The Annals of Statistics
1983, Vol. 11, No. 1, 49–58

LARGE SAMPLE BEHAVIOUR OF THE PRODUCT-LIMIT ESTIMATOR ON THE WHOLE LINE

BY RICHARD GILL

Mathematical Centre, Amsterdam

Weak convergence results are proved for the product-limit estimator on the whole line. Applications are given to confidence band construction, estimation of mean lifetime, and to the theory of q-functions. The results are obtained using stochastic calculus and in probability linear bounds for empirical processes.

1. Introduction. Let $X_1, \cdots, X_n$ be independent positive random variables with common continuous distribution function F. Independent of the X_i's, let $U_1, \cdots, U_n$ be also independent positive random variables with possibly noncontinuous and defective common distribution function G. The problem at hand is to make nonparametric inference on F based on the censored observations $(\tilde{X}_i, \delta_i)$, $i = 1, \cdots, n$, defined by

$$\tilde{X}_i = X_i \wedge U_i, \qquad \delta_i = I\{X_i \leq U_i\},$$

where $\wedge$ denotes minimum and $I\{\cdot\}$ is the indicator random variable of the specified event. Classically, F is estimated by the product-limit estimator $\hat{F}$, introduced by Kaplan and Meier (1958). Defining processes N and Y on $[0, \infty)$ by

$$N(t) = \#\{i : \tilde{X}_i \leq t, \delta_i = 1\}, \quad Y(t) = \#\{i : \tilde{X}_i \geq t\},$$

then $\hat{F}$ is given by

$$1 - \hat{F}(t) = \prod_0^t \left\{1 - \frac{dN(s)}{Y(s)}\right\}.$$

Define also the random time T by

$$T = \max_i \tilde{X}_i$$

and for any process W define the *stopped process* W^T by

$$W^T(t) = W(t \wedge T).$$

Define the important *Kaplan-Meier process* Z by

$$Z = n^{1/2} \frac{\hat{F} - F}{1 - F}.$$

Note that $\hat{F}^T = \hat{F}$, and that if $\Delta N(T) = N(T) - N(T-) = 0$, i.e. the largest observation is censored, then $\hat{F}(T) < 1$ almost surely.

Let H be the distribution function of the $\tilde{X}_i$'s, given by

$$(1 - H) = (1 - F)(1 - G),$$

and define (possibly infinite) times τ_F, τ_G and τ_H by

$$\tau_F = \sup\{t : F(t) < 1\}$$

Received September 1981; revised June 1982.

AMS 1980 *subject classification.* Primary 62G05, 62G15; secondary; 60F05, 62N05, 62P10.

Key words and phrases. Product-limit estimator, Kaplan-Meier estimator, random censorship, survival data, confidence bands, mean life-time, counting processes, martingales, stochastic integrals, weak convergence, in probability linear bounds, q-functions, censored data.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.* www.jstor.org

etc. In the sequel, by $\int_0^t$ we mean integration over the interval $(0, t]$. When we drop the limits of integration, we implicitly define a function or process $t \to \int_0^t(\cdot)$. Define also some continuous, nonnegative, nondecreasing functions Λ, C and K by

$$\Lambda(t) = \int_0^t \frac{dF(s)}{1 - F(s-)},$$

$$C(t) = \int_0^t \frac{dF(s)}{\{1 - F(s-)\}^2\{1 - G(s-)\}} = \int_0^t \frac{d\Lambda(s)}{\{1 - H(s-)\}}$$

and

$$K(t) = \frac{C(t)}{1 + C(t)},$$

where $K(t) = 1$ if $C(t) = \infty$. Note that $\Lambda(\tau_F) = \infty$, and that $C(\tau_H) = \infty$ and $K(\tau_H) = 1$ if $\tau_G \geq \tau_F$. When $\tau_G \leq \tau_F$ it is both possible that $C(\tau_H) = \infty$ and $C(\tau_H) < \infty$. (We write, e.g., $1 - F(s-)$ even though F is continuous to indicate the right extension for noncontinuous F.)

Let B be a standard Brownian motion on $[0, \infty)$ and let B^0 be a Brownian bridge on $[0, 1]$. Assuming G to be continuous, Breslow and Crowley (1974) proved a result on weak convergence of $n^{1/2}(\hat{F} - F)$ equivalent to the following theorem. The two ways in which we state it derive from Efron (1967) and Hall and Wellner (1980).

THEOREM 1.1. *For any τ such that $H(\tau-) < 1$*

$$Z = n^{1/2}\left(\frac{\hat{F} - F}{1 - F}\right) \to_{\mathscr{D}} B(C) \quad \text{in} \quad D[0, \tau] \quad \text{as} \quad n \to \infty,$$

or equivalently

$$n^{1/2}\frac{1 - K}{1 - F}(\hat{F} - F) \to_{\mathscr{D}} B^0(K) \quad \text{in} \quad D[0, \tau] \quad \text{as} \quad n \to \infty.$$

Note that $B(C)$ is a continuous Gaussian martingale, zero at time zero, with covariance function

$$\text{cov}[B(C(s)), B(C(t))] = C(s) \wedge C(t) = C(s \wedge t).$$

Breslow and Crowley's (1974) proof of Theorem 1.1 was based on approximating $n^{1/2}(\hat{F} - F)$ by an expression linear in the empirical processes N/n and Y/n, and then applying standard results on weak convergence of empirical distribution functions. Though this sounds straightforward, the proof was unavoidably complex. (Beware too of a misprint in the proof of the central theorem, where the expression three lines from below on page 447 should read $2\varepsilon + 2K\rho_T(X(1 - F)^{-2}, 0)\rho_T(\tilde{F}_N^e, \tilde{F})$.) The simple form of the limiting distribution only appears after long calculations in which complicated expressions surprisingly cancel out.

Since F and also K can be uniformly consistently estimated on $[0, \tau]$ for any τ such that $H(\tau-) < 1$, the theorem gives two obvious ways of constructing confidence bands for F on $[0, \tau]$ based on the known distributions of $\sup_{x \leq C(\tau)} |B(x)|$ and $\sup_{x \leq K(\tau)} |B^0(x)|$ respectively; see Gill (1980a) and Hall and Wellner (1980).

Clearly there is hope that the Brownian bridge version of Theorem 1.1 could be extended to $[0, \tau_H]$ giving confidence bands for F on the largest possible interval. When $K(\tau_H) = 1$ one would be able to use the distribution of $\sup_{x \leq 1} |B^0(x)|$, leading to simpler computations too. Finally, when there is no censoring ($G = 0$), such bands would reduce to the usual Kolmogorov bands for F based on the empirical distribution function $\hat{F}$. Such a result was conjectured to hold by Hall and Wellner (1980) and motivated the work presented here. Part of the conjecture remains an open question, but the techniques used will turn out to be of wider application.

Let us first define $\hat{K}$ by

$$1 - \hat{K} = 1/(1 + \hat{C}) \quad \text{and} \quad \hat{C}(t) = \int_0^t \frac{n dN(s)}{Y(s)(Y(s) - 1)}$$

(so $\hat{K}(T) = 1$ and $\hat{C}(T) = \infty$ almost surely if $\Delta N(T) = 1$). Then we shall prove

THEOREM 1.2.

(i)
$$n^{1/2}\left\{\frac{1-K}{1-F}(\hat{F} - F)\right\}^T \to_{\mathscr{D}} B^0(K) \quad \text{in} \quad D[0, \tau_H] \quad \text{as} \quad n \to \infty.$$

$$n^{1/2}\left\{\frac{1-\hat{K}}{1-\hat{F}}(\hat{F} - F)\right\}^T \to_{\mathscr{D}} B^0(K) \quad \text{in} \quad D[0, \tau_H] \quad \text{as} \quad n \to \infty,$$

(ii)

provided that

$$\int_0^{\tau_H} \frac{dF(t)}{1 - G(t-)} < \infty. \tag{1.1}$$

(*If* $\hat{F}(T) = 1$ *we interpret* $(1 - \hat{K})/(1 - \hat{F})$ *in the point* T *as equal to its value in* $T-$.)

Whether or not part (ii) of Theorem 1.2 holds without the condition (1.1) we do not know. Note that

$$1 - G \le \frac{1-K}{1-F} \le 1 \tag{1.2}$$

and that $(1 - K)/(1 - F)$ is nonincreasing. The same relationship holds between $\hat{K}$, $\hat{F}$ and $\hat{G}$, where $\hat{G}$ is the product-limit estimator of the censoring distribution G. These facts follow from the equality

$$d\left(\frac{1-F}{1-K}\right) = \int \frac{dG}{(1-F)(1-G)(1-G_-)} dF.$$

When there is no censoring $\hat{F}$ becomes the ordinary empirical distribution function of the X_i's, so apart from being "stopped at T" the theorems reduce to the classical result on weak convergence of the empirical distribution function. Theorem 1.2 (ii) gives asymptotic confidence bands for F on the random interval $[0, T]$ provided censoring is not too heavy.

The result does have some practical importance. One would be tempted to apply Theorem 1.1 after choosing τ such that $Y(\tau)$ is reasonably large. So in fact τ will not be fixed in advance. Moreover, what is felt to be "reasonably large" may well be numbers as small as 5 or 10. However, Theorem 1.1 has the implicit condition $Y(\tau) \to_{\mathscr{P}} \infty$. Thus it is not obvious that Theorem 1.1 will yield accurate approximations when applied in such a way.

In proving Theorem 1.2 it turns out that a technique is being used which leads to quite general results on *weighted* or *integrated* Kaplan-Meier processes under a natural condition.

In the next section we therefore state and prove such a general theorem, using the fact that Z^T is a martingale. In Section 3 we show that Theorem 1.2 is a corollary of this, and go on to present other applications to estimating mean lifetime and to the theory of q-functions. All our results can be easily extended to nonidentically distributed censoring variables, and in particular therefore to the model of fixed censorship. Also the results can be extended to noncontinuous F. We refer to Gill (1980a) for such a fuller treatment, where part of the material here has already appeared together with a study of two-sample tests. A brief survey of the necessary martingale and counting process theory can be found there, or in Aalen (1978). We depend heavily on this and on Aalen and Johansen (1978).

2. Main theorem

THEOREM 2.1. *Let h be a nonnegative continuous nonincreasing function on $[0, \tau_H]$ such that*

$$\int_0^{\tau_H} h(t)^2\, dC(t) < \infty. \tag{2.1}$$

Then the processes $(hZ)^T$, $(\int hdZ)^T$ and $(\int Zdh)^T$ converge jointly in $D[0, \tau_H]$ in distribution to processes $hZ^{(\infty)}$, $\int hdZ^{(\infty)}$ and $\int Z^{(\infty)}\, dh$ respectively, where

$$Z^{(\infty)} = B(C) \tag{2.2}$$

and

$$hZ^{(\infty)} = \int hdZ^{(\infty)} + \int Z^{(\infty)}\, dh. \tag{2.3}$$

REMARK. 2.2. When $C(\tau_H) = \infty$, the limiting processes here are interpreted to be equal in the point τ_H to their limits as $\tau \uparrow \tau_H$, which do exist as we shall now show. In fact the limit of $hZ^{(\infty)}$ is zero. Also we must discuss in any case what we mean by the process $\int hdZ^{(\infty)}$, which cannot be defined by pathwise Lebesgue-Stieltjes integration. Taking up the latter point first, we note that $\int hdZ^{(\infty)}$ can be defined on $[0, \tau_H)$ either by (2.3) or as a stochastic integral in the sense of Meyer (1976). By Gill (1980b) Lemma 5, 2nd part, the two definitions coincide. By (2.1), $\int hdZ^{(\infty)}$ is a square integrable martingale on $[0, \tau_H)$ which can be extended by taking limits to $[0, \tau_H]$. So it remains to show that either $hZ^{(\infty)}$ or $\int Z^{(\infty)}\, dh$ also has a limit almost surely as $t \to \tau_H$ (in which case both processes do). Now we can write

$$h(t)^2 C(t) = \int_0^{\tau_H} h^2(t) I_{[0,t]}(s)\, dC(s).$$

So when $C(\tau_H) = \infty$ and consequently by (2.1) $h(t) \downarrow 0$ as $t \uparrow \tau_H$, by dominated convergence

$$h(t)^2 C(t) \to 0 \quad \text{as} \quad t \uparrow \tau_H. \tag{2.4}$$

It follows immediately that when $C(\tau_H) = \infty$

$$h(t) Z^{(\infty)}(t) \to_{\mathscr{P}} 0 \quad \text{as} \quad t \uparrow \tau_H; \tag{2.5}$$

we must extend this to an a.s. result. By the Birnbaum-Marshall inequality (Birnbaum and Marshall, 1961) applied to the submartingale $\{Z^{(\infty)} - Z^{(\infty)}(t)\}^2$ and the nonincreasing function h^2 on $[t, \tau_H)$, we have

$$\mathscr{P}(\sup_{s \in [t, \tau_H)} [\{Z^{(\infty)}(s) - Z^{(\infty)}(t)\} h(s)]^2 \ge \varepsilon) \le \frac{1}{\varepsilon} \int_t^{\tau_H} h(s)^2\, dC(s).$$

Therefore

$$\mathscr{P}[\sup_{[t,\tau_H)} \{hZ^{(\infty)} - h(t) Z^{(\infty)}(t)\}^2 \ge 4\varepsilon] \le \frac{1}{\varepsilon} \int_t^{\tau_H} h^2\, dC + P[\{h(t) Z^{(\infty)}(t)\}^2 \ge \varepsilon]$$

$$\le \frac{1}{\varepsilon} \left(\int_t^{\tau_H} h^2\, dC + h(t)^2 C(t) \right).$$

Now let $\varepsilon_m > 0$ and $\delta_m > 0$ satisfy $\varepsilon_m \downarrow 0$ and $\sum_m \delta_m < \infty$. For each m we can now by (2.1) and

(2.4) find a t_m such that

$$\mathscr{P}[\sup_{[t_m,\tau_H)}\{hZ^{(\infty)} - h(t_m)Z^{(\infty)}(t_m)\}^2 \geq 4\varepsilon_m] \leq \delta_m.$$

Thus by the Borel-Cantelli lemma $hZ^{(\infty)}$ converges almost surely, and by (2.5) the limit is zero.

In order to prove Theorem 2.1 we first present a sequence of lemmas including a proof of Theorem 1.1.

LEMMA 2.3. *Define M by*

$$M(t) = N(t) - \int_0^t Y(s)\,d\Lambda(s).$$

Then M is a square integrable martingale on $[0, \infty]$ *with predictable variation process* $\langle M, M\rangle$ *given by*

$$\langle M, M\rangle(t) = \int_0^t Y(s)\,d\Lambda(s).$$

PROOF. See Aalen (1976), Section 5C, or Gill (1980a). □

LEMMA 2.4. *For all t*

$$\frac{1-\hat{F}(t)}{1-F^T(t)} = 1 - \int_0^t \frac{1-\hat{F}(s-)}{1-F^T(s)}\frac{dM(s)}{Y(s)}.$$

PROOF. See Aalen and Johansen (1978), or Gill (1980a). □

Since we may define $1/Y = 0$ on (T, ∞), the integrand in Lemma 2.4, $(1 - \hat{F}_-)/(1 - F^T)Y$, is a bounded predictable process on $[0, \tau]$ for any $\tau < \tau_F$. This gives us:

LEMMA 2.5. $((1-\hat{F})/(1-F))^T$ *and* $Z^T = n^{1/2}\{1 - ((1-\hat{F})/(1-F))^T\}$ *are square integrable martingales on* $[0, \tau]$ *for any* $\tau < \tau_F$, *and* $< Z^T, Z^T > (t) = \int_0^{t\wedge T} (1-\hat{F}_-)^2/(1-F)^2\, n/Y\, d\Lambda$ *for all t.*

LEMMA 2.6. *For any* $\beta \in (0, 1)$, $\mathscr{P}[1 - \hat{F}(t) \leq \beta^{-1}\{1 - F(t)\}\,\forall t \leq T] \geq 1 - \beta$.

PROOF. See Gill (1980a). (Essentially, apply Doob's inequality to the nonnegative martingale $(1-\hat{F})/(1-F^T)$.) □

LEMMA 2.7. *For any* $\beta \in (0, 1)$, $\mathscr{P}[Y(t)/n \geq \beta\{1 - H(t-)\}\,\forall t \leq T] \geq 1 - e(1/\beta)e^{-1/\beta}$.

PROOF. See Wellner (1978), Remark 1(ii). (The fact that H need not be continuous is of no consequence.) □

LEMMA 2.8. *For any τ such that* $H(\tau-) < 1$, $\sup_{t\leq\tau}|\hat{F}(t) - F(t)| \to_{\mathscr{P}} 0$ *as* $n \to \infty$.

PROOF. Since $\mathscr{P}(T < \tau) \to 0$ as $n \to \infty$, it suffices by Lemma 2.4 to show that for any $\varepsilon > 0$

$$\mathscr{P}\left[\sup_{t\leq\tau}\left|\int_0^t \frac{1-\hat{F}(s-)}{1-F^T(s)}\frac{dM(s)}{Y(s)}\right| > \varepsilon\right] \to 0 \quad \text{as } n \to \infty. \tag{2.6}$$

By Lemma 2.3, $n^{-1/2}Z^T = \int \frac{1-\hat{F}_-}{1-F}\frac{dM}{Y}$ is a square integrable martingale on $[0, \tau]$ with predictable variation process equal at time t to

$$\int_0^{t\wedge T} \frac{\{1-\hat{F}(s-)\}^2}{\{1-F(s)\}^2}\frac{d\Lambda(s)}{Y(s)}.$$

Therefore, by Lenglart's (1977) inequality applied to $(n^{-1/2}Z^T)^2$ and $n^{-1}\langle Z^T, Z^T\rangle$, it follows that for any $\eta > 0$, the left-hand side of (2.6) is bounded by

$$\frac{\eta}{\varepsilon^2} + \mathscr{P}\left[\int_0^{\tau\wedge T} \frac{\{1-\hat{F}(s-)\}^2}{\{1-F(s)\}^2}\frac{1}{Y(s)}\,d\Lambda(s) > \eta\right] \le \frac{\eta}{\varepsilon^2} + \mathscr{P}\left[\frac{\Lambda(\tau)}{\{1-F(\tau)\}^2}\frac{1}{Y(\tau\wedge T)} > \eta\right].$$

Since $Y(\tau \wedge T)/n \rightarrow_{\mathscr{P}} 1 - H(\tau-) > 0$ as $n \rightarrow \infty$, we now easily see that (2.6) holds. □

It is an open question as to whether the supremum in Lemma 2.8 may be taken over $t \le T$.

LEMMA 2.9. *Let h be a continuous, nonnegative, and nonincreasing function and let Z be a semimartingale, zero at time zero. Then for all τ*

$$\sup_{0\le t\le\tau} h(t)\,|Z(t)| \le 2\sup_{0\le t\le\tau}\left|\int_0^t h(s)\,dZ(s)\right|.$$

PROOF. Note that $\int h\,dZ$ can equivalently be interpreted here as a stochastic integral, a pathwise integral when it exists as such, and by formally integrating by parts; cf. Remark 2.2. Define $U(t) = \int_0^t h(s)\,dZ(s)$ so that for t such that $h(t) > 0$,

$$Z(t) = \int_0^t \frac{dU(s)}{h(s)} = \frac{U(t)}{h(t)} - \int_0^t U(s-)\,d\left(\frac{1}{h(s)}\right)$$

$$= \int_0^t \{U(t) - U(s-)\}\,d\left(\frac{1}{h(s)}\right).$$

Thus

$$|h(t)Z(t)| = \left|\int_0^t \{U(t) - U(s-)\}\,d\left(\frac{h(t)}{h(s)}\right)\right| \le 2\sup_{0\le s\le t}|U(s)|\left\{1 - \frac{h(t)}{h(0)}\right\}. \qquad \square$$

PROOF OF THEOREM 1.1. Recall that τ satisfies $H(\tau-) < 1$. Since $\mathscr{P}(T < \tau) \rightarrow 0$ as $n \rightarrow \infty$, it suffices to show that

$$Z^T = n^{1/2}\left(\frac{\hat{F} - F}{1 - F}\right)^T = \int \left(\frac{1-\hat{F}_-}{1-F^T}\right)\frac{n^{1/2}\,dM}{Y} \rightarrow_{\mathscr{D}} Z^{(\infty)}$$

in $D[0, \tau]$ as $n \rightarrow \infty$. By Lemma 2.5 above and Theorem V.1 of Rebolledo (1980) this is the case if for each $t \le \tau$

$$\langle Z^T, Z^T\rangle(t) = \int_0^{t\wedge T} \frac{\{1-\hat{F}(s-)\}^2}{\{1-F^T(s)\}^2}\frac{n}{Y(s)}\,d\Lambda(s) \rightarrow_{\mathscr{P}} C(t)$$

and for each $\varepsilon > 0$

$$\int_0^{\tau\wedge T} \frac{\{1-\hat{F}(s-)\}^2}{\{1-F^T(s)\}^2}\frac{n}{Y(s)}\,I\left\{\frac{1-\hat{F}(s-)}{1-F^T(s)}\frac{n^{1/2}}{Y(s)} > \varepsilon\right\}\,d\Lambda(s) \rightarrow_{\mathscr{P}} 0.$$

(Note that by continuity of F we are in the quasi-left-continuous case in which Rebolledo's (1980) strong and weak ARJ(2) conditions coincide). By Lemma 2.8 and the Glivenko-Cantelli theorem for Y/n both conditions are easily seen to hold. □

PROOF OF THEOREM 2.1. By Theorem 1.1 we certainly have weak convergence on $[0, \tau]$ for any τ such that $H(\tau-) < 1$. Also by Remark 2.2 the limiting processes do exist on $[0, \tau_H]$ and are continuous in τ_H. Thus (see Billingsley, 1968, Theorem 4.2) it suffices to prove "tightness at τ_H", i.e. we must show

$$\lim_{\tau\uparrow\tau_H} \limsup_{n\to\infty} \mathscr{P}[\sup_{\tau\le t\le T} | h(t)Z(t) - h(\tau)Z(\tau) | > \varepsilon] = 0 \qquad \forall \varepsilon > 0 \tag{2.7}$$

and

$$\lim_{\tau\uparrow\tau_H} \limsup_{n\to\infty} \mathscr{P}\left[\sup_{\tau\le t\le T} \left| \int_\tau^t h(s)\, dZ(s) \right| > \varepsilon \right] = 0 \qquad \forall \varepsilon > 0. \tag{2.8}$$

(Note that by the equality $\int Z\, dh = hZ - \int h\, dZ$ the corresponding result for $\int Z\, dh$ does not need to be explicitly verified.)

Now

$$\sup_{\tau\le t\le T} | h(t)Z(t) - h(\tau)Z(\tau) | \le \sup_{\tau\le t\le T} | h(t)(Z(t) - Z(\tau)) | + | (h(\tau) - h(\tau_H))Z(\tau) |.$$

We already know that $Z(\tau) \to_{\mathscr{D}} Z^{(\infty)}(\tau)$ as $n \to \infty$ so that

$$\limsup_{n\to\infty} \mathscr{P}\, [| \{h(\tau_H) - h(\tau)\}Z(\tau) | > \varepsilon] \le \frac{\{h(\tau) - h(\tau_H)\}^2 C(\tau)}{\varepsilon^2}.$$

If $C(\tau_H) < \infty$ this quantity converges trivially to zero as $\tau \uparrow \tau_H$. However, if $C(\tau_H) = \infty$ we must have $h(\tau_H) = 0$ by (2.1) and convergence to zero follows from (2.4). By Lemma 2.9

$$\sup_{\tau\le t\le T} | h(t)\{Z(t) - Z(\tau)\} | \le 2 \sup_{\tau\le t\le T} \left| \int_\tau^t h(s)\, dZ(s) \right|.$$

Thus (2.8) implies (2.7).

Since h is, as a process, predictable and bounded, and $Z^T - Z^T(\tau)$ is a square integrable martingale on $[\tau, \tau']$ for each τ' such that $H(\tau'-) < 1$, we have for any $\eta > 0$ by the inequality of Lenglart (1977) (cf. the proof of Lemma 2.8)

$$\begin{aligned}
&\mathscr{P}\left[\sup_{\tau\le t\le \tau'\wedge T} \left| \int_\tau^t h(s)\, dZ(s) \right| > \varepsilon \right] \\
&\quad \le \frac{\eta}{\varepsilon^2} + \mathscr{P}\left[\int_\tau^{\tau'\wedge T} \frac{\{h(s)\}^2 \{1 - \hat{F}(s-)\}^2}{\{1 - F(s)\}^2} \frac{n}{Y(s)}\, d\Lambda(s) > \eta \right] \\
&\quad \le \frac{\eta}{\varepsilon^2} + \beta + e(1/\beta)e^{-1/\beta} + \mathscr{P}\left[\int_\tau^{\tau'} \frac{\beta^{-3}\{h(s)\}^2\, d\Lambda(s)}{1 - H(s-)} > \eta \right]
\end{aligned}$$

for any $\beta \in (0, 1)$ by Lemmas 2.6 and 2.7. Letting $\tau' \uparrow \tau_H$ (or choosing $\tau' = \tau_H$ if $H(\tau_H-) < 1$) and choosing

$$\eta = \int_\tau^{\tau_H} \beta^{-3} h(s)^2\, dC(s),$$

we obtain

$$\mathscr{P}\left[\sup_{\tau\le t\le T} \left| \int_\tau^t h(s)\, dZ(s) \right| > \varepsilon \right] \le \beta^{-3}\varepsilon^{-2} \int_\tau^{\tau_H} h(s)^2\, dC(s) + \beta + e(1/\beta)e^{-1/\beta}.$$

By (2.1), and since β was arbitrary, this gives us (2.8). □

3. Applications. We first prove Theorem 1.2.

PROOF OF THEOREM 1.2. Part (i): Choose $h = 1 - K$ in Theorem 2.1. Then (2.1) holds because

$$\int_0^{\tau_H} \{1 - K(s)\}^2 \, dC(s) = \int_0^{\tau_H} \frac{dC(s)}{\{1 + C(s)\}^2} = 1 - \frac{1}{1 + C(\tau_H)} \le 1 < \infty.$$

Part (ii): Choosing $h = 1 - F$ in Theorem 2.1, we see that

$$n^{1/2}(\hat{F} - F)^T \to_{\mathscr{D}} (1 - F)Z^{(\infty)} \quad \text{in} \quad D[0, \tau_h], \tag{3.1}$$

provided that

$$\int_0^{\tau_H} \{1 - F(s)\}^2 \, dC(s) = \int_0^{\tau_H} \frac{dF(s)}{1 - G(s-)} < \infty;$$

i.e. provided that (1.1) holds. Now straightforward arguments show that

$$\sup_{0 \le t \le \tau} | \hat{K}(t) - K(t) | \to_{\mathscr{P}} 0 \qquad \text{as } n \to \infty$$

for any τ such that $H(\tau-) < 1$. Therefore, by Lemma 2.8 we certainly have weak convergence of

$$n^{1/2} \left\{ \frac{1 - \hat{K}}{1 - \hat{F}} (\hat{F} - F) \right\}^T \qquad \text{in } D[0, \tau]$$

for any τ such that $H(\tau-) < 1$. Thus to prove part (ii) it remains only to prove "tightness at τ_H" as in Theorem 2.1. Since $(1 - \hat{K})/(1 - \hat{F}) \le 1$ (*cf.* (1.2)), this follows from (3.1). □

Next we consider estimation of mean lifetime $\int_0^\infty t \, dF(t) = \int_0^\infty \{1 - F(t)\} \, dt = \mu(\infty)$, which we suppose here to be finite. Many authors mention this problem but only Susarla and Van Ryzin (1980) achieve any really general result. Even so, they are obliged to work with an estimator $\int_0^M \{1 - \hat{F}(t)\} \, dt$, where $M = M_n \uparrow \infty$ is a sequence of constants depending on the unknown F and G in a complicated way. We shall consider the estimator $\int_0^T \{1 - \hat{F}(t)\} \, dt$ and obtain a more general result under a natural condition.

Define functions μ and $\bar{\mu}$ and a process $\hat{\mu}$ by

$$\mu(t) = \int_0^t \{1 - F(s)\} \, ds \quad \text{and} \quad \bar{\mu}(t) = \mu(\infty) - \mu(t),$$

$$\hat{\mu}(t) = \int_0^t \{1 - \hat{F}(s)\} \, ds.$$

Note that

$$n^{1/2}(\hat{\mu} - \mu) = -\int n^{1/2} \left(\frac{\hat{F} - F}{1 - F} \right) (1 - F) \, dt = -\int Z \, d\bar{\mu}.$$

Thus we obtain immediately from Theorem 2.1:

THEOREM 3.1. *Suppose*

$$\int_0^{\tau_H} \bar{\mu}^2 \, dC < \infty. \tag{3.2}$$

Then

$$n^{1/2}(\hat{\mu} - \mu)^T \to_{\mathscr{D}} \int Z^{(\infty)} \, d\bar{\mu} \quad \text{in} \quad D[0, \tau_H].$$

COROLLARY 3.2. *Supppose* (3.2) *holds and furthermore*

$$n^{1/2}\,\bar{\mu}(T) \rightarrow_{\mathscr{P}} 0 \qquad \text{as } n \rightarrow \infty. \tag{3.3}$$

Then

$$n^{1/2}\{\hat{\mu}(T) - \mu(\infty)\} \rightarrow_{\mathscr{D}} -\int_0^{\tau_H} Z^{(\infty)}(s)\, d\bar{\mu}(s) = \int_0^{\tau_H} \bar{\mu}(s)\, dZ^{(\infty)}(s) =_{\mathscr{D}} N\left(0, \int_0^{\tau_H} \bar{\mu}(s)^2\, dC(s)\right).$$

PROOF. Under (3.3) we are in the situation $\tau_H = \tau_F$ and $C(\tau_H) = \infty$. Also $T \rightarrow_{\mathscr{P}} \tau_H$ as $n \rightarrow \infty$. Thus the corollary follows by Remark 2.2. □

REMARK 3.3. Suppose F is a distribution function whose residual mean life-time function $\bar{\mu}/(1 - F)$ is bounded. This covers in particular all increasing hazard rate distributions. Then (3.2) is easily seen to hold if (1.1) does. This is true in particular if

$$(1 - G) \geq c(1 - F)^{\beta} \qquad \text{close to } \tau_F \tag{3.4}$$

for some constants $c > 0$ and $\beta < 1$; i.e. when the censoring distribution is lighter in the tail than the distribution of interest.

Some straightforward calculations also show that (3.3) holds if for some $0 < \alpha < 2$ and $c' > 0$ we also have

$$(1 - H) \geq c'\bar{\mu}^{\alpha} \qquad \text{close to } \tau_F. \tag{3.5}$$

This in turn is implied by (3.4) and boundedness of $\bar{\mu}/(1 - F)$.

Finally, we sketch an application to q-functions (cf. Pyke and Shorack, 1968, Theorem 2.1, and Wellner, 1977).

THEOREM 3.4. *Suppose q is a continuous function on* $[0, 1]$ *which is positive on* $(0, 1)$, *symmetric about* ½, *nondecreasing on* $[0, ½]$, *and such that*

$$\int_0^1 \frac{dt}{\{q(t)\}^2} < \infty \tag{3.6}$$

and $(1 - t)/q(t)$ *is nonincreasing close to* 1. *Then*

$$n^{1/2}\left\{\frac{1}{q(K)}\left(\frac{1-K}{1-F}\right)(\hat{F} - F)\right\}^{T} \rightarrow_{\mathscr{D}} \frac{B^0(K)}{q(K)}$$

in $D[0, \tau_H]$ *as* $n \rightarrow \infty$.

PROOF. In Theorem 2.1 we only needed h to be nonincreasing in the neighbourhood of τ_H; and this was only needed in the case $C(\tau_H) = \infty (K(\tau_H) = 1)$. Thus by (3.6) and Theorem 2.1 with $h = (1 - K)/q(K)$ we have weak convergence on $[\tau, \tau_H]$ for any τ such that $K(\tau) > 0$. So we only need further to prove "tightness near zero". Exactly the same arguments as in Theorem 2.1 can be used again to give this result too. □

Acknowledgment. I am very grateful to Jon Wellner for the stimulating discussions which led to this paper being written.

REFERENCES

AALEN, O. O. (1976). *Statistical Inference for a Family of Counting Processes.* Institute of Mathematical Statistics, University of Copenhagen.

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

AALEN, O. O. and JOHANSEN, S. (1978). An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scand. J. Statist.* **5** 141–150.

BILLINGSLEY, P. (1968). *Weak Convergence of Probability Measures.* Wiley, New York.

BIRNBAUM, Z. W. and MARSHALL, A. W. (1961). Some multivariate Chebyshev inequalities with extensions to continuous parameter processes. *Ann. Math. Statist.* **32** 687–703.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. *Ann. Statist.* **2** 437–453.

EFRON, B. (1967). The two sample problem with censored data. Proc. 5th Berkeley Symp. **4** 831–853.

GILL, R. D. (1980a). *Censoring and Stochastic Integrals.* Mathematical Centre Tracts **124**. Mathematisch Centrum, Amsterdam.

GILL, R. D. (1980b). Nonparametric estimation based on censored observations of a Markov renewal process. *Z. Wahrsch. verw. Gebiete* **53** 97–116.

HALL, W. J. and WELLNER, J. A. (1980). Confidence bands for a survival curve from censored data. *Biometrika* **67** 133–143.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

LENGLART, E. (1977). Relation de domination entre deux processus. *Ann. Inst. Henri Poincaré* **13** 171–179.

MEYER, P. A. (1978). Un Cours sur les Intégrales Stochastiques. Séminaire de Probabilités VI, *Lecture Notes in Mathematics* **511** 245–400 Springer, Berlin.

PYKE, R. and SHORACK, G. R. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist.* **39** 755–771.

REBOLLEDO, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. verw. Gebiete* **51** 269–286.

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.

WELLNER, J. A. (1977), A martingale inequality for the empirical process. *Ann. Probab.* **5** 303–308.

WELLNER, J. A. (1978). Limit theory for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. verw. Gebiete* **45** 73–88.

MATHEMATICAL CENTRE
KRUISLAAN 413
1098 SJ AMSTERDAM
THE NETHERLANDS