The Annals of Statistics
1991, Vol. 19, No. 1, 417–442

ESTIMATING A DISTRIBUTION FUNCTION WITH TRUNCATED AND CENSORED DATA[1]

BY TZE LEUNG LAI AND ZHILIANG YING

Stanford University and University of Illinois

A minor modification of the product-limit estimator is proposed for estimating a distribution function (not necessarily continuous) when the data are subject to either truncation or censoring, or to both, by independent but not necessarily identically distributed truncation–censoring variables. Making use of martingale integral representations and empirical process theory, uniform strong consistency of the estimator is established and weak convergence results are proved for the entire observable range of the function. Numerical results are also given to illustrate the usefulness of the modification, particularly in the context of truncated data.

1. Introduction. Let $X_1, X_2, \dots,$ be i.i.d. random variables with a common distribution function F (i.e., $F(t) = P\{X_i \le t\}$). Let $\{t_i\}$ be a sequence of independent random variables independent of $\{X_i\}$. Suppose that the X_i are not completely observable and that observations are $(\tilde{X}_i, \delta_i)$, $i = 1, \dots, n$, where

$$\tilde{X}_i = X_i \wedge t_i, \qquad \delta_i = I_{\{X_i \le t_i\}}. \tag{1.1}$$

Here and in the sequel, we use $\wedge$ and $\vee$ to denote minimum and maximum, respectively, and use the notation $\#A$ to denote the number of elements of a set A. Based on these censored observations, a classical estimator of F is the product-limit estimator $\tilde{F}_n$, introduced by Kaplan and Meier (1958) and defined by

$$1 - \tilde{F}_n(t) = \prod_{s \le t} \left\{1 - \frac{\Delta N_n(s)}{Y_n(s)}\right\}, \tag{1.2}$$

where we use the convention $0/0 = 0$, $\Delta N_n(s) = N_n(s) - N_n(s-)$ and

$$N_n(s) = \#\{i \le n \colon \tilde{X}_i \le s, \delta_i = 1\} = \sum_{i=1}^{n} I_{\{X_i \le s \wedge t_i\}}, \tag{1.3}$$

$$Y_n(s) = \#\{i \le n \colon \tilde{X}_i \ge s\} = \sum_{i=1}^{n} I_{\{X_i \wedge t_i \ge s\}}. \tag{1.4}$$

Received May 1988; revised February 1990.
[1]Research supported by NSF and the Army Research Office.
AMS 1980 *subject classifications*. Primary 62E20, 62G05; secondary 60F05.
Key words and phrases. Product-limit estimator, truncated data, censoring, strong consistency, weak convergence, empirical process, stochastic integral, martingales.

Making use of stochastic integral representations and martingales, Gill (1980, 1983) established the uniform consistency and functional central limit theorem of $\tilde{F}_n$ under considerably weaker assumptions than previous authors.

Another model of incomplete observations studied in the literature is the "truncated model," which assumes the presence of truncation variables t_i such that X_i is only observable if $X_i \geq t_i$. We assume without loss of generality the case of left truncation, since a right-truncated model (in which the X_i are observable only if $X_i \leq T_i$) can be transformed into a left-truncated model by multiplying the X_i and T_i by -1. In the left-truncated model, the data consist of n observations (X_i^0, t_i^0) with $X_i^0 \geq t_i^0$. We can regard the observed sample as being generated by a larger sample of independent random variables $X_i, t_i, i = 1, \dots, m(n)$, where

$$m(n) = \inf\left\{m\colon \sum_{i=1}^{m} I_{\{X_i \geq t_i\}} = n\right\}. \tag{1.5}$$

In his study of truncated data from an application in astronomy, Lynden-Bell (1971) used nonparametric maximum likelihood arguments to derive the product-limit estimator F_n^*, which he defined from the truncated data by

$$1 - F_n^*(t) = \prod_{s \leq t} \left\{1 - \frac{\Delta L_{m(n)}(s)}{R_{m(n)}(s)}\right\}, \tag{1.6}$$

where

$$L_m(s) = \sum_{i=1}^{m} I_{\{t_i \leq X_i \leq s\}} \qquad \left[\text{so } L_{m(n)}(s) = \#\{i \leq n\colon X_i^0 \leq s\}\right], \tag{1.7}$$

$$R_m(s) = \sum_{i=1}^{m} I_{\{t_i \leq s \leq X_i\}} \qquad \left[\text{so } R_{m(n)}(s) = \#\{i \leq n\colon X_i^0 \geq s \geq t_i^0\}\right]. \tag{1.8}$$

Assuming the t_i to be i.i.d. with a common distribution function G and assuming that both F and G are continuous, Woodroofe (1985) showed that

$$\sup_{t \geq \tau_G} \left|F_n^*(t) - F_G(t)\right| \to_P 0 \quad \text{if } F(\tau_G) < 1, \tag{1.9}$$

where $\tau_G = \inf\{s\colon G(s) > 0\}$ and

$$F_G(t) = P\{X_1 \leq t | X_1 > \tau_G\}. \tag{1.10}$$

Assuming furthermore that

$$\int_{\tau_G}^{\infty} \frac{dF}{G} < \infty, \tag{1.11}$$

he also showed that $n^{1/2}(F_n^* - F_G)$ converges weakly to a Gaussian process on $[\tau_G, b]$ for every $b > \tau_G$ with $F(b) < 1$. Wang, Jewell and Tsai (1986) later gave an explicit formula for the covariance function of the limiting Gaussian process. The arguments of Woodroofe (1985) and Wang, Jewell and Tsai (1986) are similar in spirit to the approach of Breslow and Crowley (1974) in the censored case, working directly with empirical subdistribution functions and their functionals. By embedding the estimation problem in a Markov process model with a finite state space and by applying counting process theory and martingale central limit theorems to analyze F_n^*, Keiding and Gill (1990) gave an alternative derivation of these results.

The assumption of i.i.d. t_i is crucial in the asymptotic results of Woodroofe (1985) and Wang, Jewell and Tsai (1986). These results are, therefore, not applicable to the important case where the t_i need not be identically distributed, as in regression applications [cf. Wang, Jewell and Tsai (1986) and Lai and Ying (1988)]. As will be explained in Section 2, it is necessary to modify (1.6) to ensure consistency of the estimator when the t_i are not i.i.d. Specifically, choosing $c > 0$ and $0 < \alpha < 1$, we shall define the (modified) product-limit estimator $\hat{F}_n$ by

$$1 - \hat{F}_n(t) = \prod_{s \le t} \left(1 - \frac{\Delta L_{m(n)}(s)}{R_{m(n)}(s)} I_{\{R_{m(n)}(s) \ge cn^\alpha\}} \right). \tag{1.12}$$

Strong consistency of $\hat{F}_n$ will be established in Section 2, while weak convergence of the normalized process will be studied in Section 4, *without assuming F to be continuous and without assuming the t_i to be identically distributed or continuous*. Sections 3 and 4 extend the consistency and weak convergence results to the mixed model in which the data are subject to both censoring and truncation. In Section 5, we give some numerical results to illustrate the usefulness of the weight function $I_{\{R_{m(n)}(s) \ge cn^\alpha\}}$ in (1.12).

2. Uniform strong consistency of $\hat{F}_n$ and of related estimators in the truncated model. Throughout the sequel we shall let

$$G_m(s) = \frac{1}{m} \sum_{i=1}^{m} P\{t_i \le s\}, \qquad \tau_G = \inf\left\{s \colon \liminf_{m \to \infty} G_m(s) > 0\right\}. \tag{2.1}$$

In the case of censored survival data (1.1), the set $\{i \le n \colon \tilde{X}_i \ge s\}$ in (1.4) is often called the "risk set" at (age) s, so $Y_n(s)$ in (1.4) represents the size of the risk set. Note that $n^{-1}EY_n(s) = (1 - F(s-))(1 - G_n(s-))$ is monotone in s and converges to 0 only when $F(s-)$ or $G_n(s-)$ converges to 1. In the case of truncated data, an analogous "risk set" at s is given by $\{i \le m \colon X_i \ge s \ge t_i\}$ in (1.8), where $R_m(s)$ denotes the risk set size. However, $m^{-1}ER_m(s) = (1 - F(s-))G_m(s)$ is no longer monotone in s and converges to 0 if $G_m(s) \to 0$ or $F(s-) \to 1$. To ensure that $m/R_m(s)$ is not too large, particularly in those factors of (1.6) corresponding to s near τ_G, we introduce the indicator variable $I_{\{R_{m(n)}(s) \ge cn^\alpha\}}$ in (1.12). The following theorem establishes the uniform strong consistency of the (modified) product-limit estimator (1.12) as an estimate of (1.10). We shall let $G_m^{-1}(y) = \inf\{s \colon G_m(s) \ge y\}$.

THEOREM 1. *Define* $\hat{F}_n$ *by* (1.12), τ_G *by* (2.1) *and* F_G *by* (1.10). *Suppose that* $F(\tau_G) < 1$ *and*

$$\lim_{m\to\infty} \{F(\tau_G) - F([\tau_G \wedge G_m^{-1}(bm^{-(1-\alpha)})] -)\} = 0 \quad \textit{for every } b > 0. \tag{2.2}$$

Then $\sup_{-\infty<t<\infty}|\hat{F}_n(t) - F_G(t)| \to 0$ *a.s.*

REMARKS. (i) Assumption (2.2) implies that F is continuous at τ_G ($\geq -\infty$), setting $F(-\infty) = 0$. When $t_1, t_2, \dots$ are i.i.d. with a common distribution function G, $G_m = G$ and $G^{-1}(bm^{-(1-\alpha)}) \geq \tau_G$, and therefore assumption (2.2) is satisfied in this case if F is continuous, as was assumed by Woodroofe (1985) and Wang, Jewell and Tsai (1986).

(ii) Consider the following example. Suppose that $X_1, X_2, \dots$ are i.i.d. exponential random variables with density $f(x) = e^{-x}$, $x > 0$. Suppose that the t_i are independent random variables such that $P\{t_i = 1/i\} = 2^{-i} = 1 - P\{t_i = 1\}$. Then $\tau_G = 1$ and (2.2) holds. Moreover, since $\sum_{i=1}^m I_{\{t_i \leq s \leq X_i\}}$ converges a.s. to a limit $Z(s)$ with $EZ(s) = e^{-s}\sum_{i\geq 1/s} 2^{-i}$ for $s \leq 1/2$, and since $\sum_{i=1}^\infty I_{\{t_i\leq s\}}$ converges a.s. for $s \leq 1/2$ and is nondecreasing in s, it follows that $P\{\lim_{m\to\infty} R_m(s) = Z(s)$ for all $0 < s \leq 1/2\} = 1$. Using this, it can be shown that $\prod_{s\leq 1/2}\{1 - \Delta L_{m(n)}(s)/R_{m(n)}(s)\}$ converges a.s. to a random variable ($\neq 1$), and therefore Lynden-Bell's estimator (1.6) cannot converge to $F_G(t)$ in probability, for every $t \geq \tau_G = 1$.

The product-limit estimator (1.6) or its modification (1.12) tries to estimate the conditional distribution function F_G of X_1 given $X_1 > \tau_G$, without assuming knowledge of the value of τ_G. It is, however, often more natural to consider instead the problem of estimating the conditional distribution

$$F(x|u) = P\{X_1 \leq x | X_1 > u\}, \qquad x > u, \tag{2.3}$$

for various given values of u about which the sample appears to provide adequate information. Given the value of u, an obvious modification of (1.6) to estimate (2.3) is

$$F_{n,u}^*(x) = 1 - \prod_{u<s\leq x}\left\{1 - \frac{\Delta L_{m(n)}(s)}{R_{m(n)}(s)}\right\}. \tag{2.4}$$

Analogous to Theorem 1, the following theorem shows that $F(\cdot|u)$ can be consistently estimated by $F_{n,u}^*$ if $t_n^0([cn^\alpha]) \leq u \leq X_n^0([\beta n])$, where $[z]$ denotes the largest integer less than or equal to z, $t_n^0(1) \leq \cdots \leq t_n^0(n)$ denote the order statistics of the t_i^0 and $X_n^0(1) \leq \cdots \leq X_n^0(n)$ denote the order statistics of the X_i^0, $i = 1, \dots, n$.

THEOREM 2. *Define* $F(\cdot|u)$ *by* (2.3) *and* $F_{n,u}^*$ *by* (2.4). *Suppose that*

$$\liminf_{m\to\infty} \int_{-\infty}^{\infty} G_m\, dF > 0.$$

Then for every $c > 0$, $0 < \alpha < 1$ *and* $0 < \beta < 1$,

$$\sup_{t_n^0([cn^\alpha]) \le u \le X_n^0([\beta n]),\, x > u} \left| F_{n,u}^*(x) - F(x|u) \right| \to 0 \quad a.s.$$

We preface the proof of Theorems 1 and 2 by the following three lemmas. Let

$$\Lambda(t) = \int_{-\infty}^{t} \frac{dF(s)}{1 - F(s-)} \tag{2.5}$$

denote the cumulative hazard function of F. Here and in the sequel by $\int_{-\infty}^{t}$ we mean integration over the interval $(-\infty, t]$.

LEMMA 1. *For any* $0 \le p < 1$, $B > 0$ *and* $\varepsilon > 0$,

$$\sup_{(1-F(s-))G_m(s) \ge Bm^{-p}} \left| \frac{\sum_{i=1}^{m} I_{\{t_i \le s \le X_i\}}}{m(1 - F(s-))G_m(s)} - 1 \right| \to 0 \quad a.s., \tag{2.6}$$

$$\begin{aligned} &\sup_{(1-F(s-))G_m(s) \le Bm^{-p}} \left| \sum_{i=1}^{m} I_{\{t_i \le s \le X_i\}} - m(1 - F(s-))G_m(s) \right| \\ &\quad + \sup_{G_m(s) \le Bm^{-p}} \left| \sum_{i=1}^{m} I_{\{t_i \le s\}} - mG_m(s) \right| = O(m^{(1-p)/2+\varepsilon}) \quad a.s. \end{aligned} \tag{2.7}$$

PROOF. To prove (2.6), apply Corollary 1.3 of Alexander (1985) to the empirical measure defined by the independent random vectors $(t_1, X_1), \dots, (t_m, X_m)$, noting that the L_2 metric entropy with bracketing of the class of sets of the form $\{t_i \le s_1 (\text{or } t_i < s_1),\ X_i \ge s_2 (\text{or } X_i > s_2)\}$ is of logarithmic order and that $E(\sum_{i=1}^{m} I_{\{t_i \le s \le X_i\}}) = m(1 - F(s-))G_m(s)$. The desired conclusion (2.6) then follows from the exponential bound in Alexander (1985), Corollary 1.3, and the Borel–Cantelli lemma. Similarly, (2.7) follows from Alexander (1985), Theorem 2.1, and the Borel–Cantelli lemma. □

LEMMA 2. *Suppose that* $\liminf_{m \to \infty} \int_{-\infty}^{\infty} G_m\, dF = \delta > 0$. *Then*

$$\limsup_{n \to \infty} m(n)/n = 1/\delta \quad a.s.$$

In particular, if $F(\tau_G) < 1$, *then*

$$\delta \ge \sup_{a > \tau_G,\, F(a-) < 1} \left\{ (1 - F(a-)) \liminf_{m \to \infty} G_m(a) \right\} > 0.$$

PROOF. By Kolmogorov's strong law of large numbers,

$$\frac{1}{m} \sum_{i=1}^{m} \left[I_{\{X_i \ge t_i\}} - P\{X_i \ge t_i\} \right] \to 0 \quad \text{a.s.} \tag{2.8}$$

Since $m^{-1}\sum_{i=1}^{m} P\{X_i \geq t_i\} = \int_{-\infty}^{\infty} G_m \, dF$, the desired conclusion follows from (1.5) and (2.8). If $F(\tau_G) < 1$, then we can choose $a > \tau_G$ such that $1 - F(a-) > 0$. Since $a > \tau_G$, $\liminf_{m\to\infty} G_m(a) > 0$. Moreover, $\int_{-\infty}^{\infty} G_m \, dF \geq G_m(a)\int_{[a,\infty)} dF$. □

LEMMA 3. *For every a with $F(a-) < 1$ and for every $0 \leq p < 1$,*

$$\sup\left\{\left|\sum_{u\leq s\leq x} \frac{\Delta L_m(s)}{R_m(s)} - \int_{[u,x]} d\Lambda\right| : G_m^{-1}(m^{-p}) \leq u \leq x \leq a\right\} \to 0$$

$$a.s.\ (\sup\varnothing = 0).$$

PROOF. First note that

$$m^{-1}ER_m(s) = (1 - F(s-))G_m(s) \geq (1 - F(a-))G_m(u)$$

if $u \leq s \leq a$. Therefore, by (2.6), with probability 1, if $a > G_m^{-1}(m^{-p})$, then

$$\begin{aligned}\sup_{G_m^{-1}(m^{-p})\leq u\leq x\leq a}&\left|\sum_{u\leq s\leq x} \frac{\Delta L_m(s)}{R_m(s)} - \sum_{u\leq s\leq x} \frac{\Delta L_m(s)}{ER_m(s)}\right| \\ &= o\left(\sum_{G_m^{-1}(m^{-p})\leq s\leq a} \frac{\Delta L_m(s)}{ER_m(s)}\right).\end{aligned} \tag{2.9}$$

Take any $0 < \varepsilon < 1$. Let $s_0 = G_m^{-1}(m^{-p})$ and let $s_j = \inf\{s > s_{j-1}: F(s) - F(s_{j-1}) \geq \varepsilon\}$. There are only finitely many such points $s_0 < \cdots < s_{N-1}$ that are less than or equal to a; in fact, $N - 1 \leq 1/\varepsilon$, assuming that $a \geq G_m^{-1}(m^{-p})$. Redefine $s_N = a$. Note that

$$\begin{aligned}\int_{(s_{j-1},s_j)} d\Lambda &\leq (1 - F(a-))^{-1}\big(F(s_j-) - F(s_{j-1})\big) \\ &\leq \varepsilon(1 - F(a-))^{-1} \quad \text{for } 1 \leq j \leq N.\end{aligned} \tag{2.10}$$

Fix any $0 \leq j \leq k \leq N$. In view of (1.7),

$$\sum_{s\in[s_j,s_k]} \frac{\Delta L_m(s)}{ER_m(s)} = \sum_{i=1}^{m} \frac{I_{\{t_i\leq X_i,\, s_j\leq X_i\leq s_k\}}}{m(1 - F(X_i-))G_m(X_i)} = \frac{1}{m}\sum_{i=1}^{m} Z_{mi}, \tag{2.11}$$

where

$$Z_{mi} = \frac{I_{\{t_i\leq X_i,\, s_j\leq X_i\leq s_k\}}}{(1 - F(X_i-))G_m(X_i)}.$$

Since $Z_{m1}, \ldots, Z_{mm}$ are independent random variables with $|Z_{mi}| \leq (1 -$

$F(a-))^{-1}m^p$, and since

$$\sum_{i=1}^{m} EZ_{mi} = m\int_{[s_j, s_k]} \frac{dF(s)}{1-F(s-)},$$

$$\sum_{i=1}^{m} \operatorname{Var} Z_{mi} \le \frac{m}{(1-F(a-))^2}\int_{[s_j, s_k]} \frac{dF(s)}{G_m(s)} \le \frac{m^{1+p}}{(1-F(a-))^2},$$

we obtain by Lemma 3(i) of Lai and Wei (1982) that for $0 < \theta < (1-p)/2$,

$$P\left\{\left|\sum_{i=1}^{m} Z_{mi} - m\int_{[s_j, s_k]} d\Lambda\right| \ge (1-F(a-))^{-1}m^{(1+p)/2+\theta}\right\} \le 2\exp\left\{-\tfrac{1}{2}m^{2\theta}\left(1-\tfrac{1}{2}m^{\theta+(p-1)/2}\right)\right\}. \tag{2.12}$$

From (2.11), (2.12) and the Borel–Cantelli lemma, it then follows that

$$\max_{0\le j\le k\le N}\left|\sum_{s\in[s_j, s_k]} \frac{\Delta L_m(s)}{ER_m(s)} - \int_{[s_j, s_k]} d\Lambda\right| \to 0 \quad \text{a.s.} \tag{2.13}$$

A similar argument shows that (2.13) still holds with $[s_j, s_k]$ replaced by $[s_j, s_k)$ or $(s_j, s_k]$. Combining this with (2.10) gives that

$$\limsup_{m\to\infty}\left\{\sup_{s_0\le u\le x\le s_N}\left|\sum_{u\le s\le x} \frac{\Delta L_m(s)}{ER_m(s)} - \int_{[u,x]} d\Lambda\right|\right\} \le \frac{2\varepsilon}{1-F(a-)} \quad \text{a.s.,} \tag{2.14}$$

noting that for $u, x \in [s_0, s_N]$ with $u \le x$, there exist $j, k \in \{1, \dots, N\}$ with $j \le k$ such that $s_{j-1} \le u \le s_j$ and $s_{k-1} \le x \le s_k$. From (2.9) and (2.14), the desired conclusion follows by letting $\varepsilon \downarrow 0$. □

Proof of Theorem 1. By Lemma 2, we can choose $b > 0$ such that

$$P\{2b(m(n))^\alpha < cn^\alpha \text{ for all large } n\} = 1. \tag{2.15}$$

By Lemma 1, $P(\Omega_0) = 1$, where

$$\Omega_0 = \left\{\sup_{(1-F(s-))G_m(s)\ge bm^{-(1-\alpha)}}\left|\frac{R_m(s)}{m(1-F(s-))G_m(s)} - 1\right| < \frac{1}{2} \right. \\ \left. \text{and} \sup_{(1-F(s-))G_m(s)\le bm^{-(1-\alpha)}} R_m(s) < 2bm^\alpha \text{ for all large } m\right\}. \tag{2.16}$$

On Ω_0, for all large m,

$$R_m(s) \ge 2bm^\alpha \Rightarrow (1-F(s-))G_m(s) \ge bm^{-(1-\alpha)}. \tag{2.17}$$

Since $G_m(s) < bm^{-(1-\alpha)}$ if $s < G_m^{-1}(bm^{-(1-\alpha)})$, it follows from (2.17) that on Ω_0, for all large m,

$$\begin{aligned}&\int_{-\infty}^{x} \frac{I_{\{R_m(s)\geq 2bm^{\alpha}\}}}{R_m(s)}\,dL_m(s)\\ &\qquad = 0, \qquad\qquad \text{if } x < G_m^{-1}(bm^{-(1-\alpha)}),\\ &\qquad \leq \sum_{G_m^{-1}(bm^{-(1-\alpha)})\leq s\leq x} \frac{\Delta L_m(s)}{R_m(s)}, \quad \text{if } x \geq G_m^{-1}(bm^{-(1-\alpha)}).\end{aligned} \tag{2.18}$$

Take $a > d > \tau_G$ such that $F(a-) < 1$. Since $G_m(s) \geq G_m(d)$ for $s \geq d$ and since $m(n) \geq n$, we have on Ω_0,

$$\inf_{d\leq s\leq a} I_{\{R_{m(n)}(s)\geq cn^{\alpha}\}} = 1 \quad \text{for all large } n. \tag{2.19}$$

Let $d_m = \tau_G \wedge G_m^{-1}(bm^{-(1-\alpha)})$. From (2.18) and Lemma 3, it follows that on Ω_0,

$$\begin{aligned}&\limsup_{m\to\infty} \int_{-\infty}^{d} \left[I_{\{R_m(s)\geq 2bm^{\alpha}\}}/R_m(s)\right] dL_m(s)\\ &\qquad \leq \limsup_{m\to\infty} \int_{[d_m, d]} d\Lambda \quad \text{a.s.}\end{aligned} \tag{2.20}$$

Moreover, since $\liminf_{m\to\infty} G_m(d) > 0$ and $1 - F(a-) > 0$, we can apply Lemma 3 to conclude that

$$\sup_{d\leq x\leq a} \left| \int_{[d,x]} dL_m(s)/R_m(s) - \int_{[d,x]} d\Lambda \right| \to 0 \quad \text{a.s.} \tag{2.21}$$

From (2.21), it follows that with probability 1, uniformly in $d < x \leq a$,

$$\begin{aligned}&\prod_{d<s\leq x} \{1 - \Delta L_m(s)/R_m(s)\}\\ &\qquad \to \exp\{-[\Lambda^c(x) - \Lambda^c(d)]\} \prod_{d<s\leq x} \{1 - \Delta\Lambda(s)\} \quad \text{as } m \to \infty,\end{aligned} \tag{2.22}$$

where Λ^c denotes the continuous part of Λ [cf. the proof of Lemma 2 of Gill (1981)]. Combining (2.19) and (2.22) gives that with probability 1, uniformly in $d < x \leq a$,

$$\begin{aligned}&(1 - \hat{F}_n(x))/(1 - \hat{F}_n(d))\\ &\qquad \to \exp\{-[\Lambda^c(x) - \Lambda^c(d)]\} \prod_{d<s\leq x} \{1 - \Delta\Lambda(s)\} \quad \text{as } n \to \infty.\end{aligned} \tag{2.23}$$

Note that the right-hand side of (2.23) [cf. Gill (1980), page 36] is equal to

$$\begin{aligned}&(1 - F(x))/(1 - F(d))\\ &\qquad \to (1 - F(x))/(1 - F(\tau_G)) = 1 - F_G(x) \quad \text{as } d \downarrow \tau_G.\end{aligned} \tag{2.24}$$

By (2.2), $\int_{[d_m, d]} d\Lambda \to 0$ as $m \to \infty$ and $d \downarrow \tau_G$. Therefore, by (2.15) and (2.20), with probability 1,

$$\begin{aligned} &\limsup_{n\to\infty} \Big\{ \sup_{s\le d} I_{\{R_{m(n)}(s)\ge cn^\alpha\}} \Delta L_{m(n)}(s)/R_{m(n)}(s) \Big\} \\ &\quad \le \limsup_{m\to\infty} \sum_{s\le d} I_{\{R_m(s)\ge 2bm^\alpha\}} \Delta L_m(s)/R_m(s) \to 0 \quad \text{as } d \downarrow \tau_G. \end{aligned} \tag{2.25}$$

Hence, with probability 1, for all large m,

$$\begin{aligned} \log(1-\hat F_n(d)) &= \sum_{s\le d} \log\{1 - I_{\{R_{m(n)}(s)\ge cn^\alpha\}} \Delta L_{m(n)}(s)/R_{m(n)}(s)\} \\ &= -\sum_{s\le d} \big[I_{\{R_{m(n)}(s)\ge cn^\alpha\}} \Delta L_{m(n)}(s)/R_{m(n)}(s)\big](1+o(1)) \\ &\qquad\qquad \text{as } d\downarrow \tau_G. \end{aligned} \tag{2.26}$$

From (2.15), (2.25) and (2.26), it follows that with probability 1,

$$\limsup_{n\to\infty} \hat F_n(d) \to 0 \quad \text{as } d \downarrow \tau_G. \tag{2.27}$$

Combining (2.27) with (2.23) and (2.24), we obtain that with probability 1,

$$1 - \hat F_n(x) \to 1 - F_G(x) \quad \text{uniformly in } x \le a. \tag{2.28}$$

Since a can be arbitrarily chosen with $F(a-) < 1$, (2.28) implies that $\sup_x |\hat F_n(x) - F_G(x)| \to 0$ a.s. if $F^{-1}(1) = \infty$, or if F is continuous at $F^{-1}(1)$. In the case F jumps at $\tau = F^{-1}(1)$, we can simply take $a = \tau$ since $F(\tau -) < 1$. □

Proof of Theorem. In view of (2.7) and the fact that $\sup_s |m^{-1}\sum_1^m I_{\{X_i\ge s\}} - (1-F(s-))| \to 0$ a.s., we can choose $b > 0$ by Lemmas 1 and 2 so that $P(\Omega_1) = 1$, where

$$\begin{aligned} \Omega_1 &= \{3bm(n) < n \text{ and } 2b(m(n))^\alpha < cn^\alpha \text{ for all large } n\} \\ &\quad \cap \Big\{ \sup_{1-F(s-)\le b(1-\beta)} \sum_{i=1}^m I_{\{X_i\ge s\}} < 2b(1-\beta)m \text{ for all large } m \Big\} \\ &\quad \cap \Big\{ \sup_{G_m(s)\le bm^{-(1-\alpha)}} \sum_{i=1}^m I_{\{t_i\le s\}} < [2bm^\alpha] \text{ for all large } m \Big\}. \end{aligned} \tag{2.29}$$

On Ω_1, for all large m,

$$\sum_{i=1}^m I_{\{t_i\le s\}} \ge [2bm^\alpha] \Rightarrow G_m(s) \ge bm^{-(1-\alpha)} \Rightarrow s \ge G_m^{-1}(bm^{-(1-\alpha)}), \tag{2.30}$$

$$\sum_{i=1}^m I_{\{X_i\ge s\}} \ge 2b(1-\beta)m \Rightarrow 1 - F(s-) \ge b(1-\beta). \tag{2.31}$$

On Ω_1, for all large n, since $cn^\alpha > 2b(m(n))^\alpha$ and $n > 3bm(n)$, we have

$$\sum_{i=1}^{m(n)} I_{\{t_i \le t_n^0([cn^\alpha])\}} \ge \sum_{i=1}^{m(n)} I_{\{t_i \le t_n^0([cn^\alpha]),\, t_i \le X_i\}} = [cn^\alpha] \ge \left[2b(m(n))^\alpha\right],$$

$$\sum_{i=1}^{m(n)} I_{\{X_i \ge X_n^0([\beta n])\}} \ge \sum_{i=1}^{m(n)} I_{\{X_i \ge X_n^0([\beta n]),\, X_i \ge t_i\}} = n - [\beta n] + 1 \ge 2bm(n)(1-\beta),$$

and therefore by (2.30) and (2.31),

$$t_n^0([cn^\alpha]) \ge G_m^{-1}\left(b(m(n))^{-(1-\alpha)}\right), \qquad 1 - F\left(X_n^0([\beta n]) -\right) \ge b(1-\beta), \tag{2.32}$$

for all large n, on Ω_1.

Take any a such that $F(a-) < 1$. By Lemma 3,

$$\lim_{m\to\infty} \left\{ \sup_{G_m^{-1}(bm^{-(1-\alpha)}) \le u < x \le a} \left| \int_{(u,x]} dL_m(s)/R_m(s) - \int_{(u,x]} d\Lambda \right| \right\} = 0 \quad \text{a.s.} \tag{2.33}$$

From (2.33), it follows as in (2.22) that

$$\sup_{G_m^{-1}(bm^{-(1-\alpha)}) \le u < x \le a} \left| \prod_{u < s \le x} \{1 - \Delta L_m(s)/R_m(s)\} - (1-F(x))/(1-F(u)) \right| \to 0 \quad \text{a.s.} \tag{2.34}$$

[cf. the proof of Lemma 2 of Gill (1981)]. From (2.32) and (2.34), it follows that

$$\sup\left\{ \left| F_{n,u}^*(x) - F(x|u) \right| : G_m^{-1}(bm^{-(1-\alpha)}) \le u \le X_n^0([\beta n]),\, u < x \le a \right\} \to 0 \quad \text{a.s.} \tag{2.35}$$

In view of (2.32), since (2.35) holds for every a with $F(a-) < 1$, the desired conclusion follows (cf. the last part of the proof of Theorem 1). □

3. Extension to mixed models with both censoring and truncation. In this section we extend the modified product-limit estimators to the case of left-truncated and right-censored data, which often arise in biostatistical applications. Specifically, in addition to the truncation variables t_i described above, there are also censoring variables T_i such that $\{(X_i, t_i, T_i): i = 1, 2, \dots\}$ is a sequence of independent random vectors and such that X_i is independent of (t_i, T_i) for every i. The X_i are assumed to have the same distribution F, which is to be estimated, while t_i and T_i are assumed to be extended real-valued random variables (i.e., T_i and t_i may assume the values ∞ and $-\infty$ with positive probability). Thus, the truncated model studied above is a special case that corresponds to $T_i \equiv \infty$, while the classical censored model is also a special case that corresponds to $t_i \equiv -\infty$.

Letting $\tilde{X}_i = X_i \wedge T_i$ and $\delta_i = I_{\{X_i \le T_i\}}$, this mixed model assumes that $(\tilde{X}_i, \delta_i)$ is observable only when $\tilde{X}_i \ge t_i$. Thus, the data consist of n observations $(\tilde{X}_i^0, \delta_i^0, t_i^0)$ with $\tilde{X}_i^0 \ge t_i^0$. As before, we shall regard the observed sample as being generated by a larger sample of independent random variables X_i, t_i, T_i, $i = 1, \dots, m(n)$, where

$$(3.1) \qquad m(n) = \inf\left\{m: \sum_{i=1}^{m} I_{\{X_i \wedge T_i \ge t_i\}} = n\right\}.$$

Analogous to (1.3) and (1.7), and to (1.4) and (1.8), define

$$(3.2) \qquad L_m(s) = \sum_{i=1}^{m} I_{\{t_i \le X_i \le s \wedge T_i\}} \qquad \left[\text{so } L_{m(n)}(s) = \#\{i \le n: \tilde{X}_i^0 \le s,\ \delta_i^0 = 1\}\right],$$

$$(3.3) \qquad R_m(s) = \sum_{i=1}^{m} I_{\{t_i \le s \le X_i \wedge T_i\}} \qquad \left[\text{so } R_{m(n)}(s) = \#\{i \le n: \tilde{X}_i^0 \ge s \ge t_i^0\}\right].$$

With these new definitions of L_m and R_m [in place of (1.7) and (1.8)], define the product-limit estimator $F_{n,u}^*$ by (2.4) and $\hat{F}_n$ by (1.12), where $c > 0$ and $0 < \alpha < 1$ are prespecified constants. Instead of (2.1), we now defined G_m and τ_G by

$$(3.4) \qquad G_m(s) = \frac{1}{m}\sum_{i=1}^{m} P\{t_i \le s \le T_i\}, \qquad \tau_G = \inf\left\{s: \liminf_{m\to\infty} G_m(s) > 0\right\}.$$

With this definition of G_m, define F_G by (1.10). Although $G_m(s)$ is no longer monotone in s, we still use the notation

$$(3.5) \qquad G_m^{-1}(y) = \inf\{x: G_m(x) \ge y\} \qquad \left[\text{and therefore } G_m(x) \ge y \Rightarrow x \ge G_m^{-1}(y)\right].$$

Assume that $F(\tau_G) < 1$, so $-\infty \le \tau_G < F^{-1}(1)$, and define

$$(3.6) \qquad \tau^* = F^{-1}(1) \wedge \inf\left\{s > \tau_G: \liminf_{m\to\infty} G_m(s) = 0\right\} \quad (\inf\varnothing = \infty).$$

This notation will be used throughout the sequel. The following two theorems are extensions of Theorems 1 and 2, while Lemma 4 below is an extension of Lemma 1 to the present setting and can be proved by applying Corollary 1.3 and Theorem 2.1 of Alexander (1985) to the empirical measure defined by independent random vectors (t_i, T_i, X_i).

THEOREM 3. *With G_m defined by (3.4) instead of (2.1), suppose that $F(\tau_G) < F(\tau^* -)$ and that (2.2) holds, in which G_m^{-1} is defined in (3.5).*

Assume that $\liminf_{m\to\infty}(\inf_{d\le s\le a} G_m(s)) > 0$ *if* $\tau_G < d < a < \tau^*$. *Then*

$$\sup_{t<\tau^*} |\hat{F}_n(t) - F_G(t)| \to 0 \quad a.s. \tag{3.7}$$

THEOREM 4. *With* G_m *defined in* (3.4), *define* $G_m^*(s) = m^{-1}\sum_{i=1}^m P\{t_i \le T_i < s\}$ *and assume that* $\liminf_{m\to\infty} \int_{-\infty}^{\infty}(G_m + G_m^*)\,dF > 0$. *Order* $\tilde{X}_1^0, \dots, \tilde{X}_n^0$ *as* $\tilde{X}_n^0(1) \le \cdots \le \tilde{X}_n^0(n)$. *Then for every* $c > 0$, $0 < \alpha < 1$ *and* $0 < \beta < 1$,

$$\begin{aligned}\sup\Big\{|F_{n,u}^*(x) - F(x|u)| &: u \le \tilde{X}_n^0([\beta n]),\ x > u, \\ &\text{and} \min_{u<s\le x} R_{m(n)}(s) \ge cn^{\alpha}\Big\} \to 0 \quad a.s.\end{aligned} \tag{3.8}$$

LEMMA 4. *With* G_m *defined by* (3.4), *we have for any* $0 \le p < 1$, $B > 0$ *and* $\varepsilon > 0$,

$$\sup_{(1-F(s-))G_m(s)\ge Bm^{-p}} \left| \frac{\sum_{i=1}^m I_{\{t_i\le s\le X_i\wedge T_i\}}}{m(1-F(s-))G_m(s)} - 1 \right| \to 0 \quad a.s., \tag{3.9}$$

$$\begin{aligned}&\sup_{(1-F(s-))G_m(s)\le Bm^{-p}} \left| \sum_{i=1}^m I_{\{t_i\le s\le X_i\wedge T_i\}} - m(1-F(s-))G_m(s) \right| \\ &\quad + \sup_{G_m(s)\le Bm^{-p}} \left| \sum_{i=1}^m I_{\{t_i\le s\le T_i\}} - mG_m(s) \right| = O(m^{(1-p)/2+\varepsilon}) \quad a.s.\end{aligned} \tag{3.10}$$

PROOF OF THEOREM 4. In analogy with Lemma 2 and its proof, we now have with probability 1,

$$m^{-1}\sum_{i=1}^m I_{\{X_i\wedge T_i\ge t_i\}} + o(1) = m^{-1}\sum_{i=1}^m P\{X_i\wedge T_i \ge t_i\} = \int_{-\infty}^{\infty} (G_m + G_m^*)\,dF.$$

It then follows from (3.1) that

$$\limsup_{n\to\infty} m(n)/n = \left\{\liminf_{m\to\infty} \int_{-\infty}^{\infty} (G_m + G_m^*)\,dF\right\}^{-1} \ (<\infty) \quad \text{a.s.} \tag{3.11}$$

Since $G_m(s)$ is no longer monotone in s, we need to modify Lemma 3 as follows: For every a with $F(a-) < 1$ and for every $0 \le p < 1$,

$$\begin{aligned}\sup\Bigg\{\left|\sum_{u<s\le x} \frac{\Delta L_m(s)}{R_m(s)} - \int_{(u,x]} d\Lambda\right| &: u < x \le a, \\ &\inf_{u<s\le x} G_m(s) \ge m^{-p}\Bigg\} \to 0 \quad \text{a.s.}\end{aligned} \tag{3.12}$$

The proof of (3.12) is similar to that of Lemma 3. First we make use of (3.9) [in place of (2.6)] to reduce the problem to that of

$$\sup\left\{\left|\sum_{u<s\le x}\frac{\Delta L_m(s)}{ER_m(s)}-\int_{(u,x]}d\Lambda\right|: u<x\le a,\right.\tag{3.13}$$
$$\left.\inf_{u<s\le x}G_m(s)\ge m^{-p}\right\}\to 0\quad\text{a.s.}$$

To prove (3.13), it suffices to show that

$$\sup_{u<x\le a}\left|\sum_{u<s\le x}\frac{\Delta L_m(s)}{ER_m(s)}I_{\{G_m(s)\ge m^{-p}\}}\right.\tag{3.14}$$
$$\left.-\int_{(u,x]}I_{\{G_m(s)\ge m^{-p}\}}\,d\Lambda(s)\right|\to 0\quad\text{a.s.,}$$

which can be proved by arguments similar to those used to prove (2.14) in Lemma 3.

By (3.11), we can choose $b > 0$ such that (2.15) still holds. Defining Ω_0 as in (2.16) but with the new definitions of G_m and R_m, we obtain by Lemma 4 (in place of Lemma 1) that $P(\Omega_0) = 1$. On Ω_0, for all large m, (2.17) holds and therefore

$$\inf_{u<s\le x}R_m(s)\ge 2bm^{\alpha}\Rightarrow \inf_{u<s\le x}G_m(s)\ge bm^{-(1-\alpha)}.\tag{3.15}$$

Moreover, an argument similar to that used to prove (2.32) can be used to show that with probability 1,

$$1-F\big(\tilde{X}_n^0([\beta n])-\big)\ge b(1-\beta)\quad\text{for all large } n,\tag{3.16}$$

choosing $b > 0$ such that $P\{3bm(n) < n$ for all large $n\} = 1$. The rest of the proof is similar to that of Theorem 2. □

REMARKS. (i) $R_{m(n)}(s)$ is a step function with possible jumps at t_i^0 and $\tilde{X}_i^0$. Theorem 4 says that $F_{n,u}^*(x)$ provides a reliable estimate of $F(x|u)$ for $x > u$ if the risk set size $R_{m(n)}(s)$ is not too small [say, $R_{m(n)}(s) \ge cn^{\alpha}$] for all s in the interval $(u, x]$, and if u is not located at the upper tail of F [as revealed from the sample with $u \le \tilde{X}_n^0([\beta n])$]. The preceding proof shows how the requirement of adequate risk set size $R_{m(n)}(s)$ for all $s \in (u, x]$ enables us to get around the difficulties caused by the nonmonotonicity of $G_m(s)$.

(ii) In the case $t_i \equiv -\infty$, Theorem 3 gives the uniform strong consistency of the modified product-limit estimator $\hat{F}_n$ in the censored model with independent (but possibly nonidentically distributed) censoring variables T_i. Here G_m is nonincreasing in s, $\tau_G = -\infty$, $F_G = F$ and (2.2) is obviously satisfied. As pointed out by Wang (1987), for the Kaplan–Meier estimator $\tilde{F}_n$ defined in (1.2), the uniform strong consistency property (3.7) has been established in the literature for i.i.d. censoring variables T_i with a common distribution function $1 - G$ only under the assumption $F(\tau^* -) = 1$ or $G(\tau^* -) > 0$.

PROOF OF THEOREM 3. Since $\liminf_{m\to\infty}(\inf_{d\le s\le a} G_m(s)) > 0$ for all $d, a \in (\tau_G, \tau^*)$ and since $F(\tau^* -) > F(\tau_G)$, it follows that $\liminf_{m\to\infty} \int_{(\tau_G,\tau^*)} G_m\, dF > 0$. Hence, by (3.11), we can choose $b > 0$ such that (2.15) still holds. Let $d_m = \tau_G \wedge G_m^{-1}(bm^{-(1-\alpha)})$, where G_m^{-1} is defined in (3.5), and let $\tau_G < d < a < \tau^*$. By Lemma 4, $P(\Omega_0) = 1$, where Ω_0 is defined in (2.16). On Ω_0, in view of (2.17) and (3.5),

$$\sup_{s < G_m^{-1}(bm^{-(1-\alpha)})} I_{\{R_m(s)\ge 2bm^{\alpha}\}} = 0 \quad \text{for all large } m, \tag{3.17}$$

$$\inf_{d\le s\le a} I_{\{R_{m(n)}(s)\ge cn^{\alpha}\}} = 1 \quad \text{for all large } n, \tag{3.18}$$

since $m(n) \ge n$ and $\liminf_{m\to\infty}\{\inf_{d\le s\le a} G_m(s)\} > 0$. Moreover, by (3.12),

$$\sup_{d<x\le a}\left|\int_{(d,x]} dL_m(s)/R_m(s) - \int_{(d,x]} d\Lambda\right| \to 0 \quad \text{a.s.} \tag{3.19}$$

On Ω_0, it follows from (2.17) and the definition of Ω_0 that for all large m,

$$\begin{aligned} &\int_{[d_m,d]} I_{\{R_m(s)\ge 2bm^{\alpha}\}}\, dL_m(s)/R_m(s) \\ &\quad\le 2\int_{[d_m,d]} I_{\{G_m(s)\ge bm^{-(1-\alpha)}\}}\, dL_m(s)/ER_m(s). \end{aligned} \tag{3.20}$$

Making use of (3.17)–(3.20) together with (3.14), it can be shown by a straightforward modification of the proof of Theorem 1 that

$$\sup_{x\le a}|\hat{F}_n(x) - F_G(x)| \to 0 \quad \text{a.s. for every } a \in (\tau_G, \tau^*). \tag{3.21}$$

If $F(\tau^* -) = 1$, then from (3.21) and the fact that $\hat{F}_n \le 1$, the desired conclusion (3.7) follows. Now suppose that $F(\tau^* -) < 1$. Then as in (3.20), we have on Ω_0, for all large m,

$$\begin{aligned} &\sum_{a\le s<\tau^*} \Delta L_m(s) I_{\{R_m(s)\ge 2bm^{\alpha}\}}/R_m(s) \\ &\quad\le 2\sum_{a\le s<\tau^*} \Delta L_m(s) I_{\{G_m(s)\ge bm^{-(1-\alpha)}\}}/ER_m(s). \end{aligned} \tag{3.22}$$

Since $ER_m(s) = (1 - F(s -))G_m(s)$ and since $F(\tau^* -) < 1$, it can be shown by arguments similar to those used to prove (2.14) that

$$\begin{aligned} \sup_{a\le x<\tau^*}\Bigg| &\sum_{x\le s<\tau^*} \Delta L_m(s) I_{\{G_m(s)\ge bm^{-(1-\alpha)}\}}/ER_m(s) \\ &- \int_{[x,\tau^*)} I_{\{G_m(s)\ge bm^{-(1-\alpha)}\}}\, d\Lambda(s)\Bigg| \to 0 \quad \text{a.s.} \end{aligned} \tag{3.23}$$

Since $F(\tau^* -) < 1$, it follows from (3.22) and (3.23) that with probability 1,

$$(3.24)\qquad \limsup_{m\to\infty} \sum_{a\le s<\tau^*} \Delta L_m(s) I_{\{R_m(s)\ge 2bm^\alpha\}}/R_m(s) \to 0 \quad \text{as } a\uparrow\tau^*.$$

By an argument similar to (2.25) and (2.26), (3.24) implies that with probability 1,

$$\limsup_{n\to\infty}\Big\{\sup_{a\le x<\tau^*} \log\big[(1-\hat{F}_n(x))/(1-\hat{F}_n(a))\big]\Big\} \to 0 \quad \text{as } a\uparrow\tau^*,$$

which in turn implies that with probability 1,

$$(3.25)\qquad \limsup_{n\to\infty}\Big\{\sup_{a\le x<\tau^*} \big|\hat{F}_n(x)-\hat{F}_n(a)\big|\Big\} \to 0 \quad \text{as } a\uparrow\tau^*.$$

From (3.21) and (3.25), the desired conclusion (3.7) follows since $F_G(a)\to F_G(\tau^*-)$, as $a\uparrow\tau^*$. □

4. Weak convergence of normalized $\hat{F}_n$ and $F^*_{n,u}$. In this section we establish the weak convergence of $n^{1/2}(\hat{F}_n - F_G)$ to $(1-F_G)$ times a zero-mean Gaussian process with independent increments when the data are subject to both left truncation and right censoring, as considered in Section 3. Defining G_m, τ_G and τ^* as in (3.4) and (3.6), it will be assumed throughout the sequel that $F(\tau_G) < F(\tau^*-)$ (so $-\infty\le\tau_G<\tau^*\le\infty$) and that

(4.1) $G_m(s)$ converges to a limit $G(s)$ as $m\to\infty$ for every $s\in(\tau_G,\tau^*)$, the convergence being uniform on I and $\inf_{s\in I} G(s)>0$ for all compact $I\subset(\tau_G,\tau^*)$,

$$(4.2)\qquad \lim_{m\to\infty}\int_{-\infty}^{\infty}\left[G_m(s)+m^{-1}\sum_{i=1}^m P\{t_i\le T_i<s\}\right]dF(s) = p \text{ exists.}$$

Let $\tau_G<d<a<\tau^*$. By (4.1), $\inf_{d\le s\le a} G_m(s)\to\inf_{d\le s\le a} G(s)$, and therefore (3.18) holds on an event Ω_0 with $P(\Omega_0)=1$. Note that for $x>u$,

$$(4.3)\qquad \begin{aligned}(1-\hat{F}_n(x))/(1-\hat{F}_n(u)) &= \prod_{u<s\le x}\big(1-\Delta L_{m(n)}(s) I_{\{R_{m(n)}(s)\ge cn^\alpha\}}/R_{m(n)}(s)\big)\\ &= 1-F^*_{n,u}(x) \quad \text{on } \Big\{\min_{u\le s\le x} R_{m(n)}(s)\ge cn^\alpha\Big\}.\end{aligned}$$

From (3.18) and (4.3), it follows that on Ω_0, for all large n,

$$(4.4)\qquad \begin{aligned}(1-\hat{F}_n(x))/(1-\hat{F}_n(u)) &= 1-F^*_{n,u}(x)\\ &= \frac{1-F^*_{n,d}(x)}{1-F^*_{n,d}(u)} \quad \text{for } d\le u<x\le a.\end{aligned}$$

Therefore, the weak convergence of $n^{1/2}(F^*_{n,d}-F(\cdot|d))$ to a Gaussian process in $D(d,a]$, which will be established in Theorem 5(i) below, implies corresponding weak convergence of the sequence of two-parameter processes $\{n^{1/2}(F^*_{n,u}(x)-F(x|u)): d\le u<x\le a\}$. A related weak convergence result for such two-parameter processes was recently obtained by Davidsen and Jacobsen (1989) by using the weak convergence theory of two-sided stochastic integrals.

Theorem 5(iii) gives conditions under which the weak convergence of $n^{1/2}(\hat{F}_n - F_G)$ can be extended to $D(-\infty, \tau^*)$. In the case of only left truncation (i.e., $T_i \equiv \infty$), these conditions are weaker than those of Woodroofe (1985) who assumes the truncation variables t_i to be i.i.d. Thus, Theorem 5(iii) not only generalizes Woodroofe's (1985) weak convergence theorem to the mixed censorship–truncation model, but it also extends the theorem to the setting of nonidentically distributed truncation and censoring variables t_i and T_i.

THEOREM 5. *With G_m and τ_G defined in* (3.4) *and τ^* defined in* (3.6), *suppose that $F(\tau_G) < F(\tau^* -)$ and that* (4.1) *and* (4.2) *hold.*

(i) *Let $\tau_G < d < a < \tau^*$. Then as $n \to \infty$, $n^{1/2}(F^*_{n,d} - F(\cdot|d))$ converges weakly in $D(d, a]$ to $p^{1/2}(1 - F(\cdot|d))W_d$, where p is given in* (4.2) *and $\{W_d(t), d < t \le a\}$ is a zero-mean Gaussian process with independent increments and variance function*

$$\operatorname{Var}(W_d(t)) = \int_{(d,t]} \{G(s)(1 - F(s-))(1 - F(s))\}^{-1} dF(s), \qquad d < t \le a.$$

(ii) *Suppose furthermore that $(0 <)\alpha < 1/2$ is chosen in the definition* (1.12) *of $\hat{F}_n$ and that*

$$\lim_{m\to\infty} m^{1/2}\left\{F\left(\tau_G \vee \sup\left\{s \le \frac{\tau^* + \tau_G}{2} : G_m(s) < Bm^{-(1-\alpha)}\right\}\right) - F\left(\left[\tau_G \wedge G_m^{-1}(bm^{-(1-\alpha)})\right] -\right)\right\} = 0 \tag{4.5}$$

for all $B > b > 0$,

$$\limsup_{m\to\infty} \int_{[\tau_G \wedge G_m^{-1}(bm^{-(1-\alpha)}), t]} dF(s)/G_m(s) \to 0 \tag{4.6}$$

as $t \downarrow \tau_G$ for every $b > 0$,

where G_m^{-1} is defined in (3.5). *Then for every $a \in (\tau_G, \tau^*)$, $n^{1/2}(\hat{F}_n - F_G)$ converges weakly in $D(-\infty, a]$ to $p^{1/2}(1 - F_G)W$, where $\{W(t), -\infty < t \le a\}$ is a zero-mean Gaussian process with independent increments and variance function*

$$\begin{aligned} \operatorname{Var}(W(t)) &= 0, & \text{if } t \le \tau_G, \\ &= \int_{(\tau_G, t]} \{G(s)(1 - F(s-))(1 - F(s))\}^{-1} dF(s), & \text{if } t > \tau_G. \end{aligned} \tag{4.7}$$

(iii) *Suppose that in addition to the assumptions of part* (ii) *we also assume that there exist $a \in (\tau_G, \tau^*)$, $0 \le \beta < 1 - 2\alpha$ and $m_0 \ge 1$ such that F is continuous on $[a, \tau^*)$ and*

$$G_m(s) \ge \{F(\tau^* -) - F(s)\}^\beta \quad \text{if } m \ge m_0 \text{ and } \tau^* > s \ge a. \tag{4.8}$$

Then as $n \to \infty$, $n^{1/2}(\hat{F}_n - F_G)$ *converges weakly in* $D(-\infty, \tau^*)$ *to* $p^{1/2}(1 - F_G)W$, *where* $\{W(t), -\infty < t < \tau^*\}$ *is a zero-mean Gaussian process with independent increments and variance function* (4.7).

REMARKS. (A) The assumption (4.5) is stronger than (2.2), which implies that F is continuous at τ_G.

(B) When the data are only subject to left truncation (i.e., $T_i \equiv \infty$), $\tau^* = F^{-1}(1)$ and $G_m(s)$ is nondecreasing in s. Hence, if $\lim_{m\to\infty} G_m(s) = G(s)$ exists for every $s \in (\tau_G, \tau^*)$, then (4.1) and (4.2) both hold; moreover, given every $\beta > 0$ we can choose m_0 and a such that (4.8) is satisfied. Suppose that the truncation variables $t_1, t_2, \dots$ are i.i.d. with a common continuous distribution function G and that F is continuous, as Woodroofe (1985), Wang, Jewell and Tsai (1987) and Keiding and Gill (1987) have assumed. Then $G_m \equiv G$, and the condition (1.11) assumed by these authors implies both (4.6) and (4.5) since $1 - \alpha > 1/2$ and

$$(4.9) \quad F\big(G^{-1}(Bm^{-(1-\alpha)})\big) - F(\tau_G) \le \int_{\tau_G}^{G^{-1}(Bm^{-(1-\alpha)})} \{Bm^{-(1-\alpha)}/G(s)\}\, dF(s).$$

To prove Theorem 5, we make use of martingale theory and stochastic integral representations, similar to those used by Gill (1980, 1983) in the censored case. An important idea that enables us to extend the well-known martingale integral representation for the Kaplan–Meier estimator of F based on censored data to the modified product-limit estimator based on left-truncated and right-censored data is to regard the observed sample of $(\tilde{X}_i^0, \delta_i^0, t_i^0)$, $i = 1, \dots, n$, as being generated by a larger sample of independent random vectors $(\tilde{X}_i, t_i, T_i)$, $i = 1, \dots, m(n)$, where $m(n)$ is defined in (3.1). For the case of only left truncation (i.e., $T_i \equiv \infty$) by i.i.d. continuous truncation variables t_i, Keiding and Gill (1987) used a different method to develop martingale integral representations. A key idea in their approach is to embed the observation (t, X), given the constraint $t \le X$, in a Markov process $U(s)$, $s \ge 0$, with five states and to work conditional on eventual absorption in one of these states. Identifying the observed sample as a randomly stopped sequence of independent random vectors, we get around the problem of dealing with the conditional probability measures given $t_i \le X_i$.

In the case of censored data, it is well known that $N_n(t) - \int_{-\infty}^t Y_n(s)\, d\Lambda(s)$ is a martingale, where N_n and Y_n are defined in (1.3) and (1.4) and Λ is the cumulative hazard function defined in (2.5). An extension to the case of left-truncated and right-censored data is given in Lemma 5.

LEMMA 5. *Defining* L_m *and* R_m *by* (3.2) *and* (3.3), *let*

$$(4.10) \qquad M_m(t) = L_m(t) - \int_{-\infty}^t R_m(s)\, d\Lambda(s).$$

Let $\tilde{X}_i = X_i \wedge T_i$, $\delta_i = I_{\{X_i \le T_i\}}$, *and let* $\mathscr{F}(s)$ *be the complete* σ*-field generated* *by*

$$(4.11)\quad t_i,\ I_{\{t_i \le \tilde{X}_i\}},\ \delta_i I_{\{t_i \le \tilde{X}_i \le s\}},\ I_{\{t_i \le u \le \tilde{X}_i\}},\ I_{\{t_i \le \tilde{X}_i \le u\}},\qquad u \le s,\ i = 1, 2, \dots.$$

Define $m(n)$ *by* (3.1). *Then* $\{M_{m(n)}(s), \mathscr{F}(s), -\infty < s < \infty\}$ *is a martingale with predictable variation process*

$$(4.12)\qquad \langle M_{m(n)}\rangle(t) = \int_{-\infty}^{t} R_{m(n)}(s)[1 - \Delta\Lambda(s)]\, d\Lambda(s).$$

PROOF. Let $\mathscr{G}(s)$ be the complete σ-field generated by T_i, $I_{\{X_i \le u\}}$, $u \le s$, $i = 1, 2, \dots$, together with all the random variables in (4.11). Then t_i, T_i and $I_{\{t_i \le \tilde{X}_i\}}$ are measurable with respect to $\mathscr{G}_{-\infty} = \bigcap_{u=-\infty}^{\infty} \mathscr{G}(u)$. Therefore, $m(n)$ is measurable with respect to $\mathscr{G}_{-\infty}$ and $R_{m(n)}(s) = \sum_1^{m(n)} I_{\{t_i \le s \le T_i\}} I_{\{X_i \ge s\}}$ is measurable with respect to $\mathscr{G}(s-)$. Moreover, for every i, $\{I_{\{t_i \le X_i \le s \wedge T_i\}} - \int_{[t_i, s\wedge T_i]} I_{\{X_i \ge u\}}\, d\Lambda(u), \mathscr{G}(s), -\infty < s < \infty\}$ is a martingale, recalling that $X_1, X_2, \dots, (t_1, T_1), (t_2, T_2), \dots$ are independent. Since

$$\int_{[t_i, s\wedge T_i]} I_{\{X_i \ge u\}}\, d\Lambda(u) = \int_{-\infty}^{s} I_{\{X_i \wedge T_i \ge u \ge t_i\}}\, d\Lambda(u),$$

it then follows that $\{M_{m(n)}(s), \mathscr{G}(s), -\infty < s < \infty\}$ is a martingale. Since $\mathscr{F}(s) \subset \mathscr{G}(s)$ and $M_{m(n)}(t)$ is measurable with respect to $\mathscr{F}(t)$, the desired conclusion follows. □

Making use of Lemma 5, we now give the martingale integral representation that is basic to our proof of Theorem 5. Define

$$(4.13)\qquad \begin{aligned} \Lambda_n(t) &= \int_{-\infty}^{t} I_{\{R_{m(n)}(s) \ge cn^{\alpha}\}}\, d\Lambda(s),\\ F_n(t) &= 1 - \exp(-\Lambda_n^c(t)) \sum_{s \le t} (1 - \Delta\Lambda_n(s)). \end{aligned}$$

From (1.12) and (4.13) together with Proposition A.4.1 of Gill (1980), pages 153–155, it follows that if $F_n(t) < 1$, then

$$(4.14)\qquad \begin{aligned} \frac{\hat{F}_n(t) - F_n(t)}{1 - F_n(t)} &= 1 - \frac{1 - \hat{F}_n(t)}{1 - F_n(t)}\\ &= \int_{-\infty}^{t} \frac{1 - \hat{F}_n(s-)}{(1 - F_n(s-))(1 - \Delta\Lambda_n(s))}\\ &\qquad \times \left(\frac{I_{\{R_{m(n)}(s) \ge cn^{\alpha}\}}}{R_{m(n)}(s)}\, dL_{m(n)}(s) - d\Lambda_n(s) \right)\\ &= \int_{-\infty}^{t} \frac{1 - \hat{F}_n(s-)}{1 - F_n(s)} \frac{I_{\{R_{m(n)}(s) \ge cn^{\alpha}\}}}{R_{m(n)}(s)}\, dM_{m(n)}(s). \end{aligned}$$

Note that (4.14) is analogous to Gill's (1980) formula (3.2.13) for the

Kaplan–Meier estimator in the censored model. Therefore,

$$n^{1/2}(\hat{F}_n(t) - F_n(t))/(1 - F_n(t)) = U_n(t) \quad \text{if } F_n(t) < 1, \tag{4.15}$$

where

$$U_n(t) = n^{1/2} \int_{-\infty}^{t} [(1 - \hat{F}_n(s-))/(1 - F_n(s))] \times \left[I_{\{R_{m(n)}(s) \geq cn^{\alpha}\}} / R_{m(n)}(s) \right] dM_{m(n)}(s). \tag{4.16}$$

For $\tau^* > t > d > \tau_G$, since $1 - F_{n,d}^*(t) = \prod_{d < s \leq t}(1 - \Delta L_{m(n)}(s)/R_{m(n)}(s))$, we similarly have

$$n^{1/2}(F_{n,d}^*(t) - F(t|d))/(1 - F(t|d)) = U_{n,d}(t), \tag{4.17}$$

where

$$U_{n,d}(t) = n^{1/2} \int_{(d,t]} [(1 - F_{n,d}^*(s-))/(1 - F(s|d))] \times \left[I_{\{R_{m(n)}(s) > 0\}} / R_{m(n)}(s) \right] dM_{m(n)}(s). \tag{4.18}$$

PROOF OF THEOREM 5(i). By Lemma 5, $\{U_{n,d}(t), \mathscr{F}(t), -\infty < t < \infty\}$ is a martingale with predictable variation process

$$\langle U_{n,d}\rangle(t) = n \int_{(d,t]} \left\{ \frac{1 - F_{n,d}^*(s-)}{1 - F(s|d)} \right\}^2 \frac{1 - \Delta\Lambda(s)}{R_{m(n)}(s)} I_{\{R_{m(n)}(s) > 0\}} \, d\Lambda(s). \tag{4.19}$$

By Lemma 4 and (4.1), noting that $\inf_{d \leq s \leq a}\{(1 - F(s-))G(s)\} > 0$,

$$\sup_{d \leq s \leq a} \left| m/R_m(s) - \{(1 - F(s-))G_m(s)\}^{-1} \right| \to 0 \quad \text{a.s.} \tag{4.20}$$

By Theorem 4 and (4.20),

$$\sup_{d < s \leq a} \left| (1 - F_{n,d}^*(s-))/(1 - F(s-|d)) - 1 \right| \to 0 \quad \text{a.s.} \tag{4.21}$$

In view of (4.2), we obtain by the argument used in the proof of (3.11) that

$$\lim_{n \to \infty} n/m(n) = p \quad \text{a.s.} \tag{4.22}$$

Since $1 - F(s|d) = (1 - F(s - |d))(1 - \Delta\Lambda(s))$ and $d\Lambda(s) = dF(s)/(1 - F(s-))$, it follows from (4.19)–(4.22) that with probability 1, as $n \to \infty$,

$$\langle U_{n,d}\rangle(t) \to p \int_{(d,t]} \{G(s)(1 - F(s-))(1 - F(s))\}^{-1} dF(s) \quad \text{uniformly in } t \in (d, a]. \tag{4.23}$$

Hence, in view of (4.17), application of Rebolledo's (1980) martingale central limit theorem gives the desired conclusion [cf. Gill (1980), Theorem 4.2.1]. □

LEMMA 6. *Define* $U_n(t)$ *by* (4.16). *Under the assumptions* (4.1), (4.2) *and* (4.6) *of Theorem* 5(ii), *for every* $\varepsilon > 0$,

$$\limsup_{n\to\infty} P\Big\{\sup_{t\le d}|U_n(t)| \ge \varepsilon\Big\} \to 0 \quad \text{as } d \downarrow \tau_G.$$

PROOF. Take $d \in (\tau_G, \tau^*)$. As in the proof of Theorem 3, choose $b > 0$ such that (2.15) holds and define Ω_0 by (2.16). Let $d_m = \tau_G \wedge G_m^{-1}(bm^{-(1-\alpha)})$, and let $A = \Omega_0 \cap \{2b(m(n))^\alpha < cn^\alpha$ for all large $n\}$. Then $P(A) = 1$ and (2.17) holds for all large m on A. By Lemma 5, $\{U_n(t), \mathscr{F}(t), t \le d\}$ is a martingale, and on A, for all large n,

$$\begin{aligned}\langle U_n\rangle(d) &\le n(1-F_n(d))^{-2}\int_{-\infty}^{d}\Big[I_{\{R_{m(n)}(s)\ge cn^\alpha\}}/R_{m(n)}(s)\Big]\,d\Lambda(s)\\ &\le 2n(1-F_n(d))^{-2}\int_{[d_{m(n)},d]}\Big[I_{\{G_{m(n)}(s)\ge b(m(n))^{-(1-\alpha)}\}}\\ &\qquad\times\{m(n)(1-F(s-))G_{m(n)}(s)\}^{-1}\Big]\,d\Lambda(s)\\ &\le 2(1-F(d))^{-4}\int_{[d_{m(n)},d]} dF/G_{m(n)},\end{aligned}\tag{4.24}$$

since $1 - F_n \ge 1 - F$. By Lenglart's (1977) inequality, for every $\varepsilon > 0$ and $\delta > 0$,

$$P\Big\{\sup_{t\le d} U_n^2(t) \ge \varepsilon^2\Big\} \le \delta + P\{\langle U_n\rangle(d) \ge \delta\varepsilon^2\}.\tag{4.25}$$

Letting $n \to \infty$ and $d \downarrow \tau_G$ in (4.24) and (4.25) and making use of (4.7), we obtain the desired conclusion. □

LEMMA 7. *Suppose that* F *is continuous on* $[a, \tau^*)$ *for some* $a \in (\tau_G, \tau^*)$. *Then under the assumptions* (4.1), (4.2) *and* (4.8) (*with* $0 \le \beta < 1$) *of Theorem* 5(iii), *for every* $\varepsilon > 0$,

$$\limsup_{n\to\infty} P\Big\{\sup_{x\le t<\tau^*}(1-F_n(t))|U_n(t)| \ge \varepsilon\Big\} \to 0 \quad \text{as } x \uparrow \tau^*.$$

PROOF. Since F and therefore F_n also are continuous on $[a, \tau^*)$, it suffices by Lemma 2.9 of Gill (1983) (see also the proof of his Theorem 2.1) to show that for every $\varepsilon > 0$, as $n \to \infty$ and $x \uparrow \tau^*$,

$$P\Big\{\sup_{x\le t<\tau^*}|V_{n,x}(t)| \ge \varepsilon\Big\} \to 0,\tag{4.26}$$

where

$$V_{n,x}(t) = \int_{[x,t]} (1 - F_n(s))\, dU_n(s)$$

$$= n^{1/2} \int_{[x,t]} \left[(1 - \hat{F}_n(s-)) I_{\{R_{m(n)}(s) \geq cn^\alpha\}} / R_{m(n)}(s) \right] dM_{m(n)}(s).$$

Since $1 - F_n \geq 1 - F$, $F_n(x) < 1$ for $x < \tau^*$. By (4.14), $\{(1 - \hat{F}_n(t))/(1 - F_n(t)), \mathscr{F}(t), t < \tau^*\}$ is a martingale with mean 1. Therefore, by Doob's inequality, for every $0 < \delta < 1$ and $n \geq 1$,

$$P(B_{\delta,n}) \geq 1 - \delta, \quad \text{where } B_{\delta,n} = \{1 - \hat{F}_n(s) \leq \delta^{-1}(1 - F_n(s)) \text{ for all } s < \tau^*\}. \tag{4.27}$$

Choose $b > 0$ such that (2.15) holds and define the event A as in the proof of Lemma 6. For $x \in (a, \tau^*)$, $\{V_{n,x}(t), \mathscr{F}(t), x \leq t < \tau^*\}$ is a martingale by Lemma 5, and for all large n, we have on $A \cap B_{\delta,n}$,

$$\langle V_{n,x}\rangle(t) \leq 2\delta^{-2} \int_{[x,t]} \frac{(1 - F_n(s-))^2}{(1 - F(s-)) G_{m(n)}(s)} \times I_{\{(1-F(s-))G_{m(n)}(s) \geq b(m(n))^{-(1-\alpha)}\}}\, d\Lambda(s). \tag{4.28}$$

If $1 - F(\tau^* -) = \theta > 0$, then the right-hand side of (4.28) is majorized by

$$\begin{aligned} &2\delta^{-2}\theta^{-2} \int_{[x,\tau^*)} \left[G_{m(n)}(s) \right]^{-1} dF(s) \\ &\leq 2\delta^{-2}\theta^{-2} \int_{[x,\tau^*)} \{F(\tau^* -) - F(s)\}^{-\beta}\, dF(s) \to 0 \quad \text{as } x \uparrow \tau^*, \end{aligned} \tag{4.29}$$

in view of (4.8) with $\beta < 1$. Now assume that $F(\tau^* -) = 1$. Define s_m by $1 - F(s_m) = \{2cm^{-(1-\alpha)}\}^{1/(1+\beta)}$. By (4.8), for all large m,

$$\inf_{a \leq s \leq s_m} (1 - F(s)) G_m(s) \geq (1 - F(s_m))^{1+\beta} = 2cm^{-(1-\alpha)}, \tag{4.30}$$

and therefore $s_m < \tau^*$, $s_m \to \tau^*$ as $m \to \infty$. On A, for all large n, $\inf_{a \leq s \leq s_{m(n)}} R_{m(n)}(s) \geq c(m(n))^\alpha$ by (4.30), and therefore by (4.13),

$$\begin{aligned} &(1 - F_n(t))/(1 - F_n(a)) \\ &= \exp\left(-\int_a^t d\Lambda \right) = (1 - F(t))/(1 - F(a)) \quad \text{for } a \leq t \leq s_{m(n)}, \end{aligned} \tag{4.31}$$

recalling that F and F_n are continuous on $[a, \tau^*)$. For all large n, it follows from (4.28) and (4.31) that on $A \cap B_{\delta,n}$,

$$(4.32)\quad \langle V_{n,x}\rangle(t) \le 2\delta^{-2}(1-F(a))^{-2}\int_x^t dF/G_{m(n)} \quad \text{for } a < x \le t \le s_{m(n)};$$

moreover, on $A \cap B_{\delta,n}$, we obtain by (4.28) and (4.31) that for $x \ge a$ and $\tau^* > t \ge x \wedge s_{m(n)}$,

$$\begin{aligned}
\langle V_{n,x}\rangle(t) &\le 2\delta^{-2}(1-F(a))^{-2} \\
&\quad \times \left\{\int_{[x, s_{m(n)}]} dF/G_{m(n)} + b^{-1}(m(n))^{1-\alpha}(1-F(s_{m(n)}))^2 \right. \\
(4.33) &\quad \left. \times \int_{s_{m(n)}}^t I_{\{1-F(s)\ge b(m(n))^{-(1-\alpha)}\}}\, d\Lambda(s)\right\} \\
&= 2\delta^{-2}(1-F(a))^{-2}\int_{[x, s_{m(n)}]} dF/G_{m(n)} \\
&\quad + O\big((m(n))^{-(1-\alpha)\{2/(1+\beta)-1\}}\log m(n)\big).
\end{aligned}$$

Recalling that $\beta < 1$ and that $P(A \cap B_{\delta,n}) \ge 1-\delta$, we can apply Lenglart's (1977) inequality as in (4.25) to obtain the desired conclusion from (4.29), (4.32) and (4.33). □

Proof of Theorem 5(ii). Let $a \in (\tau_G, \tau^*)$. In view of Lemma 6, it follows from (3.18) and Theorem 5(i) together with an argument similar to the proof of Theorem 2.1 of Gill (1983) that $(1-F_n)U_n$ $[= n^{1/2}(\hat{F}_n - F_n)$ by (4.15)] converges weakly in $D(-\infty, a]$ to $p^{1/2}(1-F_G)W$. Hence, it remains to show that

$$(4.34)\quad \sup_{t\le a} n^{1/2}|F_n(t) - F_G(t)| \to_P 0.$$

Take $K > 0$ such that $K(1-F(a)) > 2c$ and choose $b > 0$ such that (2.15) holds. Let $d_m = \tau_G \wedge G_m^{-1}(bm^{-(1-\alpha)})$, $D_m = \tau_G \vee \sup\{s \le (\tau^* + \tau_G)/2: G_m(s) < Km^{-(1-\alpha)}\}$, and define Ω_0 by (2.16). From (4.1), it follows that $\lim_{m\to\infty} D_m = \tau_G$ and that for all large m, $G_m(s) \ge Km^{-(1-\alpha)}$ if $D_m < s \le a$. Hence, for all large m,

$$(4.35)\quad (1-F(s))G_m(s) \ge K(1-F(a))m^{-(1-\alpha)} > 2cm^{-(1-\alpha)} \quad \text{for all } s \in (D_m, a].$$

Since $m(n) \ge n$, it then follows from (4.35) that on Ω_0, for all large m,

$$(4.36)\quad \inf_{D_{m(n)} < s \le a} I_{\{R_{m(n)}(s) \ge cn^\alpha\}} = 1.$$

By (4.5),

$$(4.37)\quad n^{1/2}\int_{[d_m, D_m]} \left| I_{\{R_{m(n)}(s)\ge cn^\alpha\}} - I_{\{s>\tau_G\}}\right| d\Lambda(s) \le n^{1/2}\int_{[d_m, D_m]} d\Lambda(s) \to 0$$

as $n \to \infty$. From (3.17), (4.36) and (4.37), it follows that

$$\sup_{t \le a} n^{1/2} |\Lambda_n(t) - \Lambda(t)| \to_P 0,$$

and therefore the desired conclusion (4.34) follows. □

PROOF OF THEOREM 5(iii). In view of Lemma 7, it follows from Theorem 5(ii) together with an argument similar to the proof of Theorem 2.1 of Gill (1983) that $(1 - F_n)U_n$ $[= n^{1/2}(\hat{F}_n - F_n)]$ converges weakly in $D(-\infty, \tau^*)$ to $p^{1/2}(1 - F_G)W$. Since we have already proved (4.34), it remains to show that

$$\sup_{a \le t < \tau^*} n^{1/2} |(1 - F_n(t))/(1 - F_n(a)) - (1 - F(t))/(1 - F(a))| \to_P 0, \quad (4.38)$$

where $a \in (\tau_G, \tau^*)$ is such that (4.8) holds and F and therefore F_n also are continuous on $[a, \tau^*)$. Let $\sigma_m < \tau^*$ be a solution of the equation

$$F(\tau^* -) - F(\sigma_m) = \{2cm^{-(1-\alpha)}\}^{1/(1+\beta)}. \quad (4.39)$$

Note that in the case $F(\tau^* -) = 1$, σ_m is the same as the s_m introduced in the proof of Lemma 7. By (4.8), for all large m,

$$\inf_{a \le s \le \sigma_m} (1 - F(s))G_m(s) \ge (F(\tau^* -) - F(s))^{1+\beta} = 2cm^{-(1-\alpha)}. \quad (4.40)$$

Defining A as in the proof of Lemma 7, we obtain from (4.40) that on A,

$$\inf_{a \le s \le \sigma_m} R_{m(n)}(s) \ge c(m(n))^{\alpha} \quad \text{for all large } n,$$

and therefore as in (4.31), for all large n,

$$(1 - F_n(t))/(1 - F_n(a)) = (1 - F(t))/(1 - F(a)) \quad \text{for } a \le t \le \sigma_{m(n)}, \text{ on } A. \quad (4.41)$$

Moreover, since $\sigma_m > a$ for all large m and since F and F_m are continuous on $[a, \tau^*)$,

$$\begin{aligned} &\{F_n(\tau^* -) - F_n(\sigma_{m(n)})\}/\{1 - F_n(\sigma_{m(n)})\} \\ &\quad = 1 - \exp\left\{-\int_{(\sigma_{m(n)}, \tau^*)} I_{\{R_{m(n)}(s) \ge cn^{\alpha}\}}\, d\Lambda(s)\right\} \\ &\quad \le 1 - \exp\left\{-\int_{(\sigma_{m(n)}, \tau^*)} d\Lambda\right\} \\ &\quad = \{F(\tau^* -) - F(\sigma_{m(n)})\}/\{1 - F(\sigma_{m(n)})\}. \end{aligned} \quad (4.42)$$

Since $\beta < 1 - 2\alpha$, (4.39) implies that

$$F(\tau^* -) - F(\sigma_m) = O(m^{-(1-\alpha)/(1+\beta)}) = o(m^{-1/2}). \quad (4.43)$$

From (4.41)–(4.43), the desired conclusion (4.38) follows, recalling that $F(a) > 0$ and that $F_n(a) - F_G(a) = o_p(n^{-1/2})$ by (4.34). □

5. Some numerical results and discussion. We first give some numerical results when the data are subject only to right censorship (i.e., $t_i \equiv -\infty$) by i.i.d. censoring variables T_i with a common distribution function $1 - G$ to illustrate the difficulty with the Kaplan–Meier estimator (1.2) in the case $F(\tau^* -) < 1$ and $G(\tau^* -) = 0$ and to show how the modified product-limit estimator $\hat{F}_n$ avoids this difficulty by using the weight function $I_{\{Y_n(s) \geq cn^\alpha\}}$ with $c = 1$ and $\alpha = 1/4$, where $Y_n(s) = \sum_{i=1}^n I_{\{X_i \wedge T_i \geq s\}}$ and $N_n(s) = \sum_{i=1}^n I_{\{X_i \leq s \wedge T_i\}}$. Thus, $1 - \hat{F}_n(t) = \prod_{s \leq t}\{1 - \Delta N_n(s) I_{\{Y_n(s) \geq n^{1/4}\}}/Y_n(s)\}$, while the Kaplan–Meier estimator $\tilde{F}_n$ is given by (1.2). Let $n = 100$ and let the X_i be i.i.d. exponential random variables with mean 5, and the T_i be i.i.d. with distribution function $1 - (1 - t/5)^3$, $0 \leq t \leq 5$. Here $\tau^* = 5$ and $F(\tau^*) = 1 - e^{-1}$. The values of $\hat{F}_n(t)$ and $\tilde{F}_n(t)$ based on one simulated sample from this censored model are tabulated below and are compared with the true values $F(t)$ for different values of t. Also given are the values of $Y_n(t)$.

t	$Y_n(t)$	$\tilde{F}_n(t)$	$\hat{F}_n(t)$	$F(t)$
0.245	86	0.042	0.042	0.048
0.475	70	0.090	0.090	0.091
0.792	49	0.118	0.118	0.146
1.381	30	0.186	0.186	0.241
1.898	20	0.310	0.310	0.316
2.963	3	0.425	0.425	0.447
2.975	2	0.712	0.425	0.448
3.090	1	1	0.425	0.461

The last two rows of the table show that while $\tilde{F}_n$ increases steeply when the risk set size $Y_n(t)$ falls below 3, the modified estimator $\hat{F}_n$ remains constant by ignoring the factors in the product-limit estimator corresponding to risk set sizes less than or equal to 3. The frequency distribution of EST(4) in 100 simulations from this censored model is summarized below, where EST = $\tilde{F}_n$ or $\hat{F}_n$. The true value is $F(4) = 0.55$. Also given is the mean squared error (MSE), which is the average of $\{\text{EST}(4) - F(4)\}^2$ over the 100 simulations.

	EST = 1	0.9 > EST ≥ 0.7	0.7 > EST ≥ 0.4	0.4 > EST ≥ 0.2	Total	MSE
$\tilde{F}_n(4)$	8	10	64	18	100	0.032
$\hat{F}_n(4)$	0	1	75	24	100	0.015

While the Kaplan–Meier estimator may have difficulties only near the upper endpoint τ^* in the censored model, its analog (1.6) in the truncated model may have difficulties throughout the entire range $t \geq \tau_G$, as the following numerical results show. Let $n = 50$ and let the X_i be i.i.d. with the logistic distribution function $F(x) = (1 + e^{-x})^{-1}$, $-\infty < x < \infty$, and the truncation variables t_i be i.i.d. with distribution function $G(x) = 0.2F(x) + 0.8\{1 - \exp(-x^+)\}$. Here $\tau_G = -\infty$ and $F_G = F$. The values of the modified product-limit estimator $\hat{F}_n$, defined by (1.12) with $c = 1$ and $\alpha = 1/4$, and of Lynden-Bell's estimator F_n^* defined by (1.6), based on one simulated sample from this truncated model, are tabulated below and are compared with the values of F. Also given are the values of $R_{m(n)}(t)$, which is the "risk set size" defined in (1.8).

t	$R_{m(n)}(t)$	$F_n^*(t)$	$\hat{F}_n(t)$	$F(t)$
−2.519	1	1	0	0.075
−0.551	5	1	0.2	0.37
0.450	25	1	0.39	0.61
0.629	22	1	0.53	0.65
1.274	26	1	0.65	0.78
1.837	13	1	0.84	0.86
3.369	6	1	0.95	0.97
5.069	3	1	0.98	0.99

Note in particular the nonmonotonic oscillations of $R_{m(n)}$. The frequency distribution of EST(0) in 100 simulations from this truncated model is summarized below, where EST $= F_n^*$ or $\hat{F}_n$. Also given is the mean squared error between the estimator and the true value $F(0) = 0.5$.

	EST = 1	0.9 > EST ≥ 0.7	0.7 > EST ≥ 0.3	0.3 > EST	Total	MSE
$F_n^*(0)$	6	9	67	18	100	0.049
$\hat{F}_n(0)$	0	6	72	22	100	0.037

The preceding examples illustrate the potential instability in the product-limit estimator (1.2) or (1.6) whenever the risk set is small. For right-censored data, this is not a serious problem since the small risk set size is restricted to the right tail of the observable range of the distribution. However, for left-truncated data, the instability occurs in the left tail but is then propagated throughout the entire observable range. The presence of both left-truncation and right-censoring variables compounds the difficulties since the function $P(t_i \le s \le T_i\}$ is nonmonotonic in s, and nonidentically distributed truncation and censoring variables further accentuate this problem.

We have shown how a minor modification of the product-limit estimator avoids all these difficulties. The idea is simply to discard those factors of the product in (1.2) or (1.6) that correspond to small risk set sizes. The preceding examples show that by ignoring risk set sizes less than or equal to 3 the modified product-limit estimator becomes much more stable. Moreover, discarding those factors that correspond to risk set sizes less than cn^α with $0 < \alpha < 1$ even leads to uniform strong consistency and weak convergence (by further requiring that $\alpha < 1/2$) results for the entire observable range of the distribution function, as we have established in Theorems 1–5.

REFERENCES

ALEXANDER, K. (1985). Rates of growth for weighted empirical processes. In *Proc. Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R. A. Olshen, eds.) **2** 475–493. Wadsworth, Monterey, Calif. [Correction in *Ann. Probab.* (1987) **15** 428–430.]

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.

DAVIDSEN, M. and JACOBSEN, M. (1989). Weak convergence of two-sided stochastic integrals, with an application to models for left truncated survival data. Technical Report, Inst. Math. Statist., Univ. Copenhagen.

GILL, R. D. (1980). *Censoring and Stochastic Integrals*. Math. Centre Tracts **124**. Mathematisch Centrum, Amsterdam.

GILL, R. D. (1981). Testing with replacement and the product-limit estimator. *Ann. Statist.* **9** 853–860.

GILL, R. D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.* **11** 49–58.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

KEIDING, N. and GILL, R. D. (1990). Random truncation models and Markov processes. *Ann. Statist.* **18** 582–602.

LAI, T. L. and WEI, C. Z. (1982). A law of the iterated logarithm for double arrays of independent random variables with applications to regression and time series models. *Ann. Probab.* **10** 320–335.

LAI, T. L. and YING, Z. (1988). Linear rank statistics in regression analysis with censored or truncated data. Technical Report, Dept. Statist., Stanford Univ.

LENGLART, E. (1977). Relation de domination entre deux processus. *Ann. Inst. H. Poincaré Sect. B (N.S.)* **13** 171–179.

LYNDEN-BELL, D. (1971). A method of allowing for known observational selection in small samples applied to 3CR quasars. *Monthly Notices Roy. Astronomy Soc.* **155** 95–118.

REBOLLEDO, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. Verw. Gebiete* **51** 169–286.

WANG, J. G. (1987). A note on the uniform consistency of the Kaplan–Meier estimator. *Ann. Statist.* **15** 1313–1316.

WANG, M. C., JEWELL, N. P. and TSAI, W. Y. (1986). Asymptotic properties of the product limit estimate under random truncation. *Ann. Statist.* **14** 1597–1605.

WOODROOFE, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13** 163–177.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
CHAMPAIGN, ILLINOIS 61820