A GENERALIZED PRODUCT-LIMIT ESTIMATOR FOR WEIGHTED DISTRIBUTION FUNCTIONS BASED ON CENSORED DATA

by

G. Campbell and A. Földes*
Purdue University

Technical Report #81-38
(REVISED May, 1982)

Department of Statistics
Purdue University

September, 1981

*Mathematical Institute of the Hungarian Academy of Sciences. This research was done in part while the author was visiting the Departments of Statistics and Mathematics at Purdue University. The author would like to acknowledge the financial assistance of the two departments.

A GENERALIZED PRODUCT-LIMIT ESTIMATOR FOR WEIGHTED DISTRIBUTION FUNCTIONS BASED ON CENSORED DATA

by

G. Campbell and A. Földes

ABSTRACT

The estimation of weighted distribution functions by weighted empirical distribution functions is extended to random censoring. For a sequence $\{F_i\}$ of distributions with associated positive weights w_i, an empirical estimate $\hat{F}_n$ is proved to converge uniform almost surely to

$$\mathbb{F}_n = \left(\sum_{i=1}^{n} w_i F_i \right) / \sum_{i=1}^{n} w_i$$

on the interval $[0,T]$ as $n \to \infty$ under suitable conditions. The convergence rate of $|\hat{F}_n - \mathbb{F}_n|$ is established. In the case in which all $F_i = F$, this estimate reduces to a weighted generalization of the product-limit estimator of Kaplan and Meier [J. Amer. Statist. Assoc. 53, 457-481]. In the case of no censoring the results reduce to a special case of Singh's theorem [Ann. Probab. 3, 371-374]. In the case of equal weights $w_i = w$, a duality argument obtains a weaker result than Földes [Z. Wahrscheinlichkeitstheorie verw. Gebiete 58, 95-107] for estimation of a distribution function under variable censoring.

A GENERALIZED PRODUCT-LIMIT ESTIMATOR FOR WEIGHTED DISTRIBUTION FUNCTIONS BASED ON CENSORED DATA

by

G. Campbell and A. Földes

1. Introduction. Let $X_1, X_2, \ldots$ be independent random variables, with X_i having survival distribution function $F_i(x) = P\{X_i > x\}$. Let Y_1, $Y_2, \ldots$ denote independent censoring variables, with Y_i having survival distribution function $G_i(y) = P\{Y_i > y\}$. It is assumed $\{X_i\}$ and $\{Y_i\}$ are independent. Define

$$Z_i = \min(X_i, Y_i) \quad i=1,2,\ldots .$$

In the special case in which $F_i = F$ and $G_i = G$ for all i, Kaplan and Meier (1958) generated the product-limit estimator to estimate F. Földes, Rejtő, and Winter (1980), Földes and Rejtő (1981a) have obtained uniform almost sure rates of convergence of this estimator to F. In the more general setting of $F_i = F$ for all i but variable G's, Földes and Rejtő (1981c) and Földes (1981) have proved almost sure uniform convergence of [0,T] with specified rates for some fixed T.

The approach in this paper is more general. Let $w_i > 0$ be the weight for F_i and let $\mathbb{F}_n = (\sum_{i=1}^n w_i F_i) / \sum_{i=1}^n w_i$. With no censoring, Singh (1975) proved uniform almost sure convergence of an estimator to $\mathbb{F}_n$ on the whole real line under certain conditions. The purpose of this paper is to obtain such convergence on [0,T] for some fixed T under random censoring with $G_i = G$ for all i.

Section 2 develops the necessary notation and introduces the estimator $\hat{F}_n$. Section 3 develops five crucial lemmas. The main theorem in Section 4 proves that under certain conditions, $\sup_{-\infty<u\leq T} |\hat{F}_n(u)-\mathbb{F}_n(u)|$ converges almost surely to zero at a specified rate.

The final section discusses the special cases of this theorem. If $F_i = F$ for all i, a weighted generalization of the product-limit estimator of Kaplan and Meier (1958) is obtained. If the labels "death" and "loss" are interchanged, a weighted but weaker generalization of the variable censoring result of Földes (1981) is obtained. Two applications of the estimator are mentioned.

2. Notation. For convenience this section includes a repetition of the notation of Section 1. The variables $X_1, X_2, \ldots$ are independent, with X_i having continuous survival distribution $F_i(x) = P\{X_i > x\}$. Also, $Y_1, Y_2, \ldots$ are independent, identically distributed with survival distribution $G(y) = P\{Y > y\}$. It is assumed $\{X_i\}$ and $\{Y_i\}$ are independent. Define $Z_i = \min(X_i, Y_i)$ and let $H_i(z) = P\{Z_i > z\}$ denote the survival distribution. Note that by independence

$$H_i(z) = F_i(z)G(z).$$

For the sequence of positive weights $\{w_i\}_{i=1}^{\infty}$ let

$$W_n = \sum_{i=1}^{n} w_i \quad \text{and} \quad K_n = \max_{1\leq i\leq n} w_i.$$

Define the following weighted functions:

$$\mathbb{F}_n(u) = W_n^{-1} \sum_{i=1}^{n} w_i F_i(u) \quad ; \quad \mathbb{L}_n(u) = \sum_{i=1}^{n} w_i F_i(u) = W_n \mathbb{F}_n(u) \; ;$$

$$\mathbb{H}_n(u) = W_n^{-1} \sum_{i=1}^{n} w_i F_i(u) G(u) = G(u)\mathbb{F}_n(u);$$

$$\mathbb{M}_n(u) = \sum_{i=1}^{n} w_i F_i(u) G(u) = W_n \mathbb{H}_n(u) = G(u)\mathbb{L}_n(u).$$

Estimate $\mathbb{M}_n(u)$ by its weighted empirical

$$M_n(u) = \sum_{i=1}^{n} w_i I_{\{Z_i > u\}},$$

where I_A is 1(0) if $u \in A (u \notin A)$. Note that if V_n denotes a random variable which is Z_i with probability w_i, for $i=1,\ldots,n$, then $M_n(Z_j)/[M_n(Z_j)+w_j]$ is an estimate of $P\{V_n > Z_j | V_n \geq Z_j\}$. This leads to $\hat{F}_n$ for the estimation of $\mathbb{F}_n$

$$(2.1) \qquad \hat{F}_n(u) = \begin{cases} \prod_{j=1}^{n} \left(\frac{M_n(Z_j)}{M_n(Z_j)+w_j} \right)^{\beta_j(u)} & \text{if} \quad u \leq \max_{1 \leq i \leq n} Z_i \\ 0 & \text{otherwise} \end{cases}$$

where $\beta_j(u) = I_{\{Z_j \leq u, X_j \leq Y_j\}}$.

Additional notation is also required. Let

$$B_n(u) = \sum_{i=1}^{n} w_i \beta_i(u)$$

$$\mathbb{B}_n(u) = E(B_n(u)) = - \sum_{i=1}^{n} w_i \int_{-\infty}^{u} G(s^-) dF_i(s) = - \int_{-\infty}^{u} G(s^-) d\mathbb{L}_n(s) .$$

The additional estimator $\check{F}_n$ will prove useful:

$$(2.2) \qquad \check{F}_n(u) = \begin{cases} \prod_{j=1}^{n} \left(\frac{M_n(Z_j)+w_j}{M_n(Z_j)+w_j+K_n} \right)^{\beta_j(u)} & u \leq \max_{1 \leq i \leq n} Z_i \\ 0 & \text{otherwise} . \end{cases}$$

Finally, define

$$\varphi_n(u) = \sum_{i=1}^{n} w_i^2 F_i(u) G(u) = \sum w_i^2 H_i(u) .$$

3. Important Lemmas. A lemma due to Petrov (1975) is used extensively throughout this section.

Lemma 3.1 (Petrov (1975) p.52). If $\xi_1,\ldots,\xi_n$ are independent random variables with mean 0 with $S_n = \sum_{i=1}^{n} \xi_i$ and if $\lambda_1,\ldots,\lambda_n$ and U are positive real numbers such that

$$E(e^{u\xi_k}) \leq e^{\frac{1}{2}\lambda_k u^2} \quad \text{for } k=1,\ldots,n;\ 0 \leq u \leq U.$$

Let $\Lambda = \sum_{k=1}^{n} \lambda_k$. Then

$$\text{i)} \qquad P\{S_n > x\} \leq \begin{cases} \exp(-\frac{x^2}{2\Lambda}) & \text{if } 0 \leq x \leq \Lambda U \\ \exp(-\frac{Ux}{2}) & \text{if } x > \Lambda U \end{cases}$$

$$\text{ii)} \qquad P\{|S_n| > x\} \leq \begin{cases} 2\exp(-\frac{x^2}{2\Lambda}) & \text{if } 0 \leq x \leq \Lambda U \\ 2\exp(-\frac{Ux}{2}) & \text{if } x > \Lambda U . \end{cases}$$

Lemma 3.2. i) If $\lambda \geq 2$ and if $\mathbb{M}(T) > K_n$,

$$P\left\{\sup_{u\leq T} \frac{\mathbb{M}_n(u)}{M_n(u)} > \lambda\right\} \leq 2n\exp\left\{-\frac{(\lambda-1)}{2\lambda}\ \frac{\mathbb{M}_n(T)}{16K_n}\right\} . \tag{3.1}$$

ii) If $K_n \log n/\mathbb{M}_n(T) \to 0$, then for almost all ω there exists an $N_0(\omega)$ such that for $n \geq N_0(\omega)$

$$\frac{1}{M_n(u)} \leq \frac{2}{\mathbb{M}_n(u)} \quad \text{for all } u \leq T.$$

Proof. Let $-\infty = u_0 < u_1 < \cdots < u_{k(n)} = T$ denote a partition of $(-\infty,T]$ such that $\mathbb{M}_n(u_{i-1}) - \mathbb{M}_n(u_i^-) = \delta_n(i) \leq K_n/2$ and such that $k(n) \leq 2(n-1)$. Since $\mathbb{M}_n(-\infty) = W_n \leq nK_n$, such a partition always exists. Then

$$P\left\{\sup_{u\le T}\frac{\mathbb{M}_n(u)}{M_n(u)}>\lambda\right\}\le\sum_{i=1}^{k(n)}P\left\{\sup_{u_{i-1}\le u<u_i}\frac{\mathbb{M}_n(u)}{M_n(u)}>\lambda\right\}+P\left\{\frac{\mathbb{M}_n(T)}{M_n(T)}>\lambda\right\}$$

$$\le\sum_{i=1}^{k(n)}P\left\{\frac{\mathbb{M}_n(u_{i-1})}{M_n(u_i^-)}>\lambda\right\}+P\left\{\frac{\mathbb{M}_n(T)}{M_n(T)}>\lambda\right\},$$

the last inequality from monotone nonincreasingness of $\mathbb{M}_n$ and M_n.

Now

$$P\left\{\frac{\mathbb{M}_n(u_{i-1})}{M_n(u_i^-)}>\lambda\right\}=P\{\mathbb{M}_n(u_i^-)-M_n(u_i^-)>\mathbb{M}_n(u_i^-)-\frac{1}{\lambda}\mathbb{M}_n(u_i$$

But since $\mathbb{M}_n(u_{i-1})=\mathbb{M}_n(u_i^-)+\delta_n(i)\le\mathbb{M}_n(u_i^-)+\frac{K_n}{2}$, it follows for $\lambda\ge 2$ and $\mathbb{M}_n(u_i^-)\ge K_n$ that

$$\mathbb{M}_n(u_i^-)-\frac{1}{\lambda}\mathbb{M}_n(u_{i-1})\ge(1-\frac{1}{\lambda})\mathbb{M}_n(u_i^-)-\frac{K_n}{2\lambda}\ge(1-\frac{3}{2\lambda})\mathbb{M}_n(u_i^-)>(\frac{\lambda-1}{2\lambda})\mathbb{M}_n(u_i^-).$$

Thus,

$$(3.2)\qquad P\left\{\sup_{u\le T}\frac{\mathbb{M}_n(u)}{M_n(u)}>\lambda\right\}\le\sum_{i=1}^{k(n)}P\{\mathbb{M}_n(u_i^-)-M_n(u_i^-)>\frac{\lambda-1}{2\lambda}\mathbb{M}_n(u_i^-)\}$$

$$+P\{\mathbb{M}_n(T)-M_n(T)>(1-\frac{1}{\lambda})\mathbb{M}_n(T)\}.$$

Recall for $|x|<\frac{1}{2}$ that $1+x\le e^x\le 1+x+x^2$ so for fixed u

$$(3.3)\qquad E[\exp(sw_i(I_{\{Z_i>u\}}-H_i(u)))]\le 1+s^2w_i^2E[I_{\{Z_i>u\}}-H_i(u)]^2$$

$$\le 1+s^2w_i^2H_i(u)$$

$$\le\exp\{s^2w_i^2H_i(u)\},$$

provided $s<\frac{1}{2K_n}$. Therefore

$$E[\exp\{s\sum_{i=1}^{n}w_i[I_{\{Z_i>u\}}-H_i(u)]\}]\le\exp\{s^2\varphi_n(u)\},$$

where $\varphi_n(u) = \sum_{i=1}^{n} w_i^2 H_i(u)$. Now apply Lemma 3.1 with the identification

$U = 1/8K_n$, $\lambda_i = 2w_i^2 H_i(u)$, and $\Lambda = 2\varphi_n(u)$. Let $\xi_i = w_i[I_{\{Z_i > u\}} - H_i(u)]$ and hence $S_n = M_n(u) - \mathbb{M}_n(u)$. For s such that $|s| \leq 1/8K_n$,

$$E(e^{sw_i\xi_i}) \leq \exp\{s^2 w_i^2 H_i(u)\}$$

by (3.3). Thus

$$P\{M_n(u) - \mathbb{M}_n(u) > x\} \leq \exp\{-x/16K_n\} \text{ for } x > \varphi_n(u)/4K_n.$$

From the above and (3.2), if $\lambda \geq 2$ such that $\frac{\lambda-1}{2\lambda} \mathbb{M}_n(u_i^-) > \varphi_n(u_i)/4K_n$ and $(1-\frac{1}{\lambda})\mathbb{M}_n(T) > \frac{\varphi_n(T)}{4K_n}$,

$$P\left\{\sup \frac{\mathbb{M}_n(u)}{M_n(u)} > \lambda\right\} \leq \sum_{i=1}^{k(n)} \exp\left\{-\left(\frac{\lambda-1}{2\lambda}\right)\frac{\mathbb{M}_n(u_i^-)}{16K_n}\right\} + \exp\left\{-\left(1-\frac{1}{\lambda}\right)\frac{\mathbb{M}_n(T)}{16K_n}\right\}$$

$$\leq 2n \exp\left\{-\left(\frac{\lambda-1}{2\lambda}\right)\frac{\mathbb{M}_n(T)}{16K_n}\right\}.$$

To prove ii), if $K_n \log n/\mathbb{M}_n(T) \to 0$ then for $\varepsilon = 2^{-8}$ there exists on N_0 such that for all $n \geq N_0$

$$\mathbb{M}_n(T) \geq 2^8 K_n \log n.$$

Applying the result in i), with $\lambda = 2$,

$$\sum_{n=N_0}^{\infty} P\left\{\sup_{u \leq T} \frac{\mathbb{M}_n(u)}{M_n(u)} > 2\right\} \leq \sum_{n=N_0}^{\infty} 2n \exp\left\{-\frac{2^8 K_n \log n}{64K_n}\right\}$$

$$\leq \sum_{n=N_0}^{\infty} 2n \exp\{-4\log n\} = \sum_{n=N_0}^{\infty} 2n^{-3} < \infty.$$

Now apply the Borel-Cantelli lemma to complete the proof. □

Lemma 3.3. Assume that F_i is continuous for $i=1,2,\ldots$ and that there exists a $\gamma > 0$ such that $W_n < n^\gamma$ for all $n > N_1$. Let $1 \le \alpha \le 2$ be arbitrary and define $\nu = 2\alpha - 1$. If $K_n \log n / \mathbb{M}_n(T) \to 0$ then

$$\sup_{t \le T} \left| \int_{-\infty}^{t} \frac{1}{\mathbb{M}_n^\alpha(u)} d(B_n(u) - \mathbb{B}_n(u)) \right| = \mathcal{O}\left(\frac{\sqrt{K_n \log n}}{(\mathbb{M}_n(T))^{\nu/2}} \right) .$$

Proof. First an exponential bound for the almost sure absolute value of the integral is obtained for fixed t. Observe that

$$\int_{-\infty}^{t} \frac{1}{\mathbb{M}_n^\alpha(u)} dB_n(u) = \sum_{j=1}^{n} \frac{w_j \beta_j(u)}{\mathbb{M}_n^\alpha(Z_j)}$$

and

$$\int_{-\infty}^{t} \frac{1}{\mathbb{M}_n^\alpha(u)} d\mathbb{B}_n(u) = - \sum_{j=1}^{n} w_j \int_{-\infty}^{t} \frac{G(u^-)}{\mathbb{M}_n^\alpha(u)} dF_j(u)$$

$$= \sum_{j=1}^{n} w_j E\left(\frac{\beta_j(Z_j)}{\mathbb{M}_n^\alpha(Z_j)} \right) .$$

So let $\eta_j(u) = \beta_j(u)/\mathbb{M}_n^\alpha(u)$ and $\eta_j^*(u) = \eta_j(u) - E(\eta_j(u))$. The $\eta_j^*(u)$ are independent for $j=1,2,\ldots$ and

$$\sum_{j=1}^{n} w_j \eta_j^*(t) = \int_{-\infty}^{t} \frac{1}{\mathbb{M}_n^\alpha(u)} d(B_n(u) - \mathbb{B}_n(u)).$$

To apply Lemma 3.1, identify $\xi_i = \eta_i^*(t)$. For s such that $|s w_j \eta_j^*(t)| < \frac{1}{2}$ then

$$E(e^{s w_j \eta_j^*(t)}) \le E[1 + s w_j \eta_j^*(t) + s^2 w_j^2 \eta_j^{*2}(t)] \le 1 + s^2 w_j^2 E(\eta_j^*(t))^2 .$$

If $\beta_j(T) = 1$, then $Z_j \le T$ and $\mathbb{M}_n(Z_j) \ge \mathbb{M}_n(T)$; it follows that

$$|\eta_j^*(t)| \le |\eta_j^*(T)| \le \frac{1}{\mathbb{M}_n^\alpha(T)} \quad \text{for all } t \le T.$$

Now for any $t \leq T$

$$E((\eta_j^*(t))^2) \leq E(\eta_j^2(t)) = E\left(\frac{\beta_j(t)}{\mathbb{M}_n^{2\alpha}(Z_j)}\right) = -\int_{-\infty}^{t} \frac{G(s^-)}{\mathbb{M}_n^{2\alpha}(s)}\, dF_j(s)$$

$$\leq -\frac{1}{G^{\nu}(T)} \int_{-\infty}^{t} \frac{1}{\mathbb{L}_n^{2\alpha}(s)}\, dF_j(s).$$

Hence

$$\sum_{i=1}^{n} w_j^2 E((\eta_j^*(t)^2) \leq -\sum_{i=1}^{n} \frac{w_j^2}{G^{\nu}(T)} \int_{-\infty}^{t} \frac{1}{\mathbb{L}_n^{\nu+1}(s)}\, dF_j(s)$$

$$\leq -\frac{K_n}{G^{\nu}(T)} \int_{-\infty}^{t} \frac{1}{\mathbb{L}_n^{\nu+1}(s)}\, d\mathbb{L}_n(s)$$

$$\leq \frac{K_n}{\nu G^{\nu}(T)\mathbb{L}_n^{\nu}(T)} \leq \frac{K_n}{\mathbb{M}_n^{\nu}(T)} ,$$

since $\nu^{-1} \leq 1$. Apply Lemma 3.1 with $U = \mathbb{M}_n^{\alpha}(T)/2K_n$, $\Lambda = 2K_n/\mathbb{M}_n^{\nu}(T)$. For $0 \leq \varepsilon \leq \mathbb{M}_n^{1-\alpha}(T)$ and any $t < T$,

(3.4) $$P\{|\sum w_j \eta_j^*(t)| > \varepsilon\} \leq 2 \exp\left\{-\frac{\varepsilon^2 \mathbb{M}_n^{\nu}(T)}{4K_n}\right\} .$$

To bound the supremum over the interval $(-\infty,T]$, note that the functions

$$\rho_n(t) = \sum w_j \eta_j(t) = \int_{-\infty}^{t} \frac{1}{\mathbb{M}_n^{\alpha}(u)}\, dB_n(u) \text{ and } q_n(t) = \int_{-\infty}^{t} \frac{1}{\mathbb{M}_n^{\alpha}(u)}\, d\mathbb{B}_n(u)$$

are non-decreasing in t. Further, for $\alpha \geq 1$, $t \leq T$, and $n \geq N_0$,

$$q_n(t) = \int_{-\infty}^{t} \frac{1}{G^{\alpha}(u)\mathbb{L}_n^{\alpha}(u)} G(u^-)d\mathbb{L}_n(u) \leq \int_{-\infty}^{t} \frac{G(u^-)}{G(u)\mathbb{L}_n(u)}\, d\mathbb{L}_n(u)$$

$$\leq |\log \mathbb{L}_n(T)|,$$

for $\mathbb{M}_n(T) = G(T)\mathbb{L}_n(T) > 1$ for all $n \geq N_0$ since $K_n \log n/\mathbb{M}_n(T) \to 0$.

Let $-\infty = u_0 < u_1 < \dots < u_{L(\varepsilon)} = T$ denote a partition of $(-\infty, T]$ such that

$$q_n(u_i) - q_n(u_{i-1}) \le \frac{\varepsilon}{3} \quad i=1,2,\dots,L(\varepsilon)$$

and

$$L(\varepsilon) \le \frac{q_n(T)}{\varepsilon/3} + 1 \le \frac{3|\log \mathbb{L}_n(T)|}{\varepsilon} + 1.$$

Since $q_n(t)$ is continuous, such a partition can always be constructed.

If $\rho_n(u_{i-1}) - q_n(u_{i-1}) \le \frac{\varepsilon}{3}$ and $\rho_n(u_i^-) - q_n(u_i^-) \le \frac{\varepsilon}{3}$ then by monotonicity of $\rho_n^{(t)}$ and $q_n(t)$ for any $u_{i-1} \le t < u_i$

$$|\rho_n(t)-q_n(t)| \le \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon.$$

Consequently, if

$$\sup_{u \le T} \left|\rho_n(u)-q_n(u)\right| > \varepsilon \text{ then, for some } 0 \le i \le L(\varepsilon),$$

$$|\rho_n(u_i)-q_n(u_i)| > \frac{\varepsilon}{3} \quad \text{or} \quad |\rho_n(u_i^-)-q_n(u_i^-)| > \frac{\varepsilon}{3}.$$

Apply inequality (3.4), for any $0 < \varepsilon \le \frac{1}{\mathbb{M}_n^{\alpha-1}(T)}$ and $\mathbb{M}_n(T) > 1$,

$$(3.5) \quad P\{\sup_{u \le T}|\rho_n(u)-q_n(u)| > \varepsilon\} \le 2[2L(\varepsilon)+1]\exp\{-\varepsilon^2 \mathbb{M}_n^{\nu}(T)/36K_n\}$$

$$\le 2\left[\frac{6|\log \mathbb{L}_n(T)|}{\varepsilon} + 3\right]\exp\left\{-\frac{\varepsilon^2 \mathbb{M}_n^{\nu}(T)}{36K_n}\right\}.$$

Let

$$\varepsilon_n = \sqrt{\frac{9CK_n \log n}{\mathbb{M}_n^{\nu}(T)}} \quad \text{for some } C > 0.$$

By hypothesis $\frac{K_n \log n}{\mathbb{M}_n(T)} \to 0$ so for $n \ge N_0$, $\mathbb{M}_n(T) > 9CK_n \log n$ and

$$0 \le \varepsilon_n \le \frac{3\sqrt{CK_n \log n}}{3\sqrt{CK_n \log n}\ \mathbb{M}_n^{\frac{\nu-1}{2}}(T)} = \frac{1}{\mathbb{M}_n^{\alpha-1}(T)} .$$

Therefore inequality (3.5) must be satisfied for $n \ge N_0$. Moreover, for $n > N_2 \ge \max(N_0,N_1)$, $W_n \le n^\gamma$ for some γ and hence

$$2\left[\frac{6|\log \mathbb{L}_n(T)|}{\varepsilon_n} + 3\right]\exp\left\{-\frac{\varepsilon_n^2 \mathbb{M}_n^{\nu}(T)}{36K_n}\right\}$$

$$\le 2\left[\frac{2\gamma\sqrt{\log n}}{\sqrt{CK_n}}\ \mathbb{M}_n^{\nu/2}(T) + 3\right]\exp\left\{-\frac{C\log n}{4}\right\}$$

$$\le 2\left[\frac{2\gamma\sqrt{\log n}}{\sqrt{CK_n}}\ n^{\frac{\gamma\nu}{2}} + 3\right]\exp\left\{-\frac{C\log n}{4}\right\}$$

Therefore if we choose $C > 4(\frac{3\gamma}{2}+1) > 4(\frac{\gamma\nu}{2}+1)$ then

$$\sum_{n=N_2}^{\infty} P\{\sup_{u\le T}|\rho_n(u)-q_n(u)| > \varepsilon_n\} < +\infty$$

which implies our statement. □

<u>Lemma 3.4</u>. If F_i is continuous for $i=1,2,\dots$, and if T is such that $G(T) > 0$ then for $\beta > 1$,

$$\int_{-\infty}^{T} \frac{1}{\mathbb{M}_n^{\beta}(u)}\, d\mathbb{B}_n(u) = \mathcal{O}\left(\frac{1}{\mathbb{M}_n^{\beta-1}(T)}\right) \qquad \text{a.s.}$$

Proof. For $\beta > 1$

$$\int_{-\infty}^{T} \frac{1}{\mathbb{M}_n^{\beta}(u)}\, d\mathbb{B}_n(u) = \int_{-\infty}^{T} \frac{G(u^-)}{(\mathbb{L}_n(u)G(u))^{\beta}}\, d\mathbb{L}_n(u)$$

$$\leq \frac{1}{(G(T))^{\beta-1}} \int_{-\infty}^{T} \frac{1}{\mathbb{L}_n^{\beta}(u)}\, d\mathbb{L}_n(u)$$

$$\leq \frac{1}{G(T)^{\beta-1}} \frac{1}{\beta-1} \frac{1}{\mathbb{L}_n^{\beta-1}(T)} = \mathcal{O}\left(\frac{1}{\mathbb{M}_n^{\beta-1}(T)}\right). \qquad \square$$

Lemma 3.5. If F_i is continuous for $i=1,2,\ldots$, if T is such that $G(T) > 0$, if there exists a $\gamma > 0$ such that $W_n < n^{\gamma}$, and if $\frac{K_n \log n}{\mathbb{M}_n(T)} \to 0$ as $n \to \infty$ then

$$\sup_{u \leq T} |\hat{F}_n(u) - \check{F}_n(u)| = \mathcal{O}\left(\frac{K_n}{\mathbb{M}_n(T)}\right) \quad \text{a.s.} \tag{3.6}$$

Proof.

$$|\hat{F}_n(u)-\check{F}_n(u)| = \left| \prod_{j=1}^{n} \left(\frac{M_n(Z_j)}{M_n(Z_j)+w_j}\right)^{\beta_j(u)} - \prod_{j=1}^{n} \left(\frac{M_n(Z_j)+K_n}{M_n(Z_j)+w_j+K_n}\right)^{\beta_j(u)} \right|$$

$$\leq \sum_{j=1}^{n} \beta_j(u) \left[\frac{w_j K_n}{(M_n(Z_j)+w_j)(M_n(Z_j)+w_j+K_n)}\right]^{\beta_j(u)}$$

$$\leq K_n \sum_{j=1}^{n} \frac{w_j \beta_j(u)}{M_n^2(Z_j)} \leq K_n \int_{-\infty}^{u} \frac{1}{M_n^2(v)}\, dB_n(v).$$

Apply ii) of Lemma 3.2; then there exists an N_0 such that for $n \geq N_0$

$$|\hat{F}_n(u)-\check{F}_n(u)| \leq 4K_n \int_{-\infty}^{u} \frac{1}{\mathbb{M}_n^2(v)}\, dB_n(u). \tag{3.7}$$

Taking sups on both sides of (3.7)

$$\sup_{u\leq T} |\hat{F}_n(u)-\check{F}_n(u)| \leq 4K_n \left\{ \sup_{u\leq T} \left| \int_{-\infty}^{u} \frac{1}{\mathbb{M}_n^2(v)} \, d(B_n(v)-\mathbb{B}_n(v)) \right| + \sup_{u\leq T} \int_{-\infty}^{u} \frac{1}{\mathbb{M}_n^2(v)} \, d\mathbb{B}_n(v) \right\}$$

$$\leq 4K_n \left\{ \mathcal{O}\left(\frac{\sqrt{K_n \log n}}{\mathbb{M}_n^{3/2}(T)} \right) + \mathcal{O}\left(\frac{1}{\mathbb{M}_n(T)} \right) \right\} \quad \text{a.s.}$$

This last inequality follows from Lemma 3.3 with $\alpha = 2$ and Lemma 3.4. By hypothesis $\sqrt{\frac{K_n \log n}{\mathbb{M}_n(T)}} \to 0$ so the overall rate is $\mathcal{O}\left(\frac{K_n}{\mathbb{M}_n(T)}\right)$. □

<u>Lemma 3.6.</u> i) If $\frac{2}{\sqrt{\varphi_n(T)}} < \varepsilon < \frac{\sqrt{\varphi_n(T)}}{K_n}$ and if $\varphi_n(T) \geq K_n$ then

$$P\left\{ \sup_{u\leq T} \left| \frac{M_n(u)-\mathbb{M}_n(u)}{\sqrt{\varphi_n(u)}} \right| > \varepsilon \right\} \leq (4n+2)\exp\{-2^{-7}\varepsilon^2\} .$$

ii) If $(K_n^*)^2 \log n/\varphi_n(T) \to 0$ where $K_n^* = \max(K_n,1)$ then

$$P\left\{ \sup_{u\leq T} \left| \frac{M_n(u)-\mathbb{M}_n(u)}{\sqrt{\varphi_n(u)}} \right| = \mathcal{O}(\sqrt{\log n}) \right\} = 1. \tag{3.8}$$

<u>Proof</u>. For fixed n and T, let $-\infty = u_0 < u_1 < \ldots < u_{k(n)} = T$ denote a partition of $(-\infty,t]$ such that

$$\delta_n(i) = \mathbb{M}_n(u_{i-1}) - \mathbb{M}_n(u_i^-) \leq K_n \qquad \text{for } i=1,2,\ldots,k(n)$$

and $k(n) \leq n$. Such a partition always exists since $\mathbb{M}_n(u)$ is decreasing in u and since $\mathbb{M}_n(-\infty) = W_n \leq nK_n$. Now

$$P\left\{\sup_{u\le T} \left| \frac{M_n(u)-\mathbb{M}_n(u)}{\sqrt{\varphi_n(u)}} \right| > \varepsilon\right\}$$

$$\le \sum_{i=1}^{k(n)} P\left\{\sup_{u_{i-1}\le u<u_i} \left| \frac{M_n(u)-\mathbb{M}_n(u)}{\sqrt{\varphi_n(u)}} \right| > \varepsilon\right\} + P\left\{\frac{|M_n(T)-\mathbb{M}_n(T)|}{\sqrt{\varphi_n(T)}} > \varepsilon\right\}$$

(3.9) $$\le \sum_{i=1}^{k(n)} P\left\{\frac{\sup_{u_{i-1}\le u<u_i} |M_n(u)-\mathbb{M}_n(u)|}{\sqrt{\varphi_n(u_i^-)}} > \varepsilon\right\} + P\left\{\frac{|M_n(T)-\mathbb{M}_n(T)|}{\sqrt{\varphi_n(T)}} > \varepsilon\right\}$$

$$\le \sum_{i=1}^{k(n)} \Bigg[P\left\{|M_n(u_{i-1})-\mathbb{M}_n(u_{i-1})| > \frac{\varepsilon\sqrt{\varphi_n(u_i^-)} - \delta_n(i)}{2}\right\}$$

$$+ P\left\{|M_n(u_i^-)-\mathbb{M}_n(u_i^-)| > \frac{\varepsilon\sqrt{\varphi_n(u_i^-)} - \delta_n(i)}{2}\right\}\Bigg]$$

$$+ P\left\{|M_n(T)-\mathbb{M}_n(T)| > \varepsilon\sqrt{\varphi_n(T)}\right\} ,$$

the last inequality following from the monotonicity of $\mathbb{M}_n(u)$.

Let $\xi_i(u) = I_{\{Z_i>u\}}$, $\xi_i^*(u) = \xi_i(u) - H_i(u)$. If $|sw_i\xi_i^*(u)| < \frac{1}{2}$, then

(3.10) $$E\left(e^{sw_i\xi_i^*(u)}\right) \le 1 + s^2w_i^2E(\xi_i^*(u))^2$$

$$\le 1 + s^2w_i^2E(\xi_i^2(u)) \le \exp\{s^2w_i^2H_i(u)\} .$$

Since $|\xi_i^*(u)| \le 1$ then $|sw_i\xi_i^*(u)| \le 1$ if $0 \le s \le \frac{1}{2K_n}$. From (3.10) for such an s

$$E\left[e^{s(M_n(u)-\mathbb{M}_n(u))}\right] \le \exp\{s^2\varphi_n(u)\} .$$

In order to apply Lemma 3.1, let $U = (2K_n)^{-1}$, $\Lambda = 2\varphi_n(u)$ then for $0 \le x \le U\Lambda = \varphi_n(u)/K_n$,

(3.11) $$P\{|M_n(u)-\mathbb{M}_n(u)| > x\} \le 2\exp\{-x^2/4\varphi_n(u)\} .$$

The condition $\varepsilon < \sqrt{\varphi_n(T)}/K_n$ of the lemma implies

$$(3.12) \qquad \tfrac{1}{2}(\varepsilon\sqrt{\varphi_n(u_i^-)} - \delta_n(i)) \le \tfrac{1}{2}(\sqrt{\varphi_n(T)}\sqrt{\varphi_n(u_i^-)}/K_n - \delta_n(i))$$

$$\le \varphi_n(u_i^-)/2K_n \le \varphi_n(u_{i-1})/2K_n$$

and also

$$(3.13) \qquad \varepsilon\sqrt{\varphi_n(T)} < \varphi_n(T)/K_n.$$

Expressions (3.12) and (3.13) guarantee that (3.11) can be applied for all terms of (3.9) to obtain

$$P\left\{\sup_{u\le T} \left|\frac{M_n(u)-\mathbb{M}_n(u)}{\sqrt{\varphi_n(u)}}\right| > \varepsilon\right\} \le \sum_{i=1}^{k(n)} 2\left[\exp\left\{-\frac{(\varepsilon\sqrt{\varphi_n(u_i^-)}-\delta_n(i))^2}{16\varphi_n(u_{i-1})}\right\}\right.$$

$$\left.+\exp\left\{-\frac{(\varepsilon\sqrt{\varphi_n(u_i^-)}-\delta_n(i))^2}{16\varphi_n(u_i^-)}\right\}\right] + 2\exp\left\{-\frac{\varepsilon^2\varphi_n(T)}{4\varphi_n(T)}\right\}.$$

Since $2K_n/\sqrt{\varphi_n(T)} < \varepsilon$, it follows that $2K_n/\sqrt{\varphi_n(u_i^-)} < \varepsilon$ and hence

$$K_n^{-1}(\varepsilon\sqrt{\varphi_n(u_i^-)} - \delta_n(i)) > K_n^{-1}(\varepsilon\sqrt{\varphi_n(u_i^-)} - K_n) > \varepsilon\sqrt{\varphi_n(u_i^-)}/2K_n > 1$$

since $\varepsilon\sqrt{\varphi_n(u_i^-)} > 2K_n$. Moreover, since $\varphi_n(u_{i-1}) - \varphi_n(u_i^-) = \sum w_j^2(H(u_{i-1})-H(u_i^-)) \le K_n(\mathbb{M}_n(u_{i-1})-\mathbb{M}_n(u_i^-)) \le K_n^2$ and since $\varphi_n(T) \ge K_n$, it follows that

$$\frac{\varphi_n(u_i^-)}{\varphi_n(u_{i-1})} > \frac{\varphi_n(u_i^-)}{\varphi_n(u_i^-)+K_n} \ge \tfrac{1}{2}.$$

From (3.12) and (3.13)

$$P\left\{\sup_{u\le T} \left|\frac{M_n(u)-\mathbb{M}_n(u)}{\sqrt{\varphi_n(u)}}\right| > \varepsilon\right\} \le \sum_{i=1}^{k(n)} 2\left[\exp\left\{-\frac{\varepsilon^2\varphi_n(u_i^-)}{64\varphi_n(u_{i-1})}\right\}\right.$$

$$\left.+\exp\left\{-\frac{\varepsilon^2\varphi_n(u_i^-)}{64\varphi_n(u_i^-)}\right\}\right] + 2\exp\left\{-\frac{\varepsilon^2}{4}\right\} \le (4n+2)\exp\{-2^{-7}\varepsilon^2\}.$$

This proves i).

To prove ii), let $\varepsilon_n = \sqrt{2^9 \log n}$. Then by hypothesis there exists an N_0 such that for all $n \geq N_0$

$$\varphi_n(T) > 2^{20}(K_n^*)^2 \log n$$

Thus,

$$\frac{2}{\sqrt{\varphi_n(T)}} < \frac{2}{2^{10} K_n^* \sqrt{\log n}} = 2^{-9} \frac{1}{K_n^* \sqrt{\log n}} < 2^{-9} \frac{1}{\sqrt{\log n}}$$

and

$$\frac{\sqrt{\varphi_n(T)}}{K_n} \geq \frac{\sqrt{\varphi_n(T)}}{K_n^*} > \frac{2^{\frac{19}{2}} K_n^* \sqrt{\log n}}{K_n^*} > 2^9 \sqrt{\log n} \quad .$$

Apply i) since $2^{-9} \frac{1}{\sqrt{\log n}} < \varepsilon_n = \sqrt{2^9 \log n} < 2^9 \sqrt{\log n}$:

$$\sum_{n=N_0}^{\infty} P\left\{ \sup_{u \leq T} \left| \frac{M_n(u) - \mathbb{M}_n(u)}{\sqrt{\varphi_n(u)}} \right| > \sqrt{2^9 \log n} \right\} \leq \sum_{n=N_0}^{\infty} (4n+2)\exp\{-4 \log n\} < \infty .$$

Then ii) follows by application of the Borel-Cantelli lemma. □

4. Main Theorem. If the F_i's are continuous, if T is such that $G(T) > 0$, if there exists a $\gamma > 0$ such that $W_n < n^\gamma$, and if

$$(K_n^*)^2 \log n / \varphi_n(T) \to 0$$

where $K_n^* = \max(K_n, 1)$ then

$$P\left\{ \sup_{u \leq T} |\hat{F}_n(u) - \mathbb{F}_n(u)| = \mathcal{O}\left(\sqrt{\frac{K_n \log n}{\mathbb{M}_n(T)}} \right) \right\} = 1.$$

Proof. By Lemma 3.5,

$$\sup_{u \leq T} |\hat{F}_n(u) - \mathbb{F}_n(u)| \leq \sup_{u \leq T} |\check{F}_n(u) - \mathbb{F}_n(u)| + \sup_{u \leq T} |\hat{F}_n(u) - \check{F}(u)|$$

$$\leq \sup_{u \leq T} |\check{F}_n(u) - \mathbb{F}_n(u)| + \mathcal{O}\left(\frac{K_n}{\mathbb{M}_n(T)} \right) = \mathcal{O}\left(\sqrt{\frac{K_n \log n}{\mathbb{M}_n(T)}} \right) .$$

Using the well-known inequality $|x-y| \le |\log x - \log y|$ for $0 < x,y \le 1$,

$$|\check{F}_n(u)-\mathbb{F}_n(u)| \le |\log \check{F}_n(u)-\log \mathbb{F}_n(u)| .$$

But by a series expansion

$$\log \check{F}_n(u) = \sum_{j=1}^{n} \beta_j(u)\log\left(1 - \frac{w_j}{M_n(Z_j)+w_j+K_n}\right)$$

$$= -\sum_{j=1}^{n} \beta_j(u)\frac{w_j}{M_n(Z_j)+w_j+K_n} + \sum_{j=1}^{n} \beta_j(u)\sum_{\ell=2}^{\infty}(-1)^{\ell}\frac{1}{\ell}\left(\frac{w_j}{M_n(Z_j)+w_j+K_n}\right)^{\ell}$$

$$= D_n + E_n.$$

Observe that the condition of the theorem implies that

$$K_n \log n/\mathbb{M}(T) \to 0$$

as $n \to \infty$ since $(K_n^*)^2 \log n/\varphi_n(T) \ge (K_n^*)^2 \log n/K_n\mathbb{M}_n(T) \ge K_n \log n/\mathbb{M}_n(T)$.

For the second term, since $\frac{w_j}{M_n(Z_j)+w_j+K_n} < \frac{1}{2}$,

$$E_n < \sum_{j=1}^{n} \frac{\beta_j(u)w_j^2}{(M_n(Z_j)+w_j+K_n)^2} < K_n \int_{-\infty}^{u} \frac{1}{M_n^2(v)}\, dB_n(v).$$

By Lemmas 3.2, 3.3, and 3.4, for $n \ge N_0$, under the hypothesis of the theorem,

$$E_n \le 4K_n\left[\int_{-\infty}^{u} \frac{1}{\mathbb{M}_n^2(v)}\, d(B_n(v)-\mathbb{B}_n(v)) + \int_{-\infty}^{u} \frac{1}{\mathbb{M}_n^2(v)}\, d\mathbb{B}_n(v)\right]$$

$$\le \mathcal{O}\left(\frac{K_n\sqrt{K_n \log n}}{(\mathbb{M}_n(T))^{\frac{3}{2}}}\right) + \mathcal{O}\left(\frac{K_n}{\mathbb{M}_n(T)}\right) = \mathcal{O}\left(\frac{K_n\sqrt{K_n \log n}}{(\mathbb{M}_n(T))^{\frac{3}{2}}}\right)$$

$$= \mathcal{O}\left(\frac{K_n\sqrt{\log n}}{\mathbb{M}_n(T)}\right) .$$

$$|D_n - \log \mathbb{F}_n(u)| = \left| -\sum_{j=1}^{n} \frac{\beta_j(u)w_j}{M_n(Z_j)+w_j+K_n} + \int_{-\infty}^{u} \frac{1}{\mathbb{M}_n(v)}\, d\mathbb{B}_n(v) \right|$$

$$\leq \left| -\sum_{j=1}^{n} \frac{\beta_j(u)w_j}{M_n(Z_j)+w_j+K_n} + \sum_{j=1}^{n} \frac{\beta_j(u)w_j}{M_n(Z_j)} \right|$$

$$+ \left| -\int_{-\infty}^{u} \frac{1}{M_n(v)}\, d(B_n(v)-\mathbb{B}_n(v)) \right| + \left| -\int_{-\infty}^{u} \left(\frac{1}{M_n(v)} - \frac{1}{\mathbb{M}_n(v)} \right) d\mathbb{B}_n(v) \right|$$

$$\leq \sum_{j=1}^{n} \frac{\beta_j(u)w_j(w_j+K_n)}{M_n(Z_j)(M_n(Z_j)+w_j+K_n)} + \left| \int_{-\infty}^{u} \frac{1}{M_n(v)}\, d(B_n(v)-\mathbb{B}_n(v)) \right|$$

$$+ \left| \int_{-\infty}^{u} \frac{M_n(v)-\mathbb{M}_n(v)}{M_n(v)\mathbb{M}_n(v)}\, d\mathbb{B}_n(v) \right|$$

$$\leq 2K_n \int_{-\infty}^{u} \frac{1}{M_n^2(v)}\, dB_n(v) + \left| \int_{-\infty}^{u} \frac{1}{M_n(v)}\, d(B_n(v)-\mathbb{B}_n(v)) \right|$$

$$+ \int_{-\infty}^{u} \frac{|M_n(v)-\mathbb{M}_n(v)|}{\sqrt{\mathbb{M}_n(v)}} \frac{1}{(\mathbb{M}_n(v))^{3/2}}\, d\mathbb{B}_n(v).$$

Since $K_n \mathbb{M}_n(u) > \varphi_n(u)$, Lemma 3.6 can be applied to the last term. Then repeated application of Lemmas 3.2, 3.3 and 3.4 results in:

$$|D_n - \log \mathbb{F}_n(u)| \leq 2K_n \int_{-\infty}^{u} \frac{1}{M_n^2(v)}\, d(B_n(v)-\mathbb{B}_n(v))$$

$$+ 2K_n \int_{-\infty}^{u} \frac{1}{M_n^2(v)}\, d\mathbb{B}_n(v) + 2\mathcal{O}\left(\sqrt{K_n \log n/\mathbb{M}_n(T)}\right)$$

$$\leq \mathcal{O}\left(\frac{K_n\sqrt{K_n \log n}}{\mathbb{M}_n^{3/2}(T)} \right) + \mathcal{O}\left(\frac{K_n}{\mathbb{M}_n(T)} \right) + \mathcal{O}\left(\sqrt{\frac{K_n \log n}{\mathbb{M}_n(T)}} \right)$$

$$= \mathcal{O}\left(\sqrt{\frac{K_n \log n}{\mathbb{M}_n(T)}} \right) \qquad \text{a.s.}$$

Thus

$$\sup_{u\le T}|\hat{F}_n(u)-\mathbb{F}_n(u)| \le \sup_{u\le T}|\check{F}_n(u)-\mathbb{F}_n(u)| + \sup_{u\le T}|\hat{F}_n(u)-\check{F}_n(u)|$$

$$\le \mathcal{O}\left(\sqrt{\frac{K_n \log n}{\mathbb{M}_n(T)}}\right) + \mathcal{O}\left(\frac{K_n}{\mathbb{M}_n(T)}\right) = \mathcal{O}\left(\sqrt{\frac{K_n \log n}{\mathbb{M}_n(T)}}\right) \quad \text{a.s.}$$

The theorem is proved. □

5. Discussion and Applications. Theorem 4.1 is a limited extension of Singh (1975) to randomly censored data. If the result of Singh (1975) is specialized to positive weights w_i with Singh's weight $\alpha = 0$, then the result is that

$$\sup_u \max_{N\le n}|M_n(u)-\mathbb{M}_n(u)| = \mathcal{O}(a_n) \quad \text{a.s.,}$$

provided $a_n \ge \left(\sum_{i=1}^n w_i^2\right)^{\frac{1}{2}}$ and provided

$$\sum_{n=1}^{\infty} a_n \frac{W_n}{\left(\sum_{i=1}^n w_i^2\right)} \exp\left\{-\frac{2a_n^2}{\left(\sum_{i=1}^n w_i^2\right)}\right\} < \infty .$$

Note that the supremum is over the entire range not merely $u \le T$. If $w_i = 1$, the above result implies that

$$\sup_u \max_{N\le n}|L_n(u)-\mathbb{L}_n(u)| = \mathcal{O}(\sqrt{n\log n}) \quad \text{a.s.;}$$

the result of Theorem 4.1 here is

$$\sup_{u\le T} n|\hat{F}_n(u)-\mathbb{F}_n(u)| = \mathcal{O}\left(\frac{n\sqrt{\log n}}{\sqrt{\mathbb{M}_n(T)}}\right) \quad \text{a.s.}$$

If T is such that $\sum_{i=1}^n F_i(T) \ge Cn$, the two results agree.

As a special case of Theorem 4.1, if $w_i = 1$ and if $F_i = F$ for all i, then $\hat{F}_n$ is merely the Kaplan-Meier product-limit estimator and the rate result of the theorem is

$$\sup_{u\leq T}|\hat{F}_n(u)-F(u)| = \mathcal{O}\left(\sqrt{\frac{\log n}{n}}\right) \quad \text{a.s.}$$

This is the uniform almost sure convergence of Földes and Rejtö (1981a).

In this paper the censoring distribution denoted by G has been fixed and the F_i's vary. Suppose the labels "death" and "loss" are interchanged so that G is now the survival distribution to be estimated and the F_i's are variable censoring distributions. Let $\hat{G}_n$ denote the weighted estimator defined implicitly by the relation:

$$\hat{F}_n(u)\hat{G}_n(u) = H_n(u) = W_n^{-1}\sum_{i=1}^{n} w_i I_{\{Z_i>u\}} .$$

Then

$$(5.1)\qquad H_n(u) - \mathbb{H}_n(u) = \hat{F}_n(u)\hat{G}_n(u) - \mathbb{F}_n(u)G(u)$$

$$= (\hat{F}_n(u)-\mathbb{F}_n(u))\hat{G}_n(u) + \mathbb{F}_n(u)(\hat{G}_n(u)-G(u)).$$

It follows that

$$(5.2)\qquad \sup_{u\leq T}|\hat{G}_n(u)-G(u)| \leq \frac{1}{\mathbb{F}_n(T)}\left[\sup_{u\leq T}|H_n(u)-\mathbb{H}_n(u)| + \sup_{u\leq T}|\hat{F}_n(u)-\mathbb{F}_n(u)|\right]$$

Now apply Theorem 4.1 twice, once to the $H_n(u) - \mathbb{H}_n(u)$ term in the case in which $G(T) = 1$ (no censoring) and once to the final term. The result is

$$\sup_{u\leq T}|\hat{G}_n(u)-G(u)| = \mathcal{O}\left(\sqrt{\frac{K_n \log n}{\mathbb{L}_n(T)}}\ \frac{1}{\mathbb{F}_n(T)}\right) \quad \text{a.s.,}$$

This extends in some sense the result of Földes (1981) for variable censoring to the case of unequal weights. If $w_i = 1$ for all i,

$$\sup_{u\leq T}|\hat{G}_n(u)-G(u)| = \mathcal{O}\left(\sqrt{\frac{\log n}{\sum_{i=1}^{n} F_i(T)}} \cdot \frac{n}{\sum_{i=1}^{n} F_i(T)}\right) \quad \text{a.s.},$$

a generally much weaker rate than that of Földes (1981).

Consider the following application of this weighted estimator. Suppose that the survival functions for males and females are thought to be identical or alternately that the overall survival function is of interest. Further, the random sample contains disproportionate representation of males and females. Then one could use as weights for males and females the actual proportion of each population in the sample, and obtain the weighted estimator. Note that this estimator in general differs from the estimator

$$(p_m F_m + p_f F_f)/(p_m + p_f)$$

where p_m, p_f denote the known sampled proportions of populations and F_m,F_f the separate Kaplan-Meier (unweighted) product-limit estimators for males and females, respectively.

A second application of this weighted product-limit estimator is in the development of a nonparametric bivariate estimator of a distribution function in which the weights are probabilities that certain individuals survive to a fixed time. This estimator, the topic of another paper, would improve the estimators of Campbell and Földes (1982).

REFERENCES

Campbell, G. and Földes, A. (1982). Large-sample properties of nonparametric bivariate estimators with censored data. Coll. Math. J. Bolyai 32. Nonparametric Statistical Inference. North-Holland.

Földes, A. (1981). Strong uniform consistency of the product limit estimator under variable censoring. Z. Wahrscheinlichkeitstheorie verw. Gebiete 58, 95-107.

Földes, A. and Rejtő, L. (1981a). Strong uniform consistency of nonparametric survival curve estimators from randomly censored data. Ann. Statist. 9, 122-129.

Földes, A. and Rejtő, L. (1981b). Asymptotic properties of the nonparametric survival curve estimators under variable censoring. The First Pannonian Symposium on Mathematical Statistics. Lecture Notes in Statistics 8 (P. Revesz, ed.), 54-72. Springer-Verlag.

Földes, A. and Rejtő, L. (1981c). A LIL type result for the product limit estimator on the whole line. Z. Wahrscheinlichkeitstheorie verw. Gebiete 56, 75-86.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Am. Statist. Assoc. 53, 457-481.

Petrov, V. V. (1975). Sums of Independent Random Variables. Springer-Verlag.

Singh, R. S. (1975). On the Glivenko-Cantelli theorem for weighted empiricals based on independent random variables. Ann. Probab. 3, 371-374.