Ann Inst Stat Math (2009) 61:461–476
DOI 10.1007/s10463-007-0151-3

A class of rank-based test for left-truncated and right-censored data

Pao-sheng Shen

Received: 1 May 2006 / Revised: 14 May 2007 / Published online: 14 September 2007

Abstract A class of rank-based tests is proposed for the two-sample problem with left-truncated and right-censored data. The class contains as special cases the extension of log-rank test and Gehan test. The asymptotic distribution theory of the test is presented. The small-sample performance of the test is investigated under a variety of situations by means of Mone Carlo simulations.

Keywords Two-sample tests · Left-truncated and right-censored data

1 Introduction

In survival studies, the observed data is typically censored and/or truncated. Left truncation and right censoring together occur naturally in cohort follow-up studies (see Wang 1991). Hypothesis testing for the comparison of two groups is important in many situations. Consider the following application:

Example 1 (prevalent cohort data) Suppose that the disease population in a certain area is a representative sample from a large disease population. The target interest of a research project is to compare the natural history of the disease (such as acquired immune deficiency syndrome (AIDS)) for two subgroups of individuals who developed the disease during the calendar time period (τ_1, τ_2), $\tau_1 < \tau_2$. Consider the sampling under which all of the individuals in the area who have experienced an initial event E_1 (such as being diagnosed as HIV-positive) between τ_1 and τ_2 and have not experienced a second event E_2 (such as AIDS) are recruited at the time τ $(\tau > \tau_2)$ for a prospective follow-up study. The follow-up study is terminated at τ^* $(\tau^* > \tau)$.

P.-s. Shen (✉)
Department of Statistics, Tunghai University, Taichung 40704, Taiwan
e-mail: psshen@thu.edu.tw

For group i $(i = 1, 2)$, let T_i^* be the time from the onset of E_1 to E_2, V_i^* denote the time from onset of E_1 to τ, and C_i^* denote the time from E_1 to censoring. Left truncation arises because those individuals who have experienced E_2 as well as E_1 prior to time of recruitment (τ) are excluded from the cohort. The presence of right censoring is due to loss to follow-up, or simply to end of study at τ^*. Clearly, the calendar time of the potential censoring point must be greater than τ. Therefore, the relationship $C_i^* \geq V_i^*$ is always satisfied. When the only reason of censoring is due to the end of the study, we have $C_i^* = V_i^* + \tau^* - \tau$. Besides, when τ^* is sufficiently large, the only reason of censoring is caused by loss to follow-up.

Example 2 (data on residents of a retirement community) Consider survival data for elderly residents of a retirement community. Data on ages at death of two subgroups of individuals, who were in residence during the period τ to τ^* were reported. The life length is left-truncated because an individual must survive to a sufficient age to enter the retirement community. Individuals who die at an early age are excluded from the study. The truncation variables V_i^*'s are ages at entry, target variables T_i's are ages at death, and censoring variables C_i^*'s are ages at the end of study at τ^*. Note that the relationship $C_i^* \geq V_i^*$ is always satisfied.

The Mann–Whitney test (Mann and Whitney 1947) and the Wilcoxon test (Wilcoxon 1945) are two closely associated nonparametric two-sample tests for the case of complete data. Extensions of nonparametric two-sample tests have been developed for the right-censored data. The Gehan test (Gehan 1965) is an extension of the Mann–Whitney test that allows right-censored data. The log-rank test (Peto and Peto 1972) is an extension of the Mantel–Haenzel test (Mantel and Haenzel 1959). Based on the integrated weighted difference in Kaplan–Meier estimators (1958), Pepe and Fleming (1989) proposed a class of distance test for right-censored data. For the case of randomly left (or right) truncated data, Lagakos et al. (1988) studied a weighted log-rank test. Bilker and Wang (1989) proposed a semiparametric test for the case when truncation distribution is parameterized.

In Sect. 2, a class of rank-based test is proposed for the two-sample problem with left-truncated and right-censored data. The class contains as special cases the extension of log-rank test (Mantel 1966; Lagakos et al. 1988) and Gehan test (Gehan 1965). The asymptotic distribution theory of the test is derived using martingale theory. In Sect. 3, the small-sample performance of the test is investigated under a variety of situations by means of Mone Carlo simulations.

2 A class of tests

Let (T_i^*, C_i^*, V_i^*) $(i = 1, 2)$ be a continuous random vector from subgroup i such that (C_i^*, V_i^*) is independent of T_i^*. For subgroup i $(i = 1, 2)$, let F_i, Q_i and G_i denote the distribution function of T_i^*, C_i^* and V_i^*, respectively. For left-truncated and right-censored data, one can observe nothing if $T_i^* < V_i^*$ and observe (X_i^*, δ_i^*), with $X_i^* = \min(T_i^*, C_i^*)$ and $\delta_i^* = I_{[T_i^* \leq C_i^*]}$, if $X_i^* \geq V_i^*$. For any distribution function H denote the left and right endpoints of its support by $a_H = \inf\{t : H(t) > 0\}$ and $b_H = \inf\{t : H(t) = 1\}$, respectively. Woodroofe (1985) pointed out that if $a_{G_i} \leq \min(a_{F_i}, a_{Q_i})$ $(i = 1, 2)$ and $b_{G_i} \leq \min(b_{F_i}, b_{Q_i})$ $(i = 1, 2)$, then F_i, Q_i

and G_i are all identifiable. In this note, for technical convenience, we assume that $a_{F_1} = a_{F_2} = a_F$, and $b_{F_1} = b_{F_2} = b_F$.

For $i = 1, 2$, $j = 1, \ldots, n_i$, let $(X_{ij}, \delta_{ij}, V_{ij})$ denote the left-truncated and right-censored sample from subgroup i.

For $i = 1, 2$, let $\hat{R}_i(u) = \sum_{j=1}^{n_i} I_{[V_{ij}<u\leq X_{ij}]}$ and $N_i(u) = \sum_{j=1}^{n_i} I_{[X_{ij}\leq u, \delta_{ij}=1]}$. The nonparametric maximum likelihood estimate (NPMLE) of $F_i(x)$ (Wang 1987) is given by

$$\hat{F}_i(x) = 1 - \prod_{u\leq x}\Big[1 - \mathrm{d}\hat{\Lambda}_i(u)\Big],$$

where $\mathrm{d}\hat{\Lambda}_i(u) = \frac{\mathrm{d}N_i(u)}{\hat{R}_i(u)}$, $\mathrm{d}N_i(u) = N_i(u) - N_i(u-)$.

The main null hypothesis of interest is $H_0 : S_1(t) = S_2(t)$ for all t, and the alternative hypothesis can be two-sided $H_a : S_1(t) \neq S_2(t)$ or one-sided (e.g. $H_a : S_1(t) < S_2(t)$), where $S_i(t) = 1 - F_i(t)$ $(i = 1, 2)$, Based on $d\hat{\Lambda}_i(u)$, we define a class of rank-based statistics as

$$L_W = \sqrt{\frac{n_1 n_2}{n}}\left[\int_0^{\hat{b}} \hat{W}(t)\mathrm{d}\hat{\Lambda}_1(t) - \int_0^{\hat{b}} \hat{W}(t)\mathrm{d}\hat{\Lambda}_2(t)\right], \tag{1}$$

where $n = n_1 + n_2$, $\hat{b} = \min(\hat{b}_1, \hat{b}_2)$, where $\hat{b}_i = \sup\{t : \hat{S}_i(t) > 0\}$, $\hat{S}_i(t)$ $(i = 1, 2)$ is the product-limit estimate of $S_i(t)$ based on the sample from subgroup i, and $\hat{W}(t)$ is a random weight function. Let $\hat{R}(t) = \hat{R}_1(t) + \hat{R}_2(t)$. When $\hat{W}(t) = \frac{n}{n_1 n_2}\frac{\hat{R}_1(t)\hat{R}_2(t)}{\hat{R}(t)}$ and $V_i^* = 0$ $(i = 1, 2)$, the L_W test is reduced to the log-rank statistics proposed by Lagakos et al. (1988). When $\hat{W}(t) = \hat{R}_1(t)\hat{R}_2(t)/(n_1 n_2)$ $(i = 1, 2)$, the Eq. (1) is equivalent to

$$\sqrt{\frac{n_1 n_2}{n}}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\frac{U_{jk}}{n_1 n_2},$$

where

$$U_{jk} =_{[X_{2k}\geq X_{1j}, X_{1j}>V_{2k}, \delta_{1j}=1]} - I_{[X_{2k}\leq X_{1j}, X_{2k}>V_{1j}, \delta_{2k}=1]}.$$

In this case, when $V_i^* = 0$ $(i = 1, 2)$, the L_W test is reduced to the Gehan (1965) test.

To derive the limiting null distribution of the L_W test, we need the following assumptions:

(a) $\int_{a_F}^{t} 1/R_i^*(u) f(u)\mathrm{d}u \to 0$ as $t \to a_F$, where $R_i^*(t) = P(V_i^* \leq t \leq C_i^*)$ $(i = 1, 2)$ and $f(t)$ is the probability density function of T_i^*.

(b) Let $\mathcal{F}_i(t)$ $(i = 1, 2)$ denote the complete σ-field generated by

$$\{V_{ij}, I_{[V_{ij}<u\leq X_{ij}]}, I_{[V_{ij}<X_{ij}]}, \delta_{ij} I_{[V_{ij}<X_{ij}\leq u]}, I_{[V_{ij}<X_{ij}\leq s]}, u \leq s\}.$$

The weight function $\hat{W}(t)$ is a locally bounded predictable $\mathcal{F}_i(t-)$ process with

$$|\hat{W}(t)| \le \tau[n_i^{-1}\hat{R}_i(t)\hat{S}_i(t-)]^{\frac{1}{2}+\delta} \quad t \in (a_F, \hat{b}) \tag{2}$$

for some constants τ and δ, and

$$\sup_{a_F<t<b_F} |\hat{W}(t) - W(t)| \xrightarrow{p} 0,$$

where $W(t)$ is some deterministic function.
(c) For $i = 1, 2$ there exists a non-negative function $H_i(t)$, such that

$$\sup_{a_F<t<b_F} \left| \frac{[\hat{W}(t)]^2 I_{[\hat{R}_i(t)>0]}}{n_i^{-1}\hat{R}_i(t)} - H_i(t) \right| \xrightarrow{p} 0, \quad n_i \to \infty.$$

The following theorem derives the limiting null distribution of the L_W test.

Theorem 1 *Suppose that* $\hat{b} \to b$ *such that* $S_i(b) > 0$ *for* $i = 1, 2$. *Under the null hypothesis* $H_0 : S_1(t) = S_2(t)$, *assumptions (a), (b) and (c), then*

$$L_W \xrightarrow{d} N(0, \sigma^2),$$

where

$$\sigma^2 = p_2\left[\int_0^b H_1(t)\lambda_1(t)\mathrm{d}t\right] + p_1\left[\int_0^b H_2(t)\lambda_2(t)\mathrm{d}t\right].$$

Proof To prove Theorem 1, we make use of martingale theory and stochastic integral representation similar to those used in the censored case (see Fleming and Harrington 1991; Anderson et al. 1993).

First, for $i = 1, 2$, let $M_i(t) = N_i(t) - \int_{a_F}^t \hat{R}_i(s)\lambda_i(s)\mathrm{d}s$, where λ_i $(i = 1, 2)$ denote the hazard function of $F_i(t)$ $(i = 1, 2)$. According to the results of Lemma 5 of Lai and Ying (1991), $M_i(t), t \in (a_F, b)$ is a zero-mean martingale with respect to the filtration $\mathcal{F}_i(t)$. According to theorem 2.6.1 in Fleming and Harrington (1991), the predictable variation process of $M_i(t)$ is given by

$$\langle M_i(t)\rangle = \int_{a_F}^t \hat{R}_i(s)\lambda_i(s)\mathrm{d}s.$$

The L_W test can be written as

$$L_W = A_n + B_n,$$

where

$$A_n = \sqrt{\frac{n_1 n_2}{n}}\left\{n_1^{-1}\int_0^{\hat{b}} \hat{K}_1(t)\mathrm{d}M_1(t) - n_2^{-1}\int_0^{\hat{b}} \hat{K}_2(t)\mathrm{d}M_2(t)\right\}$$

and

$$B_n = \sqrt{\frac{n_1 n_2}{n}} \left\{ \int_0^{\hat{b}} \hat{W}(t) I_{[\hat{R}_1(t)>0]} \lambda_1(t)\mathrm{d}t - \int_0^{\hat{b}} \hat{W}(t) I_{[\hat{R}_2(t)>0]} \lambda_2(t)\mathrm{d}t \right\},$$

where $\hat{K}_i(t) = n_i \hat{W}(t) I_{[\hat{R}_i(t)>0]}/\hat{R}_i(t)$ $(i = 1, 2)$.

Let $U_i(t) = n_i^{-\frac{1}{2}} \int_0^t \hat{K}_i(s) dM_i(s)$. Then $U_i(t)$ is a martingale with predictable variation process

$$\langle U_i(t) \rangle = \int_0^t [\hat{K}_i(s)]^2 n_i^{-1} \hat{R}_i(s) \lambda_i(s)\mathrm{d}s.$$

Let $\hat{p}_i = n_i/n$. Then

$$A_n = \left[\sqrt{\hat{p}_2} U_1(\hat{b}) - \sqrt{\hat{p}_1} U_2(\hat{b}) \right].$$

Next, under the assumption (b), we have

$$\begin{aligned} E\left[\int_{a_F}^u [\hat{K}_i(t)]^2 n_i^{-1} \hat{R}_i(t) \lambda_i(t)\mathrm{d}t \right] &\le E\left[\int_{a_F}^u \frac{[\hat{W}(t)]^2 I_{[\hat{R}_i(t)>0]}}{n_i^{-1} \hat{R}_i(t)} \lambda_i(t)\mathrm{d}t \right] \\ &\le \tau^2 E\left[\int_{a_F}^u \frac{\hat{S}_i(t-)}{S_i(t)} f_i(t)\mathrm{d}t \right] \\ &\le \frac{\tau^2 F_i(u)}{S_i(u)} \to 0 \quad \text{as } u \to a_F. \end{aligned}$$

Hence, for any $\epsilon > 0$,

$$\lim_{u \downarrow a_F} \limsup_{n_i \to \infty} P\left(\int_{a_F}^u [\hat{K}_i(t)]^2 n_i^{-1} \hat{R}_i(t) \mathrm{d}\Lambda_i(t) > \epsilon \right) = 0. \tag{3}$$

Similarly, since $E[\hat{S}_i(t)/S_i(t)]$ is bounded for $t \in (u, b)$, we have

$$\begin{aligned} E\left[\int_u^{\hat{b}} [\hat{K}_i(t)]^2 n_i^{-1} \hat{R}_i(t) \lambda_i(t)\mathrm{d}t \right] &\le \tau^2 E\left[\int_u^b \frac{\hat{S}_i(t-)}{S_i(t)} f_i(t)\mathrm{d}t \right] \\ &\le \tau^2 E\left[\sup_{t \in (u,b)} \frac{\hat{S}_i(t-)}{S_i(t)} \right] \\ &\quad \times [S_i(u) - S_i(b)] \to 0 \quad \text{as } u \to b. \end{aligned}$$

Hence, any $\epsilon > 0$,

$$\lim_{u \uparrow b} \limsup_{n_i \to \infty} P\left(\int_u^b [\hat{K}_i(t)]^2 n_i^{-1} \hat{R}_i(t) \mathrm{d}\Lambda_i(t) > \epsilon \right) = 0. \tag{4}$$

Assuming $\hat{p}_i \to p_i$ $(i = 1, 2)$, it follows from (3), (4) and Theorem 5 of Lai and Ying (1991) that under H_0, $A_n \xrightarrow{d} N(0, \sigma^2)$.

Next, we show that $B_n \xrightarrow{p} 0$. Under H_0, i.e $\lambda_i(t) = \lambda(t)$ $(i = 1, 2)$, it suffices to show that

$$\lim_{n\to\infty} \sqrt{\frac{n_1 n_2}{n}} E\left[\int_0^{\hat{b}} W(t)(I_{[\hat{R}_1(t)>0]} - I_{[\hat{R}_2(t)>0]})\lambda(t)\mathrm{d}t\right] = 0.$$

Note that assumption (b) implies that there exists some constant c and ϵ such that

$$|W(t)| \le c[R_i^*(t)]^{(1/2)+\epsilon}[S_i(t-)]^{1+\epsilon} \quad t \in (a_F, b).$$

Hence,

$$\begin{aligned}
&E\left[\int_0^{\hat{b}} W(t)(I_{[\hat{R}_1(t)>0]} - I_{[\hat{R}_2(t)>0]})\lambda(t)\mathrm{d}t\right] \\
&\le E\left[\int_0^{\hat{b}} W(t)(I_{[\hat{R}_1(t)>0,\hat{R}_2(t)=0]} + I_{[\hat{R}_2(t)>0,\hat{R}_1(t)=0]})\lambda(t)\mathrm{d}t\right] \\
&\le c\int_0^b [(R_1^*(t))^{\frac{1}{2}} S(t-)P(\hat{R}_1(t) = 0) + (R_2^*(t))^{\frac{1}{2}} S(t-)P(\hat{R}_2(t) = 0)]\lambda(t)\mathrm{d}t. \\
&\le c\int_0^b [P(\hat{R}_1(t) = 0) + P(\hat{R}_2(t) = 0)]f(t)\mathrm{d}t.
\end{aligned}$$

Let $\beta_i = P(V_i^* < X_i^*)$ for $i = 1, 2$. Since $P(\hat{R}_i(t) = 0) = (1 - R_i^*(t)S(t-)/\beta_i)^{n_i}$ for $i = 1, 2$, it suffices to show that

$$\lim_{n_i\to\infty} \sqrt{n_i} \int_0^b \left(1 - \frac{R_i^*(t)S(t-)}{\beta_i}\right)^{n_i} f(t)\mathrm{d}t = 0.$$

Based on assumption (a), there exists t_0 such that for $t \in (a_F, t_0)$, $F(t) < R_i^*(t)$. Since $R_i^*(t)S(t-)$ is continuous, we can require that $R_i^*(t)S(t-)$ is increasing in (a_F, t_0).

Now,

$$\sqrt{n_i} \int_{a_F}^b \left(1 - \frac{R_i^*(t)S(t-)}{\beta_i}\right)^{n_i} f(t)\mathrm{d}t = C_{n_i} + D_{n_i},$$

where

$$C_{n_i} = \sqrt{n_i} \int_{t_0}^b \left(1 - \frac{R_i^*(t)S(t-)}{\beta_i}\right)^{n_i} f(t)\mathrm{d}t,$$

and

$$D_{n_i} = \sqrt{n_i} \int_{a_F}^{t_0} \left(1 - \frac{R_i^*(t)S(t-)}{\beta_i}\right)^{n_i} f(t)\mathrm{d}t$$

First,

$$\begin{aligned} C_{n_i} &= \sqrt{n_i} \int_{t_0}^{b} \left[\frac{P(t < V_i^* < \min(C_i^*, T_i^*))}{\beta_i}\right]^{n_i} f(t)\mathrm{d}t \\ &\le \sqrt{n_i}\left[\frac{P(t_0 < V_i^* < \min(C_i^*, T_i^*))}{\beta_i}\right]^{n_i} \to 0 \quad \text{as } n_i \to \infty. \end{aligned}$$

Next, consider a monotone sequence $t_{n_i} = F^{-1}(1/{n_i}^{0.5+\varepsilon})$, $0 < \varepsilon < 0.5$, converge to a_F. For sufficiently large n_i, we have

$$\begin{aligned} &\sqrt{n_i} \int_{a_F}^{t_{n_i}} \left(1 - \frac{R_i^*(t)S(t-)}{\beta_i}\right)^{n_i} f(t)\mathrm{d}t \\ &\le \sqrt{n_i} F(F^{-1}(1/{n_i}^{0.5+\varepsilon})) = 1/{n_i}^{\varepsilon} \to 0 \quad \text{as } n_i \to \infty, \end{aligned} \tag{5}$$

and

$$\begin{aligned} &\sqrt{n_i} \int_{t_{n_i}}^{t_0} \left(1 - \frac{R_i^*(t)S(t-)}{\beta_i}\right)^{n_i} f(t)\mathrm{d}t \le \sqrt{n_i}\left(1 - \frac{R_i^*(t_{n_i})S(t_{n_i}-)}{\beta_i}\right)^{n_i} \\ &\le \sqrt{n_i}[1 - F(t_{n_i})S(t_{n_i}-)]^{n_i} = \sqrt{n_i}\left(1 - \frac{n_i^{0.5+\varepsilon} - 1}{n_i^{1+2\varepsilon}}\right)^{n_i} \to 0 \quad \text{as } n_i \to \infty. \end{aligned} \tag{6}$$

By (5) and (6), it follows that $D_{n_i} \to 0$. Hence, we have $B_n \xrightarrow{p} 0$ and $L_W \xrightarrow{d} N(0, \sigma^2)$. The proof is completed.

Under the assumptions (b) and (c), a consistent estimator of σ^2 can be found by substituting estimators for p_i, $H_i(t)$ and $\Lambda_i(t)$ in σ^2. Thus, a consistent estimator of σ^2 is given by

$$\hat{\sigma}^2 = \hat{p}_2\left[\int_0^{\hat{b}} \hat{H}_1(t)\mathrm{d}\hat{\Lambda}_1(t)\right] + \hat{p}_1\left[\int_0^{\hat{b}} \hat{H}_2(t)\mathrm{d}\hat{\Lambda}_2(t)\right],$$

where $\hat{H}_i(t) = \frac{n_i[\hat{W}_i(t)]^2 I_{[\hat{R}_i(t)>0]}}{\hat{R}_i(t)}$ $(i = 1, 2)$. Based on Theorem 1, we have

$\lim_{n\to\infty} L_W/\hat{\sigma} \xrightarrow{d} N(0, 1)$. Given the nominal level α, for two-sided test, H_0 would be rejected if $|L_W|/\hat{\sigma} > z_{\alpha/2}$, where $z_{\alpha/2}$ is the upper $\alpha/2$ percentile of the standard normal distribution.

Note that stability constraint (2) requires that $\hat{W}_i(t)$ $(i = 1, 2)$ is a function of $\hat{R}_1(t)$ and $\hat{R}_2(t)$. When there is no truncation, (2) is similar to the constraint (3.2) of Pepe and Fleming (1989). Let $\pi_i(t) = R_i^*(t)S_i(t-) = P(V_i^* \le t \le C_i^*)S_i(t-)$ $(i = 1, 2)$. For the log-rank test, $\hat{W}(t)$ is akin to a geometric average of the two $\pi_i(t)$'s estimators, and satisfy (2). For right-censored data, Fleming and Harrington (1981) and Harrington and Fleming (1982) proposed a very general class of tests that includes, as special cases, the log-rank test. Similar to their approach, we can use the following weight functions:

$$\hat{W}(t) = [\hat{S}(t-)]^r[1 - \hat{S}(t-)]^s \hat{W}_L(t), \tag{7}$$

where $\hat{W}_L(t) = \left[\frac{n}{n_1 n_2}\right]\frac{\hat{R}_1(t)\hat{R}_2(t)}{\hat{R}(t)}, r \geq 0, s \geq 0$, and $\hat{S}(t-)$ is the product-limit estimate based on the combined sample. Note that a proper choice of (r, s) depends on what type of alternatives to expect. For example, when two survival curves are close except for the late times, a test should gives more weight to late times, i.e. $(r, s) = (0, 1)$

3 Simulations

To evaluate the performance of the L_W test, Monte Carlo simulations are carried out to study the statistical power and the type I error under a variety of situations. The properties of the L_W test are compared by using weight function of $\hat{R}_1(t)\hat{R}_2(t)/(n_1 n_2)$ (i.e. the extension of Gehan test) and (7) with $(r, s) = (0, 0), (0.5, 0.5), (1, 0), (0, 1)$. Note that the $(r, s) = (0, 0)$ corresponds to the weight function of the extension of log-rank test. The sample sizes are set at $n_1 = n_2 = 100, 200$. The number of iterations in each simulation study is 5,000.

3.1 Estimated Type I error

The goal of the first set of our simulations is to assess the performance of the L_W test under the null hypothesis. Data sets are generated in the following manner:

Situation 1 Both T_1^* and T_2^* follow an exponential distribution with mean of 5.0. Both V_1^* and V_2^* follow an exponential distribution with mean of μ_g. The C_i^* is defined by $C_i^* = D_i^* + V_i^*$, where D_i^* is independent of V_i^* and exponentially distributed with mean of μ_d. The parameter is set at the combination of $\mu_g = 0.25, 2.0, 8.0$ and $\mu_d = 0.25, 2.0, 8.0$. The estimated Type I error is calculated as the proportion of 5000 repeated random samples in which we reject the null hypothesis at 0.05 significance level. The estimated Type I error rates are presented in Table 1. Table 1 also shows the proportion of truncation $P(V_i^* > T_i^*)$ (denoted by p_t) and that of censoring $P(C_i^* < T_i^*)$ (denoted by p_c).

Situation 2 The parameters are the same as those used in Situation 1 except that V_i^* is uniformly distributed $U(0, \mu_g)$ and D_i^* is a constant with the value of μ_d. The estimated Type I error rates are presented in Table 2.

Tables 1 and 2 show that under H_0, the L_W test performs fairly well. The estimated Type I errors of all the tests are close to 0.05.

Table 1 Estimated Type I error for L_W test in Situation 1

μ_g	μ_d	n_i	p_t	p_c	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.25	0.25	100	0.05	0.91	0.052	0.054	0.035	0.054	0.052
0.25	0.25	200	0.05	0.91	0.046	0.044	0.043	0.048	0.045
0.25	2.00	100	0.05	0.68	0.054	0.055	0.038	0.063	0.039
0.25	2.00	200	0.05	0.68	0.045	0.045	0.049	0.043	0.042
0.25	8.00	100	0.05	0.37	0.060	0.055	0.043	0.060	0.058
0.25	8.00	200	0.05	0.37	0.057	0.050	0.053	0.055	0.048
2.00	0.25	100	0.29	0.68	0.056	0.042	0.039	0.058	0.046
2.00	0.25	200	0.29	0.68	0.055	0.057	0.045	0.052	0.046
2.00	2.00	100	0.29	0.51	0.042	0.044	0.048	0.039	0.041
2.00	2.00	200	0.29	0.51	0.046	0.045	0.051	0.044	0.044
2.00	8.00	100	0.29	0.28	0.055	0.038	0.055	0.042	0.039
2.00	8.00	200	0.29	0.28	0.053	0.052	0.052	0.052	0.055
8.00	0.25	100	0.62	0.36	0.050	0.044	0.033	0.046	0.051
8.00	0.25	200	0.62	0.36	0.046	0.048	0.046	0.052	0.050
8.00	2.00	100	0.62	0.27	0.044	0.044	0.044	0.042	0.048
8.00	2.00	200	0.62	0.27	0.046	0.046	0.052	0.044	0.050
8.00	8.00	100	0.61	0.15	0.046	0.056	0.040	0.042	0.060
8.00	8.00	200	0.62	0.15	0.053	0.049	0.052	0.054	0.059

3.2 Estimated statistical power

To study the statistical power of the L_W test, we consider the following six situations:

Situation 3 We consider a situation where two survival curves cross. The T_1^* follows an exponential distribution with mean of 5. The T_2^* follows an exponential ditribution with mean of 2. However, if the T_2^* is greater than 3.0, then T_2^* is re-generated to follow an exponential distribution with mean of 50. The V_i^* is generated from an exponential distribution with mean of μ_g. For $i = 1, 2$, $C_i^* = D_i^* + V_i^*$, where D_i^* is independent of V_i^* and exponentially distributed with mean of μ_d. The parameter (μ_g, μ_d) is set at the combination of $\mu_g = 0.5, 2.0, 8.0$ and $\mu_d = 20, 10, 5$. The statistical power of the L_W test in this situation is presented in Table 3.

Situation 4 The parameters are the same as those used in Situation 3 except that V_i^* is uniformly distributed $U(0, \mu_g)$ and D_i^* is a constant with the value of μ_d. The statistical power of the L_W test in this situation is presented in Table 4.

Tables 3 and 4 also show the proportion of truncation $1 - \beta_i$ (denoted by p_{t_i}) and the proportion of censoring $P(C_i^* \leq T_i^*)$ (denoted by p_{c_i}).

Tables 3 and 4 show that the cross of survival curve can leave the log-rank test (i.e $(r, s) = (0, 0)$) little power to detect the overall difference. However, when the truncation is severe and censoring is light (e.g. in Table 3: $p_{t_1} = 0.62$, $p_{t_2} = 0.71$;

Table 2 Estimated Type I error for L_W test in Situation 2

μ_g	μ_d	n_i	p_t	p_c	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.25	0.25	100	0.03	0.92	0.062	0.051	0.050	0.062	0.047
0.25	0.25	200	0.03	0.92	0.058	0.049	0.047	0.054	0.046
0.25	2.00	100	0.03	0.65	0.061	0.062	0.063	0.060	0.053
0.25	2.00	200	0.03	0.65	0.054	0.055	0.055	0.052	0.048
0.25	8.00	100	0.03	0.20	0.063	0.050	0.055	0.054	0.061
0.25	8.00	200	0.03	0.20	0.055	0.047	0.049	0.047	0.053
2.00	0.25	100	0.18	0.78	0.062	0.054	0.063	0.066	0.047
2.00	0.25	200	0.18	0.78	0.054	0.049	0.055	0.057	0.047
2.00	2.00	100	0.18	0.55	0.061	0.053	0.048	0.052	0.063
2.00	2.00	200	0.18	0.55	0.055	0.050	0.051	0.051	0.054
2.00	8.00	100	0.18	0.16	0.063	0.063	0.057	0.055	0.063
2.00	8.00	200	0.18	0.16	0.055	0.054	0.053	0.052	0.053
8.00	0.25	100	0.50	0.47	0.053	0.038	0.061	0.057	0.053
8.00	0.25	200	0.50	0.47	0.049	0.047	0.055	0.052	0.051
8.00	2.00	100	0.50	0.33	0.044	0.049	0.062	0.053	0.039
8.00	2.00	200	0.50	0.33	0.046	0.052	0.054	0.050	0.044
8.00	8.00	100	0.50	0.10	0.053	0.047	0.038	0.045	0.039
8.00	8.00	200	0.50	0.10	0.051	0.048	0.045	0.048	0.045

$p_{c_1} = 0.08$, $p_{c_2} = 0.14$), the power of the log-rank test is high since T_1^* and T_2^* (left-truncated by V_1^* and V_2^*) are observable only after the crossing of two curves. When truncation is not severe, the L_W test with weight function of $\hat{S}(t-)\hat{W}_L(t)$ (i.e. $(r, s) = (1, 0)$) has greater power than the other tests.

Situation 5 We consider a situation where two survival curves are close except for the late times. The T_1^* follows an exponential distribution with mean of 4.0. The T_2^* follows an exponential ditribution with mean of 3.5. However, if the T_2^* is greater than 5.0, then T_2^* is re-generated to follow an exponential distribution with mean of 50. The distributions of V_i^* and C_i^* are same as those used in Situation 1. The statistical power of the L_W test in this situation is presented in Table 5.

Situation 6 The parameters are the same as those used in Situation 5 except that V_i^* is uniformly distributed $U(0, \mu_g)$ and D_i^* is a constant with the value of μ_d. The statistical power of the L_W test in this situation is presented in Table 6.

Tables 5 and 6 show that the L_W test with weight function of $[1 - \hat{S}(t-)]\hat{W}_L(t)$ (i.e. $(r, s) = (0, 1)$) has greater power than the other test. This is a typical situation where a test should give more weight to late times than to early times. When truncation is not severe, the power of both log-rank and Gehan tests is low.

Table 3 Power of the L_W test in Situation 3

μ_g	μ_d	n_i	p_{t_1}	p_{c_1}	p_{t_2}	p_{c_2}	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.5	20	100	0.09	0.18	0.20	0.19	0.205	0.981	0.461	0.470	0.906
0.5	20	200	0.09	0.18	0.20	0.19	0.356	1.000	0.712	0.785	1.000
0.5	10	100	0.09	0.30	0.20	0.24	0.466	0.982	0.082	0.662	0.981
0.5	10	200	0.09	0.30	0.20	0.24	0.778	1.000	0.107	0.928	1.000
0.5	5	100	0.09	0.46	0.20	0.31	0.778	0.980	0.134	0.828	0.992
0.5	5	200	0.09	0.46	0.20	0.31	0.966	1.000	0.235	0.983	1.000
2.0	20	100	0.28	0.14	0.49	0.17	0.168	0.740	0.952	0.060	0.098
2.0	20	200	0.28	0.14	0.49	0.17	0.308	0.956	1.000	0.066	0.183
2.0	10	100	0.28	0.24	0.49	0.20	0.042	0.788	0.603	0.128	0.306
2.0	10	200	0.28	0.24	0.49	0.20	0.051	0.979	0.902	0.216	0.578
2.0	5	100	0.28	0.36	0.49	0.24	0.177	0.815	0.116	0.334	0.632
2.0	5	200	0.28	0.36	0.49	0.24	0.332	0.987	0.214	0.562	0.905
8.0	20	100	0.62	0.08	0.71	0.14	0.962	0.212	1.000	0.386	0.742
8.0	20	200	0.62	0.08	0.71	0.14	1.000	0.381	1.000	0.674	0.956
8.0	10	100	0.62	0.13	0.71	0.17	0.706	0.264	1.000	0.172	0.334
8.0	10	200	0.62	0.13	0.71	0.17	0.954	0.498	1.000	0.315	0.584
8.0	5	100	0.62	0.19	0.71	0.19	0.218	0.332	0.831	0.053	0.054
8.0	5	200	0.62	0.19	0.71	0.19	0.430	0.596	0.992	0.066	0.070

Table 4 Power of the L_W test in Situation 4

μ_g	μ_d	n_i	p_{t_1}	p_{c_1}	p_{t_2}	p_{c_2}	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.5	20	100	0.05	0.02	0.12	0.15	0.120	0.993	0.825	0.384	0.935
0.5	20	200	0.05	0.02	0.12	0.15	0.129	1.000	0.960	0.524	1.000
0.5	10	100	0.05	0.13	0.12	0.18	0.355	0.997	0.184	0.587	0.962
0.5	10	200	0.05	0.13	0.12	0.18	0.567	1.000	0.352	0.846	1.000
0.5	5	100	0.05	0.35	0.12	0.20	0.972	0.998	0.624	0.970	0.992
0.5	5	200	0.05	0.35	0.12	0.20	1.000	1.000	0.888	1.000	1.000
2.0	20	100	0.18	0.02	0.37	0.15	0.256	0.837	1.000	0.080	0.235
2.0	20	200	0.18	0.02	0.37	0.15	0.508	0.987	1.000	0.103	0.466
2.0	10	100	0.18	0.11	0.37	0.18	0.027	0.845	0.673	0.144	0.419
2.0	10	200	0.18	0.11	0.37	0.18	0.063	0.989	0.961	0.336	0.693
2.0	5	100	0.18	0.30	0.37	0.20	0.671	0.952	0.118	0.763	0.835
2.0	5	200	0.18	0.30	0.37	0.20	0.909	1.000	0.178	0.958	1.000
8.0	20	100	0.51	0.01	0.68	0.14	1.000	0.117	1.000	0.999	1.000
8.0	20	200	0.51	0.01	0.68	0.14	1.000	0.181	1.000	1.000	1.000
8.0	10	100	0.51	0.07	0.68	0.17	1.000	0.137	1.000	0.895	1.000
8.0	10	200	0.51	0.07	0.68	0.17	1.000	0.225	1.000	1.000	1.000
8.0	5	100	0.51	0.18	0.68	0.19	0.664	0.298	0.993	0.234	0.732
8.0	5	200	0.51	0.18	0.68	0.19	0.922	0.521	1.000	0.066	1.000

Table 5 Power of the L_W test in Situation 5

μ_g	μ_d	n_i	p_{t_1}	p_{c_1}	p_{t_2}	p_{c_2}	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.5	20	100	0.11	0.15	0.13	0.23	0.147	0.092	0.690	0.058	0.075
0.5	20	200	0.11	0.15	0.13	0.23	0.268	0.132	0.936	0.054	0.098
0.5	10	100	0.11	0.26	0.13	0.30	0.065	0.094	0.328	0.038	0.106
0.5	10	200	0.11	0.26	0.13	0.30	0.068	0.156	0.538	0.048	0.151
0.5	5	100	0.11	0.40	0.13	0.40	0.054	0.120	0.176	0.076	0.132
0.5	5	200	0.11	0.40	0.13	0.40	0.066	0.168	0.288	0.098	0.190
2.0	20	100	0.33	0.11	0.37	0.20	0.463	0.058	0.954	0.150	0.052
2.0	20	200	0.33	0.11	0.37	0.20	0.728	0.058	0.997	0.244	0.062
2.0	10	100	0.33	0.19	0.37	0.25	0.218	0.046	0.718	0.084	0.044
2.0	10	200	0.33	0.19	0.37	0.25	0.381	0.086	0.934	0.103	0.049
2.0	5	100	0.33	0.30	0.37	0.31	0.067	0.074	0.201	0.056	0.074
2.0	5	200	0.33	0.30	0.37	0.31	0.071	0.104	0.362	0.049	0.086
8.0	20	100	0.67	0.06	0.64	0.16	0.944	0.058	1.000	0.478	0.530
8.0	20	200	0.67	0.06	0.64	0.16	0.995	0.052	1.000	0.756	0.786
8.0	10	100	0.67	0.10	0.64	0.19	0.782	0.047	0.976	0.327	0.272
8.0	10	200	0.67	0.10	0.64	0.19	0.968	0.054	1.000	0.564	0.442
8.0	5	100	0.67	0.15	0.64	0.23	0.380	0.043	0.736	0.156	0.078
8.0	5	200	0.67	0.15	0.64	0.23	0.605	0.046	0.972	0.249	0.112

Situation 7 We consider a situation where two survival curves have proportional hazard functions. The T_1^* and T_2^* follow an exponential distribution with mean 4 and 6, respectively. The V_1^* and V_2^* follow an exponential distribution with mean of μ_g. The D_1^* and D_2^* are exponentially distributed with mean of μ_d. The parameter (μ_g, μ_d) was set at the combination of $\mu_g = 0.5, 2.0, 8.0$ and $\mu_d = 2, 8, 12$. The statistical power of the L_W test in this situation is presented in Table 7.

Situation 8 The parameters are the same as those used in Situation 7 except that V_i^* is uniformly distributed $U(0, \mu_g)$ and D_i^* is a constant with the value of μ_d. The statistical power of the L_W test in this situation is presented in Table 8.

Tables 7 and 8 show that the log-rank test has the highest power among all the tests. In this situation, the hazard rates for the two groups are proportional and earlier research has shown that the log-rank test has optimal power in right censoring situation. Given truncation proportion, the power of all the tests decrease as censoring proportion increases.

4 Application to a real data set

In this section, the proposed test is illustrated through the Channing House data from Hyde (1977, 1980). Channing House is a retirement center located in Palo Atlo, Cali-

Table 6 Power of the L_W test in Situation 6

μ_g	μ_d	n_i	p_{t_1}	p_{c_1}	p_{t_2}	p_{c_2}	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.5	20	100	0.06	0.01	0.07	0.16	0.412	0.100	0.993	0.125	0.054
0.5	20	200	0.06	0.01	0.07	0.16	0.681	0.114	1.000	0.149	0.062
0.5	10	100	0.06	0.08	0.07	0.20	0.117	0.118	0.636	0.069	0.070
0.5	10	200	0.06	0.08	0.07	0.20	0.145	0.171	0.847	0.078	0.123
0.5	5	100	0.06	0.27	0.07	0.22	0.195	0.171	0.182	0.214	0.190
0.5	5	200	0.06	0.27	0.07	0.22	0.280	0.363	0.378	0.289	0.347
2.0	20	100	0.21	0.01	0.24	0.16	0.763	0.050	1.000	0.225	0.095
2.0	20	200	0.21	0.01	0.24	0.16	0.962	0.068	1.000	0.376	0.061
2.0	10	100	0.21	0.06	0.24	0.19	0.292	0.053	0.837	0.143	0.048
2.0	10	200	0.21	0.06	0.24	0.19	0.478	0.115	1.000	0.275	0.054
2.0	5	100	0.21	0.23	0.24	0.22	0.057	0.138	0.164	0.072	0.102
2.0	5	200	0.21	0.23	0.24	0.22	0.114	0.169	0.277	0.125	0.261
8.0	20	100	0.57	0.00	0.60	0.15	1.000	0.073	1.000	0.916	0.970
8.0	20	200	0.57	0.00	0.60	0.15	1.000	0.187	1.000	0.997	1.000
8.0	10	100	0.57	0.03	0.60	0.18	0.971	0.095	1.000	0.753	0.925
8.0	10	200	0.57	0.03	0.60	0.18	1.000	0.190	1.000	0.989	1.000
8.0	5	100	0.57	0.12	0.60	0.20	0.469	0.049	0.868	0.241	0.226
8.0	5	200	0.57	0.12	0.60	0.20	0.776	0.072	0.998	0.445	0.383

fornia. Data on ages at death of 462 individuals (97 males and 365 females), who were in residence during the period January 1964 (τ) to July 1975 (τ^*), has been reported by Hyde (1977, 1980). There were 97 males (subgroup 1: 46 died, 51 were censored) and 365 females (subgroup 2: 130 died, 235 were censored). A distinctive feature of these individuals was that all were covered by a health care program provided by the center which allowed for easy access to medical care without any additional financial burden to the residents. The lifetime of interest T_i^* is the age in months at death, the truncation time V_i^* is the age in months at entry into the community, and the censoring time C_i^* is the age in months at the end of study on July 1, 1975, or the age at withdrawal from the community. For male, the risk set of the observation $X_{1j} = 777$ and $X_{1k} = 781$ are very small ($\hat{R}_1(777) = 2$ and $\hat{R}_1(781) = 1$). Thus, testing was performed conditionally given that $T_i^* > 781$ ($i = 1, 2$). We apply $\hat{T}_H$ and log-rank tests to compare the survival functions for male and female. The null hypothesis is $H_0 : S_1^*(t) = S_2^*(t)$ and the alternative hypothesis is $H_a : S_1^*(t) < S_2^*(t)$, where $S_i^*(t) = P(T_i^* > t | T_i^* > 781)$ ($i = 1, 2$). The p-values of the $\hat{L}_W$ test are computed using the weight function in Table 1. Next, we discuss about the conditions of Theorem 2.1. Since the weight function $\hat{W}_L(t)$ satisfy (2.2), it follows that conditions (b) and (c) of Theorem 1 are satisfied by the weight functions in Table 1. Next, for female subgroup, we have an observation of $(X_{2j}, \delta_{2j}, V_{2j}) = (804, 0, 746)$. Although a similar observation is not found for male subgroup, it seems reasonable to assume that

Table 7 Power of the L_W test in Situation 7

μ_g	μ_d	n_i	p_{t_1}	p_{c_1}	p_{t_2}	p_{c_2}	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.5	12	100	0.11	0.22	0.08	0.31	0.717	0.596	0.628	0.708	0.574
0.5	12	200	0.11	0.22	0.08	0.31	0.914	0.883	0.845	0.888	0.856
0.5	8	100	0.11	0.30	0.08	0.40	0.646	0.540	0.558	0.636	0.490
0.5	8	200	0.11	0.30	0.08	0.40	0.884	0.841	0.790	0.866	0.820
0.5	2	100	0.11	0.59	0.08	0.69	0.338	0.292	0.264	0.318	0.276
0.5	2	200	0.11	0.59	0.08	0.69	0.575	0.561	0.452	0.576	0.506
2.0	12	100	0.33	0.17	0.25	0.25	0.719	0.591	0.640	0.688	0.592
2.0	12	200	0.33	0.17	0.25	0.25	0.906	0.822	0.885	0.891	0.827
2.0	8	100	0.33	0.22	0.25	0.32	0.635	0.514	0.572	0.609	0.522
2.0	8	200	0.33	0.22	0.25	0.32	0.871	0.780	0.820	0.861	0.796
2.0	2	100	0.33	0.45	0.25	0.56	0.323	0.314	0.246	0.311	0.304
2.0	2	200	0.33	0.45	0.25	0.56	0.613	0.548	0.502	0.579	0.524
8.0	12	100	0.67	0.08	0.57	0.14	0.636	0.562	0.602	0.626	0.588
8.0	12	200	0.67	0.08	0.57	0.14	0.922	0.840	0.897	0.921	0.888
8.0	8	100	0.67	0.11	0.57	0.18	0.583	0.502	0.543	0.576	0.522
8.0	8	200	0.67	0.11	0.57	0.18	0.903	0.814	0.864	0.905	0.857
8.0	2	100	0.67	0.22	0.57	0.32	0.328	0.281	0.264	0.318	0.278
8.0	2	200	0.67	0.22	0.57	0.32	0.574	0.518	0.478	0.567	0.473

Table 8 Power of the L_W test in Situation 8

μ_g	μ_d	n_i	p_{t_1}	p_{c_1}	p_{t_2}	p_{c_2}	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
0.5	12	100	0.06	0.05	0.04	0.13	0.887	0.775	0.741	0.873	0.802
0.5	12	200	0.06	0.05	0.04	0.13	1.000	0.980	1.000	1.000	0.984
0.5	8	100	0.06	0.12	0.04	0.25	0.769	0.763	0.674	0.771	0.760
0.5	8	200	0.06	0.12	0.04	0.25	0.988	0.960	0.922	0.985	0.925
0.5	2	100	0.06	0.57	0.04	0.69	0.561	0.410	0.347	0.363	0.360
0.5	2	200	0.06	0.57	0.04	0.69	0.840	0.761	0.753	0.767	0.742
2.0	12	100	0.21	0.04	0.15	0.11	0.819	0.743	0.771	0.812	0.805
2.0	12	200	0.21	0.04	0.15	0.11	0.985	0.964	0.966	0.970	0.973
2.0	8	100	0.21	0.10	0.15	0.22	0.800	0.735	0.754	0.800	0.774
2.0	8	200	0.21	0.10	0.15	0.22	0.956	0.943	0.946	0.951	0.939
2.0	2	100	0.21	0.47	0.15	0.61	0.397	0.383	0.363	0.376	0.322
2.0	2	200	0.21	0.47	0.15	0.61	0.764	0.665	0.750	0.681	0.666
8.0	12	100	0.57	0.02	0.45	0.07	0.788	0.654	0.741	0.762	0.734
8.0	12	200	0.57	0.02	0.45	0.07	0.999	0.942	0.978	0.987	0.978
8.0	8	100	0.57	0.06	0.45	0.15	0.745	0.639	0.732	0.700	0.720
8.0	8	200	0.57	0.06	0.45	0.15	0.937	0.880	0.924	0.927	0.903
8.0	2	100	0.57	0.26	0.45	0.40	0.351	0.310	0.279	0.342	0.343
8.0	2	200	0.57	0.26	0.45	0.40	0.623	0.611	0.581	0.602	0.587

Table 9 p-values of $\hat{L}_W$ test for Channing House data set

	(0,0)	(1,0)	(0,1)	(0.5,0.5)	Gehan
Test statistics	−0.020	−0.008	−0.035	−0.023	−0.019
p-value	0.492	0.497	0.486	0.491	0.492

there exists t_0 such that $P(V_i^* < t < C_i^*) > 0$ for $t \in (781, t_0)$. Hence, condition (a) of Theorem 1 is satisfied, i.e. $\int_{781}^{t} 1/R_i^*(u) f^*(u)\mathrm{d}u \to 0$ as $t \to 781$, where $f^*(t)$ is the conditional probability density function of T_i^* given $T_i^* > 781$.

The results are listed in Table 9. The p-values of the $\hat{L}_W$ test indicate that there is no significant difference between the two conditional survival functions.

5 Discussions

For left-truncated and right-censored data, we have demonstrated how the asymptotic distribution of the L_W can be used to conduct the hypothesis test for the two-sample problem. Simulation has shown that the true size of of the L_W test, in moderate sample size and under varying amounts of censorship and truncation, is indeed accurately approximated by the normal significance level based on this asymptotic theory. Simulation results also show that censoring, truncation mechanism and distribution of lifetime variables all influence the relative power of the L_W test. By B_n and integration by parts, it follows that the L_W test should be consistent against stochastic ordering (i.e. $S_1(t) < S_2(t)$, see Fleming and Harrington 1991) if

$$-\int_{a_F}^{b_F} [\Lambda_1(t) - \Lambda_2(t)]\mathrm{d}W(t) > 0.$$

For the weight function (7), since $\hat{W}(t) \xrightarrow{p} [S(t-)]^r[1 - S(t-)]^s \pi_1(t)\pi_2(t)/[p_1\pi_1(t) + p_2\pi_2(t)]$, it follows that the L_W test with weight function of $(r, s) = (1, 0)$ is consistent against stochastic ordering if $\pi_i(t)$ is decreasing. On the other hand if $(r > 0, s > 0)$ or $\pi_i(t)$ is not decreasing, the consistency under stochastic ordering can fail.

Acknowledgments I would like to thank the associate editor and the two referees for their helpful comments and suggestions.

References

Anderson, P.K., Borgan, Ø., Gill, R.D., Keiding N. (1993). *Statistical models based on counting processes*. Springer.

Bilker, W.B., Wang M.-C. (1996). A semiparametric extension of the Mann–Whitney test for randomly truncated data. *Biometrka*, *52*, 10–20.

Fleming, T.R., Harrington, D.P. (1981). A class of hypothesis tests for one and two samples of censored survival data. *Communications in Statistics Theory and Methods*, *10*, 763–794.

Fleming, T.R., Harrington (1991). *Counting processes and survival analysis*. New York: Wiley.
Gehan, E.A. (1965). A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*, *52*, 203–223.
Harrington, D.P., Fleming, T.R. (1982). A class of rank test procedures for censored survival data. *Biometrika*, *69*, 553–566.
Hyde, J. (1977). Testing survival under right censoring and left truncation. *Biometrika*, *64*, 225–230.
Hyde, J. (1980). Survival analysis with incomplete observations. In R.G. Miller, B. Efron, B.W. Brown, L.E. Moses (Eds), Biostatistics casebook (pp. 31–46). New York: Wiley.
Kaplan, E.L., Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53*, 457–481.
Lagakos, S.W., Barraj, L.M., De Gruttola, V. (1988). Nonparametric analysis of truncated survival data with application to AIDS. *Biometrika*, *75*, 515–524.
Lai, T.L., Ying, Z. (1991). Estimating a distribution function with truncated and censored data. *The Annal of Statistics*, *19*, 417–422.
Mann, H.B., Whitney, D.R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, *18*, 50.
Mantel, N., Haenzel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer institute*, *22*, 719–748.
Mantel, N. (1966). Estimation of survival data and two new rank order statistics arising in its consideration. *Caner Chemotherapy Reports*, *50*, 163–170.
Pepe, M.S., Fleming, T.R. (1989). Weighted Kaplan–Meier statistics: a class of distance tests for censored data. *Biometrics*, *45*, 497–507.
Peto, R., Peto, J. (1972). Asymptotic efficient rank invariant test procedures. *Journal of the Royal Statistical Society, Series A*, *135*, 185–198.
Wang, M.-C. (1987). Product-limit estimates: a generalized maximum likelihood study. *Communications in Statistics Theory and Methods*, *6*, 3117–3132.
Wang, M.-C. (1991). Nonparametric estimation from cross-sectional survival data. *Journal of the American Statistical Association*, *86*, 130–143.
Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, *1*, 80–83.
Woodroofe, M. (1985). Estimating a distribution function with truncated data. *The Annals of Statistics*, *13*, 163–177.