Chin. Ann. of Math.
19B: 3(1998),341-348.

THE STRONG LAW FOR THE P-L ESTIMATE IN THE LEFT TRUNCATED AND RIGHT CENSORED MODEL (I)**

HE SHUYUAN*

Abstract

For the model with both left truncation and right censoring, suppose all the distributions are continuous. It is proved that the sampled cumulative hazard function Λ_n and the product-limit estimate F_n are strong consistent. For any nonnegative measurable ϕ, the almost sure convergences of $\int \phi \, d\Lambda_n$ and $\int \phi \, dF_n$ to the true values $\int \phi \, d\Lambda$ and $\int \phi \, dF$ respectively are obtained. The strong consistency of the estimator for the truncation probability is proved.

Keywords Left truncation and right censoring, Product-limit estimate, Strong law of large numbers, Reversed supermartingale

1991 MR Subject Classification 62G05, 60F15

Chinese Library Classification O211.4

§1. Introduction

Suppose that $\{X_n\}$, $\{Y_n\}$ and $\{T_n\}$ are three i.i.d. random sequences and independent one another. Let F, G and L be their right continuous distribution functions respectively. In many applications, only the right censored data of the form $(X_i \wedge Y_i, \delta_i)$ with $\delta_i = I[X_i \leq Y_i]$ are available. Here and in what follows we use $I[A]$ for the indicator of an event A and use $\wedge$, $\vee$ for the minimum and maximum respectively. Write $Z_i = X_i \wedge Y_i$. The K-M (cf. [5]) estimator of F, based on (Z_i, δ_i), $1 = 1, 2, \cdots, n$, is defined by

$$F_n^*(s) = 1 - \prod_{t \leq s} \left(1 - \frac{\#\{i : Z_i = t, \delta_i = 1\}}{\#\{i : Z_i \geq t\}} \right), \tag{1.1}$$

where an empty product is interpreted as one. It is clear that we can not estimate $F(x)$ for $x > b_G$ from the right censored data. Here and in what follows for any distribution function S, $a_S \equiv \inf\{y : S(y) > 0\}$ and $b_S \equiv \sup\{y : S(y) < 1\}$.

Another model of incomplete observation is the left truncated model, which assumes the presence of truncation variable T_i, so that (X_i, T_i) can be observed only when $X_i \geq T_i$. In this case, the data consist of n i.i.d. observation (X_i^0, T_i^0), $i = 1, 2, \cdots, n$ and the nonparametric MLE is defined by

$$F_n^0(x) = 1 - \prod_{s \leq x} \left(1 - \frac{\#\{i : X_i^0 = s\}}{\#\{i : T_i^0 \leq s \leq X_i^0\}} \right) \tag{1.2}$$

Manuscript received September 29, 1995. Revised November 21, 1996.

*Department of Probability and Statistics, Peking University, Beijing 100871, China.

**Project supported by the National Natural Science Foundation of China.

(see [10]). Based on (X_i^0, Y_i^0), $i = 1, 2, \cdots, n$, for continuous F we can estimate $F_0(x) \equiv P(X \le x | X \ge a_L)$ only (see [10]).

Mixed models with both left truncation and right censoring (LTRC) often arise in biostatistical applications such as epidemiology and individual follow-up study (cf. [9]). The LTRC model assumes that $(Z_i, \delta_i, T_i) \equiv (X_i \wedge Y_i, \delta_i, T_i)$ is observable only if $Z_i \ge T_i$. Thus, the observation consists of the i.i.d. data

$$(U_i, \eta_i, V_i), \text{ with } U_i \ge V_i, \ \eta_i = 0 \text{ or } 1, \ i = 1, 2, \cdots, n. \tag{1.3}$$

Based on (1.3) the product-limit (P-L) estimator of F is defined by

$$F_n(x) = 1 - \prod_{s \le x} \left(1 - \frac{\#\{i : U_i = s, \eta_i = 1\}}{\#\{i : V_i \le s \le U_i\}} \right). \tag{1.4}$$

Note that F_n reduces to the K-M estimator (1.1) for right censored data if $L \equiv 1$, and reduces to the product-limit estimator (1.2) for left truncated data if $G \equiv 0$. It is clear that in the LTRC model, for continuous F, only $F_0 = P(X \le x | X \ge a_L)$, $x \le b_G$ can be estimated.

For continuous F, G and L, the strong uniform consistency of F_n^* and F_n^0 are known (see e.g., [1, 8]). For the LTRC model and continuous F, let W denote the distribution function of Z. In the case of $a_W > a_L$, which insures the denominator in the expression of F_n bounded from zero (with probability 1) in a neighborhood of a_F and hence makes the study of F_n relatively easy, the uniform consistency was also proved (e.g., Theorem 2 of [3]). For the case of $a_W = a_L$, the strong uniform consistency of F_n kept unknown.

In this paper, for the LTRC model and continuous F, G and L we prove the strong uniform consistency of F_n (and the P-L estimates of G and L). In fact we will prove that for any nonnegative measurable $\phi(x)$,

$$\int \phi(x)\, dF_n(x) \to \int \phi(x)\, dF_0(x) \ \ a.s. \tag{1.5}$$

For the right ensored model, (1.5) was proved in [8] in 1993.

A natural estimate for the truncation probability $\alpha = P(Z \ge T)$ is $\alpha_n = \int L_n(s)\, dW_n(s)$, where L_n and W_n are P-L estimates of L and W. For left truncated model and continuous F and L, the weakly consistency of α_n to α was proved in [10]. In this paper, for the LTRC model we will provide a simple expression for α_n. As a corollary of the strong consistency of L_n and W_n, the strong consistency of α_n follows.

§2. Preliminaries

For the LTRC model we can image the i.i.d. data (1.3) coming from a large population (X_i, Y_i, T_i), $i = 1, 2, \cdots, m_n$ with $m_n = \inf\{m : \sum_{j=1}^{m} I[Z_j \ge T_j] = n\}$. For any right-continuous non-decreasing function S, if $\inf_s S(s) < \sup_s S(s)$, define

$$a_S = \inf\{y : S(y) > \inf_s S(s)\} \text{ and } b_S = \sup\{y : S(y) < \sup_s S(s)\},$$

$\bar{S}(x) = 1 - S(x)$, $S_-(x) = S(x-) = \lim_{s \uparrow x} S(s)$ and $S\{x\} = S(x) - S_-(x)$. To avoid the case of mathematical trivial, we assume $a_L < b_W$, where $W(x) = P(Z_i \le x) = 1 - \bar{F}(x)\bar{G}(x)$.

It is clear that $a_W = a_F \wedge a_G$ and $b_W = b_F \wedge b_G$. From the law of large numbers and the equality

$$n^{-1}\sum_{i=1}^{n} I[U_i \le u, \eta_i = \delta, V_i \le v] = \frac{m_n}{n}\frac{1}{m_n}\sum_{i=1}^{m_n} I[Z_i \le u, \delta_i = \delta, T_i \le v, T_i \le Z_i], \tag{2.1}$$

we get, for $u, v \in (-\infty, \infty)$ and $\delta = 0$ or 1,

$$P(U_i \le u, \eta_i = \delta, V_i \le v) = \alpha^{-1} P(Z_i \le u, \delta_i = \delta, T_i \le v, T_i \le Z_i), \tag{2.2}$$

where $\alpha = P(Z \ge T) = \int L\,dW > 0$. In describing the distributional properties we use (X, Y, T) to refer to any (X_i, Y_i, T_i), (Z, δ) to (Z_i, δ_i) and (U, η, V) to (U_i, η_i, V_i). The following equations can be obtained directly from (2.2):

(a1) $H_0(u) \equiv P(U \le u, \eta = 0) = \alpha^{-1} P(Y \le u, Y < X, Y \ge T) = \alpha^{-1}\int_{-\infty}^{u}(1-F)L\,dG$,

(a2) $H_1(u) \equiv P(U \le u, \eta = 1) = \alpha^{-1} P(X \le u, X \le Y, X \ge T) = \alpha^{-1}\int_{-\infty}^{u}(1-G_-)L\,dF$,

(a3) $H_2(u) \equiv H_0(u) + H_1(u) = P(U \le u) = \alpha^{-1}\int_{-\infty}^{u} L\,dW$,

(a4) $K(u) \equiv P(V \le u) = \alpha^{-1} P(T \le u, Z \ge T) = \alpha^{-1}\int_{-\infty}^{u}(1-W_-)\,dL$,

(a5) $R(u) \equiv P(V \le u \le U) = \alpha^{-1} P(T \le u \le Z) = \alpha^{-1} L(u)(1 - W_-(u))$.

Whenever it makes sense, define $G_0(u) = P(Y \le u | Y \ge a_{H_0})$ for $u \le b_{H_0}$ and $G_0(u) = G(b_{H_0})$ for $u > b_{H_0}$; $F_0(u) = P(X \le u | X \ge a_{H_1})$ for $u \le b_{H_1}$ and $F_0(u) = F(b_{H_1})$ for $u > b_{H_1}$; and $W_0(u) = P(Z \le u | Z \ge a_{H_2})$. It is obtained that

$$1 - G_0(u-) = (1 - G(u-))/P(Y \ge a_{H_0}) \text{ for } u \in [a_{H_0}, b_{H_0}],$$
$$1 - F_0(u-) = (1 - F(u-))/P(X \ge a_{H_1}) \text{ for } u \in [a_{H_1}, b_{H_1}] \text{ and}$$
$$1 - W_0(u-) = (1 - W(u-))/P(Z \ge a_{H_2}) \text{ for } u \in [a_{H_2}, \infty).$$

Write $\Lambda_j(u) = \int_{-\infty}^{u} dH_j/R$. It is obtained that for continuous F,

$$d\Lambda_0(u) = dG(u)/(1 - G_-(u)) = dG_0(u)/(1 - G_0(u-)), \text{ for } u \in [a_{H_0}, b_{H_0}],$$
$$d\Lambda_1(u) = dF(u)/(1 - F_-(u)) = dF_0(u)/(1 - F_0(u-)), \text{ for } u \in [a_{H_1}, b_{H_1}],$$
$$d\Lambda_2(u) = dW(u)/(1 - W_-(u)) = dW_0(u)/(1 - W_0(u-)), \text{ for } u \in [a_{H_2}, b_{H_2}].$$

So, Λ_0, Λ_1, Λ_2 are the cumulative hazard functions of G_0, F_0 and W_0 respectively.

Lemma 2.1. *Let F, G and L be continuous.*

(i) *If $a_G < b_F$, then* $G_0(u) = \begin{cases} P(Y \le u | Y \ge a_L), & \text{for } u \le b_W, \\ P(Y \le b_W | Y \ge a_L), & \text{for } u > b_W. \end{cases}$

(ii) *If $a_F < b_G$, then* $F_0(u) = \begin{cases} P(X \le u | X \ge a_L), & \text{for } u \le b_W, \\ P(X \le b_W | X \ge a_L), & \text{for } u > b_W. \end{cases}$

(iii) $W_0(u) = P(Z \le u | Z \ge a_L)$.

Proof. From (a1) it is seen that $a_{H_0} \ge a_G \vee a_L$ and $b_{H_0} \le b_F \wedge b_G = b_W$.

If $a_{H_0} = a_G > a_L$, then $P(a_L \le Y \le a_{H_0}) = 0$. If $a_{H_0} > a_G \vee a_L$, then $a_L > a_G$ (note that $a_L \le a_G$ implies $a_{H_0} = a_G$). Hence, $H_0(a_{H_0}-) = 0$, $H_0(a_{H_0} + \varepsilon) > 0\,(\forall\,\varepsilon > 0)$ and (a1) imply $P(a_L \le Y \le a_{H_0}) = 0$; that is, $P(Y \le u | Y \ge a_L) = P(Y \le u | Y \ge a_{H_0})$, $\forall u$.

If $b_{H_0} < b_W$, then $b_F < b_G$ (note that $b_F \ge b_G$ implies $b_{H_0} = b_W$). Hence, (a1) implies $P(b_{H_0} \le Y \le b_W) = 0$. It follows that for $u \in (b_{H_0}, b_W]$, we have $P(Y \le u | Y \ge a_{H_0}) = P(Y \le b_{H_0} | Y \ge a_{H_0})$.

(ii) and (iii) can be proved similarly.

According to [10], for the model of left truncation, only W_0 and $L_0(u) \equiv P(T \le u|T \le b_W)$ can be estimated. Hence, for the LTRC model only G_0, F_0, W_0 and L_0 can be estimated. Fortunately, we have the following

Corollary 2.2. *Let F, G and L be continuous.*

(i) $a_G \ge a_L$, $b_G \le b_F$ *imply* $G_0 = G$. (ii) $a_F \ge a_L$, $b_F \le b_G$ *imply* $F_0 = F$.

(iii) $a_W \ge a_L$ *implies* $W_0 = W$. (iv) $b_L \le b_W$ *implies* $L_0 = L$.

§3. The Strong Law for F_n, G_n and L_n

It is known that a cumulative hazard function $\Lambda_S(x)$ of S determines the distribution $S(x)$ through the algorithm (see [4])

$$S(x) = 1 - \prod_{s \le x} (1 - \Lambda_S\{s\}) \exp\left(\sum_{s \le x} \Lambda_S\{s\} - \Lambda_S(x)\right). \tag{3.1}$$

So, for continuous F_0 and W_0 we have $F_0(x) = 1 - \exp(-\Lambda_1(x)), W_0(x) = 1 - \exp(-\Lambda_2(x))$.

Now, based on (1.3), let us introduce the following sub-empiricaled functions.

$$H_{0,n}(u) = n^{-1} \sum_{i=1}^{n} I[U_i \le u, \eta_i = 0], \quad H_{1,n}(u) = n^{-1} \sum_{i=1}^{n} I[U_i \le u, \eta_i = 1],$$

$$H_{2,n}(u) = H_{0,n}(u) + H_{1,n}(u), \quad K_n(u) = n^{-1} \sum_{i=1}^{n} I[V_i \le u],$$

$$R_n(u) = n^{-1} \sum_{i=1}^{n} I[V_i \le u \le U_i] = K_n(u) - H_{2,n}(u-);$$

and the following empiricaled cumulative hazard functions:

$$\Lambda_{j,n}(u) = \int_{-\infty}^{u} \frac{dH_{j,n}}{R_n}, \quad j = 0, 1, 2. \tag{3.2}$$

Here and in what follows we use $\int_a^b$ for $\int_{(a,b]}$.

According to (3.1), the estimators of F_0 and W_0 should be

$$F_n(x) = 1 - \prod_{s \le x} \left(1 - \frac{H_{1,n}\{s\}}{R_n(s)}\right) \text{ and } W_n(x) = 1 - \prod_{s \le x} \left(1 - \frac{H_{2,n}\{s\}}{R_n(s)}\right) \tag{3.3}$$

respectively, which coincide the estimators defined by (1.4) and (1.2) and ensure that the corresponding cumulative hazard function to $\Lambda_{j,n}$, $j = 1, 2$. By symmetry, the estimates of G_0 and L_0 are defined by

$$G_n(x) = 1 - \prod_{s \le x} \left(1 - \frac{H_{0,n}\{s\}}{R_n(s)}\right) \text{ and } L_n(x) = \prod_{s > x} \left(1 - \frac{K_n\{s\}}{R_n(s)}\right) \tag{3.4}$$

respectively.

In what follows, if without further statement, F, G and L are supposed to be continuous.

Let $U_{1:n} \le U_{2:n}, \cdots, \le U_{n:n}$ be the ordered values of the U_i's and $(\eta_{1:n}, V_{1:n}), (\eta_{2:n}, V_{2:n}), \cdots, (\eta_{n:n}, V_{n:n})$ be the concomitants corresponding to the U_i's (that is, $(\eta_{j:n}, V_{j:n}) = (\eta_i, V_i)$ if and only if $U_{j:n} = U_i$). Define $\mathcal{F}_n = \sigma\{(U_{j:n}, \eta_{j:n}, V_{j:n}), (U_k, \eta_k, V_k); 1 \le j \le n, k > n\}$. Then $\forall n \ge 1$, $\mathcal{F}_{n+1} \subset \mathcal{F}_n$ and $\Lambda_{1,n}$, $\Lambda_{2,n} \in \mathcal{F}_n$. It is clear that for any $n \ge 1$, $k =$

$1, 2, \cdots, n+1$, $P(U_{n+1} = U_{k:n+1}|\mathcal{F}_{n+1}) = 1/(n+1)$ and $U_{n+1} = U_{k:n+1}$ implies

$$\begin{cases} U_{j:n+1} &= U_{j:n}, \\ \eta_{j:n+1} &= \eta_{j:n}, \\ V_{j:n+1} &= V_{j:n}, \end{cases} \text{for } j \le k-1 \text{ and } \begin{cases} U_{j:n+1} &= U_{j-1:n}, \\ \eta_{j:n+1} &= \eta_{j-1:n}, \\ V_{j:n+1} &= V_{j-1:n}, \end{cases} \text{for } j \ge k+1. \tag{3.5}$$

For any (measurable) $\phi(x) \ge 0$, define $W_{i,n} = \int \phi \, d\Lambda_{i,n}$, $i = 1, 2$. We have the following result.

Lemma 3.1. *Let $\phi(x) \ge 0$ and $i = 1$ or 2. Suppose $a_F < b_G$ for $i = 1$. Then $(W_{i,n}, \mathcal{F}_n; n \ge 1)$ is a nonnegative reverse supermartingale; that is, $\forall n \ge 1$, $W_{i,n} \in \mathcal{F}_n$ and $E(W_{i,n}|\mathcal{F}_{n+1}) \le W_{i,n+1}$.*

Proof. $\Lambda_{1,n}(x)$ is a step function and has jumps only at each $U_{j:n}$, $j = 1, \cdots, n$. So, $W_{1,n} = \sum\limits_{j=1}^{n} \phi(U_{j:n})\Lambda_{1,n}\{U_{j:n}\} \in \mathcal{F}_n$ and

$$W_{1,n} = \sum_{j=1}^{n+1} \phi(U_{j:n+1})\Lambda_{1,n}\{U_{j:n+1}\}, \text{ with } \phi(U_{j,n+1}) \in \mathcal{F}_{n+1}. \tag{3.6}$$

Therefore, it suffices to show

$$E(\Lambda_{1,n}\{U_{j:n+1}\}|\mathcal{F}_{n+1}) \le \Lambda_{1,n+1}\{U_{j:n+1}\} \text{ a.s.} \tag{3.7}$$

Write $A_k = I[U_{n+1} = U_{k:n+1}]$. It follows that $\Lambda_{1,n}\{U_{k:n+1}\}A_k = 0$ a.s. and

$$\begin{aligned} &E(\Lambda_{1,n}\{U_{j:n+1}\}|\mathcal{F}_{n+1}) \\ &= \frac{1}{n+1}\sum_{k \ne j}\left(\frac{\eta_{j:n+1}}{(n+1)R_{n+1}(U_{j:n+1}) - I[V_{k:n+1} \le U_{j:n+1} \le U_{k:n+1}]}\right) \text{ a.s.} \end{aligned}$$

Let $B_m = \{(n+1)R_{n+1}(U_{j:n+1}) = m\}$, $m = 1, 2, \cdots, n+1$. For $m > 1$, on B_m we have $\sum\limits_{k \ne j} I[V_{k:n+1} \le U_{j:n+1} \le U_{k:n+1}] = m - 1$ and it follows that

$$E(\Lambda_{1,n}\{U_{j:n+1}\}|\mathcal{F}_{n+1}) = \frac{\eta_{j:n+1}}{m} = \Lambda_{1,n+1}\{U_{j:n+1}\} \text{ a.s.}$$

On B_1, $\sum\limits_{k \ne j} I[V_{k:n+1} \le U_{j:n+1} \le U_{k:n+1}] = 0$, which implies

$$E(\Lambda_{1,n}\{U_{j:n+1}\}|\mathcal{F}_{n+1}) = \frac{n}{n+1}\eta_{j:n+1} = \Lambda_{n+1}\{U_{j:n+1}\} - \frac{\eta_{j:n+1}}{n+1}.$$

This proves (3.7) and hence the lemma for $i = 1$. The proof for $i = 2$ is similar.

By Hewitt-Savage Zero-One Law (see [2]), $\mathcal{F}_\infty = \bigcap\limits_{n \ge 1} \mathcal{F}_n$ is trivial. Hence, the following is a direct corollary of Proposition 5-3-11 of [6].

Corollary 3.1. *Let $i = 1$ or 2 and $\phi(x) \ge 0$. Suppose $a_F < b_G$ for $i = 1$.*

(i) $\xi_i = \lim\limits_{n\to\infty} EW_{i,n}$ *exists (possibly infinite).* (ii) $W_{i,n} \to \xi_i$ *a.s.*

(iii) *If $\xi_i < \infty$, then $\{W_{i,n}\}$ is uniformly integrable and $E|W_{i,n} - \xi_i| \to 0$.*

Theorem 3.1. *Let $i = 1$ or 2 and $\phi(x) \ge 0$. Suppose $a_F < b_G$ for $i = 1$. We have*

$$\lim_{n\to\infty} W_{i,n} = \int \phi(x) d\Lambda_i(x) \text{ a.s.}$$

Proof. We only give the proof for $i = 1$. The proof for $i = 2$ is similar. Let $\zeta_n(u) = (n+1)/(nR_n(u)+1) \in \mathcal{F}_n$. We have $EW_{1,n} = \int \phi E\zeta_{n-1}(u)\, dH_1$. So, by Corollary 3.1, it suffices to prove $E\zeta_n(u) \nearrow R(u)^{-1}$. By the same method used in the proof of Lemma 3.1, we have $E(\zeta_n(u)|\mathcal{F}_{n+1}) \le \zeta_{n+1}(u)$. Hence, $E\zeta_n(u) \nearrow R(u)^{-1}$ follows from $R_n(u) \to R(u)$ a.s.

For $\phi(x) = I[-\infty, x]$ it is obtained that $\forall\, x$

$$\Lambda_{i,n}(x) \to \Lambda_i(x) \text{ a.s. as } n \to \infty. \tag{3.8}$$

The fact that $\Lambda_i(x)$ is continuous and $\Lambda_{i,n}$ is monotone implies that for any $x < b_{H_i}$

$$\sup_{s \le x} |\Lambda_{i,n}(s) - \Lambda_i(s)| \to 0 \quad \text{a.s. as } \ n \to \infty.$$

Now, using (3.1) we get the following theorem.

Theorem 3.2. *Define* $\widetilde{F}_n(x) = 1 - \exp(-\Lambda_{1,n}(x))$ *and* $\widetilde{W}_n(x) = 1 - \exp(-\Lambda_{2,n}(x))$.

(i) *If* $a_F < b_G$, *then* $\sup_x |\widetilde{F}_n(x) - F_0(x)| \to 0$ *a.s. as* $n \to \infty$.

(ii) $\sup_x |\widetilde{W}_n(x) - W_0(x)| \to 0$ *a.s. as* $n \to \infty$.

Theorem 3.3. *With probability* 1, *as* $n \to \infty$

(i) $a_F < b_G$ *implies* $\sup_x |F_n(x) - F_0(x)| \to 0$, (ii) $a_G < b_F$ *implies* $\sup_x |G_n(x) - G_0(x)| \to 0$,

(iii) $\sup_x |W_n(x) - W_0(x)| \to 0$, $\sup_x |L_n(x) - L_0(x)| \to 0$.

Proof. The proof is only for F_n. The proof for W_n can be completed by considering the case $\eta_i \equiv 1$. The proof for G_n and L_n can be finished by symmetry (consider $\widetilde{X}_i = Y_i$, $\widetilde{Y}_i = X_i$ and $\widetilde{Z}_i = -T_i$, $\widetilde{T}_i = -Z_i$ respectively).

Set $\varphi_n = (\log n/n)^{1/2}$ and $d_n = \inf\{x; L(x) \ge \varphi_n\}$. We get $L(d_n) = \varphi_n \to 0$ and $d_n \to a_L$, as $n \to \infty$. Note $a_L \le a_{H_1}$, we have $\Lambda_1(d_n) \to 0$, as $n \to \infty$. $\forall a \in (a_{H_1}, b_{H_1})$, using (3.1) and $\prod a_j - \prod b_j \le \sum |a_j - b_j|$ for $a_j, b_j \in [0,1]$, we obtain

$$F_0(d_n) = 1 - \exp(-\Lambda_1(d_n)) \to 0, \ \text{ and } \ F_n(d_n) \le \Lambda_{1,n}(d_n) \to 0.$$

The product form of the estimate $\bar{F}_n(x) = 1 - F_n(x)$ can be spliced as follows. Let

$$M_n(x) = \prod_{d_n < U_i \le x} \left(1 - \frac{H_n\{U_i\}}{R_n(U_i)}\right), \ \widetilde{M}_n(x) = \exp\left(-\Lambda_1(x) + \Lambda_1(d_n)\right).$$

Then $\bar{F}_n(x) = (1 - F_n(d_n))M_n(x)$, $\bar{F}_0(x) = (1 - F_0(d_n))\widetilde{M}_n(x)$.

Hence, $|F_n(x) - F_0(x)| \le F_n(d_n) + F_0(d_n) + |M_n(x) - \widetilde{M}_n(x)|$. It suffices to prove

$$\sup_x |M_n(x) - \widetilde{M}_n(x)| \to 0 \text{ a.s.} \tag{3.9}$$

Note, $\forall a \in (a_{H_1}, b_{H_1})$ by (a5), $\inf_{d_n < x \le a} R(x) \ge \alpha^{-1}\bar{F}(a)\bar{G}(a)L(d_n)$. Applying Theorem 2.1.4B of [7], we have

$$\limsup_{n\to\infty} \sup_{d_n < x \le a} \left|\frac{R_n(x)}{R(x)} - 1\right| \le \limsup_{n\to\infty} \frac{\alpha}{\bar{F}(a)\bar{G}(a)\varphi_n}|R_n(x) - R(x)| = 0 \text{ a.s.}$$

So, $\forall x \in (a_{H_1}, b_{H_1})$, with probability 1 for large n,

$$\inf_{d_n < s \le x}(nR_n(s)) = n \inf_{d_n < s \le x}\left(\frac{R_n(s)}{R(s)} \cdot R(s)\right) \ge \frac{n}{2} \inf_{d_n < s \le x} R(s) \ge M_0\sqrt{n \log n},$$

where M_0 is a positive constant. Hence, by Taylor's formula

$$M_n(x) = \exp\left(-\Lambda_{1,n}(x) + \Lambda_1(d_n) + O(\frac{1}{\log n})\right) \to 1 - F_0(x). \ \text{ a.s.} \tag{3.10}$$

Since $M_n(x)$ is monotone in x and F_0 is continuous, the convergence in (3.10) must be uniform in x. Now (3.9) is proved by combining with the fact that

$$\sup_x |\widetilde{M}_n(x) - \bar{F}_0(x)| = \sup_x |\exp\left(-\Lambda_1(x) + \Lambda_1(d_n)\right) - \exp\left(-\Lambda_1(x)\right)| \to 0 \text{ a.s.}$$

§4. The Strong Law for $\int \phi\, dF_n$, $\int \phi\, dG_n$ and $\int \phi\, dL_n$

Lemma 4.1. *Suppose* $\phi(x) \geq 0$. *Then*

(i) $a_F < b_G$ *and* $\int \phi\, d\Lambda_1 < \infty$ *imply* $\int \phi\, dF_n \to \int \phi\, dF_0$ *a.s.*

(ii) $\int \phi\, d\Lambda_2 < \infty$ *implies* $\int \phi\, dW_n \to \int \phi\, dW_0$ *a.s.*

Proof. We give the proof for (i). By the definition of F_n, we have

$$F_n\{x\} = \bar{F}_n(x-)H_{1,n}\{x\}/R_n(x) = \bar{F}_n(x-)\Lambda_{1,n}\{x\}. \tag{4.1}$$

Now the result follows from Theorems 3.1 and 3.3.

To get a more general result, we need the following lemma.

Lemma 4.2. *For maybe non-continuous F, G and L,*

$$L_n(U_j)W_n\{U_j\} = \alpha_n H_{2,n}\{U_j\},\ j = 1, 2, \cdots, n, \tag{4.2}$$

$$\alpha_n = \int L_n\, dW_n = L_n(U_j)(1 - W_n(U_j-))/R_n(U_j),\ j = 1, 2, \cdots, n. \tag{4.3}$$

Proof. Let $U_{(1)} < U_{(2)} < \cdots < U_{(p)}$ be the distinct ordered values of $U_1, U_2, \cdots, U_n$ and $V_{(1)} < V_{(2)} < \cdots < V_{(q)}$ the distinct ordered values of $V_1, V_2, \cdots, V_n$. Then, for each fixed j

$$W_n\{U_{(j)}\} = \prod_{i:\, i<j}\Big(1 - \frac{H_{2,n}\{U_{(i)}\}}{R_n(U_{(i)})}\Big)\frac{H_{2,n}\{U_{(j)}\}}{R_n(U_{(j)})}, \quad L_n(U_{(j)}) = \prod_{i:\, V_{(i)}>U_{(j)}}\Big(1 - \frac{K_n\{V_{(i)}\}}{R_n(V_{(i)})}\Big).$$

Define

$$\xi_n(j) = \prod_{i:\, i<j}\Big(1 - \frac{H_{2,n}\{U_{(i)}\}}{R_n(U_{(i)})}\Big)\prod_{i:\, V_{(i)}>U_{(j)}}\Big(1 - \frac{K_n\{V_{(i)}\}}{R_n(V_{(i)})}\Big)\frac{1}{R_n(U_{(j)})}. \tag{4.4}$$

We have, for $j \geq 2$ $L_n(U_{(j)})W_n\{U_{(j)}\} = \xi_n(j)H_{2,n}\{U_{(j)}\}$,

$$\xi_n(j) - \xi_n(j-1) = \Big\{\prod_{i:\, i<j-1}\Big(1 - \frac{H_{2,n}\{U_{(i)}\}}{R_n(U_{(i)})}\Big)\cdot \prod_{i:\, V_{(i)}>U_{(j)}}\Big(1 - \frac{K_n\{V_{(i)}\}}{R_n(V_{(i)})}\Big)\Big\}\cdot B_n(j),$$

$$B_n(j) = \frac{R_n(U_{(j-1)}) - H_{2,n}\{U_{(j-1)}\}}{R_n(U_{(j-1)})R_n(U_{(j)})} - \frac{1}{R_n(U_{(j-1)})}\prod_l\Big(1 - \frac{K_n\{V_{(l)}\}I[U_{(j-1)} < V_{(l)} \leq U_{(j)}]}{R_n(V_{(l)})}\Big).$$

Put $h = \sum_l I[U_{(j-1)} < V_{(l)} \leq U_{(j)}]$. For $h = 0$, $R_n(U_{(j-1)}) - H_{2,n}\{U_{(j-1)}\} = K_n(U_{(j-1)}) - H_{2,n}(U_{(j)}-) = R_n(U_{(j)})$ implies $B_n(j) = 0$.

For $h > 0$, we can suppose $V'_{(1)}, V'_{(2)}, \cdots, V'_{(h)}$ are the distinct ordered values of $\{V_j; V_j \in (U_{(j-1)}, U_{(j)}]\}$, so that $U_{(j-1)} < V'_{(1)} < V'_{(2)} < \cdots < V'_{(h)} \leq U_{(j)}$. Consequently,

$$\prod_l\Big(1 - \frac{K_n\{V_{(l)}\}I[U_{(j-1)} < V_{(l)} \leq U_{(j)}]}{R_n(V_{(l)})}\Big) = \prod_{l=1}^h\Big(1 - \frac{K_n\{V'_{(l)}\}}{R_n(V'_{(l)})}\Big)$$

$$= \prod_{l=1}^h \frac{K_n(V'_{(l)}-) - H_{2,n}(V'_{(l)}-)}{R_n(V'_{(l)})} = \frac{K_n(U_{(j-1)}) - H_{2,n}(U_{(j-1)})}{R_n(U_{(j)})} = \frac{R_n(U_{(j-1)}) - H_{2,n}\{U_{(j-1)}\}}{R_n(U_{(j)})}.$$

Here, for $V'_{(h)} < U_{(j)}$ we used $R_n(V'_{(h)}) = K_n(V'_{(h)}) - H_{2,n}(V'_{(h)}-) = R_n(U_{(j)})$.

So, $B_n(j) = 0$, for $j = 2, 3, \cdots, n$, which implies (4.2). The summation over j of (4.2) and (4.4) gives (4.3).

Theorem 4.1. $\alpha_n \to \alpha_0 = \int L_0\, dW_0$ *a.s. as* $n \to \infty$.

The proof is straightforward and omitted.

If $a_W \geq a_L,\ b_W \geq b_L$, we have $W_0 = W,\ L_0 = L$. Hence α_n is a strong consistent estimate for the left truncation probability $\alpha = P(W \geq T)$ and the number of population m_n can be estimated by $\hat{m}_n = n/\alpha_n$. Otherwise, $\hat{m}_n$ will less estimate m_n, since $\alpha_0 = \alpha/(L(b_W)\bar{W}(a_L-)) > \alpha$.

Theorem 4.2. *For any nonnegative* $\phi(x)$, *with probability* 1, *as* $n \to \infty$

(i) $b_F < b_G$ *implies* $\int \phi\, dF_n \to \int \phi\, dF_0$, (ii) $\int \phi\, dW_n \to \int \phi\, dW_0$,

(iii) $b_G < b_F$ *implies* $\int \phi\, dG_n \to \int \phi\, dG_0$, (iv) $\int \phi\, dL_n \to \int \phi\, dL_0$.

Proof. We give the proof for (i). (ii) can be obtained by considering $\eta_i \equiv 1$ and (iii), (iv) can be obtained by symmetry. For any $a \in (a_{H_1}, b_{H_1})$ and $M > 0$, Lemma 4.1 implies $\liminf\limits_{n\to\infty} \int \phi\, dF_n \geq \int_{-\infty}^{a} \phi I[\phi \leq M]\, dF_0$. Let $a \uparrow b_{H_1}$ and $M \uparrow \infty$. We get $\liminf\limits_{n\to\infty} \int \phi\, dF_n \geq \int \phi\, dF_0$. So, it suffices to prove $\limsup\limits_{n\to\infty} \int \phi\, dF_n \leq \int \phi\, dF_0$ a.s., under the condition $\int \phi\, dF_0 < \infty$. In this case, $\forall a \in (a_{H_1}, b_{H_1})$, $\int_{-\infty}^{a} \phi\, d\Lambda_1 = \int_{a_{H_1}}^{a} \phi \frac{dF_0}{1-F_0} < \infty$. Hence, by Lemma 4.1

$$\limsup_{n\to\infty} \int \phi\, dF_n \leq \int_{-\infty}^{a} \phi\, dF_0 + \limsup_{n\to\infty} \int_{a}^{\infty} \phi\, dF_n.$$

Therefore, we only need to show $\limsup\limits_{n\to\infty} \int_{a}^{\infty} \phi\, dF_n \to 0$ a.s. as $a \uparrow b_{H_1}$.

In fact, (4.1), Lemma 4.2 and (a2) imply that as $n \to \infty$,

$$\int_{a}^{\infty} \phi(x) dF_n(x) = \alpha_n \int_{a}^{\infty} \frac{\phi(x)\, dH_{1,n}(x)}{L_n(x)\bar{G}_n(x-)} \leq \frac{\alpha_n}{L_n(a)\bar{G}_n(b_{H_1})} \int_{a}^{\infty} \phi(x) dH_{1,n}(x)$$

$$\to \frac{\alpha_0}{L_0(a)\bar{G}_0(b_{H_1})} \int_{a}^{\infty} \phi(x) dH_1(x) \leq \frac{\alpha_0}{\alpha L_0(a)\bar{G}_0(b_F)} \int_{a}^{\infty} \phi(x) dF_0(x) \to 0 \text{ a.s. as } a \uparrow b_{H_1}.$$

Remark 4.1. The condition $a_F < b_G$ (or $a_G < b_F$) requires nothing but we can get information from F (or G). It seems that $b_F < b_G$ ($b_G < b_F$) in (i) ((iii)) requires something more than necessity.

Corollary 4.1. *For any nonnegative* $\phi(x)$, *with probability* 1, *as* $n \to \infty$

(i) *for the right censored model if* $b_F < b_G$, $\int \phi\, dF_n^* \to \int_{-\infty}^{b_W} \phi\, dF$ *a.s.*,

(ii) *for the left truncated model* $\int \phi\, dF_n^0 \to [1 - F(a_L)]^{-1} \int_{a_L}^{\infty} \phi\, dF$ *a.s.*

Ackowledgement. The author thanks the referee for his helpful suggestion.

References

[1] Chen, K., Chao, M. & Lo, S., On strong uniform consistency of the Lynden-Bell estimator for truncated data, *Ann. Statist.,* **23** (1995), 440–449.

[2] Freedman, D., Markov chains, Springer-Verlag, New York, Heidelberg, Berlin 1983, 39–40.

[3] Gu, M. G. & Lai, T. L., Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann. Statist.,* **18** (1990), 160–189.

[4] Jacod, J., Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales, *Z. Wahrsch. verw. Gebiete.,* **31** (1975), 235–253.

[5] Kaplan, E. L. & Meier, P., Nonparametric estimation from incomplete observations, *J. Amer. Statist. Assoc.,* **53** (1958), 457–481.

[6] Neveu, J., Discrete-Parameter martingales, North-Holland, Amsterdam 1975.

[7] Serfling, R. J., Approximation theorems of mathematical statistics, Wiley, New York 1980.

[8] Stute, W. & Wang, J. L., The strong law under random censorship, *Ann. Statist.,* **21** (1993), 1591–1607.

[9] UzunoḠullari, Ü. & Wang, J. L., A comparison of hazard rate estimators for left truncated and right censored data, *Biometrika*, **79** (1992), 297–310.

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