-\gamma)\frac{c_0-\eta(F_6(v))}{1-F_5(v)}\right|\\&\le c\left[|F_3(u)-F_5(u)|+|F_3(v)-F_5(v)|+|F_4(u)-F_6(u)|+|F_4(v)-F_6(v)|\right]\end{aligned}$$

for some constant c. Then it can be shown by using the bracket entropy theorem of van der Vaart and Wellner (1996, pp. 127–159)) and the arguments similar to those used in Huang and Wellner (1995) that $\mathcal{H}$ is *P-Donsker*. Also note that $\hat{F}_{n_1},\hat{F}_{n_2}\in\mathcal{G}$ for all n sufficiently large and as $n\to\infty$, we have

$$\int\{|\hat{F}_{n_l}(u)-F_l(u)|^2+|\hat{F}_{n_l}(v)-F_l(v)|^2\}dP\longrightarrow 0$$

in probability from the strong consistency of $\hat{F}_{n_l}$ (Groeneboom and Wellner, 1992, p. 85). Thus (5) is true based on this and the uniform asymptotic equicontinuity of the empirical process resulting from the Donsker property (van der Vaart and Wellner, 1996, pp. 168–171).

It follows from (1), (4) and (5) that we have

$$\bar{U}_{n_1}=\sqrt{n_1}(Q_{n_1}-Q_1)K_{F_1,F_2}-I_n+o_p(1). \tag{6}$$

To finish the proof, next we will show that

$$I_n=\sqrt{n_1}(Q_{n_1}-Q_1)(\tilde{\theta}_{g_1,F_1})+o_p(1), \tag{7}$$

where $\tilde{\theta}_{g,F}$ is defined below. For this, note that

$$I_n = \sqrt{n_1}\int h_1(u)\frac{\eta(F_2(u))-c_0}{F_1(u)}\{\hat{F}_{n_1}(u)-F_1(u)\}du + \sqrt{n_1}\int h(u,v)\frac{\eta(F_2(v))-\eta(F_2(u))}{F_1(v)-F_1(u)}[\{\hat{F}_{n_1}(v)-F_1(v)\}-\{\hat{F}_{n_1}(u)-F_1(u)\}]dudv - \sqrt{n_1}\int h_2(v)\frac{c_0-\eta(F_2(v))}{1-F_1(v)}\{\hat{F}_{n_1}(v)-F_1(v)\}dv = \sqrt{n_1}\int g_1(t)\{\hat{F}_{n_1}(t)-F_1(t)\}dt,$$

where

$$g_1(t) = h_1(t)\frac{\eta(F_2(t))-c_0}{F_1(t)} + \int_0^t h(u,t)\frac{\eta(F_2(t))-\eta(F_2(u))}{F_1(t)-F_1(u)}du - \int_t^M h(t,v)\frac{\eta(F_2(v))-\eta(F_2(t))}{F_1(v)-F_1(t)}dv - h_2(t)\frac{c_0-\eta(F_2(t))}{1-F_1(t)}$$

with h_1 and h_2 being the marginal density functions of U_i and V_i, respectively.

Define

$$h^*(u,v) = \begin{cases} h(u,v), & \text{if } u \le v, \\ h(v,u), & \text{if } u > v, \end{cases}$$

and

$$d_F(x) = \frac{F(x)\{1-F(x)\}}{h_1(x)\{1-F(x)\}+h_2(x)F(x)}.$$

Let $\phi = \phi_{g,F}$ be the right-continuous solution to the following equation

$$\phi(x) = d_F(x)\left\{g(x) - \int_0^x \frac{\phi(x)-\phi(x')}{|F(x)-F(x')|}h^*(x',x)dx'\right\}.$$

Also define

$$\tilde{\theta}_{g,F}(u,v,\delta,\gamma) = -\delta\frac{\phi_{g,F}(u)}{F(u)} - \gamma\frac{\phi_{g,F}(v)-\phi_{g,F}(u)}{F(v)-F(u)} + (1-\delta-\gamma)\frac{\phi_{g,F}(v)}{1-F(v)}.$$

Then it follows from Groeneboom (1996, p. 149) that we have

$$\begin{aligned} I_n &= \sqrt{n_1}\int g_1(t)\{\hat{F}_{n_1}(t)-F_1(t)\}dt = \sqrt{n_1}(Q_{n_1}-Q_1)(\tilde{\theta}_{g_1,\hat{F}_{n_1}}) \\ &= \sqrt{n_1}(Q_{n_1}-Q_1)(\tilde{\theta}_{g_1,\hat{F}_{n_1}}-\tilde{\theta}_{g_1,F_1}) + \sqrt{n_1}(Q_{n_1}-Q_1)\tilde{\theta}_{g_1,F_1} \\ &= \sqrt{n_1}(Q_{n_1}-Q_1)(\tilde{\theta}_{g_1,F_1}) + o_p(1), \end{aligned}$$

which is (7). Thus based on (6) and (7), we have

$$\bar{U}_{n_1} = \sqrt{n_1}(Q_{n_1}-Q_1)\left\{K_{F_1,F_2}-\tilde{\theta}_{g_1,F_1}\right\} + o_p(1),$$

which is the first part of Theorem 1.

As pointed above, the second part of Theorem 1 can be proved similarly and in this case, we have

$$g_2(t) = h_1(t)\frac{\eta(F_1(t))-c_0}{F_2(t)} + \int_0^t h(u,t)\frac{\eta(F_1(t))-\eta(F_1(u))}{F_2(t)-F_2(u)}du - \int_t^M h(t,v)\frac{\eta(F_1(v))-\eta(F_1(t))}{F_2(v)-F_2(t)}dv - h_2(t)\frac{c_0-\eta(F_1(t))}{1-F_2(t)}.$$

References

Fang, H., Sun, J., Lee, M.-L.T., 2002. Nonparametric survival comparison for interval-censored continuous data. Statistica Sinica 12, 1073–1083.

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

Fleming, T.R., Harrington, D.P., 1991. Counting Process and Survival Analysis. John Wiley, New York.

Goggins, W.B., Finkelstein, D.M., 2000. A proportional hazards model for multivariate interval-censored failure time data. Biometrics 56, 940–943.

Groeneboom, P., 1996. Lectures on inverse problems. In: Lecture Notes in Mathematics, vol. 1648. Springer-Verlag, Berlin.

Groeneboom, P., Wellner, J.A., 1992. Information Bounds and Nonparametric Maximum Likelihood Estimation. In: DMV Seminar, Band 19, Birkhauser, New York.

Huang, J., Wellner, J.A., 1995. Asymptotic normality of the NPMLE of linear functionals for interval censored data, case I. Statistica Neerlandica 49, 153–163.

Kalbfleisch, J.D., Prentice, R.L., 2002. The Statistical Analysis of Failure Time Data, second ed. Wiley, New York.
Peto, R., Peto, J., 1972. Asymptotically efficient rank invariant test procedures. Journal of the Royal Statistical Society A 135, 185–207.
Sun, J., 1996. A nonparametric test for interval-censored failure time data with application to AIDS studies. Statistics in Medicine 15, 1387–1395.
Sun, J., 1999. A nonparametric test for current status data with unequal Censoring. Journal of the Royal Statistical Society: Series B 61, 243–250.
Sun, J., 2006. Statistical Analysis of Interval-Censored Failure Time Data. Springer, New York.
Sun, J., Zhao, Q., Zhao, X., 2005. Generalized long-rank tests for interval-censored failure time data. Scandinavian Journal of Statistics 32, 49–57.
Turnbull, B.W., 1976. The empirical distribution with arbitrarily grouped censored and truncated data. Journal of the Royal Statistical Society: Series B 38, 290–295.
van der Vaart, A.W., Wellner, J.A., 1996. Weak Convergence and Empirical Processes. Springer, New York.
Zhao, Q., Sun, J., 2004. Generalized log-rank test for mixed-censored failure time data. Statistics in Medicine 23, 1621–1629.
Zhao, X., Zhao, Q., Sun, J., Kim, J.S., 2008. Generalized log-rank tests for partly interval-censored failure time data. Biometrical Journal 50, 375–385.