The Annals of Statistics
1990, Vol. 18, No. 1, 391–404

WEAK CONVERGENCE OF A SELF-CONSISTENT ESTIMATOR OF THE SURVIVAL FUNCTION WITH DOUBLY CENSORED DATA

BY MYRON N. CHANG

University of Florida

Double censoring often occurs in collecting lifetime data and accordingly self-consistent estimators are widely employed for estimating the survival functions. In this paper we prove the weak convergence of self-consistent estimators using classical results on the Fredholm integral equation. Also, a method of calculating the asymptotic variance is presented.

1. Introduction. Let X, Y, and Z be nonnegative variables where X denotes the time of occurrence of a well-defined event such as death, and Z and Y are subject to the restriction $P(0 \le Z \le Y) = 1$. An observation on X is said to be subject to left censoring by Z and right censoring by Y if X is observable whenever X lies in the interval $[Z, Y]$. If X is outside the interval, then we know whether $X < Z$ or $Y < X$ and observe the value of Z or Y accordingly. A random sample of observations on X subject to right and left censoring by Y and Z, respectively, is called a doubly censored sample. Gehan (1965), Mantel (1967), Peto (1973), Turnbull (1974) and others have given examples of doubly censored samples encountered in practical situations.

Let (X_i, Y_i, Z_i), $i = 1, \dots, n$ be a set of n independent observations on (X, Y, Z). The available information in a doubly censored sample can be summarized using the n independent pairs (W_i, δ_i), $i = 1, \dots, n$ where

$$W_i = \max(\min(X_i, Y_i), Z_i),$$

and

$$\delta_i = \begin{cases} 1 & \text{if } Z_i \le X_i \le Y_i, \\ 2 & \text{if } X_i > Y_i, \\ 3 & \text{if } X_i < Z_i. \end{cases}$$

Turnbull (1974) constructed a self-consistent estimator (nonparametric MLE) for the survival function $S_X = P(X > t)$ using a doubly censored sample. Under a set of mild assumptions, Chang and Yang (1987) proved identifiability and strong consistency of the self-consistent estimator using direct arguments. A proof of weak convergence of the self-consistent estimator is not yet available.

Because the self-consistent estimator is an implicit solution of an estimating equation for the nonparametric MLE, it is natural to expect that an abstract implicit function theorem, such as in Reeds (1976) or in Fernholz (1983), could be

Received March 1987; revised March 1989.
AMS 1980 *subject classifications*. 62G05, 62G99.
Key words and phrases. Survival function, doubly censored data, self-consistent estimation, weak convergence, Fredholm integral equation.

used to prove weak convergence. However, a deep exploration by Gill (1989) showed that such a theorem will not be applicable generally because continuous differentiability and other conditions often fail to hold. Our objective in this paper is to continue the work of Chang and Yang (1987) to prove a weak convergence result for the self-consistent estimator by direct arguments. Instead of treating the self-consistent estimator as the implicit solution of equation (5.1) in Tsai and Crowley (1985), we use the integral equations, (2.5), relating survival functions to observable subsurvival functions developed by Chang and Yang (1987). Let $S_X^{(n)}$ be the self-consistent estimator for S_X. We shall show that if higher order terms are ignored, $\sqrt{n}(S_X^{(n)} - S_X)$ is related to the observable empirical processes through a system of Fredholm integral equations and classical results on the Fredholm integral equations can be used to prove weak convergence.

In Section 2, we state the basic assumptions, derive a system of integral equations and present a theorem on the Fredholm integral equations which is essential for proving weak convergence. Section 3 contains the main results and in Section 4 we present a method for calculating the asymptotic variance of the self-consistent estimator.

2. Assumptions and the Fredholm integral equation. Our notations are similar to Chang and Yang (1987). Let (W_i, δ_i), $i = 1, \dots, n$, defined in Section 1 be distributed as (W, δ). Let the subdistribution functions and the empirical subdistribution functions be defined as

$$Q_j(t) = P(W \le t, \delta = j), \qquad j = 1, 2, 3,$$

and

$$Q_j^{(n)}(t) = \frac{1}{n} \sum_{i=1}^{n} I[W_i \le t, \delta_i = j], \qquad j = 1, 2, 3,$$

where $I[\cdot]$ is the indicator function. Denote $S_X(t) = P(X > t)$, $S_Y(t) = P(Y > t)$ and $S_Z(t) = P(Z > t)$. We shall consider the estimation of S_X, the survival function for X.

Throughout the remainder of this paper, the following conditions will be assumed to hold.

A1. The random variables X_i and (Y_i, Z_i) are independent for each i, $i = 1, 2, \dots, n$.
A2. $P(Z \le Y) = 1$.
A3. $S_Y(t) - S_Z(t) > 0$ on $(0, \infty)$.
A4. S_X, S_Y and S_Z are continuous functions of t, on $t \ge 0$, and $0 < S_X(t) < 1$ for $t > 0$.
A5. $S_X(0) = S_Y(0) = 1$, $S_X(\infty) = S_Y(\infty) = S_Z(\infty) = 0$.
A6. There exist δ and T, $0 < \delta < T < \infty$, such that $S_Z(t) = \text{constant} < 1$ on $[0, \delta]$ and $S_Z(T) = 0$, i.e., $P(Z = 0) > 0$, $P(Z \in (0, \delta)) = 0$ and $P(Z \le T) = 1$.

Remarks. (a) Assumptions A1–A5 were used by Chang and Yang (1987) to prove identifiability and strong consistency of the self-consistent estimator.

(b) Under assumptions A2 and A4, assumption A3 is equivalent to

$$P(Z < t < Y) > 0 \quad \text{for any } t \in (0, \infty).$$

(c) The purpose of assumption A6 is to avoid singularity of certain integral equations [see (5) and (7)] encountered in the proof. It is not clear whether the results are valid without A6.

(d) The number δ in A6 could be arbitrarily small and T could be arbitrarily large.

(e) Due to symmetry of the problem, A6 could be replaced by a similar condition on Y.

(f) Assumption A6 is generally easy to verify in practice. If there is a certain lag of left censoring at time 0 and there is no left censoring after a certain time period, then assumption A6 will be satisfied.

(g) If there is no left censoring at all, i.e., if $S_Z(t) \equiv 0$, then assumption A6 is satisfied. So, weak convergence of the Kaplan–Meier estimator is a special case of our results.

Under assumptions A2 and A4, Chang and Yang (1987) derived the following system of integral equations relating the survival functions to the subdistribution functions

$$\begin{aligned} Q_1(t) &= -\int_0^t (S_Y - S_Z)\, dS_X, \\ Q_2(t) &= -\int_0^t S_X\, dS_Y, \\ Q_3(t) &= -\int_0^t (1 - S_X)\, dS_Z. \end{aligned} \tag{1}$$

Let $S_X^{(n)}$ denote the self-consistent estimator of S_X. Define $S_Y^{(n)}$ and $S_Z^{(n)}$ as in (2.8) and (2.9) of Chang and Yang (1987). Then the sample counterpart of (1) is

$$\begin{aligned} Q_1^{(n)}(t) &= -\int_0^t \left(S_Y^{(n)} - S_Z^{(n)}\right) dS_X^{(n)}, \\ Q_2^{(n)}(t) &= -\int_0^t S_X^{(n)}\, dS_Y^{(n)}, \\ Q_3^{(n)}(t) &= -\int_0^t \left(1 - S_X^{(n)}\right) dS_Z^{(n)}. \end{aligned} \tag{2}$$

Under assumptions A1–A5, Chang and Yang (1987) showed that $S_X^{(n)} \to S_X$, $S_Y^{(n)} \to S_Y$ and $S_Z^{(n)} \to S_Z$ uniformly on $[0, T]$ almost surely. Let

$$u^{(n)} = \sqrt{n}\left(S_X^{(n)} - S_X, S_Y^{(n)} - S_Y, S_Z^{(n)} - S_Z\right)^\tau \tag{3}$$

and

$$q^{(n)} = \sqrt{n}\left(Q_1^{(n)} - Q_1, Q_2^{(n)} - Q_2, Q_3^{(n)} - Q_3\right)^\tau, \tag{4}$$

where τ denotes transpose. Subtracting (1) from (2) and noting that $u_1^{(n)}(0) = u_2^{(n)}(0) = u_3^{(n)}(\infty) = 0$, we obtain

$$(5)\qquad u_1^{(n)}(t) = -\int_0^t \frac{dq_1^{(n)}}{S_Y - S_Z} - \int_0^t \frac{u_2^{(n)} - u_3^{(n)}}{S_Y - S_Z} dS_X^{(n)},$$

$$(6)\qquad u_2^{(n)}(t) = -\int_0^t \frac{dq_2^{(n)}}{S_X} - \int_0^t \frac{u_1^{(n)}}{S_X} dS_Y^{(n)}$$

and

$$(7)\qquad u_3^{(n)}(t) = \int_t^\infty \frac{dq_3^{(n)}}{1 - S_X} - \int_t^\infty \frac{u_1^{(n)}}{1 - S_X} dS_Z^{(n)},$$

where $u_i^{(n)}$ and $q_i^{(n)}$ are the ith elements of $u^{(n)}$ and $q^{(n)}$, respectively. Under assumption A6, $u_3^{(n)}(T) = 0$ almost surely. Thus, (7) can be rewritten as

$$(8)\qquad u_3^{(n)}(t) = \int_t^T \frac{dq_3^{(n)}}{1 - S_X} - \int_t^T \frac{u_1^{(n)}}{1 - S_X} dS_Z^{(n)}.$$

A more convenient representation of (5), (6) and (8) results if we use the notation

$$(9)\quad \theta^{(n)}(t) = \frac{1}{\sqrt{n}}\left(-\int_0^t \frac{u_2^{(n)} - u_3^{(n)}}{S_Y - S_Z} du_1^{(n)}, -\int_0^t \frac{u_1^{(n)}}{S_X} du_2^{(n)}, -\int_t^T \frac{u_1^{(n)}}{1 - S_X} du_3^{(n)}\right)^\tau,$$

$$(10)\quad a^{(n)}(t) = \left(-\int_0^t \frac{dq_1^{(n)}}{S_Y - S_Z}, -\int_0^t \frac{dq_2^{(n)}}{S_X}, \int_t^T \frac{dq_3^{(n)}}{1 - S_X}\right)^\tau,$$

$$\mu(ds) = -\operatorname{diag}(dS_X(s), dS_Y(s), dS_Z(s))$$

and

$$k(t, s) = (k_{ij}(t, s)),$$

where $k(t, s)$ is a 3×3 matrix with elements $k_{11} = k_{22} = k_{23} = k_{32} = k_{33} = 0$, $k_{13} = -k_{12}$,

$$k_{12}(t, s) = \frac{I[0 < s < t]}{S_Y(s) - S_Z(s)},$$

$$k_{21}(t, s) = \frac{I[0 < s < t]}{S_X(s)},$$

and

$$k_{31}(t, s) = \frac{I[t < s < T]}{1 - S_X(s)}.$$

Using this notation, (5), (6) and (8) can be rewritten as

$$(11)\qquad u^{(n)}(t) = a^{(n)}(t) + \theta^{(n)}(t) + \int_0^T \mu(ds) k(t, s) u^{(n)}(s).$$

A more compact representation of (11) is

$$(I - K)u^{(n)} = a^{(n)} + \theta^{(n)}, \tag{12}$$

where I is the identity operator and the operator K is defined as

$$Ku = \int_0^T \mu(ds)k(\cdot, s)u(s).$$

Theorem (2.1) guarantees existence of a unique solution for (12) under assumptions A1–A6. Let $D[0, T]$ denote the space of all functions on $[0, T]$ which are right continuous and have left limits. Let $\mathbf{D}[0, T] = D[0, T] \times D[0, T] \times D[0, T]$.

THEOREM 2.1. *Under assumptions* A1–A6 *there exists a resolvent kernel matrix*

$$\Gamma(t, s) = \big(\Gamma_{ij}(t, s)\big),$$

where Γ_{ij}, $i, j = 1, 2, 3$ *are bounded measurable functions on* $[0, T] \times [0, T]$ *such that for any* $a \in \mathbf{D}[0, T]$ *the integral equation*

$$(I - K)u = a$$

has the unique solution

$$u(t) = a(t) + \int_0^T \mu(ds)\Gamma(t, s)a(s),$$

or compactly,

$$u = (I + \Gamma)a.$$

PROOF. From classical results on the Fredholm integral equation, it follows that the conclusion in the theorem holds if 1 is not an eigenvalue of K, i.e., for any $u \in \mathbf{D}[0, T]$

$$(I - K)u = 0 \quad \Leftrightarrow \quad u \equiv 0 \text{ on } [0, T]. \tag{13}$$

The proof of (13) is technical and deferred to the Appendix. For a detailed discussion of the one-dimensional Fredholm integral equations see Cochran [(1972), Section 3.4]. Cochran's derivations can be translated almost line by line to the three-dimensional case. □

REMARK. For $T = +\infty$, (13) holds without assumption A6.

3. Weak convergence of $u^{(n)}$ to a Gaussian process. In Lemma 3.3, we shall prove that $\theta^{(n)}$ almost surely converges to zero uniformly on $[0, T]$. Now, each component of $\sqrt{n}\, a^{(n)}$ is a sum of n i.i.d. random variables. Therefore, $a^{(n)}$ converges to a three-dimensional Gaussian process, and the main result (Theorem 3.1), the weak convergence of $u^{(n)}$, follows from Lemma 3.3 and

$$u^{(n)} = (I + \Gamma)(a^{(n)} + \theta^{(n)}). \tag{14}$$

Let d_0 be the Skorohod metric on $D[0, T]$ as defined in Billingsley [(1968), Chapter 3], and define the metric on the product space $\mathbf{D}[0, T]$ given by

$$d(u, v) = \max(d_0(u_1, v_1), d_0(u_2, v_2), d_0(u_3, v_3)), \qquad u, v \in \mathbf{D}[0, T],$$

where $u = (u_1, u_2, u_3)^\tau$ and $v = (v_1, v_2, v_3)^\tau$. In what follows, we consider processes defined on $\mathbf{D}[0, T]$ endowed with the topology determined by the metric d.

We begin with two results (Lemmas 3.1 and 3.2) needed to establish the asymptotic negligibility of $\theta^{(n)}$.

LEMMA 3.1. *If V is a subset of $D[0, T]$ with compact closure, then almost surely*

$$\left(\int_0^t v\, d\big(S_X^{(n)} - S_X\big), \int_0^t v\, d\big(S_Y^{(n)} - S_Y\big), \int_0^t v\, d\big(S_Z^{(n)} - S_Z\big)\right)^\tau, \tag{15}$$

converges to 0 *uniformly in $t \in [0, T]$ and $v \in V$.*

PROOF. We shall prove uniform and almost sure convergence for the first component of (15). The proof for the other two is similar.

Because V has compact closure, V is totally bounded. That is, for any $\varepsilon > 0$, there exist step functions $a_1, a_2, \dots, a_k \in D[0, T]$ such that for any $v \in V$

$$d_0(v, a_i) < \varepsilon. \tag{16}$$

for some i, $1 \le i \le k$. From Billingsley [(1968), page 111] it follows that there exists a strictly increasing, continuous mapping λ of $[0, T]$ onto itself, which depends on v, such that

$$\sup_{0 \le t \le T} |\lambda(t) - t| < \varepsilon$$

and

$$\sup_{0 \le t \le T} |v(t) - a_i(\lambda(t))| < \varepsilon.$$

Furthermore, the inequality

$$\left|\int_0^t v\, d\big(S_X^{(n)} - S_X\big)\right| \le \left|\int_0^t (v - a_i\lambda)\, d\big(S_X^{(n)} - S_X\big)\right| + \left|\int_0^t a_i\lambda\, d\big(S_X^{(n)} - S_X\big)\right|$$

and the uniform convergence of $S_X^{(n)}$ to S_X imply that there exists N such that if $n > N$, then

$$\left|\int_0^t v\, d\big(S_X^{(n)} - S_X\big)\right| < 3\varepsilon$$

for all $t \in [0, T]$ and all $v \in V$ almost surely. Therefore,

$$\int_0^t v\, d\big(S_X^{(n)} - S_X\big)$$

almost surely converges to 0 uniformly in $t \in [0, T]$ and $v \in V$. □

Let $u_\alpha^{(n)}$ and $q_\alpha^{(n)}$ denote, respectively, the expressions (3) and (4) with $\sqrt{n}$ replaced by n^α, while $\theta_\alpha^{(n)}$ and $a_\alpha^{(n)}$ denote the expressions (9) and (10) with $u^{(n)}$, $q^{(n)}$ and $\sqrt{n}$ replaced by $u_\alpha^{(n)}$, $q_\alpha^{(n)}$ and n^α, respectively.

LEMMA 3.2. *For any* α, $0 \le \alpha < \frac{1}{2}$, $u_\alpha^{(n)}$ *almost surely converges to* 0 *uniformly on* $[0, T]$.

PROOF. Replacing $u^{(n)}$ and $q^{(n)}$ in (5), (6) and (8), respectively, by $u_\alpha^{(n)}$ and $q_\alpha^{(n)}$, we obtain

$$u_{\alpha 1}^{(n)}(t) = -\int_0^t \frac{dq_{\alpha 1}^{(n)}}{S_Y - S_Z} - \int_0^t \frac{u_{\alpha 2}^{(n)} - u_{\alpha 3}^{(n)}}{S_Y - S_Z} dS_X^{(n)}, \tag{17}$$

$$u_{\alpha 2}^{(n)}(t) = -\int_0^t \frac{dq_{\alpha 2}^{(n)}}{S_X} - \int_0^t \frac{u_{\alpha 1}^{(n)}}{S_X} dS_Y^{(n)}, \tag{18}$$

and

$$u_{\alpha 3}^{(n)}(t) = \int_t^T \frac{dq_{\alpha 3}^{(n)}}{1 - S_X} - \int_t^T \frac{u_{\alpha 1}^{(n)}}{1 - S_X} dS_Z^{(n)}. \tag{19}$$

Equations (17), (18) and (19) have the representation

$$(I - K) u_\alpha^{(n)} = a_\alpha^{(n)} + \theta_\alpha^{(n)}.$$

Note that, under assumption A6, $dq_{\alpha 3}^{(n)} = dS_Z^{(n)} = 0$ almost surely on $(0, \delta)$ and $S_Y - S_Z > \varepsilon > 0$ on $(0, T]$. Using integration by parts and the assumption $0 \le \alpha < \frac{1}{2}$, it is easy to show that $a_\alpha^{(n)}$ almost surely converges to 0 uniformly on $[0, T]$.

The remainder of the proof will be for a fixed sample point at which

$$a_\alpha^{(n)} \to 0 \quad \text{and} \quad \left(S_X^{(n)}, S_Y^{(n)}, S_Z^{(n)}\right) \to (S_X, S_Y, S_Z)$$

uniformly on $[0, T]$ and $dq_{\alpha 3}^{(n)} = dS_Z^{(n)} = 0$ on $(0, \delta)$.

Let

$$\|u\| = \sup_{0 \le t \le T} \max(|u_1(t)|, |u_2(t)|, |u_3(t)|).$$

We shall discuss separately the case where $\{\|u_\alpha^{(n)}\|\}$ is bounded and the case where $\{\|u_\alpha^{(n)}\|\}$ is unbounded.

CASE I. ($\{\|u_\alpha^{(n)}\|\}$ is bounded). We shall apply Theorem 14.3 in Billingsley (1968), which is an analogue of the Arzela–Ascoli theorem, to prove that $u_\alpha^{(n)}$ has a compact closure in $\mathbf{D}[0, T]$. Since $\|u_\alpha^{(n)}\|$ is bounded, $\{u_\alpha^{(n)}\}$ obviously satisfies condition (14.32) in Billingsley (1968). Because $a_\alpha^{(n)} \to 0$ uniformly on $[0, T]$, the sequence $\{a_\alpha^{(n)}\}$ satisfies condition (14.33) in Billingsley (1968). Since $S_X^{(n)} \to S_X$, $S_Y^{(n)} \to S_Y$ and $S_Z^{(n)} \to S_Z$ uniformly on $[0, T]$ and S_X, S_Y, S_Z are continuous functions, the maximum jump size of $S_X^{(n)}$, $S_Y^{(n)}$ and $S_Z^{(n)}$ converges to 0. It follows that the sequence of the second term of the right-hand side of each equation in (17)–(19) also satisfies condition (14.33) in Billingsley (1968). Consequently, Theorem 14.3 in Billingsley (1968) implies that $\{u_\alpha^{(n)}\}$ has a compact closure in $\mathbf{D}[0, T]$.

From Lemma 3.1

$$\theta_\alpha^{(n)}(t) = \left(\int_0^t \frac{u_{\alpha 2}^{(n)} - u_{\alpha 3}^{(n)}}{S_Y - S_Z} d\left(S_X^{(n)} - S_X\right), \int_0^t \frac{u_{\alpha 1}^{(n)}}{S_X} d\left(S_Y^{(n)} - S_Y\right), \int_t^T \frac{u_{\alpha 1}^{(n)}}{1 - S_X} d\left(S_Z^{(n)} - S_Z\right) \right)^T \tag{20}$$

converges to 0 uniformly on $[0, T]$. Therefore,

$$(I - K)u_\alpha^{(n)} \to 0$$

uniformly on $[0, T]$, and for any subsequence $\{u_\alpha^{(n_k)}\}$ which converges to $u_\alpha^{(0)}$ in $\mathbf{D}[0, T]$ we have

$$(I - K)u_\alpha^{(0)} = 0.$$

Theorem 2.1 implies that $u_\alpha^{(0)} \equiv 0$. Therefore, $\{u_\alpha^{(n)}\}$ converges to 0 in $\mathbf{D}[0, T]$, and it follows that $\{u_\alpha^{(n)}\}$ converges to 0 uniformly on $[0, T]$.

Case II. ($\{\|u_\alpha^{(n)}\|\}$ is unbounded). We shall show that the unboundedness of $\{\|u_\alpha^{(n)}\|\}$ will lead to a contradiction. Let $C_n = \|u_\alpha^{(n)}\|$. We can select a subsequence of $\{\|u_\alpha^{(n)}\|\}$ indexed by $\{n_k\}$ such that $C_{n_k} \to \infty$. Define a new sequence

$$v_\alpha^{(n_k)} = \frac{1}{C_{n_k}} u_\alpha^{(n_k)}.$$

As in Case I, we can prove that $\{v_\alpha^{(n_k)}\}$ converges to 0 uniformly on $[0, T]$, which contradicts

$$\|v_\alpha^{(n_k)}\| = 1. \qquad \square$$

Lemma 3.3. *The sequence of vectors* $\theta^{(n)}$ *almost surely converges to* 0 *uniformly on* $[0, T]$.

Proof. In view of (9), it is sufficient to prove that for any bounded measurable function F on $[0, T]$

$$\sup_{0 \le t \le T} \frac{1}{\sqrt{n}} \left| \int_0^t F u_i^{(n)}\, d\left(u_j^{(n)}\right) \right| \xrightarrow{P} 0, \qquad i \neq j \text{ and } i, j = 1,2,3.$$

We shall prove convergence for $i = 1$, $j = 2$. The other cases can be proved similarly. In the proof we assume $F(t) \equiv 1$ on $[0, T]$. The proof can be easily generalized to $F \not\equiv 1$.

Equation (6) implies that

$$du_2^{(n)} = -\frac{1}{S_X}\left[dq_2^{(n)} + u_1^{(n)}\, dS_Y^{(n)}\right].$$

Consequently,

$$\frac{1}{\sqrt{n}}\int_0^t u_1^{(n)}\, du_2^{(n)} = -\int_0^t \frac{S_X^{(n)} - S_X}{S_X}\, dq_2^{(n)} - \frac{1}{\sqrt{n}}\int_0^t \frac{\left(u_1^{(n)}\right)^2}{S_X}\, dS_Y^{(n)}$$

$$= -\frac{\left(S_X^{(n)}(t) - S_X(t)\right) q_2^{(n)}(t)}{S_X(t)} - \int_0^t q_2^{(n)} \frac{S_X^{(n)} - S_X}{S_X^2}\, dS_X$$

$$+ \int_0^t \frac{q_2^{(n)}}{S_X}\, d\left(S_X^{(n)} - S_X\right) - \frac{1}{\sqrt{n}}\int_0^t \frac{\left(u_1^{(n)}\right)^2}{S_X}\, dS_Y^{(n)}.$$

Strong consistency of $S_X^{(n)}$ implies that the first two terms almost surely converge to 0 uniformly on $[0, T]$. As in the proof of Theorem 4 in Breslow and Crowley (1974), the uniform and almost sure convergence to 0 of the third term can be derived from Lemma 3.1 and tightness of the empirical process $q_2^{(n)}$. Finally, Lemma 3.2 implies that the last term almost surely converges to 0 uniformly on $[0, T]$. □

We are now in a position to state the main theorem.

THEOREM 3.1. *Under assumptions* A1–A6, $u^{(n)}$ *converges weakly to a Gaussian process on* $\mathbf{D}[0, T]$.

PROOF. Theorem 2.1 guarantees the validity of (14), where Γ is a continuous mapping from $\mathbf{D}[0, T]$ into $\mathbf{D}[0, T]$. Lemma 3.3 demonstrates that $\theta^{(n)}$ is a higher order term. It follows that $(I + \Gamma)\theta^{(n)}$ almost surely converges to 0 uniformly on $[0, T]$. The weak convergence of process $a^{(n)}$ and Theorem 5.1 in Billingsley (1968) imply that $(I + \Gamma)a^{(n)}$ converges to a Gaussian process on $\mathbf{D}[0, T]$. Consequently, the conclusion in the theorem follows. □

4. Asymptotic variance of $\sqrt{n}\,(S_X^{(n)} - S_X)$. Since the asymptotic variance–covariance matrix of $u^{(n)}$ obtained using (14) involves a three-dimensional integral equation, the numerical computation may be complicated. In addition, the variance–covariance matrix of $u_2^{(n)}$ and $u_3^{(n)}$ may not be of interest in practice. In this section we shall provide a method for calculating the asymptotic variance of $u_1^{(n)} = \sqrt{n}\,(S_X^{(n)} - S_X)$ using the theory of Fredholm integral equations and the technique of influence curves. For the idea of influence curves see Huber (1981) and Reid (1981).

Substituting $u_2^{(n)}$ and $u_3^{(n)}$ in (5) by the right-hand side of (6) and (8), respectively, we have

$$\begin{aligned}
u_1^{(n)}(t) &= -\int_0^t \frac{dq_1^{(n)}}{S_Y - S_Z} + \int_0^t \left(\int_0^s \frac{dq_2^{(n)}}{S_X} \right) \frac{dS_X(s)}{(S_Y - S_Z)(s)} \\
&\quad + \int_0^t \left(\int_s^T \frac{dq_3^{(n)}}{1 - S_X} \right) \frac{dS_X(s)}{(S_Y - S_Z)(s)} \\
&\quad + \int_0^t \frac{dS_X(s)}{(S_Y - S_Z)(s)} \left[\int_0^s \frac{u_1^{(n)}}{S_X} dS_Y - \int_s^T \frac{u_1^{(n)}}{1 - S_X} dS_Z \right] + o_p^{(n)}(1) \\
&= -\int_0^t \frac{dq_1^{(n)}}{S_Y - S_Z} + \int_0^t \left(\int_s^t \frac{dS_X}{S_Y - S_Z} \right) \frac{dq_2^{(n)}(s)}{S_X(s)} \\
&\quad + \int_0^T \left(\int_0^{s \wedge t} \frac{dS_X}{S_Y - S_Z} \right) \frac{dq_3^{(n)}(s)}{1 - S_X(s)} \\
&\quad + \int_0^T u_1^{(n)}(s) \left[\frac{I[0 \le s \le t]\, dS_Y(s)}{S_X(s)} \left(\int_s^t \frac{dS_X}{S_Y - S_Z} \right) \right. \\
&\qquad \left. - \frac{dS_Z}{1 - S_X} \int_0^{s \wedge t} \frac{dS_X}{S_Y - S_Z} \right] + o_p^{(n)}(1),
\end{aligned}$$

where $s \wedge t = \min(s, t)$ and $o_p^{(n)}(1)$ almost surely converges to 0 uniformly on $[0, T]$ as $n \to \infty$.

Therefore, we have the following theorem.

THEOREM 4.1. *The process* $u_1^{(n)} = \sqrt{n}(S_X^{(n)} - S_X)$ *satisfies the following integral equation on* $[0, T]$

$$u_1^{(n)}(t) = b_n(t) + \int_0^T g(t, s, ds) u_1^{(n)}(s) + o_p^{(n)}(1), \tag{21}$$

where

$$\begin{aligned}
b_n(t) &= -\int_0^t \frac{dq_1^{(n)}}{S_Y - S_Z} + \int_0^t \left(\int_s^t \frac{dS_X}{S_Y - S_Z} \right) \frac{dq_2^{(n)}(s)}{S_X(s)} \\
&\quad + \int_0^T \left(\int_0^{s \wedge t} \frac{dS_X}{S_Y - S_Z} \right) \frac{dq_3^{(n)}(s)}{1 - S_X(s)},
\end{aligned}$$

and

$$g(t, s, ds) = \frac{I[0 \le s \le t]\, dS_Y(s)}{S_X(s)} \int_s^t \frac{dS_X}{S_Y - S_Z} - \frac{dS_Z(s)}{1 - S_X(s)} \int_0^{s \wedge t} \frac{dS_X}{S_Y - S_Z}.$$

Since (21) is a one-dimensional Fredholm integral equation, Theorem 2.1 implies that there exists a resolvent function

$$\Gamma_0(t, s, ds) = (dS_Y(s), dS_Z(s))(\gamma_1(t, s), \gamma_2(t, s))^{\tau}$$

such that

$$u_1^{(n)}(t) = b_n(t) + \int_0^T \Gamma_0(t, s, ds) b_n(s) + o_p^{(n)}(1), \tag{22}$$

where γ_i, $i = 1, 2$, are bounded and measurable functions. Define

$$F_1(t, s) = -\frac{I[0 \le s \le t]}{S_Y(s) - S_Z(s)},$$

$$F_2(t, s) = \frac{I[0 \le s \le t]}{S_X(s)} \int_s^t \frac{dS_X}{S_Y - S_Z}$$

and

$$F_3(t, s) = \frac{1}{1 - S_X(s)} \int_0^{s \wedge t} \frac{dS_X}{S_Y - S_X}.$$

Then, the first term in the right-hand side of (22) can be written as

$$b_n(t) = \int_0^T F_1(t, s)\, dq_1^{(n)}(s) + \int_0^T F_2(t, s)\, dq_2^{(n)}(s) + \int_0^T F_3(t, s)\, dq_3^{(n)}(s).$$

Therefore, the second term in the right-hand side of (22) is equal to

$$\begin{aligned}
&\int_0^T \Gamma_0(t, s, ds) \left[\int_0^T F_1(s, u)\, dq_1^{(n)}(u) + F_2(s, u)\, dq_2^{(n)}(u) + F_3(s, u)\, dq_3^{(n)}(u) \right] \\
&\quad = \int_0^T \left\{ \left[\int_0^T \Gamma_0(t, s, ds) F_1(s, u) \right] dq_1^{(n)}(u) \right. \\
&\qquad \left. + \left[\int_0^T \Gamma_0(t, s, ds) F_2(s, u) \right] dq_2^{(n)}(u) + \left[\int_0^T \Gamma_0(t, s, ds) F_3(s, u) \right] dq_3^{(n)}(u) \right\}.
\end{aligned}$$

Furthermore

$$\begin{aligned}
u_1^{(n)}(t) &= \int_0^T IC_1(t, s)\, dq_1^{(n)}(s) + \int_0^T IC_2(t, s)\, dq_2^{(n)}(s) \\
&\quad + \int_0^T IC_3(t, s)\, dq_3^{(n)}(s) + o_p^{(n)}(1),
\end{aligned}$$

where

$$IC_1(t, s) = F_1(t, s) + \int_0^T \Gamma_0(t, v, dv) F_1(v, s),$$

$$IC_2(t, s) = F_2(t, s) + \int_0^T \Gamma_0(t, v, dv) F_2(v, s)$$

and

$$IC_3(t, s) = F_3(t, s) + \int_0^T \Gamma_0(t, v, dv) F_3(v, s).$$

Thus we have

Theorem 4.2. *The process* $u_1^{(n)} = \sqrt{n}(S_X^{(n)} - S_X)$ *converges weakly to a Gaussian process with asymptotic variance*

$$\int_0^T IC_1^2(t,u)\,dQ_1(u) + \int_0^T IC_2^2(t,u)\,dQ_2(u) + \int_0^T IC_3^2(t,u)\,dQ_3(u)$$
$$-\left\{\int_0^T [IC_1(t,u)\,dQ_1(u) + IC_2(t,u)\,dQ_2(u) + IC_3(t,u)\,dQ_3(u)]\right\}^2.$$

We further note that $IC_i(t,u)$ *is the solution of the integral equation*

$$IC_i(t,u) = F_i(t,u) + \int_0^T g(t,s,ds) IC_i(s,u), \qquad i = 1,2,3.$$

APPENDIX

Proof of (13). The argument in the proof is an analogue of that of Theorem 3.2 in Chang and Yang (1987).

Assume $(I - K)u = 0$ and $u \in \mathbf{D}[0,T]$. Since u_i's are bounded, the components in Ku are absolutely continuous. Consequently, the u_i's are absolutely continuous. $(I-K)u = 0$ can be rewritten as

$$\begin{aligned} &\int_0^t (S_Y - S_Z)\,du_1 + \int_0^t (u_2 - u_3)\,dS_X = 0,\\ &\int_0^t S_X\,du_2 + \int_0^t u_1\,dS_Y = 0,\\ &\int_0^t (1 - S_X)\,du_3 - \int_0^t u_1\,dS_Z = 0,\\ &u_1(0) = u_2(0) = u_3(T) = 0. \end{aligned} \tag{23}$$

Note that $S_Y - S_Z > 0$ and $1 - S_X > 0$ on $(0,T]$. We shall use the following two results in the sequel:

(A) If $u_2 - u_3 \geq 0$ (≤ 0) on $(t_1, t_2) \subset [0,T]$, then $du_1 \geq 0$ (≤ 0) on (t_1, t_2).

(B) If $u_1 \geq 0$ (≤ 0) on $(t_1, t_2) \subset [0,T]$, then $du_2 \geq 0$ (≤ 0) and $du_3 \leq 0$ (≥ 0) on (t_1, t_2).

If $u_1 \not\equiv 0$ on $[0,T]$, there are only two cases:

Case I. There exist t_1 and t_2, $0 \leq t_1 \leq t_2 \leq T$, such that $u_1(t_1) = u_1(t_2) = 0$ and u_1 keeps the same sign on (t_1, t_2); positive, say.

Case II. u_1 keeps the same sign on $(0,T]$; positive, say.

We shall prove that each case leads to a contradiction.

Case I. From $u_1 > 0$, we have $du_2 \geq 0$ and $du_3 \leq 0$ on (t_1, t_2). If $(u_2 - u_3)$ keeps the same sign on (t_1, t_2), then u_1 is a nonincreasing or nondecreasing function on (t_1, t_2), which contradicts with the assumptions that $u_1(t_1) = u_1(t_2) = 0$ and $u_1(t) > 0$ on (t_1, t_2). So $u_2 - u_3$ doesn't keep the same sign on (t_1, t_2). From the continuity of $u_2 - u_3$, there exists a $t^* \in (t_1, t_2)$ such that $u_2(t^*) - u_3(t^*) = 0$. For any $t \in [t^*, t_2]$,

$$u_2(t) - u_3(t) = \int_{t^*}^{t} (du_2 - du_3) \geq 0.$$

From (A), u_1 is nondecreasing on (t^*, t_2), which contradicts the assumptions that $u_1(t^*) > 0$ and $u_1(t_2) = 0$.

CASE II. The assumption that $u_1 > 0$ on $(0, T]$ implies that u_2 is nondecreasing and u_3 is nonincreasing. Because of $u_2(0) = u_3(T) = 0$ and the continuity of u_2 and u_3, there exists a $t^* \in [0, T]$ such that $u_2(t^*) = u_3(t^*)$. If $u_3 \not\equiv 0$, then $t^* > 0$. For any $t \in (0, t^*)$,

$$u_2(t) - u_3(t) = -\int_{t}^{t^*} (du_2 - du_3) \leq 0.$$

Consequently, $du_1 \leq 0$ on $(0, t^*)$, which contradicts with that $u_1(0) = 0$ and $u_1(t^*) > 0$. If $u_3 \equiv 0$, by adding the first three equations in (23), we obtain

$$(S_Y(t) - S_Z(t))u_1(t) + u_2(t)S_X(t) = 0.$$

Since $u_1 > 0$, $S_Y - S_Z > 0$, $u_2 \geq 0$ and $S_X > 0$ on $(0, T]$, the left-hand side is positive.

Thus we have proved that $u_1 \equiv 0$ on $[0, T]$. From the second and the third equations in (23), it is easy to see that $u_2 \equiv u_3 \equiv 0$ on $[0, T]$. □

Acknowledgments. The paper is a continuation of my Ph.D. thesis. I would like to thank Professor Grace L. Yang for guidance, helpful discussions and encouragement during the course of my Ph.D. thesis at the University of Maryland, College Park. The author is grateful to Professor P. V. Rao and an Associate Editor for valuable suggestions which led to a considerably improved presentation of the results.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2** 437–453.

CHANG, M. N. and YANG, G. L. (1987). Strong consistency of a nonparametric estimator of the survival function with doubly censored data. *Ann. Statist.* **15** 1536–1547.

COCHRAN, J. A. (1972). *The Analysis of Linear Integral Equations*. McGraw-Hill, New York.

FERNHOLZ, L. T. (1983). *Von Mises Calculus for Statistical Functionals. Lecture Notes in Statist.* **19** Springer, Berlin.

GEHAN, E. A. (1965). A generalized two-sample Wilcoxon test for doubly censored data. *Biometrika* **52** 650–653.

GILL, R. D. (1989). Non- and semi-parametric maximum likelihood estimators and the von Mises method (part I). *Scand. J. Statist.* **16** 97–128.

HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.

MANTEL, N. (1967). Ranking procedures for arbitrarily restricted observations. *Biometrics* **23** 65–78.

PETO, R. (1973). Experimental survival curves for interval censored data. *J. Roy. Statist. Soc. Ser. C* **22** 86–91.

REEDS, J. A. (1976). On the definition of von Mises functionals. Report S-44, Dept. Statist., Harvard Univ.

REID, N. (1981). Influence functions for censored data. *Ann. Statist.* **9** 78–92.

TSAI, W. Y. and CROWLEY, J. (1985). A large sample study of generalized maximum likelihood estimators from incomplete data via self-consistency. *Ann. Statist.* **13** 1317–1334. (Correction: *Ann. Statist.* **18** 470.)

TURNBULL, B. W. (1974). Nonparametric estimation of a survivorship function with doubly censored data. *J. Amer. Statist. Assoc.* **69** 169–173.

DEPARTMENT OF STATISTICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA 32611