The Annals of Statistics
1993, Vol. 21, No. 3, 1591–1607

THE STRONG LAW UNDER RANDOM CENSORSHIP

BY W. STUTE[1] AND J.-L. WANG[2]

University of Giessen and University of California, Davis

Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables with d.f. F. We observe $Z_i = \min(X_i, Y_i)$ and $\delta_i = 1_{\{X_i \le Y_i\}}$, where $Y_1, Y_2, \dots$ is a sequence of i.i.d. censoring random variables. Denote by $\hat{F}_n$ the Kaplan–Meier estimator of F. We show that for any F-integrable function φ, $\int \varphi \, d\hat{F}_n$ converges almost surely and in the mean. The result may be applied to yield consistency of many estimators under random censorship.

1. Introduction and main results. Let $X_1, X_2, \dots$ be a sequence of independent random variables with common distribution function (d.f.) F. Assume that φ is a Borel-measurable function on the real line such that $\int |\varphi| \, dF < \infty$. Denote by F_n the empirical d.f. of $X_1, \dots, X_n$. The strong law of large numbers (SLLN) applied to $\varphi(X_1), \dots$ then yields with probability 1 (w.p.1)

$$\lim_{n \to \infty} \int \varphi(x) F_n(dx) = \int \varphi(x) F(dx). \tag{1.1}$$

Within a statistical framework (1.1) is crucial for proving consistency of the estimators of various parameters of interest. Here is a list of possible applications:

1. If $\varphi = 1_{(-\infty, t]}$, then (1.1) yields $F_n(t) \to F(t)$ w.p.1.
2. If $\varphi(x) = x^k$, then (1.1) gives the consistency of the kth moment estimators.
3. If $\varphi(x) = \exp(itx)$, we obtain pointwise convergence of the empirical characteristic function.
4. If $\varphi(x) = (x - t)1_{\{x > t\}}$ and $F(t) < 1$, then [cf. (1)], w.p.1
$$\frac{\int_{\{x>t\}} (x - t) F_n(dx)}{1 - F_n(t)} \to \frac{\int_{\{x>t\}} (x - t) F(dx)}{1 - F(t)},$$
the so-called mean residual life function (at time t).
5. If the parameter of interest θ is that value of t for which
$$\int \varphi(x, t) F(dx) = 0,$$

Received September 1991; revised October 1992.

[1]Supported by the Deutsche Forschungsgemeinschaft.

[2]Supported in part by AFOSR Grant 89-0386. Work done while on leave at the Justus–Lieberg University of Giessen.

AMS 1991 *subject classifications*. Primary 60F15; secondary 60G42, 62G30.

Key words and phrases. Censored data, SLLN, reverse supermartingale, Glivenko–Cantelli convergence.

then, under regularity assumptions on φ, (1.1) guarantees that the solution θ_n of

$$\int \varphi(x,t) F_n(dx) = 0$$

converges to θ with probability 1.

In this paper we extend (1.1) to the random censorship model. This model covers an important situation in survival analysis. To be precise, assume that $X_1, X_2, \dots$ is a sequence of nonnegative lifetimes with d.f. F. The condition that the X's are nonnegative is only for convenience and in no way limits the conclusion of our main result. Along with the X-sequence, let $Y_1, Y_2, \dots$ be a sequence of independent censoring random variables with d.f. G also being independent of the X's. We only observe the censored lifetimes $Z_i = \min(X_i, Y_i)$ together with $\delta_i = 1_{\{X_i \le Y_i\}}$, indicating the cause of death. It is our goal to estimate certain characteristics of F based on (Z_i, δ_i), $1 \le i \le n$. For notational convenience, assume that all random variables are defined on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$.

It is well known that the nonparametric maximum likelihood estimate of F is given by the Kaplan–Meier product-limit estimator (PLE) defined by

$$1 - \hat{F}_n(t) = \prod_{i=1}^{n} \left(1 - \frac{\delta_{[i:n]}}{n - i + 1}\right)^{1_{\{Z_{i:n} \le t\}}}. \tag{1.2}$$

Here $Z_{1:n} \le \cdots \le Z_{n:n}$ are the ordered Z-values, where ties within lifetimes or within censoring times are ordered arbitrarily and ties among lifetimes and censoring times are treated as if the former precedes the latter. $\delta_{[i:n]}$ is the concomitant of the ith-order statistic, that is, $\delta_{[i:n]} = \delta_j$ if $Z_{i:n} = Z_j$. Note that (1.2) is self-adjusted for tied observations and it is equivalent to another frequently adopted form, namely,

$$1 - \hat{F}_n(t) = \prod_{i=1}^{k} \left(1 - \frac{d_i}{n_i}\right)^{1_{\{Z_{(i)} \le t\}}}. \tag{1.3}$$

Here $Z_{(1)} < \cdots < Z_{(k)}$ denote the k distinct lifetimes among $\{Z_1, \dots, Z_n\}$, d_i is the number of deaths at $Z_{(i)}$ and $n_i = \sum_{j=1}^{n} 1_{\{Z_j \ge Z_{(i)}\}}$ is the number of individuals still at risk at time $Z_{(i)}$. For convenience of exposition, in this article we adopt the form (1.2) rather than (1.3).

As before, let φ be F-integrable and put

$$S_n := \int \varphi(x) \hat{F}_n(dx), \qquad n \ge 1.$$

Since $\hat{F}_n$ is a step function, it can be easily seen from (1.2) that

$$S_n = \sum_{i=1}^{n} W_{in} \varphi(Z_{i:n}),$$

where for $1 \le i \le n$,

$$W_{in} = \frac{\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left[\frac{n-j}{n-j+1} \right]^{\delta_{[j:n]}}.$$

We are concerned with the strong law of large numbers (SLLN) for S_n. In particular, we will show that S_n converges almost surely and in the mean to a constant S and then identify S. Needless to say, S turns out to be $\int \varphi(x) F(dx)$ when no censoring is present.

So far, the convergence of S_n has mainly been studied for indicators $\varphi = 1_{(-\infty, t]}$. The special structure of such a φ was crucial, via integration by parts, in the analysis so that the ordinary SLLN and Glivenko–Cantelli convergence of the empirical d.f. of the Z's and the empirical subdistribution function of $\{Z_i \le t, \delta_i = 1\}$ could then be employed. See, for example, Shorack and Wellner (1986) and Wang (1987). Application of known laws of the iterated logarithm for empirical processes also yielded rates for the convergence of $F_n(t)$ uniformly over certain (compact) intervals; cf. Földes and Rejtő (1981) and Csörgő and Horváth (1983). Gu and Lai (1990) derived (under regularity conditions) a functional law of the iterated logarithm for the Kaplan–Meier empirical process $t \to n^{1/2}[\hat{F}_n(t) - F(t)]$. This result may also be obtained on a compact interval using the almost sure representation of Lo and Singh (1986). φ's which are of bounded variation and vanish off a suitable compact set are readily dealt with upon integration by parts. For more general φ, to the best of our knowledge, we are only aware of Susarla and van Ryzin (1980). They considered the mean lifetime, that is, $\varphi(x) = x$. Their method of proof is again based on an integration by parts argument. Since φ is not of bounded variation, a truncation was needed and rather stringent conditions on the tails were introduced which are very often difficult to check.

Compared with the aforementioned papers, our methodology solely rests on the martingale properties of S_n as a process in n. This is different from the usual martingale approach in survival analysis [cf. Aalen (1978) and Gill (1980)], where for n fixed, martingale theory was applied to the Kaplan–Meier process as a process in t. Here no reduction to any empirical process is made. As a consequence, no restriction on φ (up to F-integrability) is required. Both almost sure convergence and convergence in the mean are shown.

In survival analysis, X is the lifetime of a patient due to a particular cause of interest and Y is usually the withdrawal time of the patient from the follow-up study or the time of death due to other causes in a competing risk model. Therefore, for most practical purposes, one can assume that

(1.4) F and G do not have jumps in common.

Notice that (1.4) does not exclude the possibility that either F or G are discontinuous.

Now, let H denote the d.f. of the Z's. By the independence of X_1 and Y_1,

$$(1 - H) = (1 - F)(1 - G). \tag{1.5}$$

Set

$$\tau_H := \inf\{x: H(x) = 1\}.$$

Write $F\{a\} = F(a) - F(a-)$. Finally, let A be the set of all atoms of H.

THEOREM 1.1. *Under* (1.4), *with probability* 1 *and in the mean*,

$$(1.6)\quad \lim_{n\to\infty} \int \varphi(x)\hat{F}_n(dx) = \int_{\{x\notin A,\, x<\tau_H\}} \varphi(x)F(dx) + \sum_{a_i\in A} \varphi(a_i)F\{a_i\}.$$

The right-hand side of (1.6) *may be written as*

$$(1.7)\quad \int_{\{x<\tau_H\}} \varphi(x)F(dx) + 1_{\{\tau_H\in A\}}\varphi(\tau_H)F\{\tau_H\}$$

(*where the second summand should be equal to* 0 *if* $\tau_H = \infty$).

REMARK 1. If no censoring is present, $\hat{F}_n = F_n$ and $F = H$, so that (1.6), resp. (1.7), leads to (1.1).

REMARK 2. Notice that (1.7) reduces to $\int_{\{x\le\tau_H\}}\varphi(x)F(dx)$ unless $\tau_H \notin A$ and $F\{\tau_H\} > 0$, that is, when

$$(1.8)\quad F\{\tau_H\} > 0 \quad \text{but} \quad 1 - G(\tau_{H}-) = 0.$$

Otherwise (1.7) becomes $\int_{\{x<\tau_H\}}\varphi(x)F(dx)$.

REMARK 3. From (1.7) we may also infer conditions under which S_n is a consistent estimator for $\int\varphi\, dF$:

$$(1.9)\quad \lim_{n\to\infty} \int \varphi(x)\hat{F}_n(dx) = \int \varphi(x)F(dx).$$

For this, define τ_F and τ_G similarly to τ_H. Then $\tau_H = \min(\tau_F, \tau_G)$. Of course, when $\tau_F > \tau_H$, (1.9) cannot hold in general (unless, e.g., φ is 0 on $[\tau_H, \tau_F]$). If $\tau_H = \tau_F$, however, (1.9) holds if and only if (1.8) is excluded. In particular, (1.9) holds for continuous F provided that $\tau_H = \tau_F$.

Note also that convergence in the mean implies that $\mathbb{E}S_n$ tends to (1.7). Interestingly enough, we shall prove that for $\varphi \ge 0$, $\{S_n\}_{n\ge1}$ is a reverse-time supermartingale, to the effect that $\mathbb{E}S_n$ converges from below. From a statistical point of view, this means that in general S_n is biased-downwards. For $n = 1$, for example, we get

$$\mathbb{E}S_1 = \int \varphi(X)1_{\{X\le Y\}}\, d\mathbb{P},$$

which may be strictly less than (1.7). For uncensored data, all $\mathbb{E}S_n$ are the same, namely $\int\varphi\, dF$, that is, all S_n are unbiased [which, of course, is trivial, since S_n coincides with the sample mean of $\varphi(X_1), \dots, \varphi(X_n)$].

Examples of nonnegative φ's include indicators $1_{(-\infty,t]}$ and, provided the data are nonnegative, moments and the mean residual life function. As the ordinary SLLN, Theorem 1.1 has many applications in estimation theory. For example, it may be employed to extend consistency of the M- and L-estimators to the random censorship model.

Over the last 20 years, there has also been much effort to show that (1.1) holds uniformly for a class $\mathcal{F} = \{\varphi\}$ of φ's. Most of these papers use the concept of a Vapnik–Chervonenkis class. Such a class fulfills certain combinatorial entropy conditions which enable one to bound the sup over $\varphi \in \mathcal{F}$ by the sup over some finitely many $\varphi_{n1}, \dots, \varphi_{nk_n}$, where an upper bound for k_n is obtained from the entropy condition. Then a standard exponential bound for sums of i.i.d. random variables is applied to bound the tails of $\sup_\varphi |\int \varphi \, dF_n - \int \varphi \, dF|$. A nice and readable introduction to that approach is contained in Pollard's (1984) book. In Stute (1976) an alternative concept for uniform convergence was considered, which is only based on the SLLN rather than exponential bounds. Also here, when the data are at risk of being censored, no Bernstein-type exponential bound is available so far. Theorem 1.1 may then be applied, in connection with Stute (1976), to get uniform convergence in (1.6). As a particular example, we may take for $\mathcal{F}$ the family of all indicators $\varphi_t = 1_{(-\infty,t]}$, $t \le \tau_H$. Equation (1.7) then becomes

$$\tilde{F}(t) = \begin{cases} F(t), & \text{if } t < \tau_H, \\ F(\tau_H -) + 1_{\{\tau_H \in A\}} F\{\tau_H\}, & \text{if } t \ge \tau_H. \end{cases}$$

We therefore get the following corollary.

COROLLARY 1.2. *Under* (1.4), *we have*

$$\sup_{t \le \tau_H} \left| \hat{F}_n(t) - \tilde{F}(t) \right| \to 0 \quad \textit{with probability } 1.$$

COROLLARY 1.3. *Under* (1.4), $\hat{F}_n$ *is uniformly consistent for* F *on* $(-\infty, \tau_H]$ *if and only if either* $F\{\tau_H\} = 0$ *or* $F\{\tau_H\} > 0$ *but* $G(\tau_H -) < 1$.

For the last statement, cf. (1.8). Corollary 1.3 is the a.s. version of Corollary 1 in Wang (1987) and therefore provides a positive answer to his question when there are no common jumps between the life and censoring distributions.

COROLLARY 1.4. *Under* (1.4) *and* $\int |x| F(dx) < \infty$, *we have* (*in obvious notation*) *the following results on the mean residual life function*:

$$\lim_{n \to \infty} \mathbb{E}_{\hat{F}_n}(X - t | X > t)$$

$$= \begin{cases} \mathbb{E}_F\big(1_{\{X < \tau_H\}}(X - t) | X > t\big), & \textit{when } (1.8) \textit{ holds}, \\ \mathbb{E}_F\big(1_{\{X \le \tau_H\}}(X - t) | X > t\big), & \textit{otherwise}, \end{cases} \tag{1.10}$$

*w.p.*1 *and in the mean*.

The $\mathbb{P}$-a.s. convergence in (1.10) is uniform on any compact interval $[0, T]$ for which $H(T) < 1$.

The SLLN, in nature, is an asymptotic result. We only mention here that in the proof section the interested reader may find lots of details about the finite sample behavior of S_n, both for n fixed or as n varies over time. Lemma 2.4, for example, gives an exact formula for the expectation of S_n. This formula gives an impression of what in the mean is really estimated by S_n for finite n. Its limit is presented in (2.5). For nonnegative φ, the convergence is monotone. The joint probabilistic behavior of finitely many S_n's is described by the aforementioned martingale structure.

The bias problem has been the objective of some research in the past. Gill's (1980) formula (3.2.16) yields a negative bias for $\hat{F}_n(t)$, and Mauro (1985) treats a general φ. Note that in his Theorem 3.1 the expectation of $\int \varphi \, d\hat{F}_n$ is compared with $\int \varphi \, dF$ and not with S, as it should be. Chen, Hollander and Langberg (1982) proposed a modification F_n^* of $\hat{F}_n$, which coincides with $\hat{F}_n$ left of the largest Z but jumps to 1 at $Z_{n:n}$, irrespective of whether $\delta_{[n:n]}$ is 1 or 0. Wellner (1985) provides a comparative study from which he is led to prefer $\hat{F}_n$. His Table 2 exhibits the negative bias of $\hat{F}_n$ while F_n^* is always biased upwards.

Our final remark is toward several comments one can find in the more applied work on survival analysis. Because of the technical difficulties encountered when investigating $\hat{F}_n$ on the whole real line via the cumulative hazard function, it is argued that $\hat{F}_n$ needs to be investigated only on intervals $[0, T]$ with $T < \tau_H$. We have never found such a recommendation for completely observable data. For estimation of the mean lifetime, such a restriction would lead to a serious bias, even asymptotically. Also, from a practical point of view, since τ_H is unknown, a cautious choice of (a small) T would lead to a considerable loss of information. In contrast, in Theorem 1.1, the $\hat{F}_n$-integral is computed over the whole real line.

2. Proof of Theorem 1.1. To prove (1.6), it suffices to consider nonnegative φ's. Otherwise, decompose φ into its positive and negative part. Also, in the lemmas to follow, we shall first restrict ourselves to continuous H's. It will be shown later that under (1.4) the general case may be traced back to the uniform distribution on $[0, 1]$ upon using a simple quantile transformation. Now, under a continuous H, the Z's are pairwise distinct. Recall that

$$S_n = \sum_{i=1}^{n} W_{in} \varphi(Z_{i:n}),$$

where for $1 \le i \le n$,

$$W_{in} = \frac{\delta_{[i:n]}}{n - i + 1} \prod_{j=1}^{i-1} \left[\frac{n - j}{n - j + 1} \right]^{\delta_{[j:n]}}. \tag{2.1}$$

Here and in what follows, an empty product is taken to be 1. Note that W_{in} is the jump size of $\hat{F}_n$ at $Z_{i:n}$, since under continuity there is no more than one death at that time.

The essence of our argument is to show that (for $\varphi \geq 0$) $\{S_n\}_{n \geq 1}$ is a reverse-time supermartingale with respect to an appropriate decreasing sequence of σ-fields $\{\mathscr{F}_n\}_{n \geq 1}$. It is well known that if no censoring is present, $\{S_n\}_{n \geq 1}$ (= sample means) forms a reverse-time martingale, so that the classical SLLN is an easy consequence of a proper martingale convergence theorem; cf. Neveu (1975), page 116. For censored variables the martingale property is lost in general. In view of the supermartingale property, however, we shall be able to use Proposition 5-3-11 in Neveu (1975). Application of the Hewitt–Savage zero–one law will show that the limit S is a constant. Finally, it remains to identify S.

Now, for each $n \geq 1$, put

$$\mathscr{F}_n = \sigma\big(Z_{i:n}, \delta_{[i:n]}, 1 \leq i \leq n, Z_{n+1}, \delta_{n+1}, \dots\big).$$

It is easy to see that S_n is adapted to $\mathscr{F}_n$ and that $\mathscr{F}_n \downarrow$. Set $\mathscr{F}_\infty := \bigcap_{n \geq 1} \mathscr{F}_n$. The following lemma presents two fundamental distributional facts concerning order statistics, concomitants and ranks.

LEMMA 2.1. *Let (C_i, D_i), $1 \leq i \leq n$, be independent random vectors from some bivariate distribution with conditional distribution*

$$m(y|x) = \mathbb{P}(D_i \leq y | C_i = x).$$

Let $C_{1:n} \leq \cdots \leq C_{n:n}$ denote the ordered values of the C's, and let $D_{[i:n]}$ be the ith concomitant paired with $C_{i:n}$. Then

(a) *Conditionally on $C_{1:n} \leq \cdots \leq C_{n:n}$, the concomitants are independent with*

$$\mathbb{P}\big(D_{[i:n]} \leq y | C_{i:n} = x\big) = m(y|x).$$

(b) *If the C's have a continuous d.f., then the vector $(C_{i:n}, D_{[i:n]})_{1 \leq i \leq n}$ is independent of the vector of C-ranks.*

When (C_i, D_i) has a joint bivariate density, part (a) is contained in Yang (1977). Here the more general version is needed, since when we apply Lemma 2.1, D_i will be equal to the zero–one variable δ_i (and $C_i = Z_i$).

PROOF. Denote with H the d.f. of C_i. According to Rosenblatt (1952), we have the representation

$$(C_i, D_i) = \big(H^{-1}(U_i), m^{-1}(V_i | H^{-1}(U_i))\big),$$

where (U_i, V_i), $1 \leq i \leq n$, are independent and uniformly distributed on the unit square. Accordingly,

$$D_{[i:n]} = m^{-1}\big(V_{[i:n]} | H^{-1}(U_{i:n})\big),$$

where $V_{[i:n]}$ is the ith concomitant among the V's. Observe that the vector of

V-concomitants is independent of the U-order statistics, and that $V_{[i:n]}$, $1 \le i \le n$, are independent and uniformly distributed on $[0,1]$. Furthermore, the vector of C-ranks is independent of the C-order statistics. Altogether this easily yields (a) and (b). □

The next lemma plays a crucial role in this article.

LEMMA 2.2. *For continuous H,*

$$\mathbb{E}(S_n | \mathscr{F}_{n+1}) = S_{n+1} - \frac{1}{n+1}\varphi(Z_{n+1:n+1})\delta_{[n+1:n+1]}(1-\delta_{[n:n+1]})\prod_{j=1}^{n-1}\left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n+1]}}.$$

In particular, for every nonnegative φ we have

$$\mathbb{E}(S_n | \mathscr{F}_{n+1}) \le S_{n+1},$$

that is, $\{S_n, \mathscr{F}_n\}_{n \ge 1}$ is a reverse-time supermartingale.

PROOF. Write

$$\int \varphi \, d\hat{F}_n = \sum_{i=1}^{n+1} \varphi(Z_{i:n+1})\hat{F}_n\{Z_{i:n+1}\}.$$

To prove the lemma, it suffices to show that

$$\mathbb{E}\left[\hat{F}_n\{Z_{i:n+1}\} | \mathscr{F}_{n+1}\right] = W_{i,n+1} \quad \text{for } 1 \le i \le n$$

and

$$\mathbb{E}\left[\hat{F}_n\{Z_{n+1:n+1}\} | \mathscr{F}_{n+1}\right] = W_{n+1,n+1} - \frac{1}{n+1}\delta_{[n+1:n+1]}(1-\delta_{[n:n+1]})\prod_{j=1}^{n-1}\left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n+1]}}.$$

Now,

$$\mathbb{E}\left[\hat{F}_n\{Z_{1:n+1}\} | \mathscr{F}_{n+1}\right] = \mathbb{E}\left[\hat{F}_n\{Z_{1:n}\} 1_{\{Z_{n+1} > Z_{1:n+1}\}} | \mathscr{F}_{n+1}\right] = \frac{\delta_{[1:n+1]}}{n+1} = W_{1,n+1},$$

where the last equality follows from

$$\mathbb{E}\left[1_{\{Z_{n+1} > Z_{1:n+1}\}} | \mathscr{F}_{n+1}\right] = \mathbb{P}\left[\text{Rank of } Z_{n+1} > 1 | \mathscr{F}_{n+1}\right] = \frac{n}{n+1},$$

by Lemma 2.1(b), and

$$\delta_{[1:n+1]} = \delta_{[1:n]} \quad \text{on } \{Z_{n+1} > Z_{1:n+1}\}.$$

Similarly, on $\{Z_{n+1} = Z_{k:n+1}\}$, we obtain

$$\delta_{[j:n]} = \delta_{[j:n+1]} \quad \text{if } j < k$$

and

$$\delta_{[j:n]} = \delta_{[j+1:n+1]} \quad \text{if } j \geq k.$$

Therefore, if $2 \leq i \leq n$,

$$\begin{aligned}
\mathbb{E}\big[\hat{F}_n\{Z_{i:n+1}\}|\mathscr{F}_{n+1}\big] &= \mathbb{E}\big[\hat{F}_n\{Z_{i:n+1}\}1_{\{Z_{n+1}>Z_{i:n+1}\}}|\mathscr{F}_{n+1}\big] \\
&\quad + \sum_{k=1}^{i-1}\mathbb{E}\big[\hat{F}_n\{Z_{i:n+1}\}1_{\{Z_{n+1}=Z_{k:n+1}\}}|\mathscr{F}_{n+1}\big] \\
&= \frac{1}{n+1}\Bigg\{\delta_{[i:n+1]}\prod_{j=1}^{i-1}\left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n+1]}} \\
&\quad + \sum_{k=1}^{i-1}\frac{\delta_{[i:n+1]}}{n-i+2}\prod_{j=1}^{k-1}\left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n+1]}}\prod_{j=k+1}^{i-1}\left[\frac{n-j+1}{n-j+2}\right]^{\delta_{[j:n+1]}}\Bigg\} \\
&= \frac{W_{i,n+1}}{n+1}\left\{(n-i+2)\pi_i + \sum_{k=1}^{i-1}\pi_k\left[\frac{n-k+2}{n-k+1}\right]^{\delta_{[k:n+1]}}\right\}.
\end{aligned} \tag{2.2}$$

Here

$$\pi_k = \prod_{j=1}^{k-1}\left[\frac{(n-j)(n-j+2)}{(n-j+1)^2}\right]^{\delta_{[j:n+1]}}, \qquad 1 \leq k \leq n.$$

Putting $\pi_{n+1} = 0$, it is easily seen that (2.2) also holds for $i = n+1$. If the expression in $\{\cdots\}$ is denoted Q_i, one finds

$$Q_1 = n+1$$

and for $1 \leq i \leq n-1$,

$$\begin{aligned}
Q_{i+1} - Q_i = \pi_i\Bigg[&(n-i+1)\left[\frac{(n-i)(n-i+2)}{(n-i+1)^2}\right]^{\delta_{[i:n+1]}} \\
&-(n-i+2) + \left[\frac{n-i+2}{n-i+1}\right]^{\delta_{[i:n+1]}}\Bigg].
\end{aligned}$$

The term in brackets vanishes for each choice of δ, so that

$$Q_i = n+1 \quad \text{for } 1 \leq i \leq n.$$

Since

$$Q_{n+1} - Q_n = \pi_n\{2^{\delta_{[n:n+1]}} - 2\} = -\pi_n(1 - \delta_{[n:n+1]}),$$

the proof is complete. □

From Lemma 2.2 we obtain the following lemma.

LEMMA 2.3. *Assume that* H *is continuous and* $\varphi \geq 0$. *Then*:

(i) $S = \lim_{n\to\infty} \mathbb{E}S_n$ *exists* (*possibly infinite*).
(ii) $S_n \to S$ $\mathbb{P}$-*almost surely*.
(iii) *If* $S < \infty$, *then* $\{S_n\}_{n\geq 1}$ *is uniformly integrable and* $\mathbb{E}|S_n - S| \to 0$.

PROOF. From Lemma 2.2, $\mathbb{E}S_n$ is nondecreasing in n, whence (i). By Proposition 5-3-11 in Neveu (1975), $\{S_n\}_{n\geq 1}$ converges $\mathbb{P}$-almost surely ($\mathbb{P}$-a.s.) to some $\mathscr{F}_\infty$-measurable random variable S_∞. Moreover, $S_\infty = \lim_{n\to\infty} \mathbb{E}(S_n|\mathscr{F}_\infty)$ $\mathbb{P}$-a.s. The Hewitt–Savage zero–one law implies that $\mathscr{F}_\infty$ is trivial. Hence

$$S_\infty = \lim_{n\to\infty} \mathbb{E}(S_n|\mathscr{F}_\infty) = \lim_{n\to\infty} \mathbb{E}S_n = S.$$

Part (iii) is also contained in Neveu (1975). □

We will proceed to identify the constant S. First, set

$$m(t) = \mathbb{P}(\delta = 1|Z = t),$$

the conditional probability that a death is observed given that $Z = t$. Denote, for $n \geq 1$,

$$\varphi_n(t) = \prod_{i=1}^{n}\left(1 + \frac{1 - m(Z_{i:n})}{n - i + 1}\right)^{1_{\{Z_{i:n} < t\}}}$$

and

$$g_n(t) = \mathbb{E}\varphi_n(t), \qquad g_0(t) \equiv 1.$$

LEMMA 2.4. *We have*, *for each* $n \geq 1$,

$$\mathbb{E}S_n = \mathbb{E}[\varphi(Z)m(Z)g_{n-1}(Z)].$$

PROOF. Let R_{jn} denote the rank of Z_j among $Z_1, \ldots, Z_n$. We have, by Lemma 2.1,

$$\begin{aligned}
\mathbb{E}S_n &= \mathbb{E}\left\{\sum_{i=1}^{n} \frac{\varphi(Z_{i:n})}{n-i+1}\mathbb{E}\left[\delta_{[i:n]}\prod_{j=1}^{i-1}\left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n]}}|Z_{1:n}, \ldots, Z_{n:n}\right]\right\}\\
&= \mathbb{E}\left\{\sum_{i=1}^{n} \frac{\varphi(Z_{i:n})}{n-i+1} m(Z_{i:n})\prod_{j=1}^{i-1}\left(1 - \frac{m(Z_{j:n})}{n-j+1}\right)\right\}\\
&= \mathbb{E}\left\{\sum_{i=1}^{n} \frac{\varphi(Z_{i:n})m(Z_{i:n})}{n}\prod_{j=1}^{i-1}\left(1 + \frac{1 - m(Z_{j:n})}{n-j}\right)\right\}\\
&= \mathbb{E}\left\{\sum_{i=1}^{n} \frac{\varphi(Z_i)m(Z_i)}{n}\prod_{j=1}^{n}\left(1 + \frac{1 - m(Z_j)}{n-R_{jn}}\right)^{1_{\{Z_j < Z_i\}}}\right\}\\
&= \mathbb{E}\left[\varphi(Z_1)m(Z_1)\prod_{j=1}^{n}\left(1 + \frac{1 - m(Z_j)}{n-R_{jn}}\right)^{1_{\{Z_j < Z_1\}}}\right].
\end{aligned}$$

On $\{Z_j < Z_1\}$ we have $R_{j,n} = R_{j,n-1}$, the rank of Z_j computed among $Z_2, \dots, Z_n$. Condition on Z_1 to get the result. □

We are now going to study the process $\varphi_n(t)$, t fixed, as n increases. Set

$$\mathscr{G}_n = \sigma\{Z_{1:n}, \dots, Z_{n:n}, Z_{n+1}, \dots\}.$$

Clearly, $\varphi_n(t)$ is adapted to $\mathscr{G}_n$, with $\mathscr{G}_n \downarrow$. Put $\mathscr{G}_\infty = \bigcap_{n \geq 1} \mathscr{G}_n$.

LEMMA 2.5. *For continuous H, $\{\varphi_n(t), \mathscr{G}_n\}_{n \geq 1}$ is a (nonnegative) reverse-time supermartingale.*

PROOF. We have, by Lemma 2.1,

$$\begin{aligned}
\mathbb{E}[\varphi_n(t) | \mathscr{G}_{n+1}] &= \mathbb{E}\left[\prod_{j=1}^{n}\left(1 + \frac{1 - m(Z_{j:n})}{n-j+1}\right)^{1_{\{Z_{j:n} < t\}}} | \mathscr{G}_{n+1}\right] \\
&= \sum_{k=1}^{n+1} \mathbb{E}\left[1_{\{Z_{n+1} = Z_{k:n+1}\}} \prod(\cdots) | \mathscr{G}_{n+1}\right] \\
&= \frac{1}{n+1} \sum_{k=1}^{n+1} \prod_{j=1}^{k-1}\left(1 + \frac{1 - m(Z_{j:n+1})}{n-j+1}\right)^{1_{\{Z_{j:n+1} < t\}}} \\
&\quad \times \prod_{j=k+1}^{n+1}\left(1 + \frac{1 - m(Z_{j:n+1})}{n-j+2}\right)^{1_{\{Z_{j:n+1} < t\}}}.
\end{aligned} \tag{2.3}$$

To show the lemma, we proceed by induction on n. For $n = 1$, the last equation yields

$$\begin{aligned}
\mathbb{E}[\varphi_1(t) | \mathscr{G}_2] &= \tfrac{1}{2}\left[1 + (1 - m(Z_{2:2}))1_{\{Z_{2:2} < t\}} + 1 + (1 - m(Z_{1:2}))1_{\{Z_{1:2} < t\}}\right] \\
&= 1 + \frac{1}{2} \sum_{i=1}^{2} (1 - m(Z_{i:2}))1_{\{Z_{i:2} < t\}} \\
&\leq \left(1 + \frac{1 - m(Z_{1:2})}{2} 1_{\{Z_{1:2} < t\}}\right)\left(1 + (1 - m(Z_{2:2}))1_{\{Z_{2:2} < t\}}\right) \\
&= \varphi_2(t).
\end{aligned}$$

Next assume that the assertion holds for n. Write

$$A = \prod_{j=2}^{n+2}\left(1 + \frac{1 - m(Z_{j:n+2})}{n-j+3}\right)^{1_{\{Z_{j:n+2} < t\}}}.$$

Replacing n by $n + 1$ in (2.3), we get

$$\begin{aligned}
\mathbb{E}[\varphi_{n+1}(t)|\mathscr{G}_{n+2}] &= \frac{1}{n+2}\left\{A + \sum_{k=2}^{n+2}\prod_{j=1}^{k-1}(\cdots)^{1_{\{\cdots\}}}\prod_{j=k+1}^{n+2}(\cdots)^{1_{\{\cdots\}}}\right\} \\
&= \frac{1}{n+2}\left\{A + \left(1 + \frac{1 - m(Z_{1:n+2})}{n+1}\right)^{1_{\{Z_{1:n+2}<t\}}}\right. \\
&\qquad \times \sum_{k=1}^{n+1}\prod_{j=1}^{k-1}\left(1 + \frac{1 - m(Z_{j+1:n+2})}{n-j+1}\right)^{1_{\{Z_{j+1:n+2}<t\}}} \\
&\qquad \left. \times \prod_{j=k+1}^{n+1}\left(1 + \frac{1 - m(Z_{j+1:n+2})}{n-j+2}\right)^{1_{\{Z_{j+1:n+2}<t\}}}\right\} \qquad (2.4) \\
&\le \frac{A}{n+2} + \frac{n+1}{n+2}\left(1 + \frac{1 - m(Z_{1:n+2})}{n+1}\right)^{1_{\{Z_{1:n+2}<t\}}} \\
&\qquad \times \prod_{i=1}^{n+1}\left(1 + \frac{1 - m(Z_{i+1:n+2})}{n-i+2}\right)^{1_{\{Z_{i+1:n+2}<t\}}},
\end{aligned}$$

where the last inequality follows from the induction hypothesis applied to $Z_{j+1:n+2}$ rather than $Z_{j:n+1}$. Now, (2.4) equals

$$\begin{aligned}
&\frac{A}{n+2} + \frac{n+1}{n+2}\left(1 + \frac{1 - m(Z_{1:n+2})}{n+1}\right)^{1_{\{Z_{1:n+2}<t\}}}\prod_{j=2}^{n+2}\left(1 + \frac{1 - m(Z_{j:n+2})}{n-j+3}\right)^{1_{\{Z_{j:n+2}<t\}}} \\
&= A\left(1 + \frac{1 - m(Z_{1:n+2})}{n+2}\right)^{1_{\{Z_{1:n+2}<t\}}} = \varphi_{n+2}(t),
\end{aligned}$$

as desired. □

Again, the martingale convergence theorem guarantees that $\varphi_n(t)$ converges almost surely for each fixed t. We now determine its limit.

LEMMA 2.6. *For any t such that $H(t) < 1$,*

$$\lim_{n\to\infty}\varphi_n(t) = \exp\left\{\int_{-\infty}^{t}\frac{1 - m(x)}{1 - H(x)}H(dx)\right\} \quad \mathbb{P}\text{-}a.s.$$

PROOF. Let H_n be the empirical d.f. of $Z_1, \dots, Z_n$. Then $\varphi_n(t)$ can be written as

$$\varphi_n(t) = \exp\left\{\sum_{i=1}^{n}\ln(1 + x_i)\right\},$$

where

$$x_i = \frac{1 - m(Z_i)}{n + 1 - nH_n(Z_i)} 1_{\{Z_i < t\}}$$

and $0 \le x_i \le 1$. Using the fact that

$$-x^2/2 \le \ln(1 + x) - x \le 0 \quad \text{for any } x \ge 0,$$

we obtain

$$-\frac{1}{2} \sum_{i=1}^{n} x_i^2 \le \sum_{i=1}^{n} \ln(1 + x_i) - \sum_{i=1}^{n} x_i \le 0.$$

But

$$\sum_{i=1}^{n} x_i^2 = n^{-2} \sum_{i=1}^{n} \left(\frac{1 - m(Z_i)}{1 - H_n(Z_i) + n^{-1}} \right)^2 1_{\{Z_i < t\}}$$

$$\le n^{-2} \sum_{i=1}^{n} \left(1 - H_n(t) + n^{-1}\right)^{-2} = n^{-1}\left(1 - H_n(t) + n^{-1}\right)^{-2} \to 0,$$

with probability 1, if $H(t) < 1$. Hence

$$\left| \sum_{i=1}^{n} \ln(1 + x_i) - \sum_{i=1}^{n} x_i \right| \to 0 \quad \mathbb{P}\text{-a.s.}$$

Next, consider

$$\sum_{i=1}^{n} x_i = n^{-1} \sum_{i=1}^{n} \frac{1 - m(Z_i)}{1 - H_n(Z_i) + n^{-1}} 1_{\{Z_i < t\}} = \int_{-\infty}^{t} \frac{1 - m(x)}{1 - H_n(x) + n^{-1}} H_n(dx).$$

By Glivenko–Cantelli, $H_n \to H$ uniformly with probability 1. Furthermore, the SLLN yields

$$\int_{-\infty}^{t} \frac{1 - m(x)}{1 - H(x)} [H_n(dx) - H(dx)] \to 0 \quad \mathbb{P}\text{-a.s.}$$

Since $H(t) < 1$, we therefore easily obtain [cf., e.g., Shorack and Wellner (1986), pages 304]

$$\sum_{i=1}^{n} x_i \to \int_{-\infty}^{t} \frac{1 - m(x)}{1 - H(x)} H(dx) \quad \mathbb{P}\text{-a.s.},$$

as required. □

We are now in the position to determine S, the limit of $\mathbb{E} S_n$.

LEMMA 2.7. *Suppose that H is continuous. Then*

$$\lim_{n \to \infty} \mathbb{E} S_n \equiv S = \int \varphi(t) m(t) \exp\left\{ \int_{-\infty}^{t} \frac{1 - m(x)}{1 - H(x)} H(dx) \right\} H(dt). \tag{2.5}$$

PROOF. Assume $\varphi \geq 0$ w.l.o.g. Again, by Proposition 5-3-11 in Neveu (1975) and the Hewitt–Savage zero-one law,

$$g_n(t) \equiv \mathbb{E}\varphi_n(t) = \mathbb{E}[\varphi_n(t)|\mathscr{G}_\infty] \uparrow \exp\left\{\int_{-\infty}^{t} \frac{1-m(x)}{1-H(x)} H(dx)\right\}.$$

The assertion now follows from Lemma 2.4 and the monotone convergence theorem. □

To prove Theorem 1.1 for a continuous H, it remains to handle the right-hand side of (2.5). For this note that G is continuous on $(-\infty, t]$ provided that $H(t) < 1$. In fact, for each $a \leq t$ we have by continuity of H,

$$0 = \mathbb{P}(Z = a) \geq \mathbb{P}(Y < X, Y = a) = \mathbb{P}(a < X)\mathbb{P}(Y = a).$$

Assuming that $\mathbb{P}(Y = a) > 0$, we get $\mathbb{P}(a < X) = 0$ and therefore

$$H(a) = \mathbb{P}(Z \leq a) \geq \mathbb{P}(X \leq a) = 1,$$

a contradiction. For $H(t) < 1$ we therefore get by the definition of m and independence of X and Y that

$$\begin{aligned}\int_{-\infty}^{t} \frac{1-m(x)}{1-H(x)} H(dx) &= \int_{\{Z\leq t\}} \frac{1-m(Z)}{1-H(Z)}\, d\mathbb{P} = \int_{\{Z\leq t\}} \frac{1_{\{Y<X\}}}{1-H(Z)}\, d\mathbb{P} \\ &= \int_{\{Y\leq t\}} \frac{1}{1-G(Y)}\, d\mathbb{P} = -\ln(1-G(t)).\end{aligned}$$

Thus the right-hand side of (2.5) becomes

$$\begin{aligned}\int \varphi(Z)m(Z)/(1-G(Z))\, d\mathbb{P} &= \int \varphi(X) \frac{1_{\{X\leq Y,\, X<\tau_H\}}}{1-G(X)}\, d\mathbb{P} \\ &= \int_{(-\infty, \tau_H)} \varphi(x) F(dx).\end{aligned}$$

Clearly, this equals the right-hand side of (1.6), if H is continuous.

NOTE. Because of Lemma 2.2 and Lemma 2.1, S also admits the following expansion:

$$\begin{aligned}S &= \mathbb{E}S_1 + \sum_{n=1}^{\infty} \frac{1}{n+1} \mathbb{E}\Bigg\{\varphi(Z_{n+1:n+1})m(Z_{n+1:n+1})(1-m(Z_{n:n+1})) \\ &\qquad\qquad \times \prod_{j=1}^{n-1}\left(1 - \frac{m(Z_{j:n+1})}{n-j+1}\right)\Bigg\} \\ &= \int \varphi(z)m(z)H(dz) \\ &\quad + \sum_{n=1}^{\infty} \int\cdots\int_{z_1<z_2<\cdots<z_{n+1}} \varphi(z_{n+1})m(z_{n+1})(1-m(z_n)) \\ &\quad \times \prod_{j=1}^{n-1}(n-j+1-m(z_j))H(dz_1)\ldots H(dz_{n+1}).\end{aligned}$$

We are now going to consider the general case covered by Theorem 1.1. To this end, we have to construct a sequence $\{U_i\}_{i\geq 1}$ of uniform [0, 1]-random variables such that $Z_i = H^{-1}(U_i)$ with a specified dependence structure for the pair (U_i, δ_i). Recall A, the set of (simple) atoms of H.

LEMMA 2.8. *Assume that Z has d.f. H and δ is paired with Z as before. Let V be independent of Z and uniform* $-[0,1]$. *Define*

$$U = \begin{cases} H(Z), & \text{if } Z \notin A, \\ H(a-) + [H(a) - H(a-)]V, & \text{if } Z = a \text{ for } a \in A. \end{cases}$$

Then we have:

(i) $Z = H^{-1}(U)$ *with probability* 1,
(ii) U *is uniform* $-[0,1]$,
(iii) $\tilde{m}(u) \equiv \mathbb{P}(\delta = 1 | U = u) = m(H^{-1}(u))$ *for* $0 < u < 1$,

where as before

$$m(z) = \mathbb{P}(\delta = 1 | Z = z).$$

PROOF. Straightforward. □

Suppose now that the underlying probability space $(\Omega, \mathscr{A}, \mathbb{P})$ is rich enough to carry a sequence $V_1, V_2, \dots$ of independent uniform $-[0,1]$ random variables being also independent of the X's and Y's. Define each U_i as above, so that $Z_i = H^{-1}(U_i)$. Since F and G do not have common jumps, we have (after a moment of thought)

$$S_n = \sum_{i=1}^{n} W_{in}\varphi\big(H^{-1}(U_{i:n})\big).$$

In other words, the general case may be traced back to the uniform case, if we replace φ by $\varphi \circ H^{-1}$ and m by $\tilde{m}$. From what we have proved so far, we thus get $\mathbb{P}$-almost surely and in the mean that

$$\lim_{n\to\infty} S_n = \int_0^1 \varphi \circ H^{-1}(t)\tilde{m}(t) \exp\left\{\int_0^t \frac{1 - \tilde{m}(x)}{1 - x}\, dx\right\} dt. \tag{2.6}$$

After all, to prove Theorem 1.1, it remains to identify the right-hand side of (2.6). For this, introduce the so-called cumulative hazard function for G:

$$\Lambda(y) = \int_{-\infty}^{y} \frac{G(dx)}{1 - G(x-)}, \qquad y \in \mathbb{R}.$$

The continuous part of Λ is defined by

$$\Lambda_c(y) = \Lambda(y) - \sum_{a_i \leq y;\, a_i \in A_1} \Lambda\{a_i\},$$

with $\Lambda\{a_i\} = \Lambda(a_i) - \Lambda(a_i -)$. Here A_1 denotes the set of G-atoms. As a matter of fact, $1 - G$ may be easily represented in terms of Λ_c:

$$(2.7) \qquad 1 - G(y) = e^{-\Lambda_c(y)} \prod_{a_i \le y;\, a_i \in A_1} \frac{1 - G(a_i)}{1 - G(a_i -)}, \qquad y \in \mathbb{R}.$$

Observe that each H-atom is either an F- or a G-atom. Conversely, each F- or G-atom less than τ_H is an H-atom. Since F and G do not have jumps in common,

$$(2.8) \quad m(a) = 1 \quad \text{for } a \in A - A_1 \quad \text{and} \quad m(a) = 0 \quad \text{for } a \in A \cap A_1.$$

Since $\tilde{m} = m \circ H^{-1}$, it follows that the integral in (2.6) equals

$$\int_0^1 \varphi \circ H^{-1}(t) \tilde{m}(t) \exp\left\{ \int_0^{H(H^{-1}(t))} \frac{1 - \tilde{m}(x)}{1 - x}\, dx \right\} dt.$$

Fix $0 < t < 1$ with $\tilde{m}(t) > 0$. By (2.8), the x-integral then becomes

$$\begin{aligned} &- \sum_{a_i \le H^{-1}(t);\, a_i \in A \cap A_1} \ln \frac{1 - H(a_i)}{1 - H(a_i -)} + \int_{\{Z \le H^{-1}(t),\, Z \notin A\}} \frac{1_{\{Y < X\}}}{1 - H(Z)}\, d\mathbb{P} \\ &= - \sum_{a_i \le H^{-1}(t);\, a_i \in A \cap A_1} \ln \frac{1 - G(a_i)}{1 - G(a_i -)} \\ &\quad + \int_{\{Y \le H^{-1}(t),\, Y \notin A\}} (1 - G(Y))^{-1}\, d\mathbb{P}. \end{aligned}$$

If $H^{-1}(t)$ is an H-atom, then necessarily $a_i < H^{-1}(t) \le \tau_H$. If $H^{-1}(t)$ is not an H-atom, we obtain $H^{-1}(t) < \tau_H$. Conclude that the last sum equals

$$- \sum_{a_i \le H^{-1}(t);\, a_i \in A_1} \ln \frac{1 - G(a_i)}{1 - G(a_i -)} + \int_{\{Y \le H^{-1}(t),\, Y \notin A_1\}} (1 - G(Y))^{-1}\, d\mathbb{P}.$$

By (2.7), the right-hand side of (2.6) therefore turns out to be

$$\begin{aligned} \int_0^1 \varphi \circ H^{-1}(t) \tilde{m}(t) [1 - G(H^{-1}(t))]^{-1}\, dt &= \int \varphi(Z) 1_{\{X \le Y\}} [1 - G(Z)]^{-1}\, d\mathbb{P} \\ &= \int \varphi(X) 1_{\{X \le Y\}} [1 - G(X)]^{-1}\, d\mathbb{P}, \end{aligned}$$

which is the same as (1.6). This completes the proof of Theorem 1.1. □

Acknowledgments. The comments of a referee and Associate Editor led to an improvement in the presentation of Lemma 2.2.

REFERENCES

Aalen, O. O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6** 701–726.

Chen, Y. Y., Hollander, M. and Langberg, N. A. (1982). Small sample results for the Kaplan–Meier estimator. *J. Amer. Statist. Assoc.* **77** 141–144.

CSÖRGŐ, S. and HORVÁTH, L. (1983). The rate of strong uniform consistency for the product limit estimator. *Z. Wahrsch. Verw. Gebiete* **62** 411–426.

FÖLDES, A. and REJTŐ, L. (1981). A LIL type result for the product limit estimator. *Z. Wahrsch. Verw. Gebiete* **56** 75–86.

GILL, R. D. (1980). *Censoring and Stochastic Integrals. Math. Centre Tract* **124**. Math. Centrum, Amsterdam.

GU, M. G. and LAI, T. L. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann. Probab.* **18** 160–189.

LO, S. H. and SINGH, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representations. *Probab. Theory Related Fields* **71** 455–465.

MAURO, D. (1985). A combinatoric approach to the Kaplan–Meier estimator. *Ann. Statist.* **13** 142–149.

NEVEU, J. (1975). *Discrete-Parameter Martingales*. North-Holland, Amsterdam.

POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

ROSENBLATT, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.* **23** 470–472.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

STUTE, W. (1976). On a generalization of the Glivenko–Cantelli theorem. *Z. Wahrsch. Verw. Gebiete* **35** 167–175.

SUSARLA, V. and VAN RYZIN, J. (1980). Large sample theory for an estimator of the mean survival time from censored samples. *Ann. Statist.* **8** 1002–1016.

WANG, J.-G. (1987). A note on the uniform consistency of the Kaplan–Meier estimator. *Ann. Statist.* **15** 1313–1316.

WELLNER, J. A. (1985). A heavy censoring limit theorem for the product limit estimator. *Ann. Statist.* **13** 150–162.

YANG, S. S. (1977). General distribution theory of the concomitants of order statistics. *Ann. Statist.* **5** 996–1002.

MATHEMATICAL INSTITUTE
UNIVERSITY OF GIESSEN
ARNDTSTRASSE 2
W-6300 GIESSEN
GERMANY

DIVISION OF STATISTICS
469 KERR HALL
UNIVERSITY OF CALIFORNIA
DAVIS, CALIFORNIA 95616