Strong Uniform Consistency of the Product Limit Estimator under Variable Censoring

by

Antonia Földes*

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #80-7

April 1980

(Revised June 1980)

*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.

Strong Uniform Consistency of the Product Limit Estimator under Variable Censoring

by

Antonia Földes*

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 distribution functions $G_1(x),\ldots,G_N(x)\ldots$ We suppose that $\{X_i\}$ and $\{Y_i\}$ are mutually independent.

Let

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

/[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 [5] 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

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

The case of variable censoring (i.e. Y_i's have different distributions) was discussed in [2] where the following result was proved:

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

Denote

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

$$\bar{H}_k(T) = \bar{F}(T)\bar{G}_k(T) \text{ and } \sigma_N(T) = \sum_{k=1}^{N} \bar{H}_k(T) \quad (1-\bar{H}_k(T)).$$

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

---

*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.

Theorem [2]. Suppose that

(i) The distribution functions F, $G_1,\dots,G_N\dots$ of X, $Y_1,\dots,Y_N\dots$ are continuous on $(-\infty,T]$.

(ii)

$$\frac{N^{3/4}(\log N)^{1/4}}{m_N(T)} = o(1). \tag{1.12}$$

(iii) $\{\alpha_N\}$ is a sequence of nonnegative numbers for which $0 \le \alpha_N \le \sigma_N$ and $\sum_{N=1}^{\infty} \exp\{-\frac{2}{g}\alpha_N^2\} < +\infty$.

Then

$$P\left(\sup_{-\infty\le u\le T} |F_N(u)-F(u)| = O\left(\frac{N^{3/2}\sqrt{\log N}}{m_N(m_N-\alpha_N\sigma_N)}\right)\right) = 1. \tag{1.13}$$

Recently Gill [4] proved that $\sup_{-\infty<u\le t} |F_N^*(u) - F(u)| \underset{P}{\to} 0$ if $M_N(t) \underset{P}{\to} +\infty$ and $F(t^-) < 1$. /where $\underset{P}{\to}$ denotes stochastic convergence/

We shall prove the following

Theorem 1 Suppose that

(i) F is continuous

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

Then

$$P\left(\sup_{-\infty<u<+\infty} |F_N^*(u)-F(u)| = O\left(\sqrt{\frac{\log N}{\bar{G}(N,T_F)}}\right)\right) = 1. \tag{1.4}$$

Corollary 1 Suppose that

(i) F(t) is continuous in $(-\infty,t]$

(ii) $\frac{\log N}{\bar{G}(N,t)} \to 0$.

Then

$$(1.5) \qquad P\Big(\sup_{-\infty<u\leq t} |F_N^*(u)-F(u)| = O\Big(\sqrt{\frac{\log N}{\bar{G}(N,t)}}\Big)\Big) = 1.$$

Our technic is somewhat 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 nonidentically distributed random variables, which seems to be new.

2. Definitions, notations.

In what follows we list all the necessary notations. /For the readers convenience we repeat the earlier given ones too./

(2.1) $\{X_i\}_{i=1}^{\infty}$ i.i.d. r.v.'s with $P(X_i \leq t) = F(t)$ for $i = 1,2,\ldots$, $F(t)$ is continuous. $\bar{F}(t) = 1-F(t)$.

(2.2) $\{Y_i\}_{i=1}^{\infty}$ independent sequence of r.v.'s $P(Y_i \leq t) = G_i(t)$ $i = 1,2,\ldots$ The sequence $\{X_i\}_1^{\infty}$ and $\{Y_i\}_1^{\infty}$ are mutually independent.

(2.3) $Z_i = \min\{X_i,Y_i\}$, $P(Z_i \leq t) = H_i(t)$ $i = 1,2,\ldots$

(2.4) $\bar{H}_i(t) = \bar{F}(t)\bar{G}_i(t)$ $i = 1,2,\ldots$

(2.5) $\delta_i = [X_i \leq Y_i]$ $i = 1,2,\ldots$

(2.6) $T_F = \sup\{t;\ F(t) < 1\}$, $T_{G_i} = \sup\{t;\ G_i(t) < 1\}$ $i = 1,2,\ldots$

(2.7) $T_{H_i} = \sup\{t; H_i(t) < 1\}$ $i = 1,2,\ldots$

(2.8) $M_N(t) = \sum_{k=1}^{N} [Z_k > t]$

(2.9) $m_N(t) = E(M_N(t)) = \sum_{k=1}^{N} \bar{H}_k(t)$

(2.10) $\beta_i(t) = [Z_i \leq t,\ \delta_i = 1] \qquad i = 1,2,\ldots$

(2.11) $B_N(t) = \sum_{i=1}^{N} \beta_i(t)$

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

(2.13) $\tau_N(\omega) = \max_{j\leq N}\{Z_j(\omega)\}$

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

The definition of the product limit estimator $F_N^*(t)$ is the following;

Definition 2.1.

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

In what follows (as in [2]) we will use the modified product limit estimator $F_N^o(t)$.

Definition 2.2.

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

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

Our basic tool is the following exponential bound (see Petrov [6] page 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

(3.1) $E(e^{u\xi_k}) \leq e^{(1/2)\lambda_k u^2}$ $k = 1,2,\dots,N$ for $0 \leq u \leq U$.

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

(3.3) $P(S_N > x) \leq \exp\{-\frac{x^2}{2\Lambda}\}$ if $0 \leq x \leq \Lambda U$

(3.4) $P(S_N > x) \leq \exp\{-\frac{Ux}{2}\}$ if $x \geq \Lambda U$

(3.5) $P(|S_N| > x) \leq 2\exp\{-\frac{x^2}{2\Lambda}\}$ if $0 \leq x \leq \Lambda U$

(3.6) $P(|S_N| > x) \leq 2\exp\{-\frac{Ux}{2}\}$ if $x > \Lambda U$.

For further application we need the following lemma.

Lemma 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 $A_N = \sum_{i=1}^{N} a_i$. For any $\tilde{A}_N \geq A_N$

(3.7) $P(|\sum_{i=1}^{N}\alpha_i - A_N| \geq \sqrt{\max(\tilde{A}_N,\ 2^k\log N)2^k\log N}) \leq 2\exp\{-2^{k-2}\log N\}$.

Proof: The proof is based on Lemma 3.1. Denote $\xi_i = \alpha_i - a_i$, $E(\xi_i) = 0$ $i = 1,2,\dots$

(3.8) $E(e^{u\xi_i}) \leq E(1+u\xi_i+u^2\xi_i^2) = 1+u^2E(\xi_i^2) \leq \exp(u^2E(\xi_i^2))$

if only

(3.9) $$|u\xi_i| = |u(\alpha_i - a_i)| \leq \frac{1}{2}.$$

As $|\alpha_i - a_i| \leq 1$ and $E(\xi_i^2) = a_i(1-a_i) \leq a_i$

(3.10) $E(e^{u\xi_i}) \leq \exp\{u^2 a_i\}$ if $0 \leq u \leq \frac{1}{2}$.

Using the notations of Lemma 3.1 we got

(3.11) $U = \frac{1}{2}, \qquad \lambda_i \geq 2a_i \quad i = 1,2,\ldots \qquad \Lambda \geq 2A_N.$

To prove (3.7) for an arbitrary $\tilde{A}_N \geq A_N$, let $\Lambda = 2\tilde{A}_N$, $U\Lambda = \tilde{A}_N$. Apply Lemma (3.1) for $x = \sqrt{\max\{\tilde{A}_N, 2^k \log N\} 2^k \log N}$. First suppose that $\tilde{A}_N \leq 2^k \log N$, then $x = 2^k \log N$, that is $x \geq U\Lambda = \tilde{A}_N$, therefore from (3.6)

$$P\left(\left|\sum_{i=1}^{N} \alpha_i - A_N\right| > x\right) \leq 2 \exp\left\{-\frac{2^k \log N}{2 \cdot 2}\right\}. \tag{3.12}$$

On the other hand if $\tilde{A}_N \geq 2^k \log N$ then $x = \sqrt{\tilde{A}_N 2^k \log N}$, that is $x \leq \tilde{A}_N$ hence from (3.5)

$$P\left(\left|\sum_{i=1}^{N} \alpha_i - A_N\right| > x\right) \leq 2 \exp\left\{-\frac{\tilde{A}_N 2^k \log N}{2 \cdot 2\tilde{A}_N}\right\} = 2 \exp\{-2^{k-2} \log N\} \tag{3.13}$$

(3.7) now follows from (3.12) and (3.13). □

In Lemmas 3.3, 3.4, and 3.5 we consider a sequence of independent random variables $Z_1, \ldots, Z_N, \ldots$ Let

$$M_N(u) = \sum_{i=1}^{N} [Z_i > u], \qquad m_N(u) = \sum_{i=1}^{N} P(Z_i > u)$$

$$D_N(u) = \sum_{i=1}^{N} [Z_i \leq u] \qquad d_N(u) = \sum_{i=1}^{N} P(Z_i \leq u).$$

We prove some uniform properties of $M_N(u)$ and $D_N(u)$. /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 3 lemmas we do not suppose anything about Z_i./

Lemma 3.3.

(i) If $\frac{2}{\sqrt{m_N(t)}} < \varepsilon < \sqrt{m_N(t)}$, and $m_N(t) \geq 1$ then

$$P\left(\sup_{-\infty < u \leq t} \left|\frac{M_N(u) - m_N(u)}{\sqrt{m_N(u)}}\right| > \varepsilon\right) \leq 4N \exp\{-2^{-7}\varepsilon^2\}. \tag{3.14}$$

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

$$(3.15) \quad P\left(\sup_{-\infty<u\le t} \left| \frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}} \right| = O(\sqrt{\log N})\right) = 1.$$

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

$$(3.16) \quad \delta_N(i) = m_N(u_{i-1})-m_N(u_i^-) \le 1 \qquad i = 1,2,\ldots,k(N), \text{ and } k(N) \le N-1.$$

/Both $M_N(u)$ and $m_N(u)$ are 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.

$$P\left(\sup_{-\infty<u\le t} \left| \frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}} \right| > \varepsilon\right) \le \sum_{i=1}^{k(N)} P\left(\sup_{u_{i-1}\le u<u_i} \left| \frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}} \right| > \varepsilon\right) \le$$

$$\le \sum_{i=1}^{k(N)} P\left(\frac{\sup_{u_{i-1}\le u<u_i} |M_N(u)-m_N(u)|}{\sqrt{m_N(u_i^-)}} > \varepsilon\right) + P\left(\frac{|M_N(t)-m_N(t)|}{\sqrt{m_N(t)}} > \varepsilon\right) \le$$

(3.17)

$$\le \sum_{i=1}^{k(N)} \{P(|M_N(u_{i-1})-m_N(u_{i-1})| > \frac{\varepsilon\sqrt{m_N(u_i^-)} - \delta_N(u)}{2}) +$$

$$P(|M_N(u_i^-)-m_N(u_i^-)| > \frac{\varepsilon\sqrt{m_N(u_i^-)} - \delta_N(i)}{2})\} +$$

$$+ P\left(\frac{|M_N(t)-m_N(t)|}{\sqrt{m_N(t)}} > \varepsilon\right).$$

In the last line of (3.17) we used the monotonicity of $M_N(u)$ and $m_N(u)$.

Let

$$(3.18) \qquad \frac{2}{\sqrt{m_N(t)}} < \varepsilon < \sqrt{m_N(t)}.$$

Then

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

We estimate both terms of (3.17) by Lemma 3.1. Using (3.11) and (3.18)

$$P(|M_N(u_{i-1})-m_N(u_{i-1})| > \frac{\varepsilon\sqrt{m_N(u_i^-)} - \delta_N(i)}{2}) +$$

$$(3.19) \quad + P(|M_N(u_i^-)-m_N(u_i^-)| > \frac{\varepsilon\sqrt{m_N(u_i^-)} - \delta_N(i)}{2}) \le$$

$$4 \exp\{- \frac{(\varepsilon\sqrt{m_N(u_i^-)} - \delta_N(i))^2}{4\cdot 4\cdot m_N(u_{i-1})}\} \le 4\exp\{- \frac{\varepsilon^2 m_N(u_i^-)}{2^6 m_N(u_{i-1})}\}.$$

From $m_N(t) \ge 1$ follows that

$$2m_N(u_i^-) \ge m_N(u_{i-1}) \quad \text{hence} \quad \frac{m_N(u_i^-)}{m_N(u_{i-1})} \ge \frac{1}{2} .$$

Consequently

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

$$+ 2 \exp(- \frac{\varepsilon^2}{4}) \le \quad 4N \exp\{-2^{-7}\varepsilon^2\}$$

which proves (i).

Let $\varepsilon_N = \sqrt{2^9 \log N}$. Then by condition $\frac{\log N}{m_N(t)} \to 0$ for $n \ge N_0$ $\frac{2}{\sqrt{m_N(t)}} < \varepsilon_N < \sqrt{m_N(t)}$.

Therefore

$$\sum_{N=N_0}^{\infty} P(\sup_{-\infty<u\le t}\left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right| > \sqrt{2^9 \log N}) \le$$

$$\le \sum_{N=N_0}^{\infty} (4N+2)\exp\{-4 \log N\} < + \infty.$$

Hence (ii) follows by Borel-Cantelli. □

Lemma 3.4.

(i) Suppose that for the point t, $m_N(t) \geq 2$. Then for an arbitrary $\lambda \geq 2$,

$$(3.20)\qquad P\Big(\sup_{-\infty<u\leq t} \frac{m_N(u)}{M_N(u)} > \lambda\Big) \leq N \exp\{-2^{-4}\lambda^{-2} m_N(t)\}.$$

(ii) If $\frac{\log N}{m_N(t)} \to 0$ then for almost all ω

$$(3.21)\qquad \frac{1}{M_N(u)} \leq \frac{2}{m_N(u)} \quad \text{for all } u \leq t, \text{ if } N \geq N_0(\omega), \text{ that is;}$$

$$(3.22)\qquad P\Big(\sup_{-\infty<u\leq t} \frac{m_N(u)}{M_N(u)} = O(1)\Big) = 1.$$

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 Lemma 3.2. As both $m_N(u)$ and $M_N(u)$ are monotone decreasing

$$(3.23)\qquad P\Big(\sup_{-\infty<u\leq t} \frac{m_N(u)}{M_N(u)} > \lambda\Big) \leq \sum_{i=1}^{k(N)} P\Big(\sup_{u_{i-1}\leq u<u_i} \frac{m_N(u)}{M_N(u)} > \lambda\Big) + P\Big(\frac{m_N(t)}{M_N(t)} > \lambda\Big) \leq$$
$$\leq \sum_{i=1}^{k(N)} P\Big(\frac{m_N(u_{i-1})}{M_N(u_i^-)} > \lambda\Big) + P\Big(\frac{m_N(t)}{M_N(t)} > \lambda\Big).$$

$$(3.24)\qquad P\Big(\frac{m_N(u_{i-1})}{M_N(u_i^-)} > \lambda\Big) = P\Big(\frac{m_N(u_{i-1})}{\lambda} > M_N(u_i^-)\Big) =$$
$$= P\Big(\frac{m_N(u_{i-1})}{\lambda} - m_N(u_i^-) > M_N(u_i^-) - m_N(u_i^-)\Big) =$$
$$= P\Big(m_N(u_i^-) - M_N(u_i^-) > m_N(u_i^-) - \frac{m_N(u_{i-1})}{\lambda}\Big).$$

By condition $m_N(t) \geq 2$ and (3.16) we get that

$$(3.25)\qquad 2m_N(u_{i-1}) < 3m_N(u_i^-).$$

Hence as $\lambda > 2$,

$$(3.26)\qquad m_N(u_i^-) - \frac{m_N(u_{i-1})}{\lambda} \geq m_N(u_i^-)\Big(1 - \frac{3}{2\lambda}\Big) \geq \frac{m_N(u_i^-)}{2\lambda}.$$

From (3.25) and (3.26)

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

$$\le \exp\left\{-\frac{m_N^2(u_i^-)}{4\lambda^2 \cdot 4m_N(u_i^-)}\right\} = \exp\{-2^{-4}\lambda^{-2}m_N(u_i^-)\} \le \tag{3.27}$$

$$\le \exp\{-2^{-4}\lambda^{-2}m_N(t)\}$$

where we applied again Lemma 3.1 and (3.11). It is easy to see by a similar but somewhat simpler argument, that

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

Hence

$$P\left(\sup_{-\infty<u\le t} \frac{m_N(u)}{M_N(u)} > \lambda\right) \le N \exp\{-2^{-4}\lambda^{-2}m_N(t)\} \tag{3.28}$$

which proves (i).

If $\frac{\log N}{m_N(t)} \to 0$, then for $N \ge N_1$ $m_N(t) > 2^8 \log N$, thus

$$\sum_{N\ge N_1}^{\infty} P\left(\sup_{-\infty<u\le t} \frac{m_N(u)}{M_N(u)} > 2\right) \le \sum_{N=N_0}^{\infty} N \exp\{-4 \log N\} < +\infty \tag{3.29}$$

that is, for almost every ω there exists an $N_0(\omega)$ such that for $N \ge N_0(\omega)$

$$\frac{1}{M_N(u)} \le \frac{2}{m_N(u)} \quad \text{for each } u \le t. \tag{3.30}$$

which proves (ii). □

The next lemma will not be used in this paper, we just give it for completeness.

Lemma 3.5. (i) If for an arbitrary $t(\leq +\infty)$ $2 \leq \varepsilon \leq 4d_N(t)$ then

$$(3.31)\quad P(\sup_{-\infty<u\leq t} |D_N(u)-d_N(u)| > \varepsilon) \leq (2N+1)\exp\{-\frac{\varepsilon^2}{2^6 d_N(t)}\}$$

(ii) Let $d_N^*(t) = \max\{d_N(t), 2^4 \log N\}$. Then

$$(3.32)\quad P(\sup_{-\infty<u\leq t} |D_N(u)-d_N(u)| > 4\sqrt{d_N^*(t)\log N}) \leq (4N+2)\exp\{-4 \log N\}$$

that is

$$(3.33)\quad P(\sup_{-\infty<u\leq t} |D_N(u)-d_N(u)| = O(\sqrt{d_N^*(t)\log N})) = 1.$$

Proof: Observe that both $D_N(u)$ and $d_N(u)$ are monotone nondecreasing. Let

$$-\infty = u_0 < u_1 <\dots< u_{k(N)} = t$$

be a partition of $(-\infty,t]$ for which $k(N) \leq N$

$$(3.34)\quad \delta_N(i) = d_N(u_i^-) - d_N(u_{i-1}) \leq 1 \qquad i = 1,2,\dots,k(N).$$

Since $d_N(t) \leq N$ for all t, such partition exists. By the above mentioned monotonicity

$$P(\sup_{-\infty<u\leq t} |D_N(u)-d_N(u)| > \varepsilon) \leq$$

$$(3.35)\quad \leq \sum_{i=1}^{k(N)} P(\sup_{u_{i-1}\leq u<u_i} |D_N(u)-d_N(u)| > \varepsilon) + P(|D_N(t)-d_N(t)| > \varepsilon) \leq$$

$$\leq \sum_{i=1}^{k(N)} \{P(|D_N(u_i^-)-d_N(u_i^-)| > \frac{\varepsilon-\delta_N(i)}{2}) +$$

$$+ P(|D_N(u_{i-1})-d_N(u_{i-1})| > \frac{\varepsilon-\delta_N(i)}{2})\} + P(|D_N(t)-d_N(t)| > \varepsilon).$$

By condition $\varepsilon \geq 2$

$$\frac{\varepsilon-\delta_N(i)}{2} \geq \frac{\varepsilon-1}{2} \geq \frac{\varepsilon}{4}.$$

Apply again Lemma 3.1 and (3.11) for any $\frac{\varepsilon}{4} \leq d_N(t)$

$$P(|D_N(t)-d_N(t)| > \varepsilon) \leq \sum_{i=1}^{k(N)} \{P(|D_N(u_i^-)-d_N(u_i^-)| > \frac{\varepsilon}{4}) +$$

$$P(|D_N(u_{i-1})-d_N(u_{i-1})| > \frac{\varepsilon}{4})\} + P(|D_N(t)-d_N(t)| > \frac{\varepsilon}{4}) \leq \tag{3.36}$$

$$\leq (2N+1) \exp\{- \frac{\varepsilon^2}{2^6 d_N(t)}\}$$

which proves (i).
To prove (ii) let

$$\varepsilon_N = 4\sqrt{d_N^*(t)2^4 \log N}$$

and apply Lemma 3.2 for each summand of (3.36) with $\tilde{A}_N = d_N(t)$ and $k = 4$,

$$P(\sup_{-\infty<u\leq t} |D_N(t)-d_N(t)| > 4\sqrt{d_N^*(t)2^4 \log N}) \leq (2N+1)\cdot 2 \exp\{-2^2 \log N\}.$$

Hence

$$\sum_{N=1}^{\infty} P(\sup_{-\infty<u\leq t} |D_N(u)-d_N(u)| > 4\sqrt{d_N^*(t)2^4 \log N}) < +\infty$$

and (3.33) follows by Borel-Cantelli. □

Remark 1 The last lemma is in some sense a generalization of a theorem of Singh [7]. Lemma 4.3. is a very weak generalization of Lemma 1 [9] of Wellner which deals with i.i.d. r.v.'s.

4. Lemmas

Lemma 4.1. /An elementary inequality/ (see i.e. Rényi [8] p. 517).

For arbitrary $|a_i| \leq 1$, $|c_i| \leq 1$ $i = 1,2,\dots,N$ real numbers

$$|\prod_{i=1}^{N} a_i - \prod_{i=1}^{N} c_i| \leq \sum_{i=1}^{N} |a_i - c_i|. \tag{4.1}$$

Lemma 4.2. Suppose that F is continuous. Then

$$\sup_{-\infty<u\leq t} |F_N^*(u) - F_N^0(u)| \leq \int_{-\infty}^{t} \frac{1}{(M_N(u)+1)^2} \, dB_N(u). \tag{4.2}$$

Proof: If $u > \tau_N$ then $F_N^*(u) = F_N^0(u) = 0$. From Definition 2.1 and 2.2 and Lemma 4.1, for $s \leq \tau_N$

$$|F_N^*(s) - F_N^0(s)| = |\bar{F}_N^*(x) - \bar{F}^0(s)| =$$

$$= \left| \prod_{j=1}^{N} \left(\frac{M_N(Z_j)}{M_N(Z_j)+1} \right)^{\beta_j(s)} - \prod_{j=1}^{N} \left(\frac{M_N(Z_j)+1}{M_N(Z_j)+2} \right)^{\beta_j(s)} \right| \leq$$

$$\leq \sum_{j=1}^{N} \left| \left(\frac{M_N(Z_j)}{M_N(Z_j)+1} \right)^{\beta_j(s)} - \left(\frac{M_N(Z_j)+1}{M_N(Z_j)+2} \right)^{\beta_j(s)} \right| =$$

$$= \sum_{j=1}^{N} \frac{\beta_j(s)}{(M_N(Z_j)+1)(M_N(Z_j)+2)} \leq \sum_{j=1}^{N} \frac{\beta_j(s)}{(M_N(Z_j)+1)^2} .$$

$$\sup_{-\infty<u\leq t} |F_N^*(s) - F_N^0(s)| \leq \sum_{j=1}^{N} \frac{\beta_j(t)}{(M_N(Z_j)+1)^2} = \int_{-\infty}^{t} \frac{1}{(M_N(u)+1)^2} \, dB_N(u)$$

which proves the lemma. □

Let

$$R_N(u) = \int_{-\infty}^{u} \frac{1}{M_N(s)} \, dB_N(s) \tag{4.3}$$

$$R(u) = \int_{-\infty}^{u} \frac{1}{m_N(s)} \, db_N(s). \tag{4.4}$$

Observe that for $u < T_F$

$$(4.5)\qquad R(u) = \int_{-\infty}^{u} \frac{1}{m_N(s)}\, db_N(s) = \int_{-\infty}^{u} \frac{\bar{G}(N,s^-)}{\bar{G}(N,s)\bar{F}(s)}\, dF(s) = -\log \bar{F}(u).$$

Moreover

$$(4.6)\qquad |\bar{F}_N^*(u)-\bar{F}(u)| \le |\bar{F}_N^*(u)-\bar{F}_N^o(u)|+|\bar{F}_N^o(u)-\bar{F}(u)|$$

and

$$(4.7)\qquad \bar{F}_N^o(u)-\bar{F}(u) = (e^{\log \bar{F}^o(u)}-e^{-R_N(u)}) + (e^{-R_N(u)}-e^{-R(u)}).$$

Applying the Taylor expansion for the two terms we get

$$(4.8)\qquad \begin{aligned} \bar{F}_N^o(u)-\bar{F}(u) &= e^{-R_N^*(u)}(\log \bar{F}_N^o(u)+R_N(u)) + \\ &F(u)\ (R_N(u)-R(u)) + \frac{1}{2}\, e^{-R_N^{**}(u)}(R_N(u)-R(u))^2 \end{aligned}$$

where

$$(4.9)\qquad \begin{aligned} &\min\{(-\log \bar{F}^o(u)),\ R_N(u)\} \le R_N^*(u) \le \max\{(-\log \bar{F}_N^o(u)),\ R_N(u)\} \\ &\min\{R(u),\ R_N(u)\} \le R_N^{**}(u) \le \max\{R(u),\ R_N(u)\}. \end{aligned}$$

From (4.6)-(4.9) follows, that

$$(4.10)\qquad \begin{aligned} |\bar{F}_N^*(u)-\bar{F}(u)| &\le |\bar{F}_N^*(u)-\bar{F}^o(u)| + |\log \bar{F}_N^o(u)+R_N(u)| + \\ &+ F(u)|R_N(u)-R(u)| + \frac{1}{2}\, F(u)\exp|R_N(u)-R(u)|\cdot|R_N(u)-R(u)|^2. \end{aligned}$$

Observe that

$$(4.11)\qquad \begin{aligned} R_N(u)-R(u) &= \int_{-\infty}^{u} \frac{1}{M_N(s)}\, dB_N(s) - \int_{-\infty}^{u} \frac{1}{m_N(s)}\, db_N(s) = \\ &= \int_{-\infty}^{u}\left(\frac{1}{M_N(s)} - \frac{1}{m_N(s)}\right)dB_N(s) + \int_{-\infty}^{u} \frac{1}{m_N(s)}\, d(B_N(s)-b_N(s)). \end{aligned}$$

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

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

/This sequence is well-defined if $N \geq N^*$ by the above condition/

Lemma 4.3. If F is continuous, T_N is defined by (4.12) and

(4.13) $$\frac{\log N}{\bar{G}(N,T_F)} \to 0$$

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

(4.14) $$\frac{1}{M_N(u)} \leq \frac{2}{m_N(u)} \text{ for all } u \leq T_N.$$

That is

$$P(\sup_{-\infty<u\leq T_N} \frac{m_N(u)}{M_N(u)} = O(1)) = 1.$$

Proof: We apply Lemma 3.4. (i) for the points T_N $(N > N^*)$.

By condition (4.13) we may choose an N_1 (independent from ω) such that for $N \geq N_1$

$$\bar{G}(N,T_F) > 2^{16} \log N.$$

Then for $N \geq N_1$

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

hence the condition of the Lemma is satisfied. Let $\lambda=2$ then by (3.20) and (4.15) if $N > N_1$

$$P(\sup_{-\infty<u\leq T_N} \frac{m_N(u)}{M_N(u)} > 2) \leq N \exp\{-2^{-6+8} \log N\}.$$

Hence

$$(4.16)\qquad \sum_{N\geq N_1}^{\infty} P\Big(\sup_{-\infty<u\leq T_N} \frac{m_N(u)}{M_N(u)} > 2\Big) \leq \sum_{N\geq N_1}^{\infty} N \exp\{-2^2\log N\} < \infty$$

and by Borel Cantelli follows (4.13). □

Lemma 4.4. If F is continuous and T_N is defined by (4.12) and

$$(4.17)\qquad \frac{\log N}{\bar{G}(N,T_F)} \to 0$$

then for almost all ω there exists an $N_0(\omega)$ such that, if $N \geq N_0(\omega)$ then

$$(4.18)\qquad \sup_{-\infty<u\leq T_N} \left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right| \leq \sqrt{2^9\log N}\ .$$

Proof: Apply Lemma 3.3. (i). Let $\varepsilon_N = \sqrt{2^9\log N}$. By (4.17) if $N \geq N_1$ then

$$(4.19)\qquad \bar{G}(N,T_F) > 2^{18}\log N.$$

Hence if $N \geq N_1$

$$(4.20)\qquad m_N(T_N) = \sqrt{\frac{\log N}{\bar{G}(N,T_F)}}\ \bar{G}(N,T_N) \geq \sqrt{2^{18}\log^2 N} = 2^9\log N.$$

From (4.20) follows that if $N \geq N_1$ then $\varepsilon_N \leq \sqrt{m_N(T_N)}$, and if $N \geq N_2 \geq N_1$ then $\frac{2}{\sqrt{m_N(T_N)}} \leq \varepsilon_N$ is also hold.

Hence by (3.14)

$$\sum_{N\geq N_2}^{\infty} P\Big(\sup_{-\infty<u\leq T_N} \left|\frac{M_N(u)-m_N(u)}{\sqrt{m_N(u)}}\right| > \sqrt{2^9\log N}\Big) \leq \sum_{N\geq N_2}^{\infty} 4N \exp\{-4\log N\} < +\infty$$

and our statement follows from Borel-Cantelli. □

Lemma 4.5. Suppose that F is continuous, T_N is defined by (4.12), let $1 \leq \alpha \leq 2$ arbitrary, and suppose that

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

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

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

for any $T \le T_N$.

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

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

Moreover

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

Hence introducing the notation

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

(4.26) $$\int_{-\infty}^{u} \frac{1}{m^\alpha{}_N(s)} \, d(B_N(s)-b_N(s)) = \sum_{j=1}^{N} \xi_j^*(u)$$

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

(4.27) $$e^x \le 1 + x + \frac{x^2}{2} \quad \text{if} \quad |x| \le \frac{1}{2}$$

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

if

(4.29) $$|u\xi_j^*(t)| \le \frac{1}{2}\,.$$

Observe that for $t \le T$

(4.30) $$0 < \xi_j(t) = \frac{\beta_j(t)}{(\sum_{k=1}^N \bar{H}_k(Z_j))^\alpha} \le \frac{\beta_j(T)}{(\sum_{k=1}^N \bar{H}_k(Z_j))^\alpha}\,.$$

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

(4.31) $$0 \le \xi_j(t) \le \frac{1}{(\sum_{k=1}^N \bar{H}_k(T))^\alpha} = \frac{1}{(m_N(T))^\alpha} \quad \text{for any } t \le T.$$

Hence (4.29) valid if

(4.32) $$0 \le u \le \frac{(m_N(T))^\alpha}{2}\,.$$

For any $t \le T$ we have

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

(4.33) $$= E\left(\frac{\beta_j(t)}{(\sum_{k=1}^N \bar{H}_k(Z_j))^{2\alpha}}\right) = \int_{-\infty}^t \frac{\bar{G}_j(s^-)}{(\sum_{k=1}^N \bar{H}_k(s))^{2\alpha}}\, dF(s) \le$$
$$\le \frac{1}{(\sum_{k=1}^N \bar{G}_k(T))^{2\alpha-1}} \int_{-\infty}^t \frac{\bar{G}_j(s^-)}{\bar{F}^{2\alpha}(s)(\sum_{k=1}^N \bar{G}_k(s))}\, dF(s) = \frac{1}{(\bar{G}(N,T))^{2\alpha-1}} \int_{-\infty}^t \frac{\bar{G}_j(s^-)}{\bar{F}^{2\alpha}(s)\bar{G}(N,s)}\, dF(s).$$

Hence for any $t \leq T$

$$\sum_{j=1}^{N} E(\xi_j^{*2}(t)) \leq \frac{1}{(\bar{G}(N,T))^{2\alpha-1}} \int_{-\infty}^{t} \frac{\sum_{j=1}^{N} G_j(s^-)}{\bar{F}^{2\alpha}(s)\bar{G}(N,s)} dF(s) =$$

$$(4.34) \qquad = \frac{1}{(\bar{G}(N,T))^{2\alpha-1}} \int_{-\infty}^{t} \frac{\bar{G}(N,s^-)}{\bar{F}^{2\alpha}(s)\bar{G}(N,s)} dF(s) = \frac{1}{(\bar{G}(N,T))^{2\alpha-1}} \int_{-\infty}^{t} \frac{1}{\bar{F}^{2\alpha}(s)} dF(s) =$$

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

Hence using the notations of Lemma 3.1, with

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

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

$$(4.35) \qquad P\left(\left|\sum_{j=1}^{N} \xi_j^*(t)\right| > \varepsilon\right) \leq 2\exp\left\{-\frac{\varepsilon^2 (m_N(T))^{2\alpha-1}}{4}\right\}.$$

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

$$(4.36) \qquad \eta_N(t) = \sum_{j=1}^{N} \xi_j(t) = \int_{-\infty}^{t} \frac{1}{m_N^{\alpha}(u)} dB_N(u)$$

and

$$(4.37) \qquad \ell_N(t) = \sum_{j=1}^{N} E(\xi_j(t)) = \int_{-\infty}^{t} \frac{1}{m_N^{\alpha}(u)} db_N(u)$$

are both monotone nondecreasing functions of t. Suppose that $m_N(T) > 1$ then $\ell_N(t) \leq |\log \bar{F}(t)|$. As by $m_N(T) \geq 1$, $1 \leq \alpha \leq 2$, $\ell_N(t) = \int_{-\infty}^{t} \frac{1}{m_N^{\alpha}(t)} db_N(t) \leq \int_{-\infty}^{t} \frac{1}{m_N(t)} db_N(t) = |\log \bar{F}(t)|$. For a fix $0 < \varepsilon \leq 1$ consider a partition of the interval $(-\infty,T)$

$$-\infty = u_0 < u_1 \ldots < u_{L(\varepsilon)} = T$$

such that

(i) $\ell_N(u_i) - \ell_N(u_{i-1}) < \frac{\varepsilon}{3}$ $\qquad i = 1,2,\ldots,L(\varepsilon)$

and

(4.38) $$L(\varepsilon) \leq \frac{3|\log \bar{F}(T)|}{\varepsilon} + 1.$$

Since $\ell_N(t)$ is continuous such a partition easily can be constructed.

If

$$|\eta_N(u_{i-1}) - \ell_N(u_{i-1})| \leq \frac{\varepsilon}{3} \text{ and } |\eta_N(u_i^-) - \ell_N(u_i)| < \frac{\varepsilon}{3}.$$

Then by the monotonicity of $\eta_N(t)$ and $\ell_N(t)$ and (4.38) for any $u_i \leq t < u_{i+1}$

(4.39) $$|\eta_N(t) - \ell_N(t)| \leq \frac{\varepsilon}{3} + 2\,\frac{\varepsilon}{3} = \varepsilon.$$

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

$0 \leq i \leq L(\varepsilon)$

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

Applying (4.35) - (4.37) we have* that if $m_N(T) \geq 1$

$$P(\sup_{-\infty<t\leq T} |\int_{-\infty}^{t} \frac{1}{m_N^{\alpha}(s)}\, dB_N(s) - \int_{-\infty}^{t} \frac{1}{m_N^{\alpha}(s)}\, db_N(s)| > \varepsilon) \leq$$

(4.40) $$2\cdot 2L(\varepsilon)\exp\{-\frac{\varepsilon^2 m_N^{2\alpha-1}(T)}{4\cdot 3^2}\} \leq$$

$$\leq 4(\frac{3|\log \bar{F}(T)|}{\varepsilon} + 1)\exp\{-\frac{\varepsilon^2 m_N^{2\alpha-1}(T)}{36}\}.$$

Consider now the sequence T_N defined by (4.12).

Observe that

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

if $N \geq N_1 (\geq N^*)$ by (4.21).

*A similar but weaker inequality is proved in [10].

Consequently for any $T \le T_N$, $m_N(T) > 1$, if $N \ge N_1$. Thus (4.40) valid for any $T \le T_N$, if $N \ge N_1$.

Let

(4.42) $$\varepsilon_N = \frac{\sqrt{4\cdot 36 \log N}}{(m_N(T))^{\frac{1}{2}(2\alpha-1)}} .$$

Then for any $T \le T_N$

(4.43) $$4\left(\frac{3|\log \bar{F}(T)|}{\varepsilon_N} + 1\right) \le 4\left(\frac{3|\log \bar{F}(T_N)|}{\varepsilon_N} + 1\right) \le$$

$$4\left(\frac{3 \log N}{\sqrt{4\cdot 36 \log N}} N^{\frac{2\alpha-1}{2}} + 1\right) \le N^2$$

if $N \ge N_2(\ge N_1)$ /(4.43) holds as $m_N(t) \le N$ for any t, $\bar{G}(N,T) \le N$, for any T, $\alpha \le 2$, and by the definition of T_N $|\log \bar{F}(T_N)| \le \log N$ if N is big enough/

Consequently for any $T \le T_N$ we have

(4.44) $$\sum_{N \ge N_2}^{\infty} P\left(\sup_{-\infty < t \le T} \left| \int_{-\infty}^{t} \frac{1}{m_N^{\alpha}(s)} d(B_N(s)-b_N(s)) \right| > \frac{12\sqrt{\log N}}{(m_N(T))^{\frac{1}{2}(2\alpha-1)}}\right) \le$$

$$\le \sum_{N \ge N_2}^{\infty} N^2 \exp\{-4 \log N\} < + \infty$$

which proves our statement. □

Lemma 4.6. Suppose that F is continuous, T_N is defined by (4.11), $1 < \alpha \le 2$ arbitrary, and

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

Then for almost all ω there exists an $N_0^*(\omega)$ such that for $N \ge N_0^*(\omega)$

(4.46) $$\int_{-\infty}^{T} \frac{1}{m_N^{\alpha}(u)} dB_N(u) \le \frac{2}{(\alpha-1)m_N^{\alpha-1}(T)} \qquad \text{for any } T \le T_N.$$

Proof:

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

The first term of (4.47) can be estimated by Lemma 4.5. On the other hand

$$\int_{-\infty}^{T} \frac{1}{m_N^\alpha(u)}\, db_N(u) = \int_{-\infty}^{T} \frac{\bar{G}(N,u^-)}{(\bar{F}(u)\bar{G}(N,u))^\alpha}\, dF(u) \le$$

(4.48) $$\le \frac{1}{(\bar{G}(N,T))^{\alpha-1}} \int_{-\infty}^{T} \frac{1}{\bar{F}^\alpha(u)}\, dF(u) \le \frac{1}{(\bar{G}(N,T)^{\alpha-1}(\alpha-1)\bar{F}^{\alpha-1}(T)} = \frac{1}{(\alpha-1)m_N^{\alpha-1}(T)}$$

From (4.47), (4.48) and Lemma 4.5 for almost all ω there exists an $N_0(\omega)$ ($\ge N^*$) such that for $N \ge N_0(\omega)$

$$\int_{-\infty}^{T} \frac{1}{m_N^\alpha(U)}\, dB_N(u) \le \frac{12\sqrt{\log N}}{m_N(T)^{\frac{1}{2}}(2a-1)} + \frac{1}{(\alpha-1)m_N^{\alpha-1}(T)} =$$

$$\le \frac{1}{(\alpha-1)m_N^{\alpha-1}(T)} \left(1 + \frac{12\sqrt{\log N}}{m_N(T)^{\frac{1}{2}}}\right) \qquad \text{for any } T \le T_N.$$

By condition (4.45) there exists an N_0^* ($\ge N_0$) such that if $N \ge N_0^*$ then $\frac{12\sqrt{\log N}}{m_N(T)^{\frac{1}{2}}} \le 1$, which proves (4.46). ❑

5. The strong uniform consistency theorem on the whole line

Lemma 5.1. Suppose that $F(t)$ continuous, $\frac{\log N}{G(N,T_F)} \to 0$ and suppose that

(5.1) $$\sup_{-\infty<u\le T_N} |R_N(u)-R(u)| \le \frac{2}{3} \quad \text{a.s.}$$

where T_N is defined by (4.12). Then

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

$$\le 4 \int_{-\infty}^{T_N} \frac{1}{M_N^2(s)}\, dB_N(s) + \frac{3}{2} \sup_{-\infty<u\le T_N} F(u)|R_N(u)-R(u)| \quad \text{a.s.}$$

Proof: From (4.6) - (4.10) in Section 4,

$$(5.3)\quad |\bar{F}_N^o(u) - \bar{F}(u)| \le |\log \bar{F}_N^o(u)+R_N(u)|+F(u)|R_N(u)-R(u)|+ + \frac{1}{2}F(u)e^{|R_N(u)-R(u)|}|R_N(u)-R(u)|^2.$$

By the elementary inequality

$$\frac{1}{2}e^x x^2 \le x \quad \text{for} \quad 0 \le x < \frac{2}{3}$$

and the condition of our theorem, we have for any $u \le T_N$ that

$$(5.4)\quad |\bar{F}_N^o(u)-\bar{F}(u)| \le |\log \bar{F}_N^o+R_N(u)| + \frac{3}{2}F(u)|R_N(u)-R(u)| \quad \text{a.s.}$$

Observe that, by the definition of $\bar{F}_N^o(u)$

$$|\log \bar{F}_N^o(u)+R_N(u)| = |\sum_{j=1}^{N}\beta_j(u)\log(1-\frac{1}{M_N(Z_j)+2})+R_N(u)| =$$

$$= |\int_{-\infty}^{u}\log(1-\frac{1}{M_N(s)+2})dB_N(s) + \int_{-\infty}^{u}\frac{1}{M_N(s)}dB_N(s)| =$$

$$(5.5)\quad = |\int_{-\infty}^{u}[-\sum_{\ell=1}^{\infty}\frac{1}{\ell}(2+M_N(s))^{-\ell}]dB_N(s) + \int_{-\infty}^{u}\frac{1}{M_N(s)}dB_N(s)| \le$$

$$\le \int_{-\infty}^{u}(\frac{1}{M_N(s)} - \frac{1}{2+M_N(s)})dB_N(s)| + \int_{-\infty}^{u}\frac{1}{(2+M_N(s))^2}dB_N(s)$$

/where the sum $\sum_{\ell=2}^{\infty}\frac{1}{\ell}\frac{1}{(2+M_N(s))^\ell}$ was majorized by a geometric series having quotient less than $\frac{1}{2}$ /

Hence

$$|\log \bar{F}_N^o(u)+R_N(u)| \le \int_{-\infty}^{u}\frac{2}{M_N(s)(M_N(s)+2)}dB_N(s) + \int_{-\infty}^{u}\frac{1}{(2+M_N(s))^2}dB_N(s)$$

$$(5.6)\quad \le \int_{-\infty}^{u}\frac{3}{M_N^2(s)}dB_N(s).$$

Consequently as the last integral is monotone nondecreasing in u

$$(5.7)\qquad \sup_{-\infty<u\le T_N} |\log \bar{F}_N^o(u) + R_N(u)| \le 3 \int_{-\infty}^{T_N} \frac{1}{M_N^2(s)}\, dB_N(s).$$

From Lemma 4.2, formula (4.6), and (5.4), (5.6) we get that under the condition (5.1)

$$(5.8)\qquad \sup_{-\infty<u\le T_N} |F_N^*(u)-F(u)| \le 4 \int_{-\infty}^{T_N} \frac{1}{M_N^2(s)}\, dB_N(s)+ \frac{3}{2} \sup_{-\infty<u\le T_N} F(u)|R_N(u)-R(u)| \quad \text{a.s.}$$

which proves our statement. □

Proof of Theorem 1

First observe that

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

$$+ \sup_{T_N<u<+\infty} |F_N^*(u)-F(u)| = \sup_{-\infty<u\le T_N} |F_N^*(u)-F(u)| +$$

$$(5.9)\quad + \sup_{T_N\le u<+\infty} |\bar{F}_N^*(u)-\bar{F}(u)| \le \sup_{-\infty<u\le T_N} |F_N^*(u)-F(u)| +$$

$$+ |\bar{F}_N^*(T_N)-\bar{F}(T_N)| + \bar{F}(T_N) \le 2 \sup_{-\infty<u\le T_N} |F_N^*(u)-F(u)| + \bar{F}(T_N)$$

as both $\bar{F}_N^*$ and $\bar{F}$ are monotone nonincreasing. By the definition of T_N

$$(5.10)\qquad \bar{F}(T_N) = \sqrt{\frac{\log N}{\bar{G}(N,T_F)}} \qquad \text{/for } N \ge N^* \text{ this is well defined/}$$

hence it's enough to consider

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

From Lemma 4.3 and Lemma 4.6 follows that

$$(5.11)\qquad \int_{-\infty}^{T_N} \frac{1}{M_N^2(s)}\, dB_N(s) \le 2^2 \int_{-\infty}^{T_N} \frac{1}{m_N^2(s)}\, dB_N(s) =$$

$$0\left(\frac{1}{m_N(T_N)}\right) \le 0\left(\frac{1}{\sqrt{\log N\ \bar{G}(N,T_F)}}\right) \qquad \text{a.s.}$$

as by (4.41) $\qquad m_N(T_N) \ge \sqrt{\log N\ \bar{G}(N,T_F)}, \qquad$ if $N > N_1 (\ge N^*)$.

Observe that from (4.11)

$$(5.12)\qquad \begin{aligned} |R_N(u)-R(u)| &\le \int_{-\infty}^{u} \frac{|M_N(s)-m_n(s)|}{M_N(s)m_N(s)}\, dB_N(s) + \\ &\quad \left|\int_{-\infty}^{u} \frac{1}{m_N(s)}\, d(B_N(s)-b_N(s))\right|. \end{aligned}$$

For the first term of (5.12) apply Lemma 4.3, Lemma 4.4 and Lemma 4.6 with $\alpha = \frac{3}{2}$, and for the second term apply Lemma 4.5 with $\alpha = 1$. Then for any $u \le T_N$

$$(5.13)\qquad \begin{aligned} |R_N(u)-R(u)| &\le \int_{-\infty}^{u} \frac{2|M_N(s)-m_N(s))|}{m_N^2(s)}\, dB_N(s) + \\ &\quad \sup_{-\infty<t\le u} \left|\int_{-\infty}^{t} \frac{1}{m_N(s)}\, d(B_N(s)-b_N(s))\right| \le \\ &\le 2\sqrt{2^9 \log N} \int_{-\infty}^{u} \frac{1}{m_N^{3/2}(s)}\, dB_N(s) + \\ &\quad + \sup_{-\infty<t\le u} \left|\int_{-\infty}^{t} \frac{1}{m_N(s)}\, d(B_N(s)-b_N(s))\right| \le \\ &\le 2^3\sqrt{2^9 \log N}\, \frac{1}{\sqrt{m_N(u)}} + \frac{12\sqrt{\log N}}{\sqrt{m_N(u)}} = 0\left(\sqrt{\frac{\log N}{m_N(u)}}\right) \qquad \text{a.s.} \end{aligned}$$

Hence

$$(5.14)\qquad \sup_{-\infty<u\le T_N} |R_N(u)-R(u)| = 0\left(\sqrt{\frac{\log N}{m_N(T_N)}}\right) \le 0\left(\left(\frac{\log N}{\bar{G}(N,T_F}\right)^{\frac{1}{4}}\right) \qquad \text{a.s.}$$

by (4.41). Hence $\sup_{-\infty<t\le T_N} |R_N(u)-R(u)| \le \frac{2}{3}$ a.s.

if $N \ge N_1$.

Hence by Lemma 5.1, and (5.11), (5.13)

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

$$\frac{3}{2} \sup_{-\infty<u\le T_N} F(u)|R_N(u)-R(u)| \le 0\left(\frac{1}{\sqrt{\log N\, \bar{G}(N,T_F)}}\right) +$$

$$(5.15) \quad \sup_{-\infty<u\le T_N} F(u)\, 0\left(\sqrt{\frac{\log N}{m_N(u)}}\right) \le 0\left(\frac{1}{\sqrt{\log N\, \bar{G}(N,T_F)}}\right) +$$

$$\sup_{-\infty<u\le T_N} F(u)\, 0\left(\sqrt{\frac{\log N}{F(u)\bar{G}(N,u)}}\right) \le 0\left(\frac{1}{\sqrt{\log N\, \bar{G}(N,T_F)}}\right) +$$

$$\sup_{-\infty<u\le T_N} \sqrt{F(u)}\, 0\left(\sqrt{\frac{\log N}{\bar{G}(N,u)}}\right) \le 0\left(\frac{1}{\sqrt{\log N\, \bar{G}(N,T_F)}}\right) + 0\left(\sqrt{\frac{\log N}{\bar{G}(N,T_F)}}\right) =$$

$$= 0\left(\sqrt{\frac{\log N}{\bar{G}(N,T_F)}}\right) \quad \text{a.s.}$$

From (5.9), (5.10) and (5.15) follows the theorem. □

Remark 1. From our proof it is clear that we may give a concrete bound instead of using the 0 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.

Corollary 2. If F(t) is continuous and

$$\underline{\lim} \frac{\bar{G}(N,T_F)}{N^{\alpha}} = a > 0$$

for any $0 < \alpha \leq 1$, then

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

Remark 3. Corollary 2 covers the i.i.d. censoring case (α=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 (appears in A.S. 1981).

[2] Földes, A., Rejtö, L. Asymptotic properties of the nonparametric survival curve estimators under variable censoring (to appear in Proceedings of the Pannon Conference in Statistics) (held 1979 in Bad Tatzmanndorf, Austria).

[3] Földes, A., Rejtö, L. A LIL type result for the product limit estimator on the whole line (to appear).

[4] Gill, R.D. Censoring and stochastic integrals 1979. Mathematisch Centrum Amsterdam.

[5] Kaplan, E. L., Meier, P. Nonparametric estimation from incomplete observation. J. Amer. Stat. Assoc. 58U, 457-481 (1963).

[6] Petrov. Sums of independent random variables. Springer Verlag (1975).

[7] Singh, R. S. On the Glivenko-Cantelli theorem for weighted empiricals based on independent random variables A.P. 3, 371-374.

[8] Rényi, A. Probability theory. North-Holland (1970).

[9] 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).

[10] Földes, A., Rejtö, L. Asymptotic properties of the nonparametric survival curve estimators under variable censoring. Unpublished Technical Report of the Math. Inst. of the HAS. No. 38/1979, Budapest.