The Annals of Statistics
1978, Vol. 6, No. 5, 1057-1070

NONPARAMETRIC ESTIMATION FOR NONHOMOGENEOUS MARKOV PROCESSES IN THE PROBLEM OF COMPETING RISKS

BY THOMAS R. FLEMING

University of Maryland

Consider a time-continuous nonhomogeneous Markovian stochastic process V having state space A^0. Let $A \subset A^0$ and let $P_{Aij}(\tau,t)$ be the $i \rightarrow j$ transition probability of the Markovian stochastic process V_A arising in the hypothetical situation where states $A^0 - A$ have been eliminated from the state space of V. Based upon the concept of Kaplan and Meier's product-limit estimator, a nonparametric estimator $\hat{P}_{Aij}(\tau,t)$ is formulated which is proved to be uniformly strongly consistent and asymptotically unbiased. These results generalize those by Aalen for the special case in which A^0 has one transient state.

1. Introduction and summary. Consider a continuous-time nonhomogeneous Markovian stochastic process $V \equiv \{V(t); t \in T\}$ where T is a finite interval. Assume V has state space A^0 consisting of an arbitrary but finite number of both transient and absorbing states.

In the theory of competing risks, one considers a subset A of A^0 and for i, $j \in A$ seeks to estimate $P_{Aij}(\tau, t)$ which is the $i \rightarrow j$ transition probability of the Markovian stochastic process V_A which would exist in the hypothetical situation where states $A^0 - A$ have been eliminated from the state space of V.

Rather than assuming that some of the states are eliminated from the state space of V, Hoem (1969) made the more general assumption that some of the possible transitions of V are eliminated. While attention in this paper will be devoted to the former case, the techniques employed and basic results achieved are also valid for the general situation considered by Hoem. Hoem's terminology, "partial transition probability," for $P_{Aij}(\tau, t)$ shall be adopted.

Due to a desire to formulate an approach which could be applied to nonhomogeneous Markov processes with an arbitrary finite number of transient states, the author sought to formulate a nonparametric estimator of $P_{Aij}(\tau, t)$.

Many authors, including Cutler and Ederer (1958), Berkson and Gage (1950), Gehan (1968), Elveback (1958), Chiang (1968), Littel (1952), Kimball (1960), and Kaplan and Meier (1958), proposed nonparametric estimators for $P_{Aij}(\tau, t)$ in the special case where A^0 had one transient state and two absorbing states. Kaplan and Meier's product-limit estimator has been shown to possess desirable statistical properties, including uniform strong consistency, a property not generally shared by the other estimators mentioned above.

Received May 1977; revised October 1977.

AMS 1970 *subject classifications*. Primary 62G05; Secondary 60J75, 62N05.

Key words and phrases. Competing risks, nonhomogeneous, product limit, consistency, nonparametric.

The current author, after defining in Section 2 the competing risks model generalized for an arbitrary number of transient and absorbing states, formulates in Section 3 the nonparametric estimator $\hat{P}_{Aij}(\tau, t)$ of $P_{Aij}(\tau, t)$ for the generalized model. $\hat{P}_{Aij}(\tau, t)$ is based upon the concept of the product-limit estimator and is computed from N independent identically distributed sample processes V.

It is proved in Section 4 that the bias of $\hat{P}_{Aij}(\tau, t)$ converges exponentially to zero as $N \to \infty$, and in Section 5 that $\sup_{\tau \leq t \leq t_1} N^c|\hat{P}_{Aij}(\tau, t) - P_{Aij}(\tau, t)| \to 0$ a.s. as $N \to \infty$ for $c < \frac{1}{2}$, thus implying the uniform strong consistency of the estimator. These results generalize those achieved earlier by Aalen (1978) for the special case in which A^0 has one transient state.

In Section 6, the absolute distributions are expressed in terms of the partial transition probabilities and their estimation is discussed.

2. The statistical model.

2.1. *The general model.* Consider a probability space (Ω, F, P). Fix a closed finite interval $T = [t_0, t_1]$ and consider the nonhomogeneous Markovian stochastic process $V \equiv \{V(t); t \in T\}$ where, for each fixed $t \in T$, $V(t)$ has finite state space A^0. A^0 consists of a set A_t of s transient states and a set A_d of r absorbing states. V is assumed to have almost surely left-continuous sample paths, and hence will be called left-continuous.

Furthermore, if X is any left-continuous process with right-hand limits, we will define $(X)^+$ to be the right-continuous adaptation of X.

Define $P_{ij}(\tau, t) = P(V(t) = j \mid V(\tau) = i)$, $i, j \in A^0$ and $P_i(t) = P(V(t) = i)$; $i \in A^0$.

We shall make the following assumptions:

ASSUMPTION 2.1. (Regularity assumption).

$$\lim_{h \to 0} \frac{1 - P_{ii}(\tau, \tau + h)}{h} = -\nu_{ii}(\tau) \qquad \text{for any} \quad i \in A^0,$$

$$\lim_{h \to 0} \frac{P_{ij}(\tau, \tau + h)}{h} = \nu_{ij}(\tau) \qquad \text{for any} \quad i \neq j; \quad i, j \in A^0.$$

For any $i, j \in A^0$, $\nu_{ij}(\tau)$ is continuous.

Note $\nu_{ii}(\tau) = -\sum_{j \in A^0; j \neq i} \nu_{ij}(\tau)$.

ASSUMPTION 2.2. There exists $\varphi > 0$ such that $P_i(t) > \varphi$ for any $i \in A_t \cap A$ and for any $t \in T$ where A is as defined in Section 2.2. □

Define the $(s + r) \times (s + r)$ matrix $\mathscr{V}(t)$ by $(\mathscr{V}(t))_{ij} = \nu_{ij}(t)$. Notice that if $i \in A_d$, then the ith row of $\mathscr{V}(t)$ is composed of all zeros.

Define $\beta_{ij}(\tau, t) = \int_\tau^t \nu_{ij}(s)\, ds$ for any $i, j \in A^0$, and then define the matrices $\mathscr{P}(\tau, t)$ and $B(\tau, t)$ by $(\mathscr{P}(\tau, t))_{ij} = P_{ij}(\tau, t)$ and $(B(\tau, t))_{ij} = \beta_{ij}(\tau, t)$. I denotes the identity matrix.

Since Assumption 2.1 and the Markov property imply the validity of the

Kolmogorov forward and backward differential equations,

$$\frac{\partial}{\partial t} P_{ik}(\tau, t) = \sum_{j \in A^0} P_{ij}(\tau, t)\nu_{jk}(t) .$$

Thus

$$\mathscr{P}(\tau, t) = I + \int_\tau^t \mathscr{P}(\tau, s)\, dB(\tau, s) .$$

In the further development of the model in Section 2.2, we will have need for Feller's (1940) result that for a given B there exists a unique solution to the Kolmogorov forward and backward differential equations for $\mathscr{P}$ which satisfies the properties of a probability transition matrix; namely

$$\sum_{j \in A^0} P_{ij}(s, t) \doteq 1 , \qquad 0 \leqq P_{ij}(s, t) \leqq 1 , \quad \text{and} \quad \mathscr{P}(t, t) = I .$$

2.2. *The competing risks model.* Recall that the stochastic process V under study has state space A^0. We wish to study the Markovian stochastic process V_A which arises in the hypothetical situation where the state space of V is restricted to a nonempty set A, $A \subset A^0$.

We will be interested in hypothetical partial transition probabilities of the form: $P_{Aij}(\tau, t) = P(V_A(t) = j \mid V_A(\tau) = i)$; $i, j \in A$.

For any $i, j \in A$, let the intensities $\nu_{Aij}(t)$ be defined for $P_{Aij}(\tau, t)$ exactly as the intensities $\nu_{ij}(t)$ were defined for $P_{ij}(\tau, t)$, so $\nu_{Aij}(t)$ and $\nu_{ij}(t)$ in general need not be equal.

The following is an empirical assumption discussed by Cornfield (1957), Gail (1975), and Tsiatis (1975) which is important to the development of the theory of competing risks. It, along with the result from Feller (1940) mentioned earlier, assures that the partial transition probabilities $P_{Aij}(\tau, t)$ are well defined. Empirical justification for the assumption must be considered in each particular application.

ASSUMPTION 2.3. For any $i, j \in A$ such that $i \neq j$, $\nu_{Aij}(t) = \nu_{ij}(t)$.

Consequently $\mathscr{V}_A(t)$, the matrix of intensities when the state space is restricted to A, can be obtained from $\mathscr{V}(t)$ by deleting the rows and columns of $\mathscr{V}(t)$ not corresponding to states in A, and redefining only the diagonal terms by $\nu_{Aii}(t) = -\sum_{j \in A} \nu_{ij}(t)$.

Define $B_A(\tau, t)$ by $(B_A(\tau, t))_{ij} = \beta_{Aij}(\tau, t) = \int_\tau^t (\mathscr{V}_A(s))_{ij}\, ds$ for any $i, j \in A$. Also define $\mathscr{P}_A(\tau, t)$ by $(\mathscr{P}_A(\tau, t))_{ij} = P_{Aij}(\tau, t)$ and let I_A be the identity matrix whose dimension is m which is the cardinality of A.

Since we now have the validity of the forward and backward differential equations for $\mathscr{P}_A$,

$$\mathscr{P}_A(\tau, t) = I_A + \int_\tau^t \mathscr{P}_A(\tau, s)\, dB_A(\tau, s) . \tag{2.1}$$

3. Nonparametric estimator. For purposes of estimation, we assume that we observe N independent stochastic processes $\{V_j(t);\, t \in T\}$; $j = 1, \cdots, N$, identically satisfying what has been set forth in Section 2.1.

Assume for notational simplicity that $A = \{1, 2, \cdots, m\}$.

3.1. *Counting processes.* $\mathbf{N}(t)$ is the $(s+r)$-dimensional vector whose ith component, $N_i(t)$, represents the number of the N observed processes in state i at time t.

For any $i, j \in A$, let the right-continuous transition counting process $M_{Aij}(\tau, t)$ represent the number of $i \to j$ transitions over $(\tau, t]$ if $i \neq j$; and—(number of $i \to (A - \{i\})$ transitions over $(\tau, t]$) if $i = j$. The $(m \times m)$ matrix $M_A(\tau, t)$ is then defined by

$$(M_A(\tau, t))_{ij} = M_{Aij}(\tau, t) .$$

According to the definitions of multivariate counting processes and their corresponding intensity processes, as given in Section 1 of Aalen (1977), we observe that for any fixed $A \subset A^0$,

$$\mathbf{M}_A(\tau, \cdot) \equiv (M_{A12}(\tau, \cdot), M_{A13}(\tau, \cdot), \cdots, M_{A1m}(\tau, \cdot), M_{A21}(\tau, \cdot), M_{A23}(\tau, \cdot), \cdots, M_{Am,m-1}(\tau, \cdot))$$

forms a sequence in N of multivariate counting processes such that, if $M_{Aij}(\tau, \cdot)$ has intensity process $\Lambda_{Aij}(s)$, then

$$\Lambda_{Aij}(s) = \nu_{ij}(s) N_i(s) .$$

Equip (Ω, F, P) with an increasing family of sub-σ-fields of F, $\{F_t\}$, where $F_t \equiv \sigma(V_k(s);\ t_0 \leqq s \leqq t,\ 1 \leqq k \leqq N)$. Note $\mathbf{M}_A(\tau, \cdot)$ is adapted to $\{F_t\}$.

If we define

$$L_{Aij}(\tau, t) = M_{Aij}(\tau, t) - \int_\tau^t \nu_{ij}(s) N_i(s)\, ds \tag{3.1}$$

$\forall\ 1 \leqq i, j \leqq m$, then the next lemma follows from Dolivo (1974) and Boel, Varaiya and Wong (1973).

LEMMA 3.1. *$\{L_{Aij}(\tau, t);\ 1 \leqq i, j \leqq m,\ i \neq j\}_{t \in T}$ is a collection of square integrable martingales; that is, for any $i, j \in A$, $\{L_{Aij}(\tau, t);\ \tau \leqq t \leqq t_1\}$ is a martingale such that* $\sup_{\tau \leqq t \leqq t_1} E(L_{Aij}(\tau, t))^2 < \infty$.

3.2. *Definition of* $\hat{\mathscr{P}}_A(\tau, t)$. For any $i \in A^0$, let

$$R_i(t) = [N_i(t)]^{-1} \quad \text{if} \quad N_i(t) > 0 , \quad \text{and be } 0 \text{ otherwise.}$$

Define the $(m \times m)$ diagonal matrix $\mathscr{R}_A(t)$ by $(\mathscr{R}_A(t))_{ii} = R_i(t)$ for any $i \in A$.

Based upon the concept of the product-limit estimator and the cumulative hazard estimator discussed in Kaplan and Meier (1958), Nelson (1969), and Breslow and Crowley (1974) and generalized to counting processes by Aalen (1978) we define

$$\hat{B}_A(\tau, t) = \int_\tau^t \mathscr{R}_A(s)\, dM_A(\tau, s) .$$

With this as our motivation, we propose the following left-continuous estimator of $\mathscr{P}_A(\tau, t)$, which can be shown to be based upon sufficient statistics:

$$\hat{\mathscr{P}}_A(\tau, t) = I_A + \int_\tau^{t-} \hat{\mathscr{P}}_A(\tau, s) \mathscr{R}_A(s)\, dM_A(\tau, s) . \tag{3.2}$$

It should be pointed out that $\hat{\mathscr{P}}_A(\tau, t)$ is defined recursively, is easy to calculate even for large values of N, and is intuitively appealing.

We will continue to deal throughout the remainder of this paper with estimators of parameters of the stochastic process V_A. Since A remains fixed, there should be no confusion provided by the fact that we will suppress the use of the subscript A in Sections 4 and 5.

4. Asymptotic unbiasedness of $\hat{\mathscr{P}}(\tau, t)$. Componentwise, the bias of $\hat{\mathscr{P}}(\tau, t)$ converges exponentially to zero as $N \to \infty$, as shown by the following theorem.

THEOREM 4.1. $(E\hat{\mathscr{P}}(\tau, t) - \mathscr{P}(\tau, t))_{ij} = O[e^{N \ln (1-\varphi)}]$; $i, j \in A$.

PROOF. By equation (3.2),

$$\hat{\mathscr{P}}(\tau, t + h) - \hat{\mathscr{P}}(\tau, t) = \int_{(t)-}^{(t+h)-} \hat{\mathscr{P}}(\tau, s)\mathscr{R}(s)\, dM(\tau, s) .$$

By the left-continuity of $\hat{\mathscr{P}}(\tau, s)\mathscr{R}(s)$, Assumption 2.1, and the Markov property, it follows, as in the proof of Proposition 2 in Aalen (1978) (the Appendix), that

$$\begin{aligned} \frac{\partial}{\partial t} E\hat{P}_{ij}(\tau, t) &= \textstyle\sum_{k \in A_l \cap A} \nu_{kj}(t) E\{\hat{P}_{ik}(\tau, t) R_k(t) N_k(t)\} \\ &= \textstyle\sum_{k \in A_l \cap A} \nu_{kj}(t) E\{\hat{P}_{ik}(\tau, t)\} \\ &\quad - \textstyle\sum_{k \in A_l \cap A} \nu_{kj}(t) E\{\hat{P}_{ik}(\tau, t) I_{[N_k(t)=0]}\} . \end{aligned} \tag{4.1}$$

Define the diagonal matrix $\mathscr{I}_{[N(s)=0]}$ by $(\mathscr{I}_{[N(s)=0]})_{ii} = I_{[N_i(s)=0]}$ and $\mathscr{T}(\tau, t) \equiv E\hat{\mathscr{P}}(\tau, t) - \mathscr{P}(\tau, t)$, so $\mathscr{T}(\tau, \tau) = 0$.

Then, by equation (4.1) and the fact that $(\partial/\partial t)P_{ij}(\tau, t) = \sum_{k \in A_l \cap A} P_{ik}(\tau, t)\nu_{kj}(t)$, it follows that

$$\frac{\partial}{\partial s} \mathscr{T}(\tau, s) = \mathscr{T}(\tau, s)\mathscr{V}(s) + E[\hat{\mathscr{P}}(\tau, s)\mathscr{I}_{[N(s)=0]}]\mathscr{V}(s) .$$

Hence, by the Kolmogorov backward equation,

$$\begin{aligned} \frac{\partial}{\partial s} \mathscr{T}(\tau, s)\mathscr{P}(s, t) &= \left\{\frac{\partial}{\partial s} \mathscr{T}(\tau, s)\right\} \mathscr{P}(s, t) - \mathscr{T}(\tau, s)\{-\mathscr{V}(s)\mathscr{P}(s, t)\} \\ &= -E[\hat{\mathscr{P}}(\tau, s)\mathscr{I}_{[N(s)=0]}]\mathscr{V}(s)\mathscr{P}(s, t) . \end{aligned}$$

Since for $i \in A_l$

$$\begin{aligned} P[I_{[N_i(s)=0]} = 1] &= P[N_i(s) = 0] \\ &= [1 - P_i(s)]^N \leqq e^{N \ln [1-\varphi]} , \end{aligned} \tag{4.2}$$

$$\frac{\partial}{\partial s} \mathscr{T}(\tau, s)\mathscr{P}(s, t) = O(e^{N \ln [1-\varphi]}) .$$

Integrating with respect to s over (τ, t),

$$\mathscr{T}(\tau, t) = \int_\tau^t \frac{\partial}{\partial s} \mathscr{T}(\tau, s)\mathscr{P}(s, t)\, ds .$$

Hence, the theorem follows by equation (4.2). □

5. Uniform strong consistency of $\hat{\mathscr{P}}(\tau, t)$. In this section, we prove Theorem 5.1 concerning the uniform strong consistency of $\hat{\mathscr{P}}(\tau, t)$. This result generalizes the significantly more straightforward result achieved by Aalen (1978) for the special case in which A^0 has one transient state. In this sequel, as throughout this paper, convergence refers to $N \to \infty$.

To prove the theorem, we will first prove the uniform strong consistency of $\hat{B}(\tau, t)$ in Lemma 5.8. Using a technique of Aalen (1978) this will follow easily after we have verified auxiliary Lemmas 5.1, 5.5, 5.6 and 5.7.

Define the $(m \times m)$ diagonal matrices $\Pi(t)$ and $\Pi^{-1}(t)$ as follows: for any $i \in A$, $(\Pi(t))_{ii} = P_i(t)$,

$$(\Pi^{-1}(t))_{ii} = [P_i(t)]^{-1} \quad \text{if} \quad i \in A_l \cap A\,,$$
$$= 0 \quad \text{if} \quad i \in A_d \cap A\,.$$

For any $k = 1, \cdots, N$, define the $(m \times m)$ matrix $M^k(\tau, t)$ in the following manner. For any $i, j \in A$, $(M^k(\tau, t))_{ij} = M^k_{ij}(\tau, t)$ represents the number of transitions $i \to j$ over $(\tau, t]$ for the process V_k, if $i \neq j$; and—[number of transitions $i \to (A - \{i\})$ over $(\tau, t]$ for the process V_k], if $i = j$.

In the next lemma, as throughout the paper, $\to$ a.s. denotes convergence almost surely.

LEMMA 5.1. *$(1/N)M(\tau, t) \to \int_\tau^t \Pi(s)\, dB(\tau, s)$ a.s. componentwise.*

PROOF. Since $\{M^k(\tau, t); k = 1, \cdots, N\}$ is an independent, identically distributed collection of random matrices, by Kolmogorov's strong law of large numbers componentwise we have

$$\lim_{N\to\infty} \frac{1}{N} M(\tau, t) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N M^k(\tau, t) = EM^1(\tau, t)\,.$$

The lemma now follows by equation (3.1) and Lemma 3.1. □

At this point, we will prove a lemma to be used in the proof of Lemma 5.3, which in turn will be needed in the proofs of Lemmas 5.4 and 5.5.

LEMMA 5.2. *Let $r \in \Delta \equiv \{1, 2, 3, \cdots\}$ and let $i, j \in A^0$. Then there exists $a_{2r} < \infty$ such that*

$$E(N^{-\frac{1}{2}}|L_{ij}(t_0, t_1)|)^{2r} < a_{2r} \qquad \textit{for any} \quad N \in \Delta\,.$$

PROOF. Chung (1968, page 47) has stated the next useful result:

$$\text{(5.1)}\qquad \begin{aligned} &\text{If } \ p > 0\,,\ E(|X|^p) < \infty\,, \ \text{then} \ \ x^p P(|X| > x) = o(1) \ \text{ as } \ x \to \infty\,.\\ &\text{Conversely, if } \ x^p P(|X| > x) = o(1)\,, \ \text{then}\\ &E(|X|^{p-\varepsilon}) < \infty \quad \text{for} \quad 0 < \varepsilon < p\,. \end{aligned}$$

A second immediate result which will be needed below is given in inequality (5.2): let $r \in \Delta$. Assume $\{X_n\}_{n\in\Delta}$ is an independent identically distributed collection of mean zero random variables such that $E(X_1^{2r}) < \infty$ for any $r \in \Delta$.

Then there exists $a_{2r} < \infty$ such that

$$\text{for any} \quad N \in \Delta, \quad E(N^{-\frac{1}{2}} \sum_{k=1}^{N} X_k)^{2r} < a_{2r}. \tag{5.2}$$

Now,

$$\begin{aligned} N^{-\frac{1}{2}} L_{ij}(t_0, t_1) &= N^{-\frac{1}{2}} \sum_{k=1}^{N} [M_{ij}^k(t_0, t_1) - \int_{t_0}^{t_1} \nu_{ij}(s) I_{[V_k(s)=i]}\, ds] \\ &\equiv N^{-\frac{1}{2}} \sum_{k=1}^{N} L_{ij}^k(t_0, t_1). \end{aligned}$$

The intensity process $\nu_{ij}(t) I_{[V_k(t)=i]}$ of $M_{ij}^k(t_0, t_1)$ is bounded by ν_i. Let $U_{\nu_i'}$ be a Poisson distributed random variable with parameter $(t_1 - t_0)\nu_i \equiv \nu_i'$. Clearly then,

$$|M_{ij}^k(t_0, t_1)| \quad \text{is stochastically smaller than} \quad U_{\nu_i'}. \tag{5.3}$$

Thus

$$\begin{aligned} x^{2r} P(|L_{ij}^k(t_0, t_1)| > x) &\leqq x^{2r} P(|M_{ij}^k(t_0, t_1)| + \nu_i' > x) \\ &\leqq x^{2r} P(U_{\nu_i'} > x - \nu_i') = o(1), \end{aligned}$$

where the second inequality follows from inequality (5.3), and the equality follows from equation (5.1) and the fact that the Poisson (ν_i') distribution has finite moments of all orders.

By another application of equation (5.1), we conclude

$$E(L_{ij}^k(t_0, t_1)^{2r}) < \infty \qquad \text{for any} \quad r \in \Delta.$$

The lemma now follows by inequality (5.2). □

LEMMA 5.3. *Let* $0 \leqq p < \frac{1}{2}$. *Then for any* $i, j \in A^0$,

$$\sup_{t \in T} N^p \left| \frac{1}{N} L_{ij}(t_0, t) \right| \to 0 \quad \text{a.s.}$$

PROOF. We will use the following inequality which is given in Loève (1963, page 524):

$$C_N P[\sup_{t \in T} |L_{ij}(t_0, t)| > C_N] \leqq \int_{[\sup_{t \in T} |L_{ij}(t_0, t)| > C_N]} |L_{ij}(t_0, t_1)|\, dP. \tag{5.4}$$

Fix $\alpha > 0$ and let $C_N = \alpha N^{(\frac{1}{2} + 1/2r)}$ where $r \in \Delta$. Then by equation (5.4) for any $r \in \Delta$,

$$\begin{aligned} C_N^{2r} P[\sup_{t \in T} |L_{ij}(t_0, t)| > C_N] &= C_N^{2r} P[\sup_{t \in T} (L_{ij}(t_0, t))^{2r} > C_N^{2r}] \\ &\leqq \int_{[\sup_{t \in T} |L_{ij}(t_0, t)| > C_N]} (L_{ij}(t_0, t_1))^{2r}\, dP. \end{aligned}$$

Hence by Hölder's inequality and Lemma 5.2,

$$P[\sup_{t \in T} |L_{ij}(t_0, t)| > C_N] < (a_{4r})^{\frac{1}{2}} (P[\sup_{t \in T} |L_{ij}(t_0, t)| > C_N])^{\frac{1}{2}} N^r C_N^{-2r}$$

so

$$P\left[\sup_{t \in T} N^{\frac{1}{2}(1-(1/r))} \left| \frac{1}{N} L_{ij}(t_0, t) \right| > \alpha \right] \leqq a_{4r} N^{2r} C_N^{-4r} = \frac{a_{4r}}{\alpha^{4r}} \frac{1}{N^2}.$$

Hence, by the Borel–Cantelli lemma and the above, the lemma follows. □

LEMMA 5.4. *Let $0 \leqq p < \frac{1}{2}$. Then for any $i \in A^0$,*

$$\sup_{t \in T} N^p \left| \frac{1}{N} N_i(t) - E \frac{1}{N} N_i(t) \right| \to 0 \quad \text{a.s.}$$

Suppose either $i \in A_d$ or $i \in A_l$ and A_l has cardinality one. Then the lemma would follow from Lemma 2.2 of Barlow and van Zwet (1970).

Suppose then $i \in A_l$ and A_l has cardinality greater than one. We can assume A_l has cardinality two since the method of proof can be applied with equivalent success to any higher finite cardinality. Take $A_l = \{i, k\}$.

The process denoted by $N_i(t)$ in the proof of this lemma is actually the right-continuous adaptation of $N_i(t)$, $(N_i(t))^+$. However, the validity of this lemma follows immediately from its proof for this right-continuous adaptation.

Pick any ε arbitrarily close to zero.

Because there exists $\xi < \infty$ such that for any $t \in T$ and for $j, l \in A_l$, $\nu_{jl}(t) \leqq -\nu_{jj}(t) < \xi$, we can subdivide T into a finite number of subintervals such that for any subinterval $[\alpha, \beta]$,

$$\int_\alpha^\beta -\nu_{jj}(t)\, dt < \varepsilon .$$

Clearly Lemma 5.4 will follow if we can show

$$\text{(5.5)} \qquad \sup_{\alpha \leqq t \leqq \beta} N^p \left| \frac{1}{N} N_i(t) - E \frac{1}{N} N_i(t) \right| \to 0 \quad \text{a.s.} \qquad \text{for each } [\alpha, \beta] .$$

Adopt the notation

$$M_{i\bullet}(t_0, t) = \sum_{j \in A^0; j \neq i} M_{ij}(t_0, t) \qquad \text{and} \qquad L_{i\bullet}(t_0, t) = \sum_{j \in A^0; j \neq i} L_{ij}(t_0, t) .$$

Observe that for any $\tau \leqq t$,

$$\text{(5.5a)} \qquad \begin{aligned} N_i(t) &= N_i(\tau) + M_{ki}(\tau, t) - M_{i\bullet}(\tau, t) , \\ M_{i\bullet}(\tau, t) &= L_{i\bullet}(\tau, t) + \int_\tau^t -\nu_{ii}(s) N_i(s)\, ds , \end{aligned}$$

and

$$M_{ki}(\tau, t) = L_{ki}(\tau, t) + \int_\tau^t \nu_{ki}(s) N_k(s)\, ds .$$

Take $\alpha \leqq t \leqq \beta$. Then

$$\begin{aligned} &\sup_{\alpha \leqq t \leqq \beta} N^p \left| \frac{1}{N} N_i(t) - E \frac{1}{N} N_i(t) \right| \\ &\quad \leqq N^p \left| \frac{1}{N} N_i(\alpha) - E \frac{1}{N} N_i(\alpha) \right| \\ \text{(5.6)} \qquad &\quad + \sup_{\alpha \leqq t \leqq \beta} N^p \left| \frac{1}{N} L_{ki}(\alpha, t) \right| + \sup_{\alpha \leqq t \leqq \beta} N^p \left| \frac{1}{N} L_{i\bullet}(\alpha, t) \right| \\ &\quad + \left\{ \sup_{\alpha \leqq t \leqq \beta} N^p \left| \frac{1}{N} N_k(t) - E \frac{1}{N} N_k(t) \right| \right\} \left\{ \int_\alpha^\beta \nu_{ki}(s)\, ds \right\} \\ &\quad + \left\{ \sup_{\alpha \leqq t \leqq \beta} N^p \left| \frac{1}{N} N_i(t) - E \frac{1}{N} N_i(t) \right| \right\} \left\{ \int_\alpha^\beta -\nu_{ii}(s)\, ds \right\} . \end{aligned}$$

Obtaining the inequality corresponding to equation (5.6) for $\sup_{\alpha \leqq t \leqq \beta} N^p|N^{-1}N_k(t) - EN^{-1}N_k(t)|$, and solving the system of two inequalities, one obtains:

$$
\begin{aligned}
&[(1 - \textstyle\int_\alpha^\beta -\nu_{ii}(s)\,ds)(1 - \int_\alpha^\beta -\nu_{kk}(s)\,ds) \\
&\quad - (\textstyle\int_\alpha^\beta \nu_{ik}(s)\,ds)(\int_\alpha^\beta \nu_{ki}(s)\,ds)] \sup_{\alpha \leqq t \leqq \beta} N^p \left|\frac{1}{N} N_i(t) - E\frac{1}{N} N_i(t)\right| \\
&\quad \leqq \left\{N^p \left|\frac{1}{N} N_i(\alpha) - E\frac{1}{N} N_i(\alpha)\right| + \sup_{\alpha \leqq t \leqq \beta} N^p \left|\frac{1}{N} L_{ki}(\alpha, t)\right| \right. \\
&\qquad \left. + \sup_{\alpha \leqq t \leqq \beta} N^p \left|\frac{1}{N} L_{i\bullet}(\alpha, t)\right|\right\} \{1 - \textstyle\int_\alpha^\beta -\nu_{kk}(s)\,ds\} \\
&\qquad + \left\{N^p \left|\frac{1}{N} N_k(\alpha) - E\frac{1}{N} N_k(\alpha)\right| + \sup_{\alpha \leqq t \leqq \beta} N^p \left|\frac{1}{N} L_{ik}(\alpha, t)\right| \right. \\
&\qquad \left. + \sup_{\alpha \leqq t \leqq \beta} N^p \left|\frac{1}{N} L_{k\bullet}(\alpha, t)\right|\right\} \{\textstyle\int_\alpha^\beta \nu_{ki}(s)\,ds\}\,.
\end{aligned} \tag{5.7}
$$

Note that ε was arbitrary. We can then choose it small enough such that the term in the brackets on the left-hand side of inequality (5.7) is positive.

Since the right-hand side of inequality (5.7) converges to zero a.s. by Lemma 5.3 and by Chung's (1968) Corollary 5.4.1, equation (5.5) follows. □

LEMMA 5.5. *Let* $0 \leqq p < \frac{1}{2}$. *Then for any* $i, j \in A$,

$$\sup_{t \in T} N^p \left|\frac{1}{N} M_{ij}(t_0, t) - E\frac{1}{N} M_{ij}(t_0, t)\right| \to 0 \quad \text{a.s.}$$

PROOF.

$$
\begin{aligned}
&\left|\frac{1}{N} M_{ij}(t_0, t) - E\frac{1}{N} M_{ij}(t_0, t)\right| \\
&\qquad = \left|\frac{1}{N} L_{ij}(t_0, t) + \textstyle\int_{t_0}^t \nu_{ij}(s) \left[\frac{1}{N} N_i(s) - E\frac{1}{N} N_i(s)\right] ds\right|.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&\sup_{t \in T} N^p \left|\frac{1}{N} M_{ij}(t_0, t) - E\frac{1}{N} M_{ij}(t_0, t)\right| \\
&\quad \leqq \sup_{t \in T} N^p \left|\frac{1}{N} L_{ij}(t_0, t)\right| + \left\{\sup_{t \in T} N^p \left|\frac{1}{N} N_i(t) - E\frac{1}{N} N_i(t)\right|\right\} \textstyle\int_{t_0}^{t_1} \nu_{ij}(s)\,ds\,.
\end{aligned}
$$

The lemma now follows by Lemmas 5.3 and 5.4. □

The author is indebted to the referee for the following significant simplification of the proofs given above in the important special case $p = 0$. Using the fact that a sequence of monotonic functions converging pointwise to a continuous function also converges uniformly (see Rudin [1965], Chapter 7, Example 17), Lemma 5.5 for $p = 0$ then follows immediately by the strong law of large numbers. Lemma 5.4 then follows for $p = 0$ by equation (5.5a), rendering Lemmas 5.2 and 5.3 unnecessary in this special case.

LEMMA 5.6. *Let $0 \leqq p < \frac{1}{2}$. Then componentwise*

$$\sup_{t\in T} N^p|N\mathscr{R}(t) - \Pi^{-1}(t)| \to 0 \quad \text{a.s.}$$

PROOF. Following the technique suggested by Aalen (1978), we can write the following inequality for $i \in A_l \cap A$:

$$(\sup_{t\in T} N^p|N\mathscr{R}(t) - \Pi^{-1}(t)|)_{ii} \leqq \frac{\sup_{t\in T} N^p|N^{-1}N_i(t) - P_i(t)|}{\varphi \inf_{t\in T} N^{-1}N_i(t)} I_{[\inf_{t\in T} N^{-1}N_i(t) > \frac{1}{2}\varphi]} + \frac{1}{\varphi} N^{\frac{3}{2}} I_{[\inf_{t\in T} N^{-1}N_i(t) \leqq \frac{1}{2}\varphi]} . \tag{5.8}$$

The first term on the right-hand side of equation (5.8) converges to zero a.s. by Lemma 5.4. Hence, to prove the lemma, it suffices to show that for any $\varepsilon > 0$, there exists N_0 such that

$$Q \equiv P[I_{[\inf_{t\in T} N^{-1}N_i(t) \leqq \frac{1}{2}\varphi]} = 0 \ \text{ for all } \ N \geqq N_0] \geqq 1 - \varepsilon . \tag{5.9}$$

Let $\varepsilon_0 = \min(\varepsilon, \frac{1}{2}\varphi)$. Since $\sup_{t\in T} |N^{-1}N_i(t) - P_i(t)| \to 0$ a.s., there exists N' such that

$$P\left[\sup_{t\in T}\left|\frac{1}{N} N_i(t) - P_i(t)\right| \leqq \varepsilon_0 \ \text{ for all } \ N \geqq N'\right] \geqq 1 - \varepsilon_0 ,$$

implying, since $\inf_{t\in T} P_i(t) = \varphi$,

$$Q = P\left[\inf_{t\in T} \frac{1}{N} N_i(t) > \tfrac{1}{2}\varphi \ \text{ for all } \ N \geqq N'\right] \geqq 1 - \varepsilon_0 \geqq 1 - \varepsilon .$$

Letting $N_0 = N'$, equation (5.9) follows. □

LEMMA 5.7. *Componentwise, $\Pi^{-1}(t)$ is of bounded variation over (τ, t_1).*

PROOF. By equation (2.1),

$$P_{ij}(\tau, t) = \{\delta_{ij} + \textstyle\sum_{k\neq j; k\in A_l\cap A} \int_\tau^t P_{ik}(\tau, s)\nu_{kj}(s)\, ds\} + \{\int_\tau^t P_{ij}(\tau, s)\nu_{jj}(s)\, ds\} .$$

The integrals are bounded monotone functions of t. Hence $P_{ij}(\tau, t)$ is of bounded variation. The lemma now follows by Assumption 2.2 and equation (6.1). □

LEMMA 5.8. *Let $0 \leqq p < \frac{1}{2}$. Then componentwise,*

$$\sup_{t\in T} N^p|\hat{B}(t_0, t) - B(t_0, t)| \to 0 \quad \text{a.s.}$$

PROOF.

$$\hat{B}(t_0, t) - B(t_0, t) = \int_{t_0}^t [N\mathscr{R}(s) - \Pi^{-1}(s)]\, d\frac{1}{N} M(t_0, s) + \int_{t_0}^t \Pi^{-1}(s)\left[d\frac{1}{N} M(t_0, s) - \Pi(s)\, dB(t_0, s)\right] .$$

The proof follows by Lemmas 5.1, 5.5, 5.6 and 5.7. □

THEOREM 5.1. *Let* $0 \leqq p < \frac{1}{2}$. *Componentwise,*

$$\sup_{\tau \leqq t \leqq t_1} N^p |\hat{\mathscr{P}}(\tau, t) - \mathscr{P}(\tau, t)| \to 0 \quad \text{a.s.}$$

PROOF. Define the left-continuous processes $\mathscr{Y}(\tau, t)$ and $U(\tau, t)$ such that $dN^{-p}\mathscr{Y}(\tau, t) = d(\hat{\mathscr{P}}(\tau, t) - \mathscr{P}(\tau, t))$ and

$$dU(\tau, t) = \hat{\mathscr{P}}(\tau, t)[d(\hat{B}(\tau, t) - B(\tau, t))] \tag{5.10}$$

where $\mathscr{Y}(\tau, \tau) = 0 = U(\tau, \tau)$. Then by equation (5.10),

$$\begin{aligned}(U(\tau, t))^+ &= \textstyle\int_\tau^t dU(\tau, s)\\ &= \textstyle\int_\tau^t \{d[\hat{\mathscr{P}}(\tau, s)(\hat{B}(\tau, s) - B(\tau, s))] - d\hat{\mathscr{P}}(\tau, s)[\hat{B}(\tau, s) - B(\tau, s)]\}\\ &= (\hat{\mathscr{P}}(\tau, t))^+[\hat{B}(\tau, t) - B(\tau, t)]\\ &\quad - \textstyle\int_\tau^t [\hat{\mathscr{P}}(\tau, s)(d\hat{B}(\tau, s))(\hat{B}(\tau, s) - B(\tau, s))] .\end{aligned}$$

$$\sup_{\tau \leqq t \leqq t_1} N^p |(\hat{\mathscr{P}}(\tau, t))^+[\hat{B}(\tau, t) - B(\tau, t)]| \to 0 \quad \text{a.s.}$$

componentwise by Lemma 5.8. Furthermore,

$$\begin{aligned}&\{\textstyle\int_\tau^t [\hat{\mathscr{P}}(\tau, s)(d\hat{B}(\tau, s))(\hat{B}(\tau, s) - B(\tau, s))]\}_{ij}\\ &\qquad = \textstyle\sum_{l \in A} \sum_{k \in A} \int_\tau^t \hat{P}_{ik}(\tau, s)(\hat{\beta}_{lj}(\tau, s) - \beta_{lj}(\tau, s))\, d\hat{\beta}_{kl}(\tau, s)\\ &\qquad \leqq \textstyle\sum_{l \in A} \sum_{k \in A} \sup_{\tau \leqq t \leqq t_1} |\hat{\beta}_{lj}(\tau, t) - \beta_{lj}(\tau, t)|\, |\hat{\beta}_{kl}(\tau, t_1)| .\end{aligned}$$

By Lemma 5.8, it follows that

$$\sup_{\tau \leqq t \leqq t_1} N^p |U(\tau, t)| \to 0 \quad \text{a.s.} \quad \text{componentwise.} \tag{5.11}$$

Since $dB(\tau, t) = \mathscr{V}(t)\, dt$, $d\mathscr{P}(\tau, t) = \mathscr{P}(\tau, t)\mathscr{V}(t)\, dt$, $d\hat{\mathscr{P}}(\tau, t) = \hat{\mathscr{P}}(\tau, t)\, d\hat{B}(\tau, t)$, and $N^{-p}\mathscr{Y}(\tau, t) = \hat{\mathscr{P}}(\tau, t) - \mathscr{P}(\tau, t)$ it follows by equation (5.10) that

$$d(N^{-p}\mathscr{Y}(\tau, t)) - N^{-p}\mathscr{Y}(\tau, t)\mathscr{V}(t) = dU(\tau, t) . \tag{5.12}$$

Following the same argument as that given in Section 4, we have

$$\begin{aligned}N^{-p}\mathscr{Y}(\tau, t) &= \textstyle\int_\tau^{t-} d(U(\tau, s))\mathscr{P}(s, t)\\ &= U(\tau, s)\mathscr{P}(s, t)|_{s=\tau}^{t-} - \textstyle\int_\tau^{t-} U(\tau, s)\, d\mathscr{P}(s, t)\\ &= U(\tau, t) + \textstyle\int_\tau^{t-} U(\tau, s)\mathscr{V}(s)\mathscr{P}(s, t)\, ds .\end{aligned}$$

Hence by equation (5.11) the theorem follows. □

6. Absolute distribution. For clarity, the subscript A, which has been suppressed in Sections 4 and 5, will be retained again in this section.

Our main focus has been on the partial transition probabilities $\mathscr{P}_A(\tau, t)$. Let us momentarily consider the absolute distribution defined as follows.

Let $\mathbf{P}_A(t)$, with $(\mathbf{P}_A(t))_j = P_{Aj}(t)$, be the vector whose jth component represents the probability of being in the jth state of A at time t, in the hypothetical instance when only states in A are present. Hence $\sum_{j \in A} P_{Aj}(t) = 1$.

Let $N_A(t)$ be the number of the N processes which are in some state in A at time t, so $N_A(t) = \sum_{j \in A} N_j(t)$.

For motivation, we give the following example in which the data has been artificially constructed.

EXAMPLE 6.1. Consider a sample of 10 offspring, all of whom have both parents classified as sickle cell anemia carriers. Each member of the sample is observed for 75 years, and Table 1 represents the distribution of the sample amongst the 5 states and shows all transitions which took place.

Assume one would like to estimate $P_{\{1,3,5\}3}(75)$, which is the probability of an offspring having died from sickle cell anemia within 75 years of birth in the hypothetical situation where the risk of atherosclerosis is eliminated.

A natural estimator might be $N_3(75)/N_{\{1,3,5\}}(75) = 2/2 + 2 + 2 = .33$.

This estimator $N_j(t)/N_A(t)$ in effect distributes the probability attached to the eliminated states $\{A^0 - A\}$ amongst the states in A in such a manner as to preserve their probabilistic proportion to one another.

However, inspection of the data reveals that sickle cell anemia is an early occurring disease, while death due to atherosclerosis and other causes is late occurring. Hence, those who contract and die from atherosclerosis would almost assuredly die from some other late occurring disease in the hypothetical situation where the risk of atherosclerosis is eliminated. Thus one can see from the data that $P_{\{1,3,5\}3}(75)$ is approximately .20 rather than .33.

In general then, the ratio $N_j(t)/N_A(t)$ is not a consistent estimator of $P_{Aj}(t)$, for $P_{Aj}(t)$ as defined in the theory.

On the other hand, note that since atherosclerosis is not present at birth, $P_{\{1,3,5\}3}(0) = P_3(0)$. Since $N_j(\tau)/N$ is a consistent estimator of $P_j(\tau)$, one can obtain a consistent estimator of $P_{Aj}(\tau)$ when $P_{Aj}(\tau) = P_j(\tau)$.

In this section, using results from earlier sections and making the assumption $P_{Aj}(\tau) = P_j(\tau)/\sum_{i \in A} P_i(\tau)$ for some $\tau \in T$, we will derive an estimator $\hat{P}_{Aj}(t)$ of $P_{Aj}(t)$ which will not only be consistent for $t = \tau$ but also for $t \in [\tau, t_1]$.

In the context of this example, with $A = \{1, 3, 5\}$ and $j \in A$, our estimator $\hat{P}_{Aj}(t)$ will be a consistent estimator of $P_{Aj}(t)$ for all t, since $P_{Aj}(0) = P_j(0) = P_j(0)/\sum_{i \in A} P_i(0)$. Notice from the table that $\hat{P}_{\{1,3,5\}3}(75) = .20$, which coincides with what one would desire after inspection of the data. □

TABLE 1

State	Age										
	Birth	5	20	35	40	50	55	60	65	70	75
(1) Health (no atherosclerosis present)	10	9	8	7	6	5	4	4	3	2	2
(2) Living, but atherosclerosis present	0	0	0	1	2	2	3	1	2	2	1
(3) Death from sickle cell anemia	0	1	2	2	2	2	2	2	2	2	2
(4) Death from atherosclerosis	0	0	0	0	0	0	0	2	2	2	3
(5) Death from other causes	0	0	0	0	0	1	1	1	1	2	2
$N_3(t)/N_{\{1,3,5\}}(t)$	.00	.10	.20	.22	.25	.25	.29	.29	.33	.33	.33
$\hat{P}_{\{1,3,5\}1}(t)$	1.00	.90	.80	.80	.80	.67	.67	.67	.67	.44	.44
$\hat{P}_{\{1,3,5\}3}(t)$	.00	.10	.20	.20	.20	.20	.20	.20	.20	.20	.20
$\hat{P}_{\{1,3,5\}5}(t)$	.00	.00	.00	.00	.00	.13	.13	.13	.13	.36	.36

As noted in the example above, the results in this section are valid only for the cases in which the following assumption holds.

ASSUMPTION 6.1. $P_{Aj}(\tau) = P_j(\tau)/\sum_{i \in A} P_i(\tau)$ for some $\tau \in T$. □

Note that Assumption 6.1 is often valid if $\sum_{i \in A} P_i(\tau) = 1$.

Define

$$\begin{aligned} \hat{\hat{P}}_{Aj}(t) &= N_j(t)/N_A(t) \qquad \text{if } N_A(t) > 0 \\ &= 0 \qquad \text{if } N_A(t) = 0 . \end{aligned}$$

LEMMA 6.1. *If Assumption* 6.1 *is valid at* τ, *then for* $p < \frac{1}{2}$,

$$N^p |\hat{\hat{P}}_{Aj}(\tau) - P_{Aj}(\tau)| \to 0 \quad \text{a.s.}$$

PROOF. By Assumption 6.1, if $N_A(\tau) > 0$,

$$\begin{aligned} N^p &|\hat{\hat{P}}_{Aj}(\tau) - P_{Aj}(\tau)| \\ &= N^p \left| \frac{N_j(\tau)}{N_A(\tau)} - \frac{P_j(\tau)}{\sum_{i \in A} P_i(\tau)} \right| \\ &\leqq (\textstyle\sum_{i \in A} P_i(\tau))^{-1} \left\{ N^p \left| \frac{N_j(\tau)}{N} - P_j(\tau) \right| + \frac{N_j(\tau)}{N_A(\tau)} N^p \left| \frac{N_A(\tau)}{N} - \sum_{i \in A} P_i(\tau) \right| \right\} \\ &\leqq (\textstyle\sum_{i \in A} P_i(\tau))^{-1} \left\{ N^p \left| \frac{N_j(\tau)}{N} - P_j(\tau) \right| + N^p \left| \frac{N_A(\tau)}{N} - \sum_{i \in A} P_i(\tau) \right| \right\} . \end{aligned}$$

Hence, the lemma follows from Lemma 5.4 and the fact that

$$\frac{1}{N} N_A(\tau) \to \textstyle\sum_{i \in A} P_i(\tau) > 0 \quad \text{a.s.} \qquad \square$$

For any $t \geqq \tau$,

$$\begin{aligned} \mathbf{P}_A'(t) &= \mathbf{P}_A'(\tau)\mathscr{P}_A(\tau, t) = \mathbf{P}_A'(\tau)[I_A + \textstyle\int_\tau^t \mathscr{P}_A(\tau, s)\, dB_A(\tau, s)] \\ &= \mathbf{P}_A'(\tau) + \textstyle\int_\tau^t \mathbf{P}_A'(s)\, dB_A(\tau, s) . \end{aligned} \tag{6.1}$$

It is natural to suggest $\hat{\mathbf{P}}_A'(t) = \hat{\hat{\mathbf{P}}}_A'(\tau)\hat{\mathscr{P}}_A(\tau, t)$, so

$$\hat{\mathbf{P}}_A'(t) = \hat{\hat{\mathbf{P}}}_A'(\tau) + \textstyle\int_\tau^{t-} \hat{\mathbf{P}}_A'(s)\mathscr{R}_A(s)\, dM_A(\tau, s) .$$

THEOREM 6.1. *If Assumption* 6.1 *is valid at* τ, *then*

$$\sup_{t \in [\tau, t_1]} N^p |\hat{\mathbf{P}}_A'(t) - \mathbf{P}_A'(t)| \to 0 \quad \text{a.s.} \quad \textit{componentwise.}$$

PROOF. $\hat{\mathbf{P}}_A'(t) - \mathbf{P}_A'(t) = \hat{\hat{\mathbf{P}}}_A'(\tau)\hat{\mathscr{P}}_A(\tau, t) - \mathbf{P}_A'(\tau)\mathscr{P}_A(\tau, t) = \hat{\hat{\mathbf{P}}}_A'(\tau)[\hat{\mathscr{P}}_A(\tau, t) - \mathscr{P}_A(\tau, t)] + [\hat{\hat{\mathbf{P}}}_A'(\tau) - \mathbf{P}_A'(\tau)]\mathscr{P}_A(\tau, t)$. Thus the proof is completed by Theorem 5.1 and Lemma 6.1. □

REFERENCES

AALEN, O. (1977). Weak convergence of stochastic integrals related to counting processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **38** 261–277.

AALEN, O. (1978). Nonparametric estimation of partial transition probabilities in multiple decrement models. *Ann. Statist.* **6** 534–546.

BARLOW, R. E. and VAN ZWET, W. R. (1970). Asymptotic properties of isotonic estimators for the generalized failure rate function. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.). Cambridge Univ. Press.

BERKSON, J. and GAGE, R. (1950). Calculation of survival rates for cancer. *Proc. of the Staff Meetings of the Mayo Clinic* **25** 270-286.

BOEL, R., VARAIYA, P. and WONG, E. (1973). Martingales on jump processes I: representation results. *SIAM J. Control* **13** 999-1021.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and the product limit estimates under random censorship. *Ann. Math. Statist.* **2** 437-454.

CHIANG, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics.* Wiley, New York.

CHUNG, K. L. (1968). *A Course in Probability Theory.* Harcourt, Brace and World, New York.

CORNFIELD, J. (1957). The estimation of the probability of developing a disease in the presence of competing risks. *Amer. J. Publ. Health* **47** 601-607.

CUTLER, S. and EDERER, F. (1958). Maximum utilization of the life table method in analyzing survival. *J. Chronic Diseases* **8** 699-712.

DOLIVO, F. (1974). Counting processes and integrated conditional rates: a martingale approach with application to detection. Technical Report, College of Engineering, Univ. of Michigan.

Elveback, L. (1958). Actuarial estimation of survivorship in chronic disease. *J. Amer. Statist. Assoc.* **53** 420-440.

FELLER, W. (1940). On the integrodifferential equations of purely discontinuous Markoff processes. *Trans. Amer. Math. Soc.* **48** 488-515.

FLEMING, T. R. (1978). A review of the theory of life testing with special emphasis on competing risks. Mimeographed manuscript, Mayo Clinic, Rochester, Minnesota.

GAIL, M. (1975). A review and critique of some models used in competing risk analysis. *Biometrics* **31** 209-222.

GEHAN, E. (1968). Estimating survival functions from the life table. *J. Chronic Diseases* **21** 629-644.

HOEM, J. M. (1969). Purged and partial Markov chains. *Scandinavian Actuarial Journal* **52** 147-155.

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

KIMBALL, A. W. (1960). Estimation of mortality intensities in animal experiments. *Biometrics* **16** 505-521.

LITTEL, A. S. (1952). Estimation of the T-year survival rate from follow-up studies over a limited period of time. *Hum. Biol.* **24** 87-116.

LOÈVE, M. (1963). *Probability Theory.* Van Nostrand, New York.

NELSON, W. (1969). Hazard plotting for incomplete failure data. *J. Quality Tech.* **1** 27-52.

RUDIN, W. (1965). *Principles of Mathematical Analysis.* McGraw-Hill, New York.

TSIATIS, A. (1975). A non-identifiability aspect of the problem of competing risks. *Proc. Nat. Acad. Sci.* **72** No. 1.

DEPARTMENT OF MEDICAL STATISTICS
AND EPIDEMIOLOGY
MAYO CLINIC
ROCHESTER, MINNESOTA 55901