Z. Wahrscheinlichkeitstheorie verw. Gebiete
58, 95–107 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Strong Uniform Consistency of the Product Limit Estimator under Variable Censoring

Antonia Földes*

Mathematical Institute, Reáltanoda u 13–15, 1053 Budapest, Hungary

1. Introduction

Let $X_1, \ldots, X_N \ldots$ be an i.i.d. sequence of random variables, with continuous distribution function $F(x)$, and let $Y_1, \ldots, Y_N$ be another sequence of independent random variables with right continuous distribution functions $G_1(x), \ldots, G_N(x) \ldots$. We suppose that $\{X_i\}$ and $\{Y_i\}$ are mutually independent.

Let

$$Z_i = \min\{X_i, Y_i\} \quad \text{and} \quad \delta_i = [X_i \leqq Y_i] \quad i = 1, 2, \ldots$$

($[A]$ denotes the indicator of the event A).

As it is well known, for this problem the $F_N^*(x)$ product limit estimator of Kaplan Meier [6] is maximum likelihood estimator. For i.i.d. Y_i-s it was recently proved [3], that if $F(x)$ is continuous and $G(T_F) < 1$ where $T_F = \{\sup x; F(x) < 1\}$ then

$$P\left(\sup_{-\infty < x < +\infty} |F_N^*(x) - F(x)| = O\left(\sqrt{\frac{\log\log N}{N}}\right)\right) = 1. \tag{1.1}$$

The case of variable censoring (i.e. Y_i's have different distributions) was discussed in [2].

Let $P(Z_i < x) = H_i(x)$, $\bar{F}(x) = 1 - F(x)$ and define $\bar{G}_i(x)$, $\bar{H}_i(x)\,(= \bar{F}(x)\,\bar{G}_i(x))$ similarly. Denote

$$M_N(t) = \sum_{k=1}^{N} [Z_k > t], \qquad m_N(t) = E(M_N(t)) = \bar{F}(t) \sum_{k=1}^{N} \bar{G}_k(t),$$

$$\bar{G}(N, t) = \sum_{k=1}^{N} \bar{G}_k(t).$$

---

* Mathematical Institute of the Hungarian Academy of Sciences. This research was done 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

0044-3719/81/0058/0095/$02.60

Recently Gill [5] proved that $\sup_{-\infty<u\leqq t}|F_N^*(u)-F(u)|\xrightarrow[P]{}0$ if $M_N(t)\xrightarrow[P]{}+\infty$ and $F(t^-)<1$ (where $\xrightarrow[P]{}$ denotes stochastic convergence).

We shall prove the following

Theorem 1. *Suppose that F is continuous and*

$$\frac{\log N}{\bar{G}(N,T_F)}\to 0. \tag{1.2}$$

Then

$$P\left(\sup_{-\infty<u<+\infty}|F_N^*(u)-F(u)|=O\left(\sqrt{\frac{\log N}{\bar{G}(N,T_F)}}\right)\right)=1. \tag{1.3}$$

Remark. To get a better insight into the meaning of the theorem, denote by $G_{(N)}(t)=\frac{1}{N}\sum_{i=1}^{N}G_i(t)$ the average of the censoring distributions. Then condition (1.2) implies that $T_F<T_{G(N)}$ for all but finitely many N (in particular $T_F<\infty$). Theorem 1 can also be stated as

$$\overline{\lim_{N\to\infty}}\sqrt{1-G_{(N)}(T_F)}\sqrt{\frac{N}{\log N}}\sup_{-\infty<u<+\infty}|F_N^*(u)-F(u)|\leqq K \quad \text{a.s.}.$$

Consequently if $\overline{\lim_{N\to\infty}}\, G_{(N)}(T_F)<1$ then the rate is $O(\sqrt{\log N/N})$, if not, then the $\sqrt{1-G_{(N)}(T_F)}$ factor worsen this rate.

Corollary 1. *Suppose that $F(t)$ is continuous in $(-\infty,t]$, and $\log N/\bar{G}(N,t)\to 0$, then*

$$P\left(\sup_{-\infty<u\leqq t}|F_N^*(u)-F(u)|=O\left(\sqrt{\frac{\log N}{\bar{G}(N,t)}}\right)\right)=1.$$

Corollary 2. *If $F(t)$ is continuous and $\underline{\lim}(\bar{G}(N,T_F)/N^\alpha)=a>0$, for any $0<\alpha\leqq 1$, then*

$$P\left(\sup_{-\infty<u<\infty}|F_N^*(u)-F(u)|=O\left(\sqrt{\frac{\log N}{N^\alpha}}\right)\right)=1.$$

Our technic is similar to the paper [3]. Theorem 1 gives a much stronger result under less restrictive conditions than the above mentioned theorem in [2]. It is important to emphasize that in course of proving Theorem 1 we need some theorems on strong uniform behaviour of empirical distributions of non-identically distributed random variables, which seems to be new.

2. Definitions, notations

In what follows we list some more notations.

$$\beta_i(t) = [Z_i \leqq t, \delta_i = 1] \qquad i = 1, 2, \ldots, \tag{2.1}$$

$$B_N(t) = \sum_{i=1}^{N} \beta_i(t), \qquad b_N(t) = E(B_N(t)) = \sum_{i=1}^{N} \int_{-\infty}^{t} \bar{G}_i(u^-)\, dF(u), \tag{2.2}$$

$$\tau_N(\omega) = \max_{j \leqq N} \{Z_j(\omega)\}. \tag{2.3}$$

The definition of the product limit estimator $F_N^*(t)$ (in case of continuous F), and its' modification $F_N^0(t)$ are the following

$$\bar{F}_N^*(t) = \begin{cases} \prod_{j=1}^{N} \left(\dfrac{M_N(Z_j)}{M_N(Z_j)+1}\right)^{\beta_j(t)} & \text{if } t \leqq \tau_N(\omega) \\ 0 & \text{if } t > \tau_N(\omega), \end{cases}$$

$$\bar{F}_N^0(t) = \begin{cases} \prod_{j=1}^{N} \left(\dfrac{M_N(Z_j)+1}{M_N(Z_j)+2}\right)^{\beta_j(t)} & \text{if } t \leqq \tau_N(\omega) \\ 0 & \text{if } t > \tau_N(\omega). \end{cases}$$

3. Uniform Properties of Empirical Distribution of Nonidentically Distributed Random Variables

Our basic tool is the following exponential bound (see Petrov [7] p. 52).

Lemma 3.1. *Let $\xi_1, \ldots, \xi_N$ be a sequence of independent random variables. $S_N = \sum_{i=1}^{N} \xi_i$. Suppose that there exist $\lambda_1, \lambda_2, \ldots, \lambda_N$ and U positive real numbers such that*

$$E(e^{u\xi_k}) \leqq e^{(1/2)\lambda_k u^2} \qquad k = 1, 2, \ldots, N \text{ for } 0 \leqq u \leqq U. \tag{3.1}$$

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

$$P(S_N > x) \leqq \exp\left\{-\frac{x^2}{2\Lambda}\right\} \quad \text{if } 0 \leqq x \leqq \Lambda U, \tag{3.2}$$

$$P(S_N > x) \leqq \exp\left\{-\frac{Ux}{2}\right\} \quad \text{if } x \geqq \Lambda U, \tag{3.3}$$

$$P(|S_N| > x) \leqq 2\exp\left\{-\frac{x^2}{2\Lambda}\right\} \quad \text{if } 0 \leqq x \leqq \Lambda U, \tag{3.4}$$

$$P(|S_N| > x) \leqq 2\exp\left\{-\frac{Ux}{2}\right\} \quad \text{if } x > \Lambda U. \tag{3.5}$$

Remark 3.2. Let $\{\alpha_i\}_{i=1}^{\infty}$ be a sequence of independent Bernoulli variables $P(\alpha_i=1)=a_i$, $P(\alpha_i=0)=1-a_i$. Let $A_N=\sum_{i=1}^{N} a_i$. Denote $\xi_i=\alpha_i-a_i$, $E(\xi_i)=0$, $i=1,2,\ldots$

$$E(e^{u\xi_i})\leqq E(1+u\,\xi_i+u^2\,\xi_i^2)=1+u^2\,E(\xi_i^2)\leqq \exp(u^2\,E(\xi_i^2)) \tag{3.6}$$

if only

$$|u\,\xi_i|=|u(\alpha_i-a_i)|\leqq \tfrac{1}{2}.$$

As $|\alpha_i-a_i|\leqq 1$ and $E(\xi_i^2)=a_i(1-a_i)\leqq a_i$, $E(e^{u\xi_i})\leqq \exp\{u^2\,a_i\}$ if $0\leqq u\leqq \frac{1}{2}$. Using the notations of Lemma 3.1 we got

$$U=\tfrac{1}{2},\qquad \lambda_i\geqq 2a_i\qquad i=1,2,\ldots\qquad \Lambda\geqq 2A_N. \tag{3.7}$$

In what follows we prove some uniform properties of $M_N(u)$. (Though we use the same $M_N(u)$ notation here as in the rest of the paper for the empirical of $Z_i=\min(X_i, Y_i)$, in these 2 theorems we do not suppose anything about the minimum structure of Z_i $\{Z_i\}$ are independent but non-identically distributed.)

Theorem 2. (i) *If* $\dfrac{2}{\sqrt{m_N(t)}}<\varepsilon<\sqrt{m_N(t)}$, *and* $m_N(t)\geqq 1$ *then*

$$P\left(\sup_{-\infty<u\leqq t}\left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right|>\varepsilon\right)\leqq 4N\exp\{-2^{-7}\varepsilon^2\}. \tag{3.8}$$

(ii) *If* $\dfrac{\log N}{m_N(t)}\to 0$, *then*

$$P\left(\sup_{-\infty<u\leqq t}\left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right|=O(\sqrt{\log N})\right)=1. \tag{3.9}$$

Proof. For fixed N and t let $-\infty=u_0<u_1<\ldots<u_{k(N)}=t$ be a partition of $(-\infty, t]$ such that

$$\delta_N(i)=m_N(u_{i-1})-m_N(u_i^-)\leqq 1\qquad i=1,2,\ldots,k(N),\ \text{and}\ k(N)\leqq N-1 \tag{3.10}$$

$$(m_N(u)\ \text{is right continuous, hence}\ m_N(u)=m_N(u^+)).$$

Since $m_N(u)$ is monotone decreasing and $m_N(-\infty)=N$, such partition always exists.

$$\begin{aligned}
P\left(\sup_{-\infty<u\leqq t}\left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right|>\varepsilon\right)&\leqq \sum_{i=1}^{k(N)} P\left(\sup_{u_{i-1}\leqq u<u_i}\left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right|>\varepsilon\right)\\
&\leqq \sum_{i=1}^{k(N)} P\left(\frac{\sup_{u_{i-1}\leqq u<u_i}|M_N(u)-m_N(u)|}{\sqrt{m_N(u_i^-)}}>\varepsilon\right)+P\left(\frac{|M_N(t)-m_N(t)|}{\sqrt{m_N(t)}}>\varepsilon\right)\\
&\leqq \sum_{i=1}^{k(N)}\left\{P\left(|M_N(u_{i-1})-m_N(u_{i-1})|>\frac{\varepsilon\sqrt{m_N(u_i^-)}-\delta_N(i)}{2}\right)\right.\\
&\quad\left.+P\left(|M_N(u_i^-)-m_N(u_i^-)|>\frac{\varepsilon\sqrt{m_N(u_i^-)}-\delta_N(i)}{2}\right)\right\}+P\left(\frac{|M_N(t)-m_N(t)|}{\sqrt{m_N(t)}}>\varepsilon\right).
\end{aligned} \tag{3.11}$$

In the last line of (3.11) we used the monotonicity of $M_N(u)$ and $m_N(u)$.

Let $\frac{2}{\sqrt{m_N(t)}}<\varepsilon<\sqrt{m_N(t)}$. Under condition (i) on ε

$$0\leqq\frac{\varepsilon\sqrt{m_N(u_i^-)}-\delta_N(i)}{2}\leqq\frac{m_N(u_i^-)}{2}\leqq m_N(u_{i-1}).$$

We estimate both terms of (3.11) by Lemma 3.1. Using (3.7) and condition (i)

$$\begin{aligned}&P\left(|M_N(u_{i-1})-m_N(u_{i-1})|>\frac{\varepsilon\sqrt{m_N(u_i^-)}-\delta_N(i)}{2}\right)\\&\quad+P\left(|M_N(u_i^-)-m_N(u_i^-)|>\frac{\varepsilon\sqrt{m_N(u_i^-)}-\delta_N(i)}{2}\right)\\&\quad\leqq 4\exp\left\{-\frac{(\varepsilon\sqrt{m_N(u_i^-)}-\delta_N(i))^2}{4\cdot 4\cdot m_N(u_{i-1})}\right\}\leqq 4\exp\left\{-\frac{\varepsilon^2 m_N(u_i^-)}{2^6 m_N(u_{i-1})}\right\}.\end{aligned}$$

From $m_N(t)\geqq 1$ follows that $2m_N(u_i^-)\geqq m_N(u_{i-1})$ hence $\frac{m_N(u_i^-)}{m_N(u_{i-1})}\geqq\frac{1}{2}$. Consequently

$$\begin{aligned}P\left(\sup_{-\infty<u\leqq t}\left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right|>\varepsilon\right)&\leqq k(N)\cdot 4\cdot\exp\{-2^{-7}\varepsilon^2\}+2\exp\left(-\frac{\varepsilon^2}{4}\right)\\&\leqq 4N\exp\{-2^{-7}\varepsilon^2\}\end{aligned}$$

which proves (i).

Leting $\varepsilon_N=\sqrt{2^9\log N}$, we obtain, that if $(\log N/m_N(t))\to 0$, then (3.9) holds by Borel-Cantelli. □

Theorem 3. *Suppose that for the point t, $m_N(t)\geqq 2$. Then for an arbitrary $\lambda\geqq 2$,*

$$P\left(\sup_{-\infty<u\leqq t}\frac{m_N(u)}{M_N(u)}>\lambda\right)\leqq N\exp\left\{-2^{-4}\left(1-\frac{1}{\lambda}\right)^2 m_N(t)\right\}. \tag{3.12}$$

Proof. For fixed N and t let $-\infty=u_0<u_1<\ldots<u_{k(N)}=t$ be the same partition of $(-\infty,t]$ as in Theorem 2. As both $m_N(u)$ and $M_N(u)$ are monotone decreasing

$$\begin{aligned}P\left(\sup_{-\infty<u\leqq t}\frac{m_N(u)}{M_N(u)}>\lambda\right)&\leqq\sum_{i=1}^{k(N)}P\left(\sup_{u_{i-1}\leqq u<u_i}\frac{m_N(u)}{M_N(u)}>\lambda\right)+P\left(\frac{m_N(t)}{M_N(t)}>\lambda\right)\\&\leqq\sum_{i=1}^{k(N)}P\left(\frac{m_N(u_{i-1})}{M_N(u_i^-)}>\lambda\right)+P\left(\frac{m_N(t)}{M_N(t)}>\lambda\right),\end{aligned}$$

$$\begin{aligned}P\left(\frac{m_N(u_{i-1})}{M_N(u_i^-)}>\lambda\right)&=P\left(\frac{m_N(u_{i-1})}{\lambda}>M_N(u_i^-)\right)\\&=P\left(\frac{m_N(u_{i-1})}{\lambda}-m_N(u_i^-)>M_N(u_i^-)-m_N(u_i^-)\right)\\&=P\left(m_N(u_i^-)-M_N(u_i^-)>m_N(u_i^-)-\frac{m_N(u_{i-1})}{\lambda}\right).\end{aligned}$$

By condition $m_N(t) \geqq 2$ and (3.10) we get that $2m_N(u_{i-1}) < 3m_N(u_i^-)$. Hence as $\lambda > 2$,

$$m_N(u_i^-) - \frac{m_N(u_{i-1})}{\lambda} \geqq m_N(u_i^-)\left(1 - \frac{3}{2\lambda}\right) \geqq \frac{m_N(u_i^-)(\lambda-1)}{2\lambda}.$$

Consequently

$$P\left(\frac{m_N(u_{i-1})}{M_N(u_i^-)} > \lambda\right) \leqq P\left(m_N(u_i^-) - M_N(u_i^-) > \frac{m_N(u_i^-)(\lambda-1)}{2\lambda}\right)$$

$$\leqq \exp\left\{-\frac{m_N^2(u_i^-)(\lambda-1)^2}{4\lambda^2 \cdot 4m_N(u_i^-)}\right\} = \exp\left\{-2^{-4}\left(1-\frac{1}{\lambda}\right)^2 m_N(u_i^-)\right\}$$

$$\leqq \exp\left\{-2^{-4}\left(1-\frac{1}{\lambda}\right)^2 m_N(t)\right\}$$

where we applied again Lemma 3.1 and (3.7). It is easy to see by a similar but somewhat simpler argument, that

$$P\left(\frac{m_N(t)}{M_N(t)} > \lambda\right) \leqq \exp\left\{-2^{-2}\left(1-\frac{1}{\lambda}\right)^2 m_N(t)\right\}.$$

Being $k(N) \leqq N-1$, (3.12) follows. □

This theorem is a generalization of Lemma 1 of Wellner [10], which deals with i.i.d. r.v.-s.

4. Lemmas

Suppose throughout the rest that F is continuous.

Lemma 4.1.

$$\sup_{-\infty < u \leqq t} |F_N^*(u) - F_N^0(u)| \leqq \int_{-\infty}^{t} \frac{1}{(M_N(u)+1)^2} dB_N(u). \tag{4.1}$$

Proof. The same as Lemma 2.2. in [3]. Let

$$R_N(u) = \int_{-\infty}^{u} \frac{1}{M_N(s)} dB_N(s), \quad \text{and} \quad R(u) = \int_{-\infty}^{u} \frac{1}{m_N(s)} db_N(s). \tag{4.2}$$

Observe that for $u < T_F$

$$R(u) = \int_{-\infty}^{u} \frac{1}{m_N(s)} db_N(s) = \int_{-\infty}^{u} \frac{\bar{G}(N, s^-)}{\bar{G}(N, s)\bar{F}(s)} dF(s) = -\log \bar{F}(u). \tag{4.3}$$

By Taylor expansion, as in [3] (formulae (2.1)-(2.3), (2.9)-(2.10)) we get that

$$|\bar{F}_N^*(u) - \bar{F}(u)| \leqq |\bar{F}_N^*(u) - \bar{F}^0(u)| + |\log \bar{F}_N^0(u) + R_N(u)| + \bar{F}(u)|R_N(u) - R(u)| + \tfrac{1}{2}\bar{F}(u)\exp|R_N(u) - R(u)| \cdot |R_N(u) - R(u)|^2. \tag{4.4}$$

Observe that

$$R_N(u)-R(u)=\int_{-\infty}^{u}\frac{1}{M_N(s)}dB_N(s)-\int_{-\infty}^{u}\frac{1}{m_N(s)}db_N(s)$$

$$=\int_{-\infty}^{u}\left(\frac{1}{M_N(s)}-\frac{1}{m_N(s)}\right)dB_N(s)+\int_{-\infty}^{u}\frac{1}{m_N(s)}d(B_N(s)-b_N(s)). \qquad (4.5)$$

Suppose that $\frac{\log N}{\bar{G}(N,T_F)}\to 0$ (hence T_F is finite), and consider the following sequence of points: $T_1, T_2, \ldots, T_N \ldots$ defined by the equation

$$\bar{F}(T_N)=\sqrt{\frac{\log N}{\bar{G}(N,T_F)}}. \qquad (4.6)$$

(This sequence is well-defined if $N \geqq N^*$ by the above condition.)

Lemma 4.2. *If T_N is defined by* (4.6) *and* (1.2) *holds then for almost all ω there exists an $N_0(\omega)$ such that if $N>N_0(\omega)$ then*

$$\frac{1}{M_N(u)}\leqq\frac{2}{m_N(u)} \quad \textit{for all} \quad u\leqq T_N. \qquad (4.7)$$

Proof. We apply Theorem 3 for the points T_N $(N<N^*)$.

By condition (1.2) we may choose an N_1 (independent from ω) such that for $N\geqq N_1$ $\bar{G}(N,T_F)>2^{16}\log N$. Then for $N\geqq N_1$

$$m_N(T_N)=\bar{G}(N,T_N)\,\bar{F}(T_N)=\sqrt{\frac{\log N}{\bar{G}(N,T_F)}}\,\bar{G}(N,T_N)\geqq\sqrt{\log N\,\bar{G}(N,T_F)}\geqq 2^8\log N$$

hence the condition of Theorem 3 is satisfied. Leting $\lambda=2$, the result follows by standard Borel-Cantelli argument. □

Lemma 4.3. *If T_N is defined by* (4.6) *and* (1.2) *holds then for almost all ω there exists an $N_0(\omega)$ such that, if $N\geqq N_0(\omega)$ then*

$$\sup_{-\infty<u\leqq T_N}\left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right|\leqq\sqrt{2^9\log N}. \qquad (4.8)$$

Proof. Apply Theorem 2 with $\varepsilon_N=\sqrt{2^9\log N}$. (As in Lemma 4.2 it's easy to check that the conditions of Theorem 2 holds if N is big enough.) □

Lemma 4.4. *Suppose that* (1.2) *holds, T_N is defined by* (4.6) *and let $1\leqq\alpha\leqq 2$ arbitrary.*

Then for almost all ω there exists an $N_0(\omega)$ such that, for $N>N_0(\omega)$,

$$\sup_{-\infty\leqq t\leqq T}\left|\int_{-\infty}^{t}\frac{1}{m_N^{\alpha}(u)}d(B_N(u)-b_N(u))\right|\leqq\frac{12\sqrt{\log N}}{(m_N(T))^{\frac{1}{2}(2\alpha-1)}} \qquad (4.9)$$

for any $T\leqq T_N$.

Proof. We prove the statement in two steps. At first we give an exponential bound for fix t and then estimate the sup in $(-\infty, T]$. First observe, that

$$\int_{-\infty}^{u} \frac{1}{m_N^\alpha(s)} dB_N(s) = \sum_{j=1}^{N} \frac{\beta_j(u)}{m_N^\alpha(Z_j)} = \sum_{j=1}^{N} \frac{\beta_j(u)}{\left(\sum_{k=1}^{N} \bar{H}_k(Z_j)\right)^\alpha}. \tag{4.10}$$

Moreover

$$\int_{-\infty}^{u} \frac{1}{m_N^\alpha(s)} db_N(s) = \sum_{j=1}^{N} \int_{-\infty}^{u} \frac{\bar{G}_j(s^-)}{\left(\sum_{k=1}^{N} \bar{H}_k(s)\right)^\alpha} dF(s) = \sum_{j=1}^{N} E\left(\frac{\beta_j(u)}{\left(\sum_{k=1}^{N} \bar{H}_k(Z_j)\right)^\alpha}\right).$$

Hence introducing the notations

$$\xi_j(u) = \frac{\beta_j(u)}{\left(\sum_{k=1}^{N} \bar{H}_k(Z_j)\right)^\alpha} \quad \text{and} \quad \xi_j^*(u) = \xi_j(u) - E(\xi_j(u)), \tag{4.11}$$

$$\int_{-\infty}^{u} \frac{1}{m_N^\alpha(s)} d(B_N(s) - b_N(s)) = \sum_{j=1}^{N} \xi_j^*(u) \tag{4.12}$$

where $\xi_j^*(u)$ $j=1, \dots, N$ are independent nonidentically distributed zero mean random variables. At first we estimate the probability $P\left(|\sum_1^N \xi_j^*(t)| > \varepsilon\right)$ by Lemma 3.1 and then we estimate $P\left(\sup_{-\infty < t \leq u} |\sum_1^N \xi_j^*(t)| > \varepsilon\right)$. Using the elementary inequality

$$e^x \leq 1 + x + \frac{x^2}{2} \quad \text{if} \quad |x| \leq \frac{1}{2},$$

$$E(e^{u\xi_j^*(t)}) \leq E(1 + u\xi_j^*(t) + u^2 \xi_j^{*2}(t)) = 1 + u^2 E(\xi_j^{*2}(t)) \leq e^{u^2 E(\xi_j^{*2}(t))}$$

if

$$|u\xi_j^*(t)| \leq \tfrac{1}{2}. \tag{4.13}$$

Observe that for $t \leq T$

$$0 \leq \xi_j(t) = \frac{\beta_j(t)}{\left(\sum_{k=1}^{N} \bar{H}_k(Z_j)\right)^\alpha} \leq \frac{\beta_j(T)}{\left(\sum_{k=1}^{N} \bar{H}_k(Z_j)\right)^\alpha}. \tag{4.14}$$

Moreover, if $Z_j \leq T$ then $\beta_j(T) \leq 1$ and $\bar{H}_k(Z_j) \geq \bar{H}_k(T)$. On the other hand if $Z_j > T$ then $\beta_j(T) = 0$. Consequently

$$0 \leq \xi_j(t) \leq \frac{1}{\left(\sum_{k=1}^{N} \bar{H}_k(T)\right)^\alpha} = \frac{1}{(m_N(T))^\alpha} \quad \text{for any} \quad t \leq T.$$

Hence (4.13) valid if $0 \leq u \leq \frac{(m_N(T))^\alpha}{2}$.

For any $t \leqq T$ we have

$$\sum_{j=1}^{N} E(\xi_j^{*2}(t)) \leqq \sum_{j=1}^{N} E(\xi_j^2(t))$$

$$= \sum_{j=1}^{N} E\left(\frac{\beta_j^2(t)}{\left(\sum_{k=1}^{N} \bar{H}_k(Z_j)\right)^{2\alpha}}\right) = \sum_{j=1}^{N} E\left(\frac{\beta_j(t)}{\left(\sum_{k=1}^{N} \bar{H}_k(Z_j)\right)^{2\alpha}}\right)$$

$$\leqq \frac{1}{\left(\sum_{k=1}^{N} \bar{G}_k(T)\right)^{2\alpha-1}} \int_{-\infty}^{t} \frac{\sum_{j=1}^{N} \bar{G}_j(s^-)}{\bar{F}^{2\alpha}(s)\left(\sum_{k=1}^{N} \bar{G}_k(s)\right)} dF(s)$$

$$\leqq \frac{1}{(\bar{G}(N,T))^{2\alpha-1}} \int_{-\infty}^{t} \frac{\bar{G}(N,s^-)}{\bar{F}^{2\alpha}(s)\,\bar{G}(N,s)} dF(s)$$

$$= \frac{1}{(2\alpha-1)(\bar{G}(N,T))^{2\alpha-1}} \left(\frac{1}{(\bar{F}(t))^{2\alpha-1}} - 1\right)$$

$$\leqq \frac{1}{(\bar{G}(N,T)\,\bar{F}(T))^{2\alpha-1}} = \frac{1}{(m_N(T))^{2\alpha-1}}.$$

Hence using the notations of Lemma 3.1, with

$$U = \frac{(m_N(T))^\alpha}{2}, \qquad \Lambda = \sum_{j=1}^{N} \lambda_j = \frac{2}{(m_N(T))^{2\alpha-1}}, \qquad U\Lambda = 1$$

we have for any $0 \leqq \varepsilon \leqq 1$ and any $t \leqq T$

$$P\left(\left|\sum_{j=1}^{N} \xi_j^*(t)\right| > \varepsilon\right) \leqq 2\exp\left\{-\frac{\varepsilon^2 (m_N(T))^{2\alpha-1}}{4}\right\}.$$

To estimate the supremum in $(-\infty, T)$ observe that

$$\eta_N(t) = \sum_{j=1}^{N} \xi_j(t) = \int_{-\infty}^{t} \frac{1}{m_N^\alpha(u)} dB_N(u), \qquad l_N(t) = \sum_{j=1}^{N} E(\xi_j(t)) = \int_{-\infty}^{t} \frac{1}{m_N^\alpha(u)} db_N(u)$$

are both monotone nondecreasing functions of t. Suppose that $m_N(T) > 1$ then $l_N(t) \leqq |\log \bar{F}(t)|$. $\left(\text{As by } m_N(T) \geqq 1,\ 1 \leqq \alpha \leqq 2,\ l_N(t) = \int_{-\infty}^{t} \frac{1}{m_N^\alpha(t)}\, db_N(t) \leqq \int_{-\infty}^{t} \frac{1}{m_N(t)}\, db_N(t) = |\log \bar{F}(t)|.\right)$ For a fix $0 < \varepsilon < 1$ consider a partition of the interval $(-\infty, T)$ $\;-\infty = u_0 < u_1 \ldots < u_{L(\varepsilon)} = T$ such that

$$l_N(u_i) - l_N(u_{i-1}) < \frac{\varepsilon}{3} \quad i = 1, 2, \ldots, L(\varepsilon) \quad \text{and} \quad L(\varepsilon) \leqq \frac{3|\log \bar{F}(T)|}{\varepsilon} + 1. \tag{4.14}$$

Since $l_N(t)$ is continuous such a partition easily can be constructed. If

$$|\eta_N(u_{i-1})-l_N(u_{i-1})|\leqq\frac{\varepsilon}{3} \quad\text{and}\quad |\eta_N(u_i^-)-l_N(u_i)|\leqq\frac{\varepsilon}{3},$$

then by the monotonicity of $\eta_N(t)$ and $l_N(t)$ and (4.14) for any $u_i\leqq t<u_{i+1}$,

$$|\eta_N(t)-l_N(t)|\leqq\frac{\varepsilon}{3}+2\frac{\varepsilon}{3}=\varepsilon.$$

Consequently, if $\sup_{-\infty<t\leqq T}|\eta_N(t)-(-\log\bar{F}(t))|>\varepsilon$ then for some $0\leqq i\leqq L(\varepsilon)$

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

Thus we have[1] that if $m_N(T)\geqq 1$

$$\begin{aligned} P\left(\sup_{-\infty<t\leqq T}\left|\int_{-\infty}^{t}\frac{1}{m_N^\alpha(s)}dB_N(s)-\int_{-\infty}^{t}\frac{1}{m_N^\alpha(s)}db_N(s)\right|>\varepsilon\right) &\leqq 2\cdot 2L(\varepsilon)\exp\left\{-\frac{\varepsilon^2 m_N^{2\alpha-1}(T)}{4\cdot 3^2}\right\}\\ &\leqq 4\left(\frac{3|\log\bar{F}(T)|}{\varepsilon}+1\right)\exp\left\{-\frac{\varepsilon^2 m_N^{2\alpha-1}(T)}{36}\right\}. \end{aligned} \tag{4.15}$$

Consider now the sequence T_N defined by (4.6). Observe that

$$m_N(T_N)=\bar{G}(N,T_N)\,\bar{F}(T_N)=\bar{G}(N,T_N)\sqrt{\frac{\log N}{\bar{G}(N,T_F)}}\geqq\sqrt{\log N\,\bar{G}(N,T_F)}>1 \tag{4.16}$$

if $N\geqq N_1(\geqq N^*)$ by (1.2).

Consequently for any $T\leqq T_N$, $m_N(T)>1$, if $N\geqq N_1$. Thus (4.15) is valid for any $T\leqq T_N$, if $N\geqq N_1$.

Let $\varepsilon_N=\dfrac{\sqrt{4\cdot 36\log N}}{(m_N(T))^{\frac{1}{2}(2\alpha-1)}}$. Then for any $T\leqq T_N$

$$4\left(\frac{3|\log\bar{F}(T)|}{\varepsilon_N}+1\right)\leqq 4\left(\frac{3|\log\bar{F}(T_N)|}{\varepsilon_N}+1\right)\leqq 4\left(\frac{3\log N}{\sqrt{4\cdot 36\log N}}N^{\frac{2\alpha-1}{2}}+1\right)\leqq N^2$$

if $N\geqq N_2(\geqq N_1)$ (as $m_N(t)\leqq N$ for any t, $\bar{G}(N,T)\leqq N$, for any T, $\alpha\leqq 2$, and by the definition of T_N $|\log\bar{F}(T_N)|\leqq\log N$, if N is big enough). Consequently for any $T\leqq T_N$ we have

$$\begin{aligned} &\sum_{N\geqq N_2}^{\infty}P\left(\sup_{-\infty<t\leqq T}\left|\int_{-\infty}^{t}\frac{1}{m_N^\alpha(s)}d(B_N(s)-b_N(s))\right|>\frac{12\sqrt{\log N}}{(m_N(T))^{\frac{1}{2}(2\alpha-1)}}\right)\\ &\quad\leqq\sum_{N\geqq N_2}^{\infty}N^2\exp\{-4\log N\}<+\infty \end{aligned}$$

which proves our statement. □

[1] A similar but weaker inequality is proved in [4]

Lemma 4.5. *If T_N is defined by* (4.6), $1<\alpha\leqq 2$ *arbitrary, and* (1.2) *holds, then for almost all ω there exists an $N_0^*(\omega)$ such that for $N\geqq N_0^*(\omega)$*

$$\int_{-\infty}^{T}\frac{1}{m_N^\alpha(u)}\,dB_N(u)\leqq\frac{2}{(\alpha-1)\,m_N^{\alpha-1}(T)}\qquad \textit{for any } T\leqq T_N.$$

Proof.

$$\int_{-\infty}^{T}\frac{1}{m_N^\alpha(u)}\,dB_N(u)=\int_{-\infty}^{T}\frac{1}{m_N^\alpha(u)}\,d(B_N(u)-b_N(u))+\int_{-\infty}^{T}\frac{1}{m_N^\alpha(u)}\,db_N(u). \tag{4.17}$$

The first term of (4.17) can be estimated by Lemma 4.4. On the other hand

$$\begin{aligned}\int_{-\infty}^{T}\frac{1}{m_N^\alpha(u)}\,db_N(u)&=\int_{-\infty}^{T}\frac{\bar G(N,u^-)}{(\bar F(u)\,\bar G(N,u))^\alpha}\,dF(u)\\&\leqq\frac{1}{(\bar G(N,T))^{\alpha-1}}\int_{-\infty}^{T}\frac{1}{\bar F^\alpha(u)}\,dF(u)\\&\leqq\frac{1}{(\bar G(N,T))^{\alpha-1}(\alpha-1)\,\bar F^{\alpha-1}(T)}=\frac{1}{(\alpha-1)\,m_N^{\alpha-1}(T)}.\end{aligned}\tag{4.18}$$

From (4.17), (4.18) and Lemma 4.4 for almost all ω there exists an $N_0(\omega)$ ($\geqq N^*$) such that for $N\geqq N_0(\omega)$, for any $T\leqq T_N$

$$\begin{aligned}\int_{-\infty}^{T}\frac{1}{m_N^\alpha(u)}\,dB_N(u)&\leqq\frac{12\sqrt{\log N}}{m_N(T)^{\frac12(2\alpha-1)}}+\frac{1}{(\alpha-1)\,m_N^{\alpha-1}(T)}\\&\leqq\frac{1}{(\alpha-1)\,m_N^{\alpha-1}(T)}\left(1+\frac{12\sqrt{\log N}}{m_N(T)^{\frac12}}\right).\end{aligned}$$

By condition (1.2) there exists an N_0^* ($\geqq N_0$) such that if $N\geqq N_0^*$ then $\dfrac{12\sqrt{\log N}}{m_N(T)^{\frac12}}\leqq 1$, which proves the Lemma. □

Proof of Theorem 1. First observe that

$$\begin{aligned}\sup_{-\infty<u<+\infty}|F_N^*(u)-F(u)|&\leqq\sup_{-\infty<u\leqq T_N}|F_N^*(u)-F(u)|+\sup_{T_N<u<+\infty}|F_N^*(u)-F(u)|\\&\leqq\sup_{-\infty<u\leqq T_N}|F_N^*(u)-F(u)|+\sup_{T_N\leqq u<+\infty}|\bar F_N^*(u)-\bar F(u)|\leqq\sup_{-\infty<u\leqq T_N}|F_N^*(u)-F(u)|\\&\quad+|\bar F_N^*(T_N)-\bar F(T_N)|+\bar F(T_N)\leqq 2\sup_{-\infty<u\leqq T_N}|F_N^*(u)-F(u)|+\bar F(T_N)\end{aligned}$$

as both $\bar F_N^*$ and $\bar F$ are monotone nonincreasing. By the definition of T_N it's enough to consider

$$\sup_{-\infty<u\leqq T_N}|F_N^*(u)-F(u)|.$$

Using Lemma 4.1 and applying the same argument for (4.4) which was used in [3] (Lemma 2.2, (2.7) in Lemma 2.3, and Lemma 2.5) we get that under conditions (1.2) if,

$$\sup_{-\infty<u\leqq T_N}|R_N(u)-R(u)|\leqq\tfrac{2}{3}\quad\text{a.s.,}\tag{5.1}$$

then

$$\sup_{-\infty<u\leqq T_N} |F_N^*(u)-F(u)| \leqq 4 \int_{-\infty}^{T_N} \frac{1}{M_N^2(s)}\, dB_N(s)$$

$$+2 \sup_{-\infty<u\leqq T_N} \bar{F}(u)\,|R_N(u)-R(u)| \quad \text{a.s.}. \tag{5.2}$$

From Lemma 4.2 and Lemma 4.5 follows that

$$\int_{-\infty}^{T_N} \frac{1}{M_N^2(s)}\, dB_N(s) \leqq 2^2 \int_{-\infty}^{T_N} \frac{1}{m_N^2(s)}\, dB_N(s) = O\left(\frac{1}{m_N(T_N)}\right) \leqq O\left(\frac{1}{\sqrt{\log N\, \bar{G}(N, T_F)}}\right) \quad \text{a.s.} \tag{5.3}$$

as by (4.16)

$$m_N(T_N) \geqq \sqrt{\log N\, \bar{G}(N, T_F)}, \qquad \text{if } N > N_1 (\geqq N^*).$$

Considering the difference $|R_N(u)-R(u)|$ apply Lemma 4.2, 4.3 and 4.5 with $\alpha = \frac{3}{2}$ for the first term of (4.5) and for the second term apply Lemma 4.4 with $\alpha = 1$. Then for any $u \leqq T_N$

$$|R_N(u)-R(u)| \leqq \int_{-\infty}^{u} \frac{2\,|M_N(s)-m_N(s))|}{m_N^2(s)}\, dB_N(s) + \sup_{-\infty<t\leqq u} \left| \int_{-\infty}^{t} \frac{1}{m_N(s)}\, d(B_N(s)-b_N(s)) \right|$$

$$\leqq 2\sqrt{2^9 \log N} \int_{-\infty}^{u} \frac{1}{m_N^{3/2}(s)}\, dB_N(s) + \sup_{-\infty<t\leqq u} \left| \int_{-\infty}^{t} \frac{1}{m_N(s)}\, d(B_N(s)-b_N(s)) \right|$$

$$\leqq 2^3 \sqrt{2^9 \log N}\, \frac{1}{\sqrt{m_N(u)}} + \frac{12\sqrt{\log N}}{\sqrt{m_N(u)}} = O\left(\sqrt{\frac{\log N}{m_N(u)}}\right) \quad \text{a.s.}. \tag{5.4}$$

Hence by (4.16)

$$\sup_{-\infty<u\leqq T_N} |R_N(u)-R(u)| = O\left(\sqrt{\frac{\log N}{m_N(T_N)}}\right) \leqq O\left(\left(\frac{\log N}{\bar{G}(N, T_F)}\right)^{\frac{1}{4}}\right) \quad \text{a.s.}. \tag{5.5}$$

Hence (5.1) holds if $N \geqq N_1$.

From (5.2), (5.3), (5.5)

$$\sup_{-\infty<u\leqq T_N} |F_N^*(u)-F(u)| \leqq O\left(\frac{1}{\sqrt{\log N\, \bar{G}(N, T_F)}}\right) + \sup_{-\infty<u\leqq T_N} \bar{F}(u)\, O\left(\sqrt{\frac{\log N}{m_N(u)}}\right)$$

$$\leqq O\left(\frac{1}{\sqrt{\log N\, \bar{G}(N, T_F)}}\right) + \sup_{-\infty<u\leqq T_N} \bar{F}(u)\, O\left(\sqrt{\frac{\log N}{\bar{F}(u)\, \bar{G}(N, u)}}\right) \leqq O\left(\frac{1}{\sqrt{\log N\, \bar{G}(N, T_F)}}\right)$$

$$+ \sup_{-\infty<u\leqq T_N} \sqrt{\bar{F}(u)}\, O\left(\sqrt{\frac{\log N}{\bar{G}(N, u)}}\right) \leqq O\left(\frac{1}{\sqrt{\log N\, \bar{G}(N, T_F)}}\right) + O\left(\sqrt{\frac{\log N}{\bar{G}(N, T_F)}}\right)$$

$$= O\left(\sqrt{\frac{\log N}{\bar{G}(N, T_F)}}\right) \quad \text{a.s.}$$

and the theorem follows. □

Remark 1. From our proof it is clear that we may give a concrete bound instead of using the O symbol. But this bound would be very crude.

Remark 2. Corollary 1 easily follows from Theorem 1. For this it's enough to observe that all of the lemmas and statements are valid for $(-\infty, t]$ using conditions of the corollary instead of the conditions of Theorem 1.

Remark 3. Corollary 2 covers the i.i.d. censoring case ($\alpha = 1$) and gives slightly weaker result then (1.1).

References

1. Földes, A., Rejtö, L.: Strong uniform consistency for nonparametric survival estimators from randomly censored data. Ann. Statist. **9**, 122–129 (1981)
2. Földes, A., Rejtö, L.: Asymptotic properties of the nonparametric survival curve estimators under variable censoring. The First Pannonian Symposium on Mathematical Statistics. Lecture Notes in Statistic **8**, 55–71. Berlin-Heidelberg-New York: Springer 1981
3. Földes, A., Rejtö, L.: A LIL type result for the product limit estimator on the whole line. Z. Wahrscheinlichkeitstheorie verw. Gebiete **56**, 75–86 (1981)
4. Földes, A., Rejtö, L.: Asymptotic properties of the nonparametric survival curve estimators under variable censoring. Unpublished Technical Report of the Math. Inst. of the Hungarian Academy of Sciences No. 38/1979, Budapest
5. Gill, R.D.: Censoring and stochastic integrals. Mathematisch Centrum Amsterdam, 1979
6. Kaplan, E.L., Meier, P.: Nonparametric estimation from incomplete observation. J. Amer. Statist. Assoc. **58** U, 457–481 (1963)
7. Petrov, V.V.: Sums of independent random variables. Berlin-Heidelberg-New York: Springer 1975
8. Singh, R.S.: On the Glivenko-Cantelli theorem for weighted empiricals based on independent random variables. Ann. Probability **3**, 371–374 (1975)
9. Rényi, A.: Probability theory. Amsterdam: North-Holland 1970
10. Wellner, Jon A.: Limit theorems for the ratio of the empirical distribution functions to the true distribution function. Z. Wahrscheinlichkeitstheorie verw. Gebiete **45**, 73–88 (1978)

Received June 13, 1980; in revised form May 2, 1981