行政院國家科學委員會專題研究計畫　成果報告

在左截取和右設限資料下二元變數存活函數之非參數估計值

計畫類別：個別型計畫
計畫編號：NSC94-2118-M-029-002-
執行期間：94年08月01日至95年07月31日
執行單位：東海大學統計學系

計畫主持人：沈葆聖

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 95年8月9日

Nonparametric Bivariate Estimation with Left-Truncated and Right-Censored Data

Pao-sheng Shen

Department of Statistics
Tunghai University
Taichung, Taiwan 40704
psshen@mail.thu.edu.tw

Abstract

In this paper we consider estimating the bivariate observations where one of the components is subject to left truncation and right censoring and the other is subject to right censoring only. Two types of nonparametric estimators are proposed. One is in the form of inverse-probability-weighted average (Satten and Datta (2001)) and the other is a generalization of Dabrowska's (1988) estimator. The two are then compared based on their empirical performances.

Key Words: Bivariate distribution; truncation; censoring.

1. Introduction

In survival or reliability studies, the observed data is typically censored and/or truncated. Left truncation and right censoring together naturally occur in cohort follow-up studies (see van der Laan (1996), Gürler and Gijbels (1996)). In this paper we consider the situation of bivariate observations where one of the components is subject to left truncation and right censoring and the other is subject to right censoring only. Consider the following application. In pediatric AIDS cohort studies, a group of pregnant women who are infected with HIV but have not yet developed AIDS are selected. Suppose that the infection time (denoted by T_s) can be accurately determined. The recruitment starts at time T_0 and the follow-up is terminated at time T_e. Let U_1 be the incubation time between T_s and development of AIDS and let $V_1 = T_0 - T_s$ if $T_s < T_0$ and $V_1 = 0$ if $T_s \geq T_0$. Then only those women for whom the incubation time $U_1 \geq V_1$ are observed. Let U_2 denote the time from birth, T_b, to development of AIDS for babies. Let $C_1 = T_e - T_s$ and $C_2 = T_e - T_b$. Thus (U_1, U_2) constitutes a bivariate data where U_1 is subject to left truncation and both U_1 and U_2 are subject to right censoring due to termination of study. This problem is a case of the following bivariate problem covered in this paper. Let $(U_1, U_2, C_1, C_2, V_1)$ be i.i.d. random vectors such that

(C_1, C_1, V_1) is independent of (U_1, U_2) and $P(V_1 < C_1) = 1$. Let $X_i = \min\{U_i, C_i\}$ $(i = 1, 2)$ and let δ_i $(i = 1, 2)$ be one if $X_i = U_i$ $(i = 1, 2)$ and zero otherwise. Under left truncation and right censoring models one observes nothing if $U_1 < V_1$ and observes $(X_1, \delta_1, X_2, \delta_2)$ if $U_1 \geq V_1$.

Special Case: $C_2 = \infty$

In this case, only one of the components is subject to left truncation and right censoring, i.e., second component is always observed. Consider the following application. In hemophilia AIDS-data sets the time of infection T_s can be quite accurately determined. A database will cover patients from, say 1978, till 1995, and hence a patient with a longer survival time will have a larger probability of being part of the sample than a patient with a short survival time. Let U_1 be the time between T_s and death and let $V_1 = 1978 - T_s$ if $T_s < 1978$ and $V_1 = 0$ if $T_s \geq 1978$. Then a patient will only be part of the sample if $U_1 \geq V_1$. Let $C_1 = 1995 - T_s$ denote the the time from T_s to the end of study and U_2 denote the time between T_s and AIDS. All patients are followed until the development of AIDS. However, some patients are still alive at the end of the study. In this case, only U_1 is subject to left-truncation and right-censoring.

In Section 2 and 3, two types of nonparametric estimators are proposed. The first one is in the form of inverse-probability-weighted (IPW) average (Satten and Datta (2001)). The second one is a generalization of Dabrowska's (1988) estimator.

2. Inverse-Probability-Weighted (IPW) Estimators

2.1 The Estimator

For the univariate random censoring model, Satten and Datta (2001) showed that the Kaplan-Meier (1958) estimator (known as a NPMLE) of survival function can be expressed as an IPW average (see Robins (1993, 2000)). For the univariate random truncation and censoring model, Shen (2003) showed that the truncation NPMLE (see Woodroofe (1985)) and the censoring-truncation NPMLE (see Wang (1987)) of survival function can also be expressed as IPW averages. Now, let $F_{uu}(u_1, u_2)$, $F_{cc}(c_1, c_2)$, and $F_{vc}(v_1, c_2)$ denote the bivariate distribution functions of (U_1, U_2), (C_1, C_2), and (V_1, C_2), respectively. Let $(X_{1i}, \delta_{1i}, X_{2i}, \delta_{2i}, V_{1i})$ $(i = 1, \ldots, n)$ denote the left-truncated and right-censored sample. We consider the IPW estimator of bivariate distribution functions $F_{uu}(u_1, u_2)$ by simultaneously estimating the three distribution functions as follows:

$$\hat{F}_{uu}(u_1, u_2) =$$

$$\left[\sum_{i=1}^{n} \frac{\delta_{1i}\delta_{2i}}{\hat{S}_{cc}(X_{1i}-, X_{2i}-) - \hat{S}_{vc}(X_{1i}, X_{2i}-)}\right]^{-1} \sum_{i=1}^{n} \frac{\delta_{1i}\delta_{2i} I_{[X_{1i} \le u_1, X_{2i} \le u_2]}}{\hat{S}_{cc}(X_{1i}-, X_{2i}-) - \hat{S}_{vc}(X_{1i}, X_{2i}-)}, \quad (2.1)$$

$$\hat{F}_{cc}(c_1, c_2) = \left[\sum_{i=1}^{n} \frac{1 - \delta_{2i}}{\hat{S}_{uu}(V_{1i}-, X_{2i}-)}\right]^{-1} \sum_{i=1}^{n} \frac{(1 - \delta_{1i})(1 - \delta_{2i}) I_{[X_{1i} \le c_1, X_{2i} \le c_2]}}{\hat{S}_{uu}(X_{1i}-, X_{2i}-)}, \quad (2.2)$$

and

$$\hat{F}_{vc}(v_1, c_2) = \left[\sum_{i=1}^{n} \frac{1 - \delta_{2i}}{\hat{S}_{uu}(V_{1i}-, X_{2i}-)}\right]^{-1} \sum_{i=1}^{n} \frac{(1 - \delta_{2i}) I_{[\tilde{V}_{1i} \le v_1, X_{2i} \le c_2]}}{\hat{S}_{uu}(V_{1i}-, X_{2i}-)}, \quad (2.3)$$

where $\hat{S}_k(x, y) = 1 - \hat{F}_k(x, \infty) - \hat{F}_k(\infty, y) + \hat{F}_k(x, y)$ for $k = uu,\ cc$ and vc.

The following arguments provide a justification of using $\hat{F}_{uu}(u_1, u_2)$, $\hat{F}_{cc}(c_1, c_2)$ and $\hat{F}_{vc}(v_1, c_2)$.

Let $p = P(U_1 \ge V_1)$. Consider the subdistribution function

$$W_{uu}(u_1, u_2) = P(X_{1i} \le u_1, \delta_{1i} = 1, X_{2i} \le u_2, \delta_{2i} = 1)$$

$$= P(U_1 \le u_1, U_1 \le C_1, U_2 \le u_2, U_2 \le C_2 | U_1 \ge V_1)$$

$$= p^{-1} \int_0^{u_2} \int_0^{u_1} [S_{cc}(x-, y-) - S_{vc}(x, y-)] F_{uu}(dx, dy),$$

where $F_{uu}(dx, dy) = \frac{F(x,y)}{\partial x \partial y} dxdy$ if F_{uu} is absolutely continuous in both components at (x, y); $F_{uu}(dx, dy) = \frac{F(x,\Delta y)}{\partial x} dx$ if F_{uu} is continuous in its first, but not its second argument at (x, y) and a similar definition holds for the other cases. Thus, we have

$$F_{uu}(du_1, du_2) = p \frac{W_{uu}(du_1, du_2)}{S_{cc}(u_1-, u_2-) - S_{vc}(u_1, u_2-)}. \quad (2.4)$$

When $S_{cc}(c_1, c_2)$, $S_{vc}(v_1, c_2)$ and p are known, $F_{uu}(u_1, u_2)$ can be estimated by

$n^{-1} p \sum_{i=1}^{n} \frac{\delta_{1i}\delta_{2i} I_{[X_{1i} \le u_1, X_{2i} \le u_2]}}{S_{cc}(X_{1i}-, X_{2i}-) - S_{vc}(X_{1i}, X_{2i}-)}$. Let $u_1 = u_2 = \infty$. It follows that p can be estimated by $n\left[\sum_{i=1}^{n} \frac{\delta_{1i}\delta_{2i}}{S_{cc}(X_{1i}-, X_{2i}-) - S_{vc}(X_{1i}, X_{2i}-)}\right]^{-1}$. This justifies the use of the estimator $\hat{F}_{uu}(u_1, u_2)$. Similarly, the justification of using $\hat{S}_{vc}$ can be obtained by considering the subdistribution function

$$W_{vc}(v_1, c_2) = P(V_{1i} \le v_1, X_{2i} \le c_2, \delta_{2i} = 0)$$

$$= p^{-1} P(C_2 \le c_2, C_2 \le U_2, V_1 \le v_1, V_1 \le U_1) = p^{-1} \int_0^{c_2} \int_0^{v_1} S_{uu}(x-, y-) F_{vc}(dx, dy).$$

Thus, we have

$$F_{vc}(dv_1, dc_2) = p\frac{W_{vc}(dv_1, dc_2)}{S_{uu}(v_1-, c_2-)}.$$

When $S_{uu}(u_1, u_2)$ and p are known, $F_{vc}(v_1, c_2)$ can be estimated by $n^{-1}p\sum_{i=1}^{n}\frac{(1-\delta_{2i})I_{[V_{1i}\le v_1, X_{2i}\le c_2]}}{S_{uu}(V_{1i}-, X_{2i}-)}$. Let $v_1 = c_2 = \infty$. It follows that p can be estimated by $n\Big[\sum_{i=1}^{n}\frac{1-\delta_{2i}}{S_{uu}(V_{1i}-, X_{2i}-)}\Big]^{-1}$. This justifies the use of the estimator $\hat{F}_{vc}(v_1, c_2)$. The justification of using $\hat{F}_{cc}(c_1, c_2)$ is simlar to that of $\hat{F}_{vc}(v_1, c_1)$ and is omitted.

2.2 Special Case: $C_2 = \infty$

2.2.1. Equivalence of $\hat{F}_{uu}$ and $\tilde{F}_{uu}$

In this case, one can consider the probability

$$R(x) = P(V_{1i} \le x \le X_{1i}) = p^{-1}[G(x) - Q(x-)]S_{uu}(x-, 0), \tag{2.5}$$

where $G(x) = P(V_1 \le x)$ and $Q(x) = P(C_1 \le x)$ denote the distribution function of V_1 and C_1, respectively. Expression (2.4) and (2.5) motivate an alternative esimator of $F_{uu}(u_1, u_2)$ as follows.

$$\tilde{F}_{uu}(u_1, u_2) = \sum_{i=1}^{n}\frac{\delta_{1i}I_{[U_{1i}\le u_1, U_{2i}\le u_2]}[1 - \hat{F}_{pl}(U_{1i}-)]}{R_n(U_{1i})}, \tag{2.6}$$

where $\hat{F}_{pl}(u_1)$ is the NPMLE of $F(u_1, \infty)$ (see Wang (1987)) and given by

$$\hat{F}_{pl}(u_1) = 1 - \prod_{x\le u_1}\Big[1 - \frac{N_F(dx)}{R_n(x)}\Big],$$

where $R_n(x) = N_G(x) - N_F(x-)$, $N_F(dx) = N_F(x) - N_F(x-)$, $N_F(x) = \sum_{i=1}^{n}\delta_{1i}I_{[U_{1i}\le x]}$ and $N_G(x) = \sum_{i=1}^{n}I_{[V_{1i}\le x]}$. Note that the estimator $\tilde{F}_{uu}(u_1, u_2)$ has the same form as the estimator proposed by Gijbels and G$\ddot{u}$rler (1996, 1998). Instead of assuming $P(C_1 > V_1) = 1$, they assumed that V_1 and C_1 and the vector (U_1, U_2) are mutually independent. When $C_2 = \infty$, (2.1), (2.2) and (2.3) are reduced to

$$\hat{F}_{uu}(u_1, u_2) = \Big[\sum_{i=1}^{n}\frac{\delta_{1i}}{\hat{G}(X_{1i}) - \hat{Q}(X_{1i}-)}\Big]^{-1}\sum_{i=1}^{n}\frac{\delta_{1i}I_{[X_{1i}\le u_1, X_{2i}\le u_2]}}{\hat{G}(X_{1i}) - \hat{Q}(X_{1i}-)}, \tag{2.7}$$

$$\hat{Q}(c_1) = \Big[\sum_{i=1}^{n}\frac{1}{\hat{S}_{uu}(V_{1i}-, 0)}\Big]^{-1}\sum_{i=1}^{n}\frac{(1-\delta_{1i})I_{[X_{1i}\le c_1]}}{\hat{S}_{uu}(X_{1i}-, 0)}, \tag{2.8}$$

and

$$\hat{G}(v_1) = \Big[\sum_{i=1}^{n}\frac{1}{\hat{S}_{uu}(V_{1i}-, 0)}\Big]^{-1}\sum_{i=1}^{n}\frac{I_{[V_{1i}\le v_1]}}{\hat{S}_{uu}(V_{1i}-, 0)}. \tag{2.9}$$

When $C_2 = \infty$, the equivalence of $\tilde{F}_{uu}(u_1, u_2)$ and $\hat{F}_{uu}(u_1, u_2)$ can be established by considering the estimation of p. For all x such that $R_n(x) > 0$, expression (2.5) suggests an estimator $\hat{p}(x) = n[\hat{G}(x) - \hat{Q}(x-)][\hat{S}_{uu}(x-, o)]/R_n(x)$. Besides, (2.7) and (2.8) suggest two alternative estimators of p, namely, $\hat{p}(\hat{S}_{uu}) = n\left[\sum_{i=1}^{n} \frac{1}{\hat{S}_{uu}(V_{1i}-,0)}\right]^{-1}$ and $\hat{p}(\hat{H}) = n\left[\sum_{i=1}^{n} \frac{\delta_{1i}}{\hat{H}(U_{1i})}\right]^{-1}$, where $\hat{H}(x) = \hat{G}(x) - \hat{Q}(x-)$. The following Lemma establishes the equivalence of the three estimators.

Lemma 2.1.
Let $A_d = \{k : \delta_{1k} = 1\}$. Suppsose that $R_n(U_{1j}) > 0$ for all j and the largest observation is not censored, then $\hat{p}(U_{1j}) = \hat{p}(\hat{H}) = \hat{p}(\hat{S}_{uu})$ for all $j \in A_d$.

Proof:

First, it is easily shown that when the largest observation is not censored

$$\int [\hat{G}(x) - \hat{Q}(x-)]\hat{F}_{uu}(dx, \infty) = \int (\hat{S}_{uu}(x-, 0))d[\hat{G}(x) - \hat{Q}(x-)].$$

By (2.7)-(2.9), we have $\int [\hat{G}(x) - \hat{Q}(x-)]\hat{F}_{uu}(dx, \infty) = (n_d/n)\hat{p}(\hat{H})$ and

$\int (\hat{S}_{uu}(x-, 0))d[\hat{G}(x) - \hat{Q}(x-)] = (n_d/n)\hat{p}(\hat{S}_{uu})$, where $n_d = \sum_{i=1}^{n} \delta_{1i}$. Thus, $\hat{p}(\hat{H}) = \hat{p}(\hat{S}_{uu})$. Next, by Theorem 3.1 of Shen (2003), for any $j \in A_d$, $\hat{F}_{uu}(dU_{1j}, \infty) = \hat{F}_{pl}(dU_{1j}) = \hat{F}_{pl}(U_{1j}) - \hat{F}_{pl}(U_{1j}-)$. Hence, for any $j \in A_d$, we have

$$\left[\sum_{i=1}^{n} \frac{\delta_{1i}}{\hat{G}(U_{1i}) - \hat{Q}(U_{1i}-)}\right]^{-1} \frac{1}{[\hat{G}(U_{1j}) - \hat{Q}(U_{1j}-)]} = \frac{[1 - \hat{F}_{pl}(U_{1j}-)]}{R_n(U_{1j})}.$$

Thus, $\hat{p}(U_{1j}) = \hat{p}(\hat{H})$. The proof is completed.

By Lemma 2.1 and (2.6), it follows that

$$\tilde{F}_{uu}(u_1, u_2) = \sum_{i=1}^{n} \frac{\delta_{1i}\hat{p}(U_{1i})I_{[U_{1i} \le u_1, U_{2i} \le u_2]}}{n[\hat{G}(U_{1i}) - \hat{Q}(U_{1i}-)]} = \frac{\hat{p}(\hat{H})}{n} \sum_{i=1}^{n} \frac{\delta_{1i}I_{[U_{1i} \le u_1, U_{2i} \le u_2]}}{\hat{G}(U_{1i}) - \hat{Q}(U_{1i}-)} = \hat{F}_{uu}(u_1, u_2).$$

2.2.2 Asymptotic Properties of $\hat{F}_{uu}$

Assuming V_1 and C_1 and the vector (U_1, U_2) are mutually independent, G$\ddot{u}$rler and Gijbels (1996, 1998) derived the asymptotic properties of $\tilde{F}_{uu}$ via a strong i.i.d. representation. Using the weighted average form given in (2.7) and assuming $P(C_1 > V_1) = 1$, we can establish consistent properties of $\hat{F}_{uu}$. First, we consider the asymptotic properties of $\hat{Q}$ and $\hat{G}$.

Lemma 2.2.

(i) $\hat{Q}(c_1) - Q(c_1) = \phi_{q1}(c_1) + \phi_{q2}(c_1) + o_p(n^{-1/2})$,

where $\phi_{q1}(c_1) = pn^{-1}\sum_{i=1}^{n} \psi_q(X_{1i}, V_{1i}, \delta_{1i}, c_1)$ and $\phi_{q2}(c_1) = pn^{-1}\sum_{i=1}^{n} \zeta_q(X_{1i}, \delta_{1i}, c_1)$,

$\psi_q(X_{1i}, V_{1i}, \delta_{1i}, c_1) = \int_0^\infty \frac{\xi_f(X_{1i},V_{1i},\delta_{1i},x-)}{S_{uu}^2(x-,0)}(Q(c_1)dW_g(x) - I_{[x\le c_1]})dW_q(x)$,

$W_q(x) = P(X_{1i} \le x, \delta_{1i} = 0)$, $W_g(x) = P(V_{1i} \le x)$, $\xi_f(X_{1i}, V_{1i}, \delta_{1i}, x)$

$= -S_{uu}(x, 0)\left[\frac{I_{[X_{1i}\le x,\delta_{1i}=1]}}{R(x)} + \int_0^x \frac{I_{[X_{1i}\le s,\delta_{1i}=1]}}{R^2(s)}dR(s) - \int_0^x \frac{I_{[V_{1i}\le s\le X_{1i}]}}{R^2(s)}W_{uu}(ds, \infty)\right]$, and

$\zeta_q(X_{1i}, \delta_{1i}, c_1) = \frac{(1-\delta_{1i})(I_{[X_{1i}\le c_1]}-Q(c_1))}{S_{uu}(X_{1i}-,0)}$.

(ii) $\hat{G}(v_1) - G(v_1) = \phi_{g1}(v_1) + \phi_{g2}(v_1) + o_p(n^{-1/2})$,

where $\phi_{g1}(v_1) = pn^{-1}\sum_{i=1}^{n} \psi_g(X_{1i}, V_{1i}, \delta_{1i}, v_1)$ and $\phi_{g2}(v_1) = pn^{-1}\sum_{i=1}^{n} \zeta_g(V_{1i}, v_1)$,

where $\psi_g(X_{1i}, V_{1i}, \delta_{1i}, v_1) = \int_0^\infty \frac{\xi_f(X_{1i},V_{1i},\delta_{1i},x)}{S_{uu}^2(x-,0)}(G(v_1) - I_{[x\le v_1]})dW_g(x)$, and

$\zeta_g(V_{1i}, v_1) = \frac{I_{[V_{1i}\le v_1]}-G(v_1)}{S_{uu}(V_{1i}-,0)}$.

Proof:

The proof can be obtained using Lemma 4.1-4.3 of Wang (1991) and is omitted.

Lemma 2.3.

$\hat{F}_{uu}(u_1, u_2) - F_{uu}(u_1, u_2) = \phi_{f1}(u_1, u_2) + \phi_{f2}(u_1, u_2) + o_p(n^{-1/2})$,

where $\phi_{f1}(u_1, u_2) = p^2n^{-1}\sum_{i=1}^{n} \psi_f(X_{1i}, X_{2i}, V_{1i}, \delta_{1i}, u_1, u_2)$, and

$\phi_{f2}(u_1, u_2) = p^2n^{-1}\sum_{i=1}^{n} \zeta_f(X_{1i}, X_{2i}, \delta_{1i}, u_1, u_2)$,

where $\psi_f(X_{1i}, X_{2i}, V_{1i}, \delta_{1i}, u_1, u_2) = \int_0^\infty\int_0^\infty \Bigg\{\frac{[\psi_g(X_{1i},V_{1i},\delta_{1i},x)+\zeta_g(V_{1i},x)]}{H^2(x)} -$

$\frac{[\psi_q(X_{1i},V_{1i},\delta_{1i},x-)+\zeta_q(X_{1i},\delta_{1i},x-)]}{H^2(x)}\Bigg\}(F_{uu}(u_1, u_2) - I_{[x\le u_1,y\le u_2]})W_{uu}(dx, dy)$,

and $\zeta_f(X_{1i}, X_{2i}, \delta_{1i}, u_1, u_2) = \frac{I_{[X_{1i}\le u_1,X_{2i}\le u_2]}\delta_{1i}-F_{uu}(u_1,u_2)\delta_{1i}}{H(X_{1i})}$.

Proof:

The proof can be obtained using Lemma 2.1 and 2.2 and is omitted.

It can be easily shown that $E[\zeta_q(X_{1i},\delta_{1i},c_1)] = E[\zeta_f(X_{1i},X_{2i},\delta_{1i},u_1,u_2] =$

$= E[\zeta_g(V_{1i},v_1)] = E[\psi_q(X_{1i},V_{1i},\delta_{1i},c_1)] = E[\psi_g(X_{1i},V_{1i},\delta_{1i},v_1)] = 0.$

By Lemma 2.2 and 2.3, we have $E[\psi_f(X_{1i},X_{2i},V_{1i},\delta_{1i},u_1,u_2)] = 0$. By the multivariate central limit theorem, it follows that the finite-dimensional distribution of $\sqrt{n}(\hat{F}(u_1,u_2) - F(u_1,u_2))$ converges weakly to the multivariate normal distribution $N(\mathbf{0},\Sigma_f)$.

3. Dabrowska's Estimatior

In this Section, motivated by Dabrowska (1988), we propose an alternative estimator of $F_{uu}(u_1,u_2)$. In the univariate models, it is well known that the hazard function and the distribution function determine each other in a unique way. In the bivariate case, however, there have been several definitions of the hazard function or failure rate. Dabrowska (1988) presented a nice representation of the bivariate distribution function in terms of the three component bivariate hazard vector. The hazard vector is given by $\Lambda(u_1,u_2) = (\Lambda_{10}(du_1,u_2),\Lambda_{01}(u_1,du_2),\Lambda_{11}(du_1,du_2))$, where $\Lambda_{11}(du_1,du_2) = \frac{S_{uu}(du_1,du_2)}{S_{uu}(u_1-,u_2-)}$, $\Lambda_{10}(du_1,u_2) = \frac{-S_{uu}(du_1,u_2)}{S_{uu}(u_1-,u_2)}$, and $\Lambda_{01}(u_1,du_2) = \frac{-S_{uu}(u_1,du_2)}{S_{uu}(u_1,u_2-)}$. By Propositon 2.1 of Dabrowska (1988), for (u_1,u_2) such that $S_{uu}(u_1,u_2) > 0$, we have

$$S_{uu}(u_1,u_2) = S_{uu}(u_1,0)S_{uu}(0,u_2)\prod_{y\le u_2}\prod_{x\le u_1}[1 - L(dx,dy)],$$

where $L(dx,dy) = \frac{\Lambda_{10}(dx,y-)\Lambda_{01}(x-,dy)-\Lambda_{11}(dx,dy)}{[1-\Lambda_{10}(dx,y-)][1-\Lambda_{01}(x-,dy)]}$. Define $R(u_1-,u_2-) = P(V_{1i} \le u_1 \le X_{1i}, u_2 \le X_{2i})$. Hence,

$$S_{uu}(u_1-,u_2-) = p\frac{R(u_1-,u_2-)}{S_{cc}(u_1-,u_2-) - S_{vc}(u_1,u_2-)}. \tag{3.1}$$

By (2.4) and (3.1), we have $\Lambda_{11}(du_1,du_2) = \frac{F_{uu}(du_1,du_2)}{S_{uu}(u_1-,u_2-)} = \frac{W_{uu}(du_1,du_2)}{R(u_1-,u_2-)}$. Similarly, $\Lambda_{10}(du_1,u_2) = \frac{-S_{uu}(du_1,u_2)}{S_{uu}(u_1-,u_2)} = \frac{W_{u0}(du_1,u_2)}{R(u_1-,u_2)}$, where $R(u_1-,u_2) = P(V_{1i} \le u_1 \le X_{1i}, u_2 < X_{2i})$ and $W_{u0}(u_1,u_2) = P(X_{1i} \le u_1, \delta_{1i} = 1, X_{2i} > u_2)$, and $\Lambda_{01}(u_1,du_2) = \frac{-S_{uu}(u_1,du_2)}{S_{uu}(u_1,u_2-)} = \frac{W_{0u}(u_1,du_2)}{R(u_1,u_2-)}$, where $R(u_1,u_2-) = P(V_{1i} < u_1 < X_{1i}, u_2 \le X_{2i})$ and $W_{0u}(u_1,u_2) = P(X_{2i} \le u_2, \delta_{2i} = 1, X_{1i} > u_1 > V_{1i})$.

Define $\hat{W}_{uu}(u_1,u_2) = n^{-1}\sum_{i=1}^n I_{[X_{1i}\le u_1,\delta_{1i}=1,X_{2i}\le u_2,\delta_{2i}=1]}$,

$\hat{W}_{u0}(u_1,u_2) = \frac{1}{n}\sum_{i=1}^n I_{[X_{1i}\le u_1,\delta_{1i}=1,X_{2i}\ge u_2]}$, $\hat{W}_{0u}(u_1,u_2) = \frac{1}{n}\sum_{i=1}^n I_{[X_{2i}\le u_2,\delta_{2i}=1,X_{1i}\ge u_1]}$

and $\hat{R}(u_1-, u_2-) = n^{-1}\sum_{i=1}^{n} I_{[V_{1i}\leq u_1 \leq X_{1i}, X_{i2}\geq u_2]}$. A natural candidate for an estimaor of $S_{uu}(u_1, u_2)$ is provided by

$$\hat{S}^D_{uu}(u_1, u_2) = \hat{S}^1_{uu}(u_1)\hat{S}^2_{uu}(u_2)\prod_{y\leq u_2}\prod_{x\leq u_1}[1-\hat{L}(dx, dy)],$$

where

$$\hat{L}(dx, dy) = \frac{\hat{\Lambda}_{10}(dx, y-)\hat{\Lambda}_{01}(x-, dy) - \hat{\Lambda}_{11}(dx, dy)}{[1-\hat{\Lambda}_{10}(dx, y-)][1-\hat{\Lambda}_{01}(x-, dy)]},$$

$$\hat{\Lambda}_{11}(dx, dy) = \frac{\hat{W}_{uu}(dx, dy)}{\hat{R}(x-, y-)},\ \hat{\Lambda}_{10}(dx, y-) = \frac{\hat{W}_{u0}(dx, y-)}{\hat{R}(x-, y-)}\ , \hat{\Lambda}_{01}(x-, dy) = \frac{\hat{W}_{0u}(x, dy)}{\hat{R}(x-, y-)},$$

$$\hat{S}^1_{uu}(u_1) = \prod_{x\leq u_1}[1-\hat{\Lambda}_{10}(dx, 0)], \hat{S}^2_{uu}(u_2) = \prod_{y\leq u_2}[1-\hat{\Lambda}_{01}(0, dy)].$$

Note that $\hat{S}^1_{uu}(u_1)$ is the univariate product limit estimate for left-truncated and right-censored data (see Tsai, Jewell and Wang (1987), Lai and Ying (1991)) and $\hat{S}^2_{uu}(u_2)$ is the Kaplan-Meier (1958) estimate. By proposition 4.1 of Dabrowska (1988) and the uniform consistency of $\hat{S}^1_{uu}$ (Gijbels and Wang (1993)) and $\hat{S}^2_{uu}$, it is enough to show the consistency of $\hat{S}^D_{uu}$.

4. Simulation Results

In this section, a simulation study is conducted to examine the performances of the $\hat{S}_{uu}(u_1, u_2)$ and $\hat{S}^D_{uu}(u_1, u_2)$. The (U_1, U_2)'s are i.i.d. bivariate exponential distributed with survival function: $S_{uu}(u_1, u_2) = e^{-(u_1+u_2)-\max(u_1, u_2)}$. The V_1's are i.i.d. exponential distributed with survival function $\bar{G}(v_1) = 1 - G(v_1) = e^{-\lambda_g v_1}$. We consider the following two cases:

Case 1:

The U_1's are subject to left truncation and right censoring and the U_2 is subject to right censoring only. The C_1's are defined as $C_1 = D + V_1 + C_2$, such that $P(V_1 < C_1) = 1$, where D's and C_2 are both exponential distributed with survival function $S_D(x) = e^{-\lambda_d x}$ and $\bar{Q}(c_2) = e^{-\lambda_q c_2}$, respectively. With the choice of $\lambda_d = 1.0$, $\lambda_g = 2,\ 7$ and $\lambda_q = 0.5, 1$, it covers a wide range from heavy to light truncation, and light to heavy censoring.

Case 2 ($C_2 = 0$):

Only the U_1's are subject to left truncation and right censoring. The C_1's are defined as $C_1 = D + V_1$ such that $P(V_1 < C_1) = 1$. With the choice of $\lambda_d = 1.0,\ 5.0$, and

$\lambda_g = 1.0, 2.0, 4.0$, it covers a wide range from heavy to light truncation, and light to heavy censoring.

The sample sizes are set at $n = 100$ and 200, and the replication is 5000 times. Tables 1 thruogh 4 show the biases, standard deviations (denoted by std), and the ratio of the squared root of mean squared error of $\hat{S}^D_{uu}$ to that of $\hat{S}_{uu}$ (denoted by $\sqrt{\frac{mse(\hat{S}^D_{uu})}{mse(\hat{S}_{uu})}}$) at $S_{uu}(0.25, 0.193) = 0.5$ and $S_{uu}(0.708, 0.193) = 0.2$. Tables 1 and 2 (case 1) also list the probability of truncation (denoted by $q = 1 - p$) and probability of censoring $p_c = 1 - P(U_1 \leq C_1, U_2 \leq C_2)$. Similarly, Tables 3 and 4 (case 2) list the probability of truncation q and probability of censoring $p_c = P(C_1 \leq U_1)$. Besides, Table 5 (case 2) lists the ratio $\sqrt{\frac{mse(\hat{S}^D_{uu})}{mse(\hat{S}_{uu})}}$ at a 3×3 grid of values of (u_1, u_2) for parameter values, $\lambda_g = \lambda_d = 1.0$ and $\lambda_g = 4.0, \lambda_d = 5.0$.

Table 1. Simulation results for bias, std. and $\sqrt{mse}$ (case 1:$\hat{S}_{uu}(0.25, 0.193) = 0.5$)

					bias		std		
λ_g	λ_q	q	p_c	n	$\hat{S}_{uu}$	$\hat{S}^D_{uu}$	$\hat{S}_{uu}$	$\hat{S}^D_{uu}$	$\sqrt{\frac{mse(\hat{S}^D_{uu})}{mse(\hat{S}_{uu})}}$
2.0	0.5	0.50	0.29	100	-0.0614	0.0623	0.2034	0.1169	0.624
2.0	0.5	0.50	0.29	200	0.0429	0.0613	0.1385	0.0595	0.589
2.0	1.0	0.50	0.48	100	-0.0388	0.0657	0.2164	0.0817	0.477
2.0	1.0	0.50	0.48	200	0.0209	0.0565	0.1477	0.0763	0.636
7.0	0.5	0.22	0.27	100	-0.0584	0.0695	0.1306	0.0756	0.718
7.0	0.5	0.22	0.27	200	-0.0080	0.0596	0.0819	0.0447	0.905
7.0	1.0	0.22	0.43	100	-0.0627	0.0707	0.1280	0.1074	0.902
7.0	1.0	0.22	0.43	200	-0.0112	0.0605	0.0849	0.0388	0.840

Table 2. Simulation results for bias, std. and $\sqrt{mse}$ (case 1:$\hat{S}_{uu}(0.708, 0.193) = 0.2$)

					bias		std		
λ_g	λ_q	q	p_c	n	$\hat{S}_{uu}$	$\hat{S}_{uu}^D$	$\hat{S}_{uu}$	$\hat{S}_{uu}^D$	$\sqrt{\frac{mse(\hat{S}_{uu}^D)}{mse(\hat{S}_{uu})}}$
2.0	0.5	0.50	0.29	100	-0.0408	0.0588	0.2055	0.0632	0.412
2.0	0.5	0.50	0.29	200	-0.0312	0.0359	0.1347	0.0338	0.357
2.0	1.0	0.50	0.48	100	-0.0194	0.0498	0.2119	0.0669	0.392
2.0	1.0	0.50	0.48	200	0.0137	0.0585	0.1758	0.0349	0.385
7.0	0.5	0.22	0.27	100	-0.0126	0.0703	0.1047	0.0379	0.757
7.0	0.5	0.22	0.27	200	-0.0091	0.0598	0.0722	0.0242	0.886
7.0	1.0	0.22	0.43	100	-0.0470	0.0579	0.0851	0.0459	0.760
7.0	1.0	0.22	0.43	200	-0.0338	0.0520	0.0669	0.0340	0.828

Table 3. Simulation results for bias, std. and $\sqrt{mse}$ (case 2:$\hat{S}_{uu}(0.25, 0.193) = 0.5$)

					bias		std		
λ_g	λ_q	q	p_c	n	$\hat{S}_{uu}$	$\hat{S}_{uu}^D$	$\hat{S}_{uu}$	$\hat{S}_{uu}^D$	$\sqrt{\frac{mse(\hat{S}_{uu}^D)}{mse(\hat{S}_{uu})}}$
1.0	1.0	0.67	0.11	100	0.0011	0.0483	0.0940	0.0998	1.169
1.0	1.0	0.67	0.11	200	0.0071	0.0697	0.0764	0.0829	1.412
1.0	5.0	0.67	0.24	100	-0.0158	0.0596	0.1166	0.1078	1.049
1.0	5.0	0.67	0.24	200	0.0083	0.0481	0.1034	0.0984	1.056
2.0	1.0	0.50	0.17	100	0.0152	0.0743	0.0770	0.1074	1.664
2.0	1.0	0.50	0.17	200	0.0076	0.0694	0.0721	0.0955	1.629
2.0	5.0	0.50	0.36	100	-0.0040	0.0645	0.1125	0.0879	0.969
2.0	5.0	0.50	0.36	200	0.0053	0.0608	0.1004	0.0844	1.035
4.0	1.0	0.33	0.22	100	-0.0112	0.0707	0.0788	0.0667	1.221
4.0	1.0	0.33	0.22	200	0.0073	0.0615	0.0594	0.0516	1.341
4.0	5.0	0.33	0.48	100	-0.0400	0.0529	0.1144	0.0937	0.888
4.0	5.0	0.33	0.48	200	-0.0167	0.0715	0.1037	0.0619	0.901

Table 4. Simulation results for bias, sd. and $\sqrt{mse}$ (case 2:$\hat{S}_{uu}(0.708, 0.193) = 0.2$)

					bias		std		
λ_g	λ_q	q	p_c	n	$\hat{S}_{uu}$	$\hat{S}_{uu}^D$	$\hat{S}_{uu}$	$\hat{S}_{uu}^D$	$\sqrt{\frac{mse(\hat{S}_{uu}^D)}{mse(\hat{S}_{uu})}}$
1.0	1.0	0.67	0.11	100	-0.0048	0.0506	0.0837	0.0794	1.124
1.0	1.0	0.67	0.11	200	0.0097	0.0463	0.0535	0.0321	1.034
1.0	5.0	0.67	0.24	100	0.0052	0.0447	0.1203	0.1052	0.949
1.0	5.0	0.67	0.24	200	0.0014	0.0401	0.1013	0.0825	0.905
2.0	1.0	0.50	0.17	100	-0.0012	0.0647	0.0812	0.0605	1.091
2.0	1.0	0.50	0.17	200	0.0052	0.0505	0.0747	0.0581	1.028
2.0	5.0	0.50	0.36	100	-0.0043	0.0493	0.1126	0.0689	0.752
2.0	5.0	0.50	0.36	200	0.0070	0.0476	0.0979	0.0672	0.833
4.0	1.0	0.33	0.22	100	-0.0199	0.0897	0.0913	0.0503	0.836
4.0	1.0	0.33	0.22	200	0.0085	0.0598	0.0861	0.0382	0.821
4.0	5.0	0.33	0.48	100	-0.0238	0.0672	0.1294	0.0813	0.784
4.0	5.0	0.33	0.48	200	-0.0075	0.0569	0.1170	0.1014	0.916

Table 5. Simulation results of $\sqrt{mse(\hat{S}_{uu}^D)/mse(\hat{S}_{uu})}$
($\lambda_g = \lambda_d = 1.0$, $\lambda_g = 4.0 \& \lambda_d = 5.0$) for $n = 100$

	$u_2 = 0.193$	$u_2 = 0.250$	$u_2 = 0.708$
$u_1 = 0.193$	1.039, 0.627	0.935, 0.745	1.321, 1.019
$u_1 = 0.250$	1.169, 0.888	1.052, 0.783	1.003, 0.840
$u_1 = 0.708$	1.124, 0.754	0.851, 0.602	0.745, 0.568

Based on the results of Tables 1 and 2 (case 1), we conclude that:

(i) The bias of $\hat{S}_{uu}$ is smaller than that of $\hat{S}_{uu}^D$ for all the cases considered.

(ii) The standard deviation of $\hat{S}_{uu}^D$ is smaller than that of $\hat{S}_{uu}^D$ for all the cases considered. In terms of $\sqrt{mse}$, $\hat{S}_{uu}^D$ is dominating. One explanation for the results is that the estimator $\hat{S}_{uu}$ is based on the data with $\delta_{1i} = \delta_{2i} = 1$, which makes the estimator less efficient.

Based on the results of Tables 3 through 5 (case 2), we conclude that:

(i) The bias of $\hat{S}_{uu}$ is smaller than that of $\hat{S}_{uu}^D$ for all the cases considered.

(ii) In most of the simulated cases, the $\sqrt{mse}$ of $\hat{S}_{uu}$ is larger than that of $\hat{S}_{uu}^D$. When censoring is light and truncation is severe (e.g., $\lambda_g = \lambda_d = 1.0$, $p_c = 0.11$, $q = 0.67$), the *std* and $\sqrt{mse}$ of $\hat{S}_{uu}$ can be smaller than those of $\hat{S}_{uu}^D$. However, when censoring is not light (e.g., $\lambda_g = 4.0$, $\lambda_d = 5.0$, $p_c = 0.48$), the situation is reverse and the estimator $\hat{S}^D_{uu}$ is a better choice than the $\hat{S}_{uu}$ estimator.

The results of Table 3 through 5 agree with those of G*ü*rler (1996, 1997), where he proposed several nonparametric estimators for the special case of $C_1 = C_2 = \infty$. The estimator that performed best (better than $\hat{S}^D_{uu}$) has the same form as $\tilde{S}_{uu}$, which is equivalent to $\hat{S}_{uu}$.

References

Dabrowska, D. M. Kaplan-Meier estimate on the plane. The Annals of Statistics, **1988**, *18*, 308-325.

Gijbels, I. and G*ü*rler, $\ddot{U}$. Covariance function of a bivariate distribution function estimator for left truncated and right censored data. Discussion paper NO. 9703, **1996**, Institute of Statistics, Catholic University of Louvain, Louvain-la-Neuve, Belgium.

Gijbels, I. and Wang, J. L. (1993). Strong representations of the survival function estimator for truncated and censored data with applications. *J. Mult. Anal.* **47**, 210-229.

G*ü*rler, $\ddot{U}$. Bivariate estimation with right truncated data. *J. Amer. Statist. Assoc.*, **1996**, *91*, 1152-1165.

Gijbels, I. and G*ü*rler, $\ddot{U}$. Covariance function of a bivariate distribution function estimator for left truncated and right censored data. *Statistica Sinica*, **1998**, 1219-1232.

G*ü*rler, $\ddot{U}$. Bivariate distribution and hazard functions when a component is randomly truncated. J. Multivariate Anal., **1997**, *60*, 20-47.

Kaplan, E. L.; Meier, P. Nonparametric estimation from incomplete observations, J. Amer. Statist. Assoc., **1958**, *53*, 457-481.

Lai, T. L. and Ying, Z. Estimating a distribution function with truncated and censored data. Ann. Statist., **1991** *19*, No. 1, 417-442.

Robins, J. M. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers, in Proceedings of the American Statistical Association- Biopharmaceutical Section, Alexaandria, VA: ASA, **1993**, pp. 24-33.

Robins, J. M. and Finkelstein, D. Correcting for non-compliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted(IPCW) log-rank tests. *Biometrices*, **2000**, *56*, 779-788.

Shen, P-S. The product-limit estimate as an inverse-probability-weighted average. *Communi. in Statist.*, *Part A- Theory and Methods*, **2003**, *32*, 1119-1133.

Satten, G. A. and Datta S. The Kaplan-Meier estimator as an inverse-probability-of-censoring weighted average. *Amer. Statist.*, **2001**, *55*, 207-210.

Tsai, W.-Y., Jewell, N. P. and Wang, M.-C. A note on the product-limit estimator under right censoring and left truncation. *Biometrika* **1987**, *74* 883-886.

Van der laan, M. J. Nonparametric estimation of the bivariate survival function with truncated data. *J. Multivariate Anal.*, **1996**, *58*, 107-131.

Wang, M.-C. Product-limit estimates: a generalized maximum likelihood study. *Communi. in Statist.*, *Part A- Theory and Methods*, **1987**, *6*, 3117-3132.

Woodroofe, M. Estimating a distribution function with truncated data. **1985**, *The Annals of Statistics*, *13*, 163-177.