The Annals of Statistics
1981, Vol. 9, No. 1, 122–129

STRONG UNIFORM CONSISTENCY FOR NONPARAMETRIC SURVIVAL CURVE ESTIMATORS FROM RANDOMLY CENSORED DATA

BY ANTÓNIA FÖLDES AND LIDIA REJTÖ

Mathematical Institute of Hungarian Academy of Science, Budapest

Let $X_1, \cdots, X_n$ be i.i.d. $P(X > u) = F(u)$ and $Y_1, \ldots, Y_n$ be i.i.d. $P(Y > u) = G(u)$, where both F and G are unknown continuous distributions. For $i = 1, \ldots, n$ set $\delta_i = 1$ if $X_i \le Y_i$ and 0 if $X_i > Y_i$ and $Z_i = \min\{X_i, Y_i\}$. One way to estimate F from the observations (Z_i, δ_i) $i = 1, \ldots, n$ is by means of the *product limit* (PL) estimator, F_n^* (Kaplan-Meier, [8]).

In this paper it is shown that F_n^* is uniformly almost sure consistent with rate $O(\sqrt{\log n}/\sqrt{n})$, that is

$$P(\sup_{0\le u\le T}|F_n^*(u) - F(u)| = O(\sqrt{\log n/n})) = 1.$$

Assuming that F is distributed according to a Dirichlet process (Ferguson, [3]) with parameter α, Susarla and Van Ryzin ([11]) obtained the Bayes estimator F_n^α of F.

In the present paper a similar result is established for the Bayes estimator, namely:

$$P(\sup_{0\le u\le T}|F_n^\alpha(u) - F(u)| = O(\sqrt{(\log n)^{1+\gamma}}/\sqrt{n})) = 1 \qquad (\gamma > 0).$$

1. Introduction and summary. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be i.i.d. sequences of nonnegative random variables. Let the two sequences be independent of each other. The statistician has available only the data

$$Z_i = \min\{X_i, Y_i\},$$
$$\delta_i = 1 \quad \text{if} \quad X_i \le Y_i \qquad\qquad i = 1, \cdots, n.$$
$$\quad = 0 \quad \text{if} \quad X_i > Y_i,$$

Let $P(X > t) = F(t)$, $P(Y > t) = G(t)$ and $P(Z > t) = H(t)$, $(H(t) = F(t)G(t))$. An important problem of survival analysis is the estimation of the distribution function F.

Recently the properties of two types of estimators were investigated. One of them is the product limit (P.L.) estimator of Kaplan and Meier [8], the other is the nonparametric Bayesian estimator introduced by V. Susarla and J. Van Ryzin [11].

In an earlier paper [13], B. B. Winter and the authors proved the uniform consistency of the P.L. In a second joint paper [6] there was proved the uniform consistency of the P.L. with rate of $o((\log n)^{1/2}/n^{1/4})$. The results of the above two papers were proved without any continuity conditions on F and G. In case of arbitrary F and discrete G (having finitely many jump points), the uniform consistency of the P.L. with rate factor $O(\sqrt{\log n}/\sqrt{n})$ was proved in [5].

In [10] Susarla and Van Ryzin proved for the Bayesian estimator the pointwise consistency with rate factor $O(\log n/\sqrt{n})$ and the pointwise mean square consistency with rate factor $O(1/n)$; for arbitrary F and continuous G. In [9] Phadia and Van Ryzin proved that the pointwise expected square difference between the Bayes estimator with a Dirichlet process prior and the P.L. estimator for a survival function based on censored data is $O(n^{-2})$.

For arbitrary F and continuous G this result implies the pointwise consistency and the pointwise mean-square consistency of the P.L. estimator with the analogous rate factors.

In the present paper the uniform (sup norm) consistency of the P.L. estimator with rate

Received October 1978; revised August 1979.

AMS 1970 *subject classifications.* Primary 60F15; secondary 62G05.

Key words and phrases. Product limit, survival distribution, strong uniform consistency, censored data, Bayes estimator.

factor $O(\sqrt{\log n}/\sqrt{n})$ is proved. Furthermore, by a sharpening of the lemma of [9] of Phadia and Van Ryzin we get the almost same result for the Bayes estimator too. These results are valid for continuous F and G.

Recently Odd Aalen [1] investigated this problem in a more general context. In that paper a slightly weaker result than Theorem 3.2 is proved, namely the rate factor is $O(\log n/\sqrt{n})$, but the smoothness (see page 534) is more stringent than the condition A1 of the present paper.

Note that Theorem 3.1. is the analog of Lemma 2 of [2] (except that in [2] F need not be continuous) concerning the uncensored case. The convergence rate of Theorem 3.2 is slightly weaker than the best known rate for the law of iterated logarithm in the uncensored case.

2. Definitions, notations and assumptions. Let us introduce the following notations: $[A]$ the indicator function of the event A, $Z_i = \min\{X_i;\ Y_i\}$, $i = 1, \cdots, n$, $\tau_n = \max_{1\le j\le n}\{Z_j\}$,

$$\delta_i = 1 \quad \text{if} \quad X_i \le Y_i$$

$$= 0 \quad \text{if} \quad X_i > Y_i, \qquad i = 1, \cdots, n,$$

$$N^+(u) = \{Z_j > u, j = 1, \cdots, n\}.$$

The main assumptions are the following:

(A1). $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are i.i.d. nonnegative random variables with right-sided fixed unknown continuous distribution functions F and G respectively. The two sequences are independent of each other.

(A2). Suppose that at the point $0 < T < +\infty$

$$\min\{H(T), 1 - H(T)\} \ge \delta,$$

where $0 < \delta < 1/2$. The usual *definition of the product limit estimator* of F in case of continuous F and G is

$$F_n^*(u) = \prod_{j=1}^n \left(\frac{N^+(Z_j)}{N^+(Z_j)+1}\right)^{[\delta_j=1, Z_j\le u]} \quad \text{if} \quad u < \tau_n,$$

$$= 0 \qquad \text{if} \quad u \ge \tau_n.$$

3. Results for the P.L. estimator. This section contains the main theorems and the lemmas. The proofs are postponed to the next section.

THEOREM 3.1. *Under the conditions* (A1), (A2) *for each* $1 > \epsilon > 12/n\delta^4$

$$P(\sup_{0\le u\le T}|F_n^*(u) - F(u)| > \epsilon) \le \frac{24d_0}{\epsilon}\exp\{-dn\epsilon^2\delta^6\}$$

where $d_0 = 3\max\{8; 10c_0\}$ *and* $d = 1/18$ *are universal constants, and* c_0 *is the constant of Lemma* 2 *of* [2].

THEOREM 3.2. *Under the conditions* (A1), (A2) with probability 1

$$\sup_{0\le u\le T}|F_n^*(u) - F(u)| = O(\sqrt{\log n}/\sqrt{n}).$$

Throughout the paper the supremum is taken over the set $[0, T]$. In what follows it will be denoted simply by sup.

Before the lemmas are given, we need some calculations. Let us denote by A_n the set

$$A_n = \{\omega;\ \max_{1\le j\le n} Z_j(\omega) > T\}. \tag{3.1}$$

Observe that

$$P(\bar{A}_n) \le (1-\delta)^n \le e^{-n\delta}. \tag{3.2}$$

Thus

$$(3.3)\qquad P(\sup|F_n^*(u)-F(u)|>\epsilon) \le P(\sup|F_n^*(u)-F(u)|>\epsilon\,|\,A_n)P(A_n)+e^{-n\delta}.$$

Therefore it is enough to deal with the first term of the right-hand side.

On the set A_n we may consider

$$(3.4)\qquad \log F_n^*(u)=\textstyle\sum_{j=1}^n[\delta_j=1,\ Z_j\le u]\log\{N^+(Z_j)/(N^+(Z_j)+1)\}$$

for all $0\le u\le T$. Let

$$(3.5)\qquad \beta_j(u)=[\delta_j=1,\ Z_j\le u],\qquad j=1,\ldots,n.$$

Using the logarithmic expansion let us consider the following decomposition of (3.4) (similar decomposition is used in [12] (3.2))

$$(3.6)\qquad \log F_n^*(u)=R_{n,1}(u)+R_{n,2}(u)+R_{n,3}(u)$$

where

$$(3.7)\qquad R_{n,1}(u)=-\frac{1}{n}\textstyle\sum_{j=1}^n\beta_j(u)H^{-1}(Z_j),$$

$$(3.8)\qquad R_{n,2}(u)=-\textstyle\sum_{j=1}^n\beta_j(u)\sum_{l=2}^\infty\frac{1}{l}(1+N^+(Z_j))^{-l},$$

$$(3.9)\qquad R_{n,3}(u)=-\frac{1}{n}\textstyle\sum_{j=1}^n\beta_j(u)\{n(1+N^+(Z_j))^{-1}-H^{-1}(Z_j)\}.$$

After an easy computation it can be seen that $E\,R_{n,1}(u)=\log F(u)$. Lemmas 3.1–3.3 state that for every fixed $u\in[0,T]$, $|R_{n,1}(u)-\log F(u)|$, $|R_{n,2}(u)|$ and $|R_{n,3}(u)|$ are small outside of a set of exponentially small measure. Hence the analogous statement holds true for $|\log F_n^*(u)-\log F(u)|$ under the condition A_n. Lemma 3.4 gives an exponential bound to $|F_n^*(u)-F(u)|$.

LEMMA 3.1. *Under the assumptions* (A1), (A2)

$$P(|R_{n,1}(u)-\log F(u)|>\epsilon)<2e^{-2n\epsilon^2\delta^2}$$

for every fixed $u\in[0,T]$.

COROLLARY 3.1. *Suppose that* $n\delta>1$, (A1) (A2) *hold, then*

$$P(|R_{n,1}(u)-\log F(u)|>\epsilon\,|\,A_n)\le 4e^{-2n\epsilon^2\delta^2}$$

for every fixed $u\in[0,T]$.

LEMMA 3.2. *Suppose that* (A1) *and* (A2) *hold and* $1>\epsilon>1/(n\delta^4)$ *then*

$$P(|R_{n,2}(u)|>\epsilon\,|\,A_n)\le 4e^{-2n\delta^2/g}$$

for every fixed $u\in[0,T]$.

LEMMA 3.3. *Suppose that* (A1) *and* (A2) *hold, and* $1>\epsilon>4/(n\delta^3)$. *Then for every fixed* $u\in[0,T]$

$$P(|R_{n,3}(u)|>\epsilon\,|\,A_n)\le 2c\,e^{-\delta^6n\epsilon^2/6}$$

where c is universal constant.

LEMMA 3.4. *Suppose that* (A1) *and* (A2) *hold and* $1>\epsilon>12/(n\delta^4)$. *Then for every fixed* $u\in[0,T]$

$$P(|F_n^*(u)-F(u)|>\epsilon\,|\,A_n)\le d_0e^{-d\delta^6n\epsilon^2}$$

where d_0 and d are universal constants.

4. Proof of theorems and lemmas.

PROOF OF THEOREM 3.1. Let us choose a partition $0 = \eta_0 < \eta_1 < \cdots < \eta_{L(\epsilon)} = T$ *of* $[0; T]$ in such a way that

(a) $F(\eta_{i-1}) - F(\eta_i) \le \epsilon/3, \quad i = 1, \dots, L,$
(b) $L(\epsilon) \le 6/\epsilon$.

If $|F_n^*(\eta_{i-1}) - F(\eta_{i-1})| < \epsilon/3$ and $|F_n^*(\eta_i) - F(\eta_i)| < \epsilon/3$ and $\eta_{i-1} \le x < \eta_i$ then $|F_n^*(x) - F(x)| < \epsilon/3 + 2(\epsilon/3) = \epsilon$. Hence if $\sup|F_n^*(x) - F(x)| > \epsilon$ then for some $0 \le j \le L$; $|F_n^*(\eta_j) - F(\eta_j)| \ge \epsilon/3$. Therefore applying (3.3) and Lemma 3.4

$$P(\sup_{0\le u\le T}|F_n^*(u) - F(u)| > \epsilon)$$

$$\le 2L(\epsilon)\sup_{0\le j\le L} P(|F_n^*(\eta_j) - F(\eta_j)| > \frac{\epsilon}{3}|A_n) + e^{-n\delta}$$

$$\le \frac{12}{\epsilon} d_0 e^{-d\delta^6 n\epsilon^2} + e^{-n\delta} \le \frac{24}{\epsilon} d_0 e^{-d\delta^6 n\epsilon^2}$$

which proves the theorem.

PROOF OF THEOREM 3.2. For $n > n_0(\delta) > 1$ choose

$$\epsilon_n = \sqrt{\frac{2\log n}{nd\delta^6}}.$$

Then Theorem 3.1 is applicable, hence

$$P\left(\sup|F_n^*(u) - F(u)| > \sqrt{\frac{2\log n}{nd\delta^6}}\right) \le 24\, d_0 \sqrt{\frac{nd\delta^6}{2\log n}}\frac{1}{n^2} = \frac{24 d_0\sqrt{d\delta^6}}{n^{3/2}\sqrt{2\log n}}.$$

Therefore

$$\sum_{n=1}^{\infty} P\left(\sup|F_n^*(u) - F(u)| > \sqrt{\frac{2\log n}{nd\delta^6}}\right) < +\infty.$$

Now, the theorem follows from the Borel-Cantelli lemma.

PROOF OF LEMMA 3.1. Observe that $R_{n,1}(u)$ is the sum of n independent identically distributed random variables. The expectation of the summands of $R_{n,1}$ is

$$E(-\beta_1(u)H^{-1}(Z_1)) = -E([\delta_1 = 1, Z_1 \le u]H^{-1}(Z_1))$$

$$= -\int_0^u \frac{G(t)}{H(t)}\, d(1 - F(t)) = \int_0^u \frac{-d(1 - F(t))}{F(t)} = \log F(u).$$

Moreover, for every fixed $u \in [0, T]$ the rv's are bounded.

$$0 \le |\beta_i(u)H^{-1}(Z_i)| \le H^{-1}(T) \le \frac{1}{\delta} \qquad i = 1, \cdots, n.$$

Then Theorem 2 of paper [7] of Hoeffding is applicable. Hence

$$P\left(\left|-\frac{1}{n}\sum_{j=1}^{n}\beta_j(u)H^{-1}(Z_j) - \log F(u)\right| > \epsilon\right) \le 2e^{-2n\epsilon^2\delta^2}.$$

PROOF OF COROLLARY 3.1. Using (3.2) and the condition $n\delta > 1$ we have that

$$P(A_n) = 1 - P(\bar{A}_n) > 1 - e^{-n\delta} > 1/2. \tag{4.1}$$

PROOF OF LEMMA 3.2. It is easy to see that on the set A_n

$$\beta_j(u) \sum_{l=2}^{\infty} \frac{1}{l} (1 + N^+(Z_j))^{-l} \le \frac{\beta_j(u)}{(N^+(Z_j) + 1)^2} \le \frac{1}{(N^+(T) + 1)^2}.$$

Therefore by (4.1)

$$(4.2) \quad P(|R_{n,2}(u)| > \epsilon \mid A_n) \le P\left(\frac{n}{(N^+(T) + 1)^2} > \epsilon\right) \Big/ P(A_n) \le 2P\left(\frac{n}{(N^+(T) + 1)^2} > \epsilon\right).$$

Applying the Bernstein inequality (see, e.g., [10] page 387) to the binomial random variable $N^+(T)$

$$(4.3) \quad \begin{aligned} &P(|N^+(T) - nH(T)| > nH(T)(1 - H(T))) \\ &\le 2\exp\{-(2/9)nH(T)(1 - H(T))\} \le 2\exp\{-(2/9)n\delta^2\}. \end{aligned}$$

Let us define a set

$$(4.4) \quad B_n = \{|N^+(T) - nH(T)| > nH(T)(1 - H(T))\}.$$

Then using (4.3)

$$(4.5) \quad P(B_n) \le 2e^{-(2/9)n\delta^2}.$$

From (4.2), (4.4) and (4.5) it follows that

$$P(|R_{n,2}(u)| > \epsilon \mid A_n) \le 2\left\{P\left(\frac{n}{(N^+(T) + 1)^2} > \epsilon \mid \bar{B}_n\right) + 2e^{-(2/9)n\delta^2}\right\}.$$

On the set $\bar{B}_n$

$$(4.6) \quad nH^2(T) < N^+(T) < nH(T)(2 - H(T))$$

hence

$$\frac{n}{(N^+(T) + 1)^2} < \frac{1}{nH^4(T)} < \frac{1}{n\delta^4}.$$

Using the condition $n\epsilon > 1/\delta^4$, the statement of the lemma immediately follows.

PROOF OF LEMMA 3.3. Using (4.1) we get

$$\begin{aligned} P(|R_{n,3}(u)| > \epsilon \mid A_n) &\le P\left(\frac{1}{n} \sum_{j=1}^{n} \beta_j(u) \sup \left| \left(\frac{n}{N^+(t) + 1} - \frac{1}{H(t)}\right| > \epsilon \mid A_n\right) \\ &\le 2P\left(\sup \left| \frac{n}{N^+(t) + 1} - \frac{1}{H(t)} \right| > \epsilon\right) \\ &\le 2P\left(\frac{n}{(N^+(T) + 1)H(T)} \sup \left| \frac{N^+(t)}{n} - H(t) + \frac{1}{n} \right| > \epsilon\right). \end{aligned}$$

Arguing similarly as in the proof of Lemma 3.2, using (4.4) and (4.5) it follows that

$$(4.7) \quad \begin{aligned} &P(|R_{n,3}(u)| > \epsilon \mid A_n) \\ &\le 4e^{-(2/9)n\delta^2} + 2P\left(\frac{n}{(N^+(T) + 1)H(T)} \sup \left| \frac{N^+(t)}{n} - H(t) + \frac{1}{n} \right| > \epsilon \mid \bar{B}_n\right). \end{aligned}$$

Using the inequality (4.6) it follows that, on the set $\bar{B}_n$

$$\frac{n}{(N^+(T) + 1)H(T)} < \frac{1}{H^3(T)} < \frac{1}{\delta^3}.$$

Hence the second term of the right-hand side of (4.7) can be estimated by

$$(4.8)\quad P\left(\sup\left|\frac{N^+(t)}{n}-H(t)+\frac{1}{n}\right|>\epsilon\delta^3\mid\bar{B}_n\right)\le P\left(\frac{1}{n}>\frac{\epsilon\delta^3}{2}\;\Big|\;\bar{B}_n\right)+P\left(\sup\left|\frac{N^+(t)}{n}-H(t)\right|>\frac{\epsilon\delta^3}{2}\;\Big|\;\bar{B}_n\right).$$

Under the condition of Lemma 3.3 the first term of the right-hand side of (4.8) is zero, moreover

$$(4.9)\quad P(\bar{B}_n)>1-2e^{-(2/9)n\delta^2}>1-2e^{-8/9}=c_1\ge 0,2.$$

Using Lemma 2 of paper [2] of Dvoretzky, Kiefer and Wolfowitz, and (4.9) we get that

$$(4.10)\quad P\left(\sup\left|\frac{N^+(t)}{n}-H(t)+\frac{1}{n}\right|>\epsilon\delta^3\mid\bar{B}_n\right)\le c_2e^{-(1/2)n\delta\epsilon^2\delta^6}.$$

Here $c_2=5c_0$ where c_0 is an universal constant of the above Lemma 2. Now the statement of the lemma follows immediately from (4.7) and (4.10).

PROOF OF LEMMA 3.4. An easy consequence of Lemmas 3.2–3.3 and Corollary 3.1 is that if $1>\epsilon>12/n\delta^4$, then for every fixed $u\in[0,T]$

$$(4.11)\quad \begin{aligned}P(|\log F_n^*(u)-\log F(u)|>\epsilon\mid A_n)&\le P\left(|R_{n,1}(u)-\log F(u)|>\frac{\epsilon}{3}\;\Big|\;A_n\right)\\&+P\left(|R_{n,2}(u)|>\frac{\epsilon}{3}\;\Big|\;A_n\right)+P\left(|R_{n,3}(u)|>\frac{\epsilon}{3}\;\Big|\;A_n\right)\le d_0e^{-dn\epsilon^2\delta^6}\end{aligned}$$

and d_0, d are universal constants. If $0<x\le 1$ and $0<y\le 1$ then $|x-y|\le|\log x-\log y|$. Hence on the set A_n

$$(4.12)\quad |F_n^*(u)-F(u)|\le|\log F_n^*(u)-\log F(u)|.$$

Therefore the statement of Lemma 3.4 follows from (4.11) and (4.12).

5. Consequences to the Bayesian estimator. In this section we show that the analog of Theorem 3.2 is valid for the Bayesian estimator of F. We prove this result by a sharpening of the lemma of Phadia and Van Ryzin [9].

Now we list definitions, notations and assumptions in connection with the Bayesian estimator.

Let α be an arbitrary nonnull positive measure on $(0;+\infty)$. We shall make the following assumption:

(A3). $\alpha(u)=\alpha((u;+\infty))>0$ for all $u\in[0;T]$.

Let us denote by $Z_{(1)}\le Z_{(2)}\le\cdots\le Z_{(n)}$ the ordering of the sample $Z_1,\cdots,Z_n$ and $\delta_{(i)}$ $(i=1,\cdots,n)$ denotes the δ corresponding $Z_{(i)}$ in the original sample.

DEFINITION. If $1-F$ is assumed to be a random distribution function with Dirichlet process prior with parameter α, the Bayes estimator of $F(u)$ (when G is assumed to be continuous) has been obtained by Susarla and Van Ryzin [11] as

$$(5.1)\quad F_n^\alpha(u)=\begin{cases}\dfrac{N^+(u)+\alpha(u)}{n+\alpha(0)} & \text{if } 0<u\le Z_{(1)}\\[2ex] \dfrac{N^+(u)+\alpha(u)}{n+\alpha(0)}\prod_{j=1}^{i}\left(\dfrac{N^+(Z_{(j)})+\alpha[Z_{(j)}]+1}{N^+(Z_{(j)})+\alpha[Z_{(j)}]}\right)^{[\delta_{(j)}=0]} & \text{if } Z_{(i)}\le u<Z_{(i+1)}\\ & i=1,\cdots n-1\\ =0 & \text{if } Z_{(n)}<u\end{cases}$$

where $\alpha[Z_i) = \alpha([Z_i; +\infty))$. Then the P.L. estimator may be written as

$$
(5.2)\quad
\begin{aligned}
F_n^*(u) &= 1 && \text{if } 0 \le u \le Z_{(1)} \\
&= \frac{N^+(u)}{n} \prod_{j=1}^{i} \left(\frac{N^+(Z_{(j)}) + 1}{N^+(Z_{(j)})} \right)^{[\delta_{(j)}=0]} && \text{if } Z_{(i)} \le u < Z_{(i+1)} \\
& && i = 1, \cdots n-1 \\
&= 0 && \text{if } Z_{(n)} \le u.
\end{aligned}
$$

LEMMA 5.1. *Under the assumptions* (A1)–(A3)

$$E(\sup|F_n^*(u) - F_n^\alpha(u)|)^2 = O\left(\frac{1}{n^2}\right).$$

PROOF. Let $u \in [Z_{(i)}; Z_{(i+1)})$ $(i = 0, \cdots, n-1)$ where $Z_{(0)} = 0$. Then it was proved by Phadia and Van Ryzin ([9] see formula (2.8)) that

$$|F_n^*(u) - F_n^\alpha(u)|^2 \le \frac{4n\alpha^2(0)}{(n-i)^3}.$$

It follows that the inequality

$$(5.3)\quad \sup_{Z_{(i)} \le u \le Z_{(i+1)}} |F_n^*(u) - F_n^\alpha(u)|^2 \le \frac{4n\alpha^2(0)}{(n-i)^3}.$$

holds for $i = 0, 1, \cdots, n$. Now

$$
(5.4)\quad
\begin{aligned}
&E(\sup|F_n^*(u) - F_n^\alpha(u)|)^2 \\
&\qquad = \int_{A_n} \sup|F_n^*(u) - F_n^\alpha(u)|^2\, dP + \int_{\bar{A}_n} \sup|F_n^*(u) - F_n^\alpha(u)|^2\, dP
\end{aligned}
$$

where A_n was defined by (3.1). Using (3.2) it is enough to estimate the first term of the right-side of (5.4). Applying (5.3) on the set A_n, we get that

$$(5.5)\quad \sup|F_n^*(u) - F_n^\alpha(u)|^2 \le \frac{4n\alpha^2(0)}{N^+(T)^3}.$$

Using the set B_n, defined by (4.4) and the inequalities (4.5), (4.6), (5.5)

$$
\begin{aligned}
&\int_{A_n} \sup|F_n^*(u) - F_n^\alpha(u)|^2\, dP \\
&\qquad = \int_{A_n \cap B_n} \sup|F_n^*(u) - F_n^\alpha(u)|^2\, dP + \int_{A_n \cap \bar{B}_n} \sup|F_n^*(u) - F_n^\alpha(u)|^2\, dP \\
&\qquad \le P(B_n) + \int_{A_n \cap \bar{B}_n} \frac{4n\alpha^2(0)}{N^+(T)^3}\, dP \\
&\qquad \le 2e^{-(2/9)n\delta^2} + \frac{1}{n^2}\frac{4\alpha(0)}{\delta^6}.
\end{aligned}
$$

Thus

$$E(\sup|F_n^*(u) - F_n^\alpha(u)|)^2 \le e^{-n\delta} + 2e^{-(2/9)n\delta^2} + \frac{1}{n^2}\frac{4\alpha(0)}{\delta^6}$$

which proves the lemma.

THEOREM 5.1. *Under the assumptions* (A1)–(A3) *the Bayesian estimator is uniformly almost* *surely consistent with rate factor* $O((\sqrt{\log n})^{1+\gamma}/\sqrt{n})$ *(for any* $\gamma > 0$*) on the interval* $[0; T]$, *namely*

$$\sup |F_n^\alpha(u) - F(u)| = O\left(\sqrt{\frac{(\log n)^{1+\gamma}}{n}}\right)$$

with probability 1.

PROOF.

$$\sup |F_n^\alpha(u) - F(u)| \leq \sup |F_n^*(u) - F(u)| + \sup |F_n^\alpha(u) - F_n^*(u)|.$$

Since $n^2 E(\sup |F_n^\alpha(u) - F_n^*(u)|)^2 \leq q$ (where q is a constant) and

$$\sum_{n=1}^{\infty} \frac{1}{n(\log n)^{1+\gamma}} < +\infty \qquad \text{(for any } \gamma > 0)$$

$$\sup |F_n^\alpha(u) - F_n^*(u)| = O\left(\sqrt{\frac{(\log n)^{1+\gamma}}{n}}\right) \quad \text{a.s.}$$

by the Borel-Cantelli lemma. Now the theorem follows from Theorem 3.2.

REFERENCES

[1] AALEN, O. (1978). Nonparametric estimation of partial transition probabilities in multiple decrement models. *Ann. Statist.* **6** 534–545.
[2] DVORETZKY, A., KIEFER, J. and WOLFOWITZ, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.
[3] FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.
[4] FERGUSON, T. S. and PHADIA, E. G. (1979). Bayesian nonparametric estimation based on censored data. *Ann. Statist.* **7** 163–186.
[5] FÖLDES, A., REJTÖ, L. (1980). Remark to the paper "Strong consistency properties of the PL estimator." *Periodica Math. Hungar.* **11** 251–254.
[6] FÖLDES, A., REJTÖ, L. and WINTER, B. B. (1980). Strong consistency properties of nonparametric estimators for randomly censored data (Part I). *Periodica Math. Hungar.* **11** 233–250.
[7] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.
[8] KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.
[9] PHADIA, E. G. and VAN RYZIN, J. (1980). A note on convergence rates for the PL estimator. *Ann. Statist.* **8** 673–678.
[10] RÉNYI, A. (1970). *Probability Theory.* Akadémiai Kiadó, Budapest.
[11] SUSARLA, V. and VAN RYZIN, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. *J. Amer. Statist. Assoc.* **72** 897–902.
[12] SUSARLA, V. and VAN RYZIN, J. (1978). Large sample theory for a Bayesian nonparametric survival curve estimator based on censored samples. *Ann. Statist.* **6** 755–768.
[13] WINTER, B. B., FÖLDES, A. and REJTÖ, L. (1978). Glivenko-Cantelli theorems for the PL estimate. *Problems of Control and Information Theory* **7** 213–225.

MATHEMATICAL RESEARCH INSTITUTE
HUNGARIAN ACADEMY OF SCIENCE
BUDAPEST 1053
REÁLTANODA U. 13–15.