The Annals of Statistics
1981, Vol. 9, No. 4, 853–860

TESTING WITH REPLACEMENT AND THE PRODUCT LIMIT ESTIMATOR

By R. D. Gill

Mathematical Centre, Amsterdam

Let $X_1, X_2, \cdots$ be a sequence of i.i.d. nonnegative rv's with nondegenerate df F. Define $\tilde{N}(t) = \#\{j: X_1 + \cdots + X_j \le t\}$. In "testing with replacement" (also known as "renewal testing") n independent copies of $\tilde{N}$ are observed each over the time interval $[0, \tau]$ and we are interested in nonparametric estimation of F based on these observations. We prove consistency of the product limit estimator as $n \to \infty$ for arbitrary F, and weak convergence in the case of integer valued X_i. We state the analogue of this result for continuous F and briefly discuss the similarity of our results with those for the product limit estimator in the model of "random censorship."

1. Introduction. Let $X_1, X_2, \cdots$ be nonnegative i.i.d. random variables (rv's) with nondegenerate right continuous distribution function (df) F. Define $T_0 = 0$, $T_j = \sum_{i=1}^{j} X_i$, $j = 1, 2, \cdots$ and $\tilde{N}(t) = \#\{j \ge 1: T_j \le t\}$. We consider nonparametric estimation of F based on the first n of an infinite sequence of independent realisations of $\tilde{N}$, each observed over the fixed time interval $[0, \tau]$, $0 < \tau < \infty$. This situation, known as testing with replacement or renewal testing, might arise when light bulbs are lifetested in a large number n of sockets, failed bulbs being replaced immediately by new ones.

We call an X_j such that $T_j \le \tau$ an uncensored observation: if $T_{j-1} \le \tau < T_j$ we call X_j censored, for we only observe in this case that X_j takes an unknown value strictly greater than $\tau - T_{j-1}$. An X_j for which $T_{j-1} > \tau$ is not observed at all.

If $n \to \infty$ the empirical df based on the uncensored observations is inconsistent. An obvious alternative estimator of F is the product limit estimator of Kaplan and Meier (1958) which also takes account of the censored observations; it is introduced in Section 3 of this paper. Crow and Shimi (1972) also consider nonparametric estimation of F for this type of lifetesting and give references to examples of parametric analysis of renewal testing in the literature. They, however, work in the case in which F is known to have a monotone failure rate.

In the next section we use renewal theory to establish a relation between the expected values of processes related to the empirical distribution functions of the censored and uncensored observations in the case of a single renewal process, $n = 1$. In Section 3 we prove strong consistency as $n \to \infty$ of the product limit estimator with arbitrary F. In the final section we apply maximum likelihood theory to prove weak convergence in the discrete case (the case of integer valued X_i). We also state a weak convergence result for continuous F. It turns out that the limiting distributions involved are identical to those arising in the usual models of random and fixed censorship (see e.g. Breslow and Crowley (1974) and Meier (1975)); we briefly discuss the reason for this phenomenon as well as its practical consequences.

We conclude the present section with a summary of notation and conventions; important definitions are given at the beginning of Sections 2 and 3. Let A and B be two extended-real-valued functions on $(-\infty, \infty)$. If A has left-hand limits everywhere we write A_- for the

Received June, 1978; revised July, 1980.

AMS 1980 *subject classifications.* 62G05, 60K10.

Key words and phrases. Testing with replacement, renewal testing, product limit estimator, Kaplan-Meier estimator, censoring.

function defined by $A_-(t) = A(t-)$; we define A_+ similarly. $A(-\infty)$ and $A(\infty)$ denote the limits as $t \to -\infty$ and $+\infty$ respectively of $A(t)$. We write ΔA for the function $A_+ - A_-$ and $\bar{A}$ for the function $A(\infty) - A_-$ if these are defined. So for a df F, $\bar{F} = 1 - F_-$. If the set function $\mu((s, t]) = A(t+) - A(s+)$ generates a σ-finite measure on the Borel sets of $(-\infty, \infty)$, we write $\int_I B(s)\, dA(s)$ for the Lebesgue-Stieltjes integral of B with respect to this measure over the interval I. We denote by $\int B\, dA$ the function defined by

$$\left(\int B\, dA\right)(t) = \int_{(-\infty,t]} B(s)\, dA(s) \tag{1}$$

if this exists.

All this notation is also applied to the sample paths of stochastic processes. For a stochastic process $X = \{X(t)\}_{t\in(-\infty,\infty)}$ we write $\mathscr{E}X$ for the function defined by

$$(\mathscr{E}X)(t) = \mathscr{E}(X(t)) \tag{2}$$

if $\mathscr{E}(X(t))$ exists for all t.

Some miscellaneous points of notation are χ_I for the indicator function of an interval I, and $\to_{\mathscr{D}}$ for convergence in distribution. $\mathscr{N}(\mu, \Sigma)$ denotes the multivariate normal distribution with mean vector μ and covariance matrix Σ. The symbols $\wedge$ and $\vee$ denote minimum and maximum respectively. We make frequent use of the convention $0/0 = 0$.

2. Observation of a single renewal process on $[0, \tau]$.

Consider a single renewal process $\tilde{N}$ as defined in the introduction. The distribution function F of interrenewal times satisfies $F(0-) = 0$, $F(0) < 1$ but is otherwise arbitrary. Recall that $\bar{F} = 1 - F_-$ and define $S = 1 - F$ and $H = \int \bar{F}^{-1}\, dF$. We also define F^{k*} to be the k-fold convolution of F with itself; $F^{0*} = \chi_{[0,\infty)}$. Define $U = \chi_{[0,\infty)} + \mathscr{E}\tilde{N} = \sum_{j=0}^{\infty} F^{j*}$. It is well known (see e.g. Prabhu (1965) Chapter 5, Theorem 2.1) that $\tilde{N}(\tau)$ is not only almost surely finite but in fact has finite moments of all orders. We shall always neglect the event $\{\tilde{N}(\tau) = \infty\}$.

We next define two processes N and Y which together record the censored and uncensored observations which result when $\tilde{N}$ is observed on $[0, \tau]$:

$$N(t) := \#\{j \geq 1: X_j \leq t \quad \text{and} \quad T_j \leq \tau\} \tag{3}$$

the number of uncensored observations less than or equal to t;

$$Y(t) := \#\{j \geq 1: X_j \geq t \quad \text{and} \quad T_{j-1} \leq (\tau - t) \wedge \tau\}, \tag{4}$$

the number of censored or uncensored observations which are known to take a value greater than or equal to t. For each t, define

$$M(t) = Y(-\infty) - Y(t+) - N(t),$$

the number of censored observations censored at a value $\leq t$. N and Y are nonnegative integer valued, N nondecreasing and right continuous and Y nonincreasing and left continuous. If $T_{j-1} = \tau$ for some j, we still count X_j as an observation, so there is always exactly one censored observation. (Other conventions as to what is observed at time τ are possible, and in fact lead to the same general results.) Note also that $Y(t) = Y(0) = \tilde{N}(\tau) + 1$ and $N(t) = 0$ for $t < 0$; and $Y(t) = 0$ and $N(t) = N(\tau) = \tilde{N}(\tau)$ for $t > \tau$. So N and Y are dominated by the rv $\tilde{N}(\tau) + 1$ which has finite moments of all orders.

LEMMA 1.

$$\mathscr{E}N = \int \mathscr{E}Y\, dH.$$

PROOF. For $t \geq 0$

$$\mathscr{E}Y(t) = \textstyle\sum_{i=1}^{\infty} P(T_{i-1} \leq \tau - t \quad \text{and} \quad X_i \geq t)$$

$$= \sum_{i=1}^{\infty} F^{(i-1)*}(\tau - t)\bar{F}(t)$$
$$= U(\tau - t)\bar{F}(t)$$

and

$$\begin{aligned}\mathscr{E}N(t) &= \textstyle\sum_{j=1}^{\infty} \sum_{i=1}^{j} P(X_i \le t,\ T_j \le \tau \quad \text{and} \quad T_{j+1} > \tau)\\ &= \textstyle\sum_{j=1}^{\infty} jP(X_1 \le t,\ T_j \le \tau \quad \text{and} \quad T_{j+1} > \tau)\\ &= \textstyle\sum_{j=1}^{\infty} j(P(X_1 \le t \quad \text{and} \quad T_j \le \tau) - P(X_1 \le t \quad \text{and} \quad T_{j+1} \le \tau)).\end{aligned}$$

But

$$P(X_1 \le t \quad \text{and} \quad T_j \le \tau) = \int_{[0,t]} F^{(j-1)*}(\tau - s)\, dF(s)$$

and

$$\textstyle\sum_{j=1}^{\infty} jF^{j*}(\tau) \le U^{2*}(\tau) \le U(\tau)^2 < \infty.$$

So

$$\begin{aligned}\mathscr{E}N(t) &= \textstyle\sum_{j=1}^{\infty} \int_{[0,t]} F^{(j-1)*}(\tau - s)\, dF(s)\\ &= \int_{[0,t]} U(\tau - s)\, dF(s)\\ &= \int_{[0,t]} \mathscr{E}Y(s)\, dH(s). \qquad \square\end{aligned}$$

3. Strong consistency of the product limit estimator. Consider an infinite sequence of independent copies of the renewal process $\tilde{N}$, and let N_n be the sum of the first n independent realisations of N, similarly for Y_n and M_n. So given n renewal processes observed on $[0, \tau]$, $N_n(t)$ is the number of uncensored observations less than or equal to t, $Y_n(t)$ the number of censored or uncensored observations known to be greater than or equal to t, and $M_n(t)$ the number of censored observations less than or equal to t.

The product limit estimator of F is defined by

$$\hat{F}_n(t) = 1 - \textstyle\prod_{s \le t} \left(1 - \dfrac{\Delta N_n(s)}{Y_n(s)}\right) \tag{5}$$

where, by our convention $0/0 = 0$, only s for which $Y_n(s) \ge \Delta N_n(s) > 0$ give rise to a factor not equal to 1 in the product. We also define

$$\hat{S}_n = 1 - \hat{F}_n.$$

We shall consider the product limit estimator as a function of the random functions $n^{-1}N_n$ and $n^{-1}Y_n$. Using the techniques of the Glivenko-Cantelli theorem it is easy to show that a.s. $n^{-1}N_n(t)$ and $n^{-1}Y_n(t)$ converge uniformly in t to their respective expectations $\mathscr{E}N(t)$ and $\mathscr{E}Y(t)$; recall that $N(t),\ Y(t) \le \tilde{N}(\tau) + 1$ for all t, and $\mathscr{E}\tilde{N}(\tau) < \infty$. We shall extend the definition of $\hat{F}_n$ in a continuous way so that it is also defined as a function of $(\mathscr{E}N, \mathscr{E}Y)$.

Suppose G_1 and G_2 are bounded, nondecreasing, right continuous functions on $(-\infty, \infty)$ such that $G_i(-\infty) = 0$. Denote the space of such pairs (G_1, G_2) as $\mathscr{G}$. Think of $G_1(t)$ and $G_2(t)$ as being the number of uncensored and censored observations less than or equal to t. Define $G = G_1 + G_2$. We have $\bar{G}_i(t) = G_i(\infty) - G_i(t-)$ and $\bar{G}(t) = \bar{G}_1(t) + \bar{G}_2(t)$. G_1 plays the role of N and $\bar{G}$ that of Y. For $(G_1, G_2) \in \mathscr{G}$ define

$$\Phi(G_1, G_2)(t) = \textstyle\prod_{s \le t} \left(1 - \dfrac{\Delta G_1(s)}{\bar{G}(s)}\right) \exp\left(-\displaystyle\int_{-\infty}^{t} \frac{dG_{1c}(s)}{\bar{G}(s)}\right) \tag{6}$$

where G_{1c} is the continuous part of G_1 and by convention $\exp(-\infty) = 0$. The infinite product is defined as the limit as $n \to \infty$ of the product of the first n of the (at most) countably many terms unequal to 1, taken in any enumeration of them. Note that

$$\hat{S}_n = \Phi(n^{-1}N_n, n^{-1}M_n) \tag{7}$$

and that $\Phi(G_1, G_2)$ is a right continuous nonnegative nonincreasing function on $(-\infty, \infty)$ with $\Phi(G_1, G_2)(-\infty) = 1$. Our definition extends that of Peterson (1977) who required G_1 and G_2 to have no jumps in common. He stated but did not prove that Φ is continuous on $\mathscr{G}$, which is the content of the following lemma.

LEMMA 2. *Let ρ_σ be the supremum metric on $(-\infty, \sigma]$. Let $(G_1, G_2) \in \mathscr{G}$ be fixed and let $\sigma > 0$ satisfy $\bar{G}(\sigma) > 0$. Then $\rho_\sigma(\Phi(G_1, G_2), \Phi(G_1', G_2')) \to 0$ as $\max(\rho_\sigma(G_1, G_1'), \rho_\sigma(G_2, G_2')) \to 0$.*

PROOF. Let $H = \int \bar{G}^{-1}\, dG_1$. H is right continuous, nondecreasing. Also $H(-\infty) = 0$, $H(\sigma) < \infty$ and for $s < \sigma$, $0 \le \Delta H(s) < 1$; in fact it is easily shown that $\sup_{s \in (-\infty, \sigma)} \Delta H(s) < 1$. (It is possible that $\Delta H(\sigma) = 1$). Then

$$\Phi(G_1, G_2)(t) = \exp[-H(t) + \textstyle\sum_{s \le t} \{\Delta H(s) + \log(1 - \Delta H(s))\}] \tag{8}$$

and

$$\Phi(G_1, G_2)(\sigma) = \Phi(G_1, G_2)(\sigma-)(1 - \Delta H(\sigma)). \tag{9}$$

Absolute convergence of the sum in (8), and thereby also of the product in (6), is proved below. We first show that the mapping $(G_1, G_2) \to H$ is continuous. For $t \le \sigma$ we find

$$\begin{aligned}
H(t) - H'(t) &= \int_{(-\infty, t]} \frac{dG_1(s)}{\bar{G}(s)} - \int_{(-\infty, t]} \frac{dG_1'(s)}{\bar{G}'(s)} = \int_{(-\infty, t)} \frac{dG_1(s) - dG_1'(s)}{\bar{G}(s)} \\
&\quad + \frac{\Delta G_1(t) - \Delta G_1'(t)}{\bar{G}(t)} + \int_{(-\infty, t]} \left(\frac{1}{\bar{G}(s)} - \frac{1}{\bar{G}'(s)}\right) dG_1'(s) \\
&= \left[\frac{G_1(s) - G_1'(s)}{(\bar{G})_+(s)}\right]_{-\infty}^{t-} + \frac{\Delta G_1(t) - \Delta G_1'(t)}{\bar{G}(t)} \\
&\quad - \int_{(-\infty, t)} (G_1(s) - G_1'(s))\, d\left(\frac{1}{\bar{G}(s)}\right) + \int_{(-\infty, t]} \frac{\bar{G}'(s) - \bar{G}(s)}{\bar{G}(s)\bar{G}'(s)}\, dG_1'(s) \\
&= \frac{G_1(t) - G_1'(t)}{\bar{G}(t)} - \int_{(-\infty, t)} (G_1(s) - G_1'(s))\, d\left(\frac{1}{\bar{G}(s)}\right) \\
&\quad + \int_{(-\infty, t]} \frac{\bar{G}'(s) - \bar{G}(s)}{\bar{G}(s)\bar{G}'(s)}\, dG_1'(s);
\end{aligned}$$

so, provided that $\bar{G}(\sigma) - \rho_\sigma(\bar{G}, \bar{G}') > 0$,

$$\rho_\sigma(H, H') \le 2\frac{\rho_\sigma(G_1, G_1')}{\bar{G}(\sigma)} + \frac{\rho_\sigma(\bar{G}, \bar{G}')(G_1(\sigma) + \rho_\sigma(G_1, G_1'))}{\bar{G}(\sigma)(\bar{G}(\sigma) - \rho_\sigma(\bar{G}, \bar{G}'))}$$

which proves continuity.

It remains to show that Φ is continuous as a function of H. For fixed H define $\mathscr{T} = \mathscr{T}(\delta) = \{t < \sigma: \Delta H(t) > \delta\}$, $\delta > 0$. So $\#\mathscr{T}(\delta) < \delta^{-1}H(\sigma)$. Let $\epsilon > 0$ satisfy $2\epsilon < 1 - \sup_{s<\sigma} \Delta H(s)$ and suppose $\rho_\sigma(H, H') < \epsilon$. Now $s \in (-\infty, \sigma)\backslash\mathscr{T} \Rightarrow \Delta H'(s) \le \delta + 2\epsilon$. For all s, $|\Delta H(s) - \Delta H'(s)| < 2\epsilon$. By Taylor expansion

$$0 \le -\{\Delta H(s) + \log(1 - \Delta H(s))\} \le \frac{1}{2}\frac{\Delta H(s)^2}{(1 - \Delta H(s))^2}$$

so for $t < \sigma$

$$(10)\qquad \Big|\sum_{s\in(-\infty,t]\setminus\mathcal{T}} \{\Delta H(s) + \log(1 - \Delta H(s))\}\Big| \le \frac{1}{2}\sum_{s\in(-\infty,\sigma)\setminus\mathcal{T}} \frac{\Delta H(s)^2}{(1-\Delta H(s))^2} \le \frac{1}{2}\,\delta H(\sigma)(1 - \sup_{s<\sigma}\Delta H(s))^{-2} = R_1(\delta).$$

(Putting $\delta = 1$ proves absolute convergence of the sum in (8).) Similarly

$$(11)\qquad \Big|\sum_{s\in(-\infty,t]\setminus\mathcal{T}} \{\Delta H'(s) + \log(1 - \Delta H'(s))\}\Big| \le \tfrac{1}{2}(\delta + 2\epsilon)(H(\sigma) + \epsilon)(1 - \sup_{s<\sigma}\Delta H(s) - 2\epsilon)^{-2} = R_2(\delta, \epsilon).$$

By the relations

$$|\log x - \log y| = \int_{x\wedge y}^{x\vee y} s^{-1}\,ds \le \frac{|x-y|}{x\wedge y}$$

we obtain

$$|\{\Delta H(s) + \log(1-\Delta H(s))\} - \{\Delta H'(s) + \log(1 - \Delta H'(s))\}| \le |\Delta H(s) - \Delta H'(s)|\cdot\{1 + (1 - \Delta H(s) \vee \Delta H'(s))^{-1}\}.$$

So for $t < \sigma$

$$(12)\qquad \sum_{s\in\mathcal{T}\cap(-\infty,t]} |\{\Delta H(s) + \log(1-\Delta H(s))\} - \{\Delta H'(s) + \log(1-\Delta H'(s))\}| \le \delta^{-1}H(\sigma)2\epsilon(1 + (1 - \sup_{s<\sigma}\Delta H(s) - 2\epsilon)^{-1}) = R_3(\delta, \epsilon).$$

Therefore combining (10), (11) and (12)

$$\sup_{t<\sigma}\Big|\sum_{s\le t}\{\Delta H(s) + \log(1-\Delta H(s))\} - \sum_{s\le t}\{\Delta H'(s) + \log(1 - \Delta H'(s))\}\Big| \le R_1(\delta) + R_2(\delta,\epsilon) + R_3(\delta,\epsilon),$$

the right-hand member of which can be made arbitrarily small by choice of δ and ϵ. In view of (8) this shows continuity of $\log\Phi$ as a function of H uniformly for $t \in (-\infty, \sigma)$. Because $\sup_{s<\sigma}\Delta H(s) < 1$ the exponent in (8) is bounded and the same is true for H'; hence Φ is also continuous in H uniformly for $t \in (-\infty, \sigma)$ and by (9) uniformly for $t \in (-\infty, \sigma]$. □

COROLLARY. *For any* df F

$$\Phi(F, 0)(t) = \prod_{s\le t}\left(1 - \frac{\Delta F(s)}{\bar F(s)}\right)\exp\left(-\int_{-\infty}^{t}\frac{dF_c(s)}{\bar F(s)}\right) = 1 - F(t)$$

when F_c denotes the continuous part of F.

REMARK. The result of this corollary is a generalization of a result contained in some informal remarks in Cox (1972) page 188 and in Peterson (1977) Lemma 2.1. Jacod (1975) Lemma 3.5 (proved in Jacod (1973)) and Liptser and Shiryayev (1978) Lemma 18.8 give a result which is essentially the same and in fact is a special case of a result on semimartingales due to Doléans-Dade (1970) Theorem 1.

PROOF OF COROLLARY. A df F can be arbitrarily well approximated by a step (distribution) function making a finite number of jumps; and for such a df the result is trivial. So by Lemma 2 the result holds for t such that $\bar F(t) > 0$. If for some s, $\Delta F(s) = \bar F(s) > 0$, it is easy to check that it now holds for all t; and otherwise it is easy to check that it holds for all t by taking limits as $t_n \uparrow s_0$ where $s_0 = \sup\{t: \bar F(t) > 0\}$. □

Now $(\mathscr{E}N, \mathscr{E}M) \in \mathscr{G}$. Fix a $\sigma > 0$ such that $\mathscr{E}Y(\sigma) > 0$; this condition is equivalent to $\sigma \le \tau$ and $\bar{F}(\sigma) > 0$. Combining the continuity of Φ with the Glivenko-Cantelli theorem for $n^{-1}N_n$, $n^{-1}Y_n$ gives

$$\sup_{t\in[0,\sigma]} |\hat{S}_n(t) - \Phi(\mathscr{E}N, \mathscr{E}M)(t)| \to 0 \quad \text{as } n \to \infty \quad \text{a.s.}$$

It remains to evaluate $\Phi(\mathscr{E}N, \mathscr{E}M)$. By Lemma 1,

$$\Delta\mathscr{E}N(s) = \mathscr{E}Y(s)\frac{\Delta F(s)}{\bar{F}(s)} \qquad \text{and} \qquad d(\mathscr{E}N)_c(s) = \mathscr{E}Y(s)\frac{dF_c(s)}{\bar{F}(s)};$$

so for $t \le \sigma$, using the fact that $\overline{N + M} = Y$,

$$\Phi(\mathscr{E}N, \mathscr{E}M)(t) = \prod_{s\le t}\left(1 - \frac{\Delta F(s)}{\bar{F}(s)}\right)\exp\left(-\int_{-\infty}^{t}\frac{dF_c(s)}{\bar{F}(s)}\right)$$

$$= S(t) \quad \text{by the corollary to Lemma 2.}$$

Define $\sigma_0 = \sup\{\sigma \le \tau : \bar{F}(\sigma) > 0\}$. If $\bar{F}(\sigma_0) > 0$, consistency has been proved on $[0, \sigma_0]$ and so in effect on $[0, \tau]$. Otherwise it has been proved on $[0, \sigma]$ for all $\sigma < \sigma_0$ and $F(\sigma) \uparrow 1$ as $\sigma \uparrow \sigma_0$. Because $\hat{F}_n$ is increasing and bounded above by 1 it is easy to extend consistency to $[0, \sigma_0]$ and so in effect to $[0, \tau]$; this proves

THEOREM 1.

$$\sup_{t\in[0,\tau]} |\hat{F}_n(t) - F(t)| \to 0 \qquad as \quad n \to \infty \quad \text{a.s.}$$

(which was conjectured in the closing remarks of Kaplan and Meier (1958)).

4. Weak convergence. Till further notice we suppose the X_i take values in $\mathbb{N} = \{0, 1, \cdots\}$. All time variables s, t etc. and the fixed τ are supposed to be in $\mathbb{N}$. Hence

$$\hat{F}_n(t) = 1 - \prod_{s=0}^{t}\left(1 - \frac{\Delta N_n(s)}{Y_n(s)}\right). \tag{13}$$

Recalling that $H = \int \bar{F}^{-1}\, dF$, we see that

$$\Delta H(t) = \begin{cases} P(X_i = t \mid X_i \ge t) & \text{if} \quad P(X_i \ge t) > 0 \\ 0 & \text{otherwise} \end{cases}$$

and that

$$F(t) = 1 - \prod_{s=0}^{t} (1 - \Delta H(s)). \tag{14}$$

Let $(p_0, \cdots, p_T) = (\Delta H(0), \cdots, \Delta H(\tau))$. As the likelihood function, given a single observation of $\tilde{N}$ on $[0, \tau]$, equals $\prod_{t=0}^{\tau} p_t^{\Delta N(t)}(1 - p_t)^{Y(t)-\Delta N(t)}$, the log likelihood function for $(p_0, \cdots, p_\tau)$ based on our observations is

$$L(p_0, \cdots, p_\tau) = \log\left(\prod_{t=0}^{\tau} p_t^{\Delta N_n(t)}(1 - p_t)^{Y_n(t)-\Delta N_n(t)}\right).$$

We have for $p_s, p_t \in (0, 1)$

$$\frac{\partial L}{\partial p_t} = \frac{\Delta N_n(t)}{p_t} - \frac{Y_n(t) - \Delta N_n(t)}{1 - p_t}$$

and

$$\frac{\partial^2 L}{\partial p_s \partial p_t} = \begin{cases} -\dfrac{\Delta N_n(t)}{p_t^2} - \dfrac{Y_n(t) - \Delta N_n(t)}{(1 - p_t)^2} & s = t \\ 0 & s \ne t. \end{cases}$$

Disregarding the cases $p_t = 0$ and $p_t = 1$ for the moment we find that the maximum likelihood estimator of $(p_0, \cdots, p_\tau)$ is $(\Delta N_n(0)/Y_n(0), \cdots, \Delta N_n(\tau)/Y_n(\tau))$. Now for $n = 1$ and for t such that $0 < p_t < 1$ using Lemma 1

$$\mathscr{E}\left(\frac{-\partial^2 L}{\partial p_t^2}\right) = \frac{\mathscr{E}\Delta N(t)}{p_t^2} + \frac{\mathscr{E}Y(t) - \mathscr{E}\Delta N(t)}{(1-p_t)^2}$$
$$= \frac{\mathscr{E}Y(t)}{p_t} + \frac{\mathscr{E}Y(t) - p_t\mathscr{E}Y(t)}{(1-p_t)^2}$$
$$= \frac{\mathscr{E}Y(t)}{p_t(1-p_t)};$$
$$\mathscr{E}\left(-\frac{\partial^2 L}{\partial p_s \partial p_t}\right) = 0 \qquad \text{for } s \neq t.$$

Therefore by Cramér (1946) Section 33.3 we find that

$$\left\{n^{1/2}\left(\frac{\Delta N_n(t)}{Y_n(t)} - \Delta H(t)\right)\right\}_{t=0,1,\cdots,\tau} \to_{\mathscr{D}} \mathscr{N}(0, \Sigma) \tag{15}$$

where

$$(\Sigma)_{st} = \begin{cases} \dfrac{\Delta H(t)(1-\Delta H(t))}{\mathscr{E}Y(t)} & s = t \\ 0 & s \neq t. \end{cases}$$

It is easy to verify that this remains valid if $p_t = 0$ or $p_t = 1$ for some t (recalling the convention $0/0 = 0$). Finally a standard Taylor series argument based on (13), (14) and (15) (see e.g. Witting and Nölle (1970) Theorem 2.10(d) gives us

THEOREM 2. *Suppose the X_i's are integer valued. Then*

$$\{n^{1/2}(\hat{F}_n(t) - F(t))\}_{t=0,1,\cdots,\tau} \to_{\mathscr{D}} \mathscr{N}(0, \Psi) \qquad as \quad n \to \infty$$

where

$$(\Psi)_{st} = S(s)S(t) \textstyle\sum_{u=0}^{s\wedge t} \dfrac{\Delta H(u)}{(1-\Delta H(u))\mathscr{E}Y(u)}.$$

Also by Cramér (1946) Section 33.3 we may conclude that

$$\mathscr{E}\left(\frac{\partial L}{\partial p_t}\right) = 0$$
$$\mathscr{E}\left(\frac{\partial L}{\partial p_s}\frac{\partial L}{\partial p_t}\right) = -\mathscr{E}\left(\frac{\partial^2 L}{\partial p_s \partial p_t}\right)$$

which gives us

$$\mathscr{E}(\Delta N(t) - Y(t)\Delta H(t)) = 0 \tag{16}$$

(already known from Lemma 1) and

$$\mathscr{E}((\Delta N(s) - Y(s)\Delta H(s))(\Delta N(t) - Y(t)\Delta H(t))) = \begin{cases} 0 & s \neq t \\ \Delta H(t)(1-\Delta H(t))\mathscr{E}Y(t) & s = t. \end{cases} \tag{17}$$

In Gill (1978) we prove an analogous result to (17) in the case of F continuous, which after calculations on the lines of Breslow and Crowley (1974) gives us

THEOREM 3. *Suppose that F is continuous and let $\sigma < \tau$ satisfy $F(\sigma) < 1$. Then*

$$\{n^{1/2}(\hat{F}_n(t) - F(t))\}_{t\in[0,\sigma]} \to_{\mathscr{D}} Z \qquad in \quad D[0, \sigma]$$

(*see* e.g. *Billingsley* (1968)) *where Z is a continuous zero mean Gaussian process with*

$$\mathrm{Cov}(Z(s), Z(t)) = S(s)S(t) \int_0^{s\wedge t} \frac{dH(u)}{\mathscr{E}Y(u)}.$$

PROOF. See Gill (1978) Theorem 11. □

The significance of Theorems 2 and 3 is that the form of the limiting distributions is identical to that obtained in the case of random censorship, so that e.g. confidence band procedures there are also valid in testing with replacement (see for instance Gillespie and Fisher (1979), Hall and Wellner (1980) or Gill (1980a)). This comes about because in random censorship (16) and (17) are also valid; this is very easy to verify. In Gill (1980a) and (1980b) we go into this matter in more detail.

REFERENCES

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

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.

COX, D. R. (1972). Regression models and life-tables. *J. Roy. Statist. Soc. Ser B* **34** 187–200 (with discussion).

CRAMÉR, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ.

CROW, L. H. and SHIMI, I. N. (1972). Maximum likelihood estimation of life-distributions from renewal testing. *Ann. Math. Statist.* **43** 1827–1838.

DOLÉANS-DADE, C. (1970). Quelques applications de la formule de changement de variables pour les semimartingales. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **16** 181–194.

GILL, R. D. (1978). Testing with Replacement and the Product Limit Estimator. Report SW 57/78 (first printing). Dept. Math. Statist. Mathematical Centre, Amsterdam.

GILL, R. D. (1980a). Censoring and Stochastic Integrals. M. C. Tract *124*, Mathematical Centre, Amsterdam.

GILL, R. D. (1980b). Nonparametric Estimation based on Censored Observations of a Markov Renewal Process. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **53** 97–116.

GILLESPIE, M. J. and FISHER, L. (1979). Confidence bands for the Kaplan-Meier survival curve estimate. *Ann. Statist.* **7** 920–924.

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

JACOD, J. (1973). On the stochastic intensity of a random point process over the half-line, Tech. Report *51*, Dept. Statist. Princeton Univ.

JACOD, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **31** 235–253.

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

LIPTSER, R. S. and SHIRYAYEV, A. N. (1978). *Statistics of Random Processes II: Applications.* Springer, Berlin.

MEIER, P. (1975). Estimation of a distribution function from incomplete observations. 67–87. In: *Perspectives in Probability and Statistics.* (J. Gani, ed.). Applied Prob. Trust, Sheffield.

PETERSON JR., A. V. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. *J. Amer. Statist. Assoc.* **72** 854–858.

PRABHU, N. U. (1965). *Stochastic Processes: Basic Theory and its Applications.* MacMillan, New York.

WITTING, H. and NÖLLE, G. (1970). *Angewandte Mathematische Statistik: Optimale Finite und Asymptotische Verfahren.* Teubner, Stuttgart.

MATHEMATISCH CENTRUM
KRUISLAAN 413
AMSTERDAM, NETHERLANDS