The Annals of Statistics
1978, Vol. 6, No. 5, 1071–1079

ASYMPTOTIC DISTRIBUTION RESULTS IN COMPETING RISKS ESTIMATION

BY THOMAS R. FLEMING

University of Maryland

Consider a time-continuous nonhomogeneous Markovian process V having state space A^0. For $A \subset A^0$ and $i, j \in A$, $P_{Aij}(\tau, t)$ is the $i \to j$ transition probability of the Markovian process V_A which arises in the hypothetical situation where states $A^0 - A$ have been eliminated from the state space of V. Let $\hat{P}_{Aij}(\tau, t)$ be the generalized product-limit estimator of $P_{Aij}(\tau, t)$. It is shown that the vector consisting of components in $\{N^{\frac{1}{2}}(\hat{P}_{Aij}(\tau, t) - P_{Aij}(\tau, t)) : i, j \in A;\ i \neq j\}$ converges weakly to a vector of dependent Gaussian processes. The structure of this limiting vector process is studied. Finally these results are applied to the estimation of certain biometric functions.

1. Introduction and summary. Consider a probability space (Ω, F, P) and a time-continuous nonhomogeneous Markovian stochastic process $V \equiv \{V(t) : t \in T = [t_0, t_1]\}$ with left-continuous sample paths. V has finite state space A^0 consisting of s transient and r absorbing states.

The transition probabilities of V, given by $P_{ij}(\tau, t) = P(V(t) = j | V(\tau) = i)$ for $i, j \in A^0$, are assumed to have continuous intensity functions $\nu_{ij}(t)$ such that for $i \neq j$, $P_{ij}(\tau, \tau + h) = h\nu_{ij}(\tau) + o(h)$ and $P_{ii}(\tau, \tau + h) = 1 + h\nu_{ii}(\tau) + o(h)$. The cumulative intensity function is given by $\beta_{ij}(\tau, t) = \int_\tau^t \nu_{ij}(s)\, ds$.

Let $A \subset A^0$ and $i, j \in A$. We will be interested in $P_{Aij}(\tau, t)$ which is the $i \to j$ transition probability of the Markovian stochastic process V_A which arises in the hypothetical situation where states $A^0 - A$ have been eliminated from the state space of V.

More generally, as suggested by Hoem (1969), one could be interested in the $i \to j$ transition probabilities arising when certain transitions of V rather than certain states of V are eliminated. Techniques employed and basic results achieved in this paper are also valid for this general situation. We shall adopt Hoem's terminology "partial transition probability" for $P_{Aij}(\tau, t)$.

If $\nu_{Aij}(t)$ is the intensity function corresponding to the partial transition probability, $P_{Aij}(\tau, t)$ then by the following assumption, $P_{Aij}(\tau, t)$ is well defined.

ASSUMPTION 1.1.

$$\nu_{Aij}(t) = \nu_{ij}(t) \qquad \text{for any} \quad i \neq j; \qquad i, j \in A .$$
$$\nu_{Aii}(t) = -\sum_{j \in A; j \neq i} \nu_{ij}(t) \qquad \text{for any} \quad i \in A .$$

Based upon the concept of Kaplan and Meier's (1958) product limit estimator,

Received May 1977; revised October 1977.
AMS 1970 *subject classifications*. Primary 62E20; Secondary 60J75, 62N05.
Key words and phrases. Competing risks, product-limit, nonparametric, nonhomogeneous, weak convergence, expected survival time.

a nonparametric estimator $\hat{P}_{Aij}(\tau, t)$ of $P_{Aij}(\tau, t)$ was formulated by Fleming (1978).

In Section 3 of this paper, it is shown that the vector consisting of components in $\{N^{\frac{1}{2}}(\hat{P}_{Aij}(\tau, t) - P_{Aij}(\tau, t)) : i \neq j;\ i, j \in A\}$ converges weakly to a vector of dependent Gaussian processes as $N \to \infty$, and in Section 4 the covariance structure of this limiting vector process is studied.

These results rely upon those by Aalen (1977, 1978) concerning the function $\beta_{ij}(\tau, t)$. He relied upon recent results concerning the decomposition of counting processes, stochastic integrals, and weak convergence of martingales.

Finally in Section 5 the above results are applied to the estimation of certain biometric functions which in the literature have been pointed out to be of interest.

2. Nonparametric estimator of $P_{Aij}(\tau, t)$. For purposes of estimation 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 1, in addition to the following assumption.

ASSUMPTION 2.1. Let $P_i(\tau) \equiv P(V_j(\tau) = i)$ for $i \in A^0$. There exists $\varphi > 0$ such that $P_i(\tau) > \varphi$ for any $i \in A_l \cap A$ and for any $\tau \in T$, where A_l is the set of nonabsorbing states in A^0.

We may now define $\pi_i^{-1}(t) = (P_i(t))^{-1}$ if $i \in A_l$, and $\pi_i^{-1}(t) = 0$ if $i \in A^0 - A_l$.

$\mathbf{N}(t)$ 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. $R_i(t)$ is defined to be $[N_i(t)]^{-1}$ if $N_i(t) > 0$, and to be zero if $N_i(t) = 0$. Then if we let m be the cardinality of A, the $(m \times m)$ diagonal matrix $\mathscr{R}_A(t)$ is defined by $(\mathscr{R}_A(t))_{ii} = R_i(t)$ for any $i \in A$.

The upcoming lemma, whose validity follows essentially from the fact that $N_i(t)$ is binomially distributed, will be useful in Section 3.

LEMMA 2.1. *For every integer n and for any $i \in A_l \cap A$, there exists K_n, where $K_n < \infty$, such that $E[NR_i(t)]^n \leqq K_n$ for every $t \in T$, and $N = 1, 2, \cdots$. In other words, one can bound $E(NR_i(t))^n$ uniformly in t by a term which is independent of the sample size N.*

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

$\hat{\beta}_{Aij}(\tau, t)$ is the generalized cumulative hazard estimator (see Fleming (1978)) given by $\hat{\beta}_{Aij}(\tau, t) = \int_\tau^t R_i(s)\, dM_{Aij}(\tau, s)$.

This, through the use of Kolmogorov's forward differential equation, leads to the generalized product limit estimator given by

$$\hat{P}_{Aij}(\tau, t) = (\hat{\mathscr{P}}_A(\tau, t))_{ij} = (I_A + \textstyle\int_\tau^t \hat{\mathscr{P}}_A(\tau, s)\mathscr{R}_A(s)\, dM_A(\tau, s))_{ij}$$

where I_A is an $(m \times m)$ identity matrix.

For notational simplicity, the subscript A will be suppressed throughout the remainder of this paper, except on the process V_A. Furthermore, we will assume $A = \{1, 2, \cdots, m\}$.

3. Weak convergence. Let $D(T)$ be the space of functions on the interval T which have discontinuities of only the first kind; that is, each function in $D(T)$ has left- and right-hand limits everywhere in T. Let d be the Skorohod metric on $D(T)$. In the following sections, the term "weak convergence" will be used with respect to the product metric d_n on the product space $D^n(T)$ for appropriate values of n, and will be denoted $\Rightarrow$. Convergence in probability will be with respect to d_n, except when referring to random variables, in which case it will be in the usual sense and will be denoted by $\rightarrow_p$. Almost sure convergence of random variables will be denoted by $\rightarrow_{\text{a.s.}}$.

We will have frequent need for the following processes. Assume $\{W_{ij}: 1 \leqq i, j \leqq m, i \neq j\}$ is a collection of independent Wiener processes. For any $1 \leqq i, j \leqq m$; $i \neq j$; define

$$\xi_{ij}(\tau, t) = \int_\tau^t g_i(s)\, dW_{ij}(s) \qquad \text{where} \quad g_i^2(t) = \pi_i^{-1}(t)\nu_{ij}(t)\,.$$

From their definition, we see that $\{\xi_{ij}: 1 \leqq i, j \leqq m;\ i \neq j\}$ is a collection of independent Gaussian processes, each with continuous sample paths and independent increments, such that

$$E\xi_{ij}(\tau, t) = 0 \qquad \text{and} \qquad \operatorname{Var} \xi_{ij}(\tau, t) = \int_\tau^t \nu_{ij}(s)\pi_i^{-1}(s)\, ds\,.$$

Finally define

$$\boldsymbol{\xi} = (\xi_{12}, \xi_{13}, \cdots, \xi_{1m}, \xi_{21}, \xi_{23}, \cdots, \xi_{m,m-1})'$$

and

$$\xi_{ii} = -\sum_{j=1; j\neq i}^m \xi_{ij} \qquad \text{for any} \quad i = 1, \cdots, m\,.$$

3.1. *Weak convergence of* $\boldsymbol{\chi}$. Define $\beta_{ij}^*(\tau, t) = \int_\tau^t \nu_{ij}(s) I_{[N_i(s)\neq 0]}\, ds$ and $X_{ij}^*(\tau, t) = N^{\frac{1}{2}}(\hat\beta_{ij}(\tau, t) - \beta_{ij}^*(\tau, t))$ for any $1 \leqq i, j \leqq m$.

The following is a result of Theorem 6.4 of Aalen (1978).

LEMMA 3.1. *Let* $\boldsymbol{\chi}^* = (X_{12}^*, X_{13}^*, \cdots, X_{1m}^*, X_{21}^*, X_{23}^*, \cdots, X_{m,m-1}^*)'$. *Then* $\boldsymbol{\chi}^* \Rightarrow \boldsymbol{\xi}$.

Now define $X_{ij}(\tau, t) = N^{\frac{1}{2}}(\hat\beta_{ij}(\tau, t) - \beta_{ij}(\tau, t))$ for any $1 \leqq i, j \leqq m$. We will need the next elementary lemma.

LEMMA 3.2. $\sup_{\tau\leq t\leq t_1} |X_{ij}^*(\tau, t) - X_{ij}(\tau, t)| \rightarrow_p 0$ *for any* $1 \leqq i, j \leqq m, i \neq j$.

PROOF. $\sup_{\tau\leq t\leq t_1} |X_{ij}^*(\tau, t) - X_{ij}(\tau, t)| = \sup_{\tau\leq t\leq t_1} N^{\frac{1}{2}} \int_\tau^t \nu_{ij}(s) I_{[N_i(s)=0]}\, ds \leqq (t_1 - \tau)\nu_i \sup_{\tau\leq t\leq t_1} N^{\frac{1}{2}} I_{[N_i(t)=0]} = (t_1 - \tau)\nu_i N^{\frac{1}{2}} I_{[\inf_{t\in T} N_i(t)=0]}$, where $\nu_i = \sup_{t\in T} -\nu_{ii}(t)$. The result now follows from Lemma 5.4 of Fleming (1978). □

If $\boldsymbol{\chi} = (X_{12}, X_{13}, \cdots, X_{m,m-1})'$ Lemmas 3.1 and 3.2 imply the following result:

LEMMA 3.3. $\boldsymbol{\chi} \Rightarrow \xi$.

3.2. *Weak convergence of* vec $\mathscr{Y}(\tau, t)$. Assume $\mathscr{X}$ is an arbitrary matrix with dimension $u \times v$. Define vec $\mathscr{Z}$ to be the vector formed by combining rows of $\mathscr{Z}$; that is, if $(\mathscr{Z})_{ij} = z_{ij}$, then

$$\text{vec}\,\mathscr{Z} = (z_{11}, z_{12}, \cdots, z_{1v}, z_{21}, \cdots, z_{2v}, \cdots, z_{u1}, \cdots, z_{uv})' .$$

Define the $(m \times m)$ dimensional matrices $\mathscr{P}$, $\mathscr{V}$, Ξ, $\mathscr{X}$, and $\mathscr{X}^*$ by $(\mathscr{P}(\tau, t))_{ij} = P_{ij}(\tau, t)$; $(\mathscr{V}(s))_{ij} = \nu_{ij}(s)$; $(\Xi)_{ij} = \xi_{ij}$; $(\mathscr{X})_{ij} = X_{ij}$; $(\mathscr{X}^*)_{ij} = X_{ij}^*$ for any $1 \leqq i, j \leqq m$.

Define $\mathscr{Y}(\tau, t) = N^{\frac{1}{2}}(\hat{\mathscr{P}}(\tau, t) - \mathscr{P}(\tau, t))$. We will now prove the key result of this paper.

THEOREM 3.1. vec $\mathscr{Y}(\tau, t) \Rightarrow$ vec $\int_\tau^t \mathscr{P}(\tau, s)\, d\Xi(\tau, s)\mathscr{P}(s, t)$ *where the integrals are stochastic integrals in the quadratic mean.*

PROOF. Equip (Ω, F, P) with an increasing family of sub-σ-fields of F, $\{F_t, t \in T\}$ where $F_t = \sigma(V_k(s) : t_0 \leqq s \leqq t, 1 \leqq k \leqq N)$. Martingales with which we will be dealing are adapted to $\{F_t\}$.

By equation (5.12) of Fleming (1978)

$$d\mathscr{Y}(\tau, s) - \mathscr{Y}(\tau, s)\mathscr{V}(s) = \hat{\mathscr{P}}(\tau, s)\, d\mathscr{X}(\tau, s) .$$

Following a derivation similar to that given in Section 4 of that source we have

$$\begin{aligned} (\mathscr{Y}(\tau, t))^+ &= \int_\tau^t \hat{\mathscr{P}}(\tau, s)\, d\mathscr{X}(\tau, s)\mathscr{P}(s, t) \\ &= \int_\tau^t \mathscr{P}(\tau, s)\, d\mathscr{X}(\tau, s)\mathscr{P}(s, t) \\ &\quad + \int_\tau^t \mathscr{Y}(\tau, s)\, d[N^{-\frac{1}{2}}\mathscr{X}^*(\tau, s)]\mathscr{P}(\tau, s) \\ &\quad + N^{-\frac{1}{2}} \int_\tau^t \mathscr{Y}(\tau, s)\, d[\mathscr{X}(\tau, s) - \mathscr{X}^*(\tau, s)]\mathscr{P}(s, t) \end{aligned} \tag{3.1}$$

where these integrals and those that follow until otherwise mentioned are still Lebesgue-Stieltjes integrals, and where $\mathscr{Y}^+$ is the right-continuous adaptation of the process $\mathscr{Y}$.

Since componentwise $N^{-\frac{1}{2}}\mathscr{Y}(\tau, s)$ and $\mathscr{P}(s, t)$ are bounded in absolute value by one, the third term on the right-hand side of equation (3.1) converges to zero in probability componentwise by Lemma 3.2.

Next observe that

$$\begin{aligned} &\left(\int_\tau^t \mathscr{Y}(\tau, s)\, d(N^{-\frac{1}{2}}\mathscr{X}^*(\tau, s))\mathscr{P}(s, t)\right)_{il} \\ &\quad = \sum_{j=1}^m \sum_{k=1}^m \int_\tau^t N^{\frac{1}{2}}[\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s)]P_{kl}(s, t)R_j(s) \\ &\qquad \times d[M_{jk}(\tau, s) - \int_\tau^s \nu_{jk}(u)N_j(u)\, du] . \end{aligned} \tag{3.2}$$

Now, by Hölder's inequality,

$$\begin{aligned} &E[N(\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s))^2 P_{kl}^2(s, t)R_j(s)\nu_{jk}(s)] \\ &\quad \leqq P_{kl}^2(s, t)\nu_{jk}(s)[E|\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s)|^4]^{\frac{1}{2}}[E(NR_j(s))^2]^{\frac{1}{2}} . \end{aligned} \tag{3.3}$$

The right-hand side of equation (3.3) converges to zero by Theorem 5.1 of Fleming (1978), Lemma 2.1, and the fact that $|\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s)| \leqq 1$. Hence we have shown that

$$[N^{\frac{1}{2}}(\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s))P_{kl}(s, t)R_j(s)]^2\nu_{jk}(s)N_j(s) \rightarrow_p 0 . \tag{3.4}$$

It follows from Prososition 3 of Doléans-Dadé and Meyer (1970) that the Lebesgue–Stieltjes integral in equation (3.2) coincides with the Meyer stochastic integral (see e.g., Meyer (1971)).

The Meyer stochastic integral is well defined since $\{N^{\frac{1}{2}}[\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s)]P_{kl}(s, t)R_j(s): s \in [\tau, t]\}$ is a left-continuous process adapted to $\{F_t\}$ such that

$$E\{\int_\tau^t [N^{\frac{1}{2}}(\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s))P_{kl}(s, t)R_j(s)]^2 \nu_{jk}(s)N_j(s)\, ds\} < \infty .$$

Thus by Theorem 2.1 of Aalen (1977), and equation (3.4),

$$\{\int_\tau^t N^{\frac{1}{2}}[\hat{P}_{ij}(\tau, s) - P_{ij}(\tau, s)]P_{kl}(s, t)R_j(s) \\ \times d[M_{jk}(\tau, s) - \int_\tau^s \nu_{jk}(u)N_j(u)\, du]: 1 \leqq j, k \leqq m\}$$

is a collection of square integrable martingales, each of which converges to 0 in probability.

Hence, the second term on the right-hand side of equation (3.1) converges to 0 in probability, componentwise.

Next observe that, by integration by parts,

$$\begin{aligned}\text{vec} \int_\tau^t \mathscr{P}(\tau, s)\, d[\mathscr{X}(\tau, s)]\mathscr{P}(s, t) &= \text{vec}\,\{\mathscr{P}(\tau, t)\mathscr{X}(\tau, t) - \int_\tau^t ([d\mathscr{P}(\tau, s)]\mathscr{X}(\tau, s)\mathscr{P}(s, t)) \\ &\quad - \int_\tau^t \mathscr{P}(\tau, s)\mathscr{X}(\tau, s)\, d\mathscr{P}(s, t)\} ,\end{aligned}$$

which is a continuous mapping from $D^{m(m-1)}$ to D^{m^2}.

Hence, by Theorem 5.1 of Billingsley (1968), Lemma 3.3, and equation (3.1),

$$\begin{aligned}\text{vec}\, \mathscr{Y}(\tau, \cdot) \Rightarrow \text{vec}\,\{&\mathscr{P}(\tau, \cdot)\Xi(\tau, \cdot) - \int_\tau^t ([d\mathscr{P}(\tau, s)]\Xi(\tau, s)\mathscr{P}(s, t)) \\ &- \int_\tau^t \mathscr{P}(\tau, s)\Xi(\tau, s)\, d\mathscr{P}(s, t)\} .\end{aligned}$$

The proof of the theorem is completed through an integration by parts using equation (5.3.7) of Cramér and Leadbetter (1967). The Lebesgue–Stieltjes integrals appearing in the limiting vector above have the same value as stochastic integrals in the quadratic mean, thus justifying the use of that equation.

4. Covariance considerations for the limiting vector process.

4.1. *Covariance structure of* vec $\int_\tau^t \mathscr{P}(\tau, s)[d\Xi(\tau, s)]\mathscr{P}(s, t)$. For future statistical applications, we need to investigate the covariance structure of vec $\int_\tau^t \mathscr{P}(\tau, s)[d\Xi(\tau, s)]\mathscr{P}(s, t)$.

Define $\Psi(\tau, t) = \int_\tau^t \mathscr{P}(\tau, s)[d\Xi(\tau, s)]\mathscr{P}(s, t)$, and $\psi_{ij}(\tau, t) = (\Psi(\tau, t))_{ij}$ for any $1 \leqq i, j \leqq m$.

Thus,

$$\begin{aligned}\psi_{ij}(\tau, t) &= \sum_{\alpha=1}^m \sum_{\beta=1}^m \int_\tau^t P_{i\alpha}(\tau, s)P_{\beta j}(s, t)\, d\xi_{\alpha\beta}(\tau, s) \\ &= \sum_{\alpha=1}^m \sum_{\beta=1;\beta\neq\alpha}^m \int_\tau^t P_{i\alpha}(\tau, s)[P_{\beta j}(s, t) - P_{\alpha j}(s, t)]\, d\xi_{\alpha\beta}(\tau, s) .\end{aligned}$$

Clearly $\psi_{ij}(\tau, \cdot)$ is a mean zero Gaussian process.

Let $1 \leqq i, j, k, l \leqq m$, $\tau < s$, and $t < u$. By the independence of the $\{\xi_{ij}; i \neq j\}$

we have

$$\text{Cov}\,[\psi_{ij}(\tau, s), \psi_{kl}(t, u)] \tag{4.1}$$
$$= \sum_{\alpha=1}^m \sum_{\beta=1;\beta\neq\alpha}^m \int_t^u \int_\tau^s P_{i\alpha}(\tau, w)[P_{\beta j}(w, s) - P_{\alpha j}(w, s)]P_{k\alpha}(t, v) \times [P_{\beta l}(v, u) - P_{\alpha l}(v, u)]\, d\,\text{Cov}\,(\xi_{\alpha\beta}(\tau, w), \xi_{\alpha\beta}(t, v))\,.$$

Since $\{\xi_{ij};\, i \neq j\}$ have independent increments,

$$\text{Cov}\,(\xi_{\alpha\beta}(\tau, s), \xi_{\alpha\beta}(t, u)) \quad \text{equals} \quad 0 \qquad \text{if} \quad \tau \leqq s \leqq t \leqq u\,, \tag{4.2}$$

and equals

$$\int_t^{(s\wedge u)} \nu_{\alpha\beta}(y)\pi_\alpha^{-1}(y)\, dy \qquad \text{if} \quad \tau \leqq t \leqq (s \wedge u)\,,$$

where $(s \wedge u) \equiv \min(s, u)$.

Hence the next lemma follows immediately from equations (4.1) and (4.2).

LEMMA 4.1. *For any* $1 \leqq i, j, k, l \leqq m$:

$$\text{Cov}\,(\psi_{ij}(\tau, s), \psi_{kl}(t, u)) = 0 \qquad \textit{if} \quad \tau \leqq s \leqq t \leqq u \tag{4.3}$$

and

$$\text{Cov}\,(\psi_{ij}(\tau, s), \psi_{kl}(t, u)) \tag{4.4}$$
$$= \sum_{\alpha=1}^m \sum_{\beta=1;\beta\neq\alpha}^m \int_t^{(s\wedge u)} P_{i\alpha}(\tau, v)[P_{\beta j}(v, s) - P_{\alpha j}(v, s)]P_{k\alpha}(t, v) \times [P_{\beta l}(v, u) - P_{\alpha l}(v, u)]\nu_{\alpha\beta}(v)\pi_\alpha^{-1}(v)\, dv \qquad \textit{if} \quad \tau \leqq t \leqq (s \wedge u)\,.$$

COROLLARY 4.1. *For any* $1 \leqq i, j \leqq m$, *and* $\tau \leqq s \leqq t$,

$$\text{Cov}\,(\psi_{ij}(\tau, s), \psi_{ij}(\tau, t))$$
$$= \sum_{\alpha=1}^m \sum_{\beta=1;\beta\neq\alpha}^m \int_\tau^s P_{i\alpha}^2(\tau, u)[P_{\beta j}(u, s) - P_{\alpha j}(u, s)] \times [P_{\beta j}(u, t) - P_{\alpha j}(u, t)]\nu_{\alpha\beta}(u)\pi_\alpha^{-1}(u)\, du\,.$$

4.2. *Covariance estimation.* We propose the following natural estimator for $\text{Cov}\,(\psi_{ij}(\tau, s), \psi_{kl}(t, u))$ where $1 \leqq i, j, k, l \leqq m$, and $\tau \leqq t \leqq (s \wedge u)$. Setting $\gamma = (s \wedge u)$, define

$$\widehat{\text{Cov}}\,(\psi_{ij}(\tau,s), \psi_{kl}(t, u))$$
$$= \sum_{\alpha=1}^m \sum_{\beta=1;\beta\neq\alpha}^m \int_t^\gamma \hat{P}_{i\alpha}(\tau,v)[\hat{P}_{\beta j}(v, s) - \hat{P}_{\alpha j}(v, s)]\hat{P}_{k\alpha}(t, v) \times [\hat{P}_{\beta l}(v, u) - \hat{P}_{\alpha l}(v, u)]NR_\alpha^2(v)\, dM_{\alpha\beta}(\tau, v)\,.$$

In the next theorem, it is proved that this estimator is uniformly strongly consistent.

THEOREM 4.1. *For any* $\tau \leqq t$, *and* $1 \leqq i, j, k, l \leqq m$,

$$\sup_{s,u\in[t,t_1]} \{\widehat{\text{Cov}}\,[\psi_{ij}(\tau, s), \psi_{kl}(t, u)] - \text{Cov}\,[\psi_{ij}(\tau, s), \psi_{kl}(t, u)]\} \rightarrow_{\text{a.s.}} 0\,.$$

PROOF. In this proof "uniform" means over $\{(s, u, v): t \leqq v \leqq \gamma = (s \wedge u);\, s, u \leqq t_1\}$. $\widehat{\text{Cov}}\,[\psi_{ij}(\tau, s), \psi_{kl}(t, u)] - \text{Cov}\,[\psi_{ij}(t, s), \psi_{kl}(t, u)] = \sum_{\alpha=1;\alpha\in A_l}^m \sum_{\beta=1;\beta\neq\alpha}^m \{\int_t^\gamma (\hat{\theta}-\theta)R_\alpha(v)\, dM_{\alpha\beta}(\tau, v) + \int_t^\gamma \theta[R_\alpha(v)dM_{\alpha\beta}(\tau, v) - d\beta_{\alpha\beta}(\tau, v)]\}$

where

$$\theta = P_{i\alpha}(\tau, v)[P_{\beta j}(v, s) - P_{\alpha j}(v, s)]P_{k\alpha}(t, v)[P_{\beta l}(v, u) - P_{\alpha l}(v, u)]\pi_\alpha^{-1}(v) .$$

By Theorem 5.1, Lemma 5.6, and Lemma 5.8 of Fleming (1978) it follows that $\hat{\theta} - \theta \rightarrow_{\text{a.s.}} 0$ uniformly and $|\hat{\beta}_{\alpha\beta}(t, v) - \beta_{\alpha\beta}(t, v)| \rightarrow_{\text{a.s.}} 0$ uniformly, which together imply that $\int_t^\gamma (\hat{\theta} - \theta)R_\alpha(v)\, dM_{\alpha\beta}(\tau, v) \rightarrow_{\text{a.s.}} 0$ uniformly.

Noting that $|\theta|$ is uniformly bounded by φ^{-1}, and that Lemma 5.7 of Fleming (1978) implies θ is of bounded variation, it then follows through an integration by parts and from another use of the result in Lemma 5.8 that

$$\int_t^\gamma \theta[R_\alpha(v)dM_{\alpha\beta}(\tau, v) - d\beta_{\alpha\beta}(\tau, v)] \rightarrow_{\text{a.s.}} 0 \quad \text{uniformly.} \qquad \square$$

5. Some biometric functions. For the stochastic process V_A, let $e_{ij}(t, t_1)$ be the expected period of time spent in state j within the interval (t, t_1), conditionally for a process in state i at time t. Furthemore, let $e_i(t, t_1)$ be the expected "survival time" (i.e., time spent in states in A_l) in (t, t_1), conditionally for a process in state i at time t.

Fix and Neyman (1951) emphasized the importance of estimating biometric functions such as those just defined. In the special case where A^0 has one transient and one absorbing state, Chiang (1968) introduced a nonparametric estimator of $e_{ij}(t, t_1)$, which was later generalized to the case of an arbitrary number of absorbing states by Yang (1976), who at the same time investigated its distributional properties.

In this section we will inspect estimators for $e_{ij}(t, t_1)$ and $e_i(t, t_1)$ in the general situation where the number of absorbing and transient states in A^0 is arbitrary but finite. These estimators are conceptually based upon the product-limit estimator while those presented by Yang which are defined when A^0 has one transient state are based upon the closely related empirical cumulative hazard estimator discussed by Nelson (1969) and Breslow and Crowley (1974).

5.1. $e_{ij}(t, t_1)$.

$$\begin{aligned} e_{ij}(t, t_1) &= \int_t^{t_1} P(V_A(s) = j \mid V_A(t) = i)\, ds \\ &= \int_t^{t_1} P_{ij}(t, s)\, ds \qquad \text{for any} \quad 1 \leqq i, j \leqq m . \end{aligned}$$

If $\xi(t, t_1)$ is an $(m \times m)$ matrix defined by $(\xi(t, t_1))_{ij} = e_{ij}(t, t_1)$ for any $1 \leqq i, j \leqq m$, then

$$\begin{aligned} N^{\frac{1}{2}}(\hat{\xi}(t, t_1) - \xi(t, t_1)) &= \int_t^{t_1} N^{\frac{1}{2}}(\hat{\mathscr{P}}(t, s) - \mathscr{P}(t, s))\, ds \\ &= \int_t^{t_1} \mathscr{Y}(t, s)\, ds . \end{aligned}$$

As usual, the integrals above are Lebesgue–Stieltjes. We now need the next lemma, whose validity follows from direct verification.

LEMMA 5.1. *The map* $H\colon D^{m^2}[t_0, t_1] \rightarrow D^{m^2}[t_0, t_1]$ *defined by* $H(\operatorname{vec} \Psi)(t) = \operatorname{vec} \int_t^{t_1} \Psi(t, s)\, ds$ *for* $\operatorname{vec} \Psi \in D^{m^2}[t_0, t_1]$ *is a continuous map with respect to* d_{m^2}.

Hence by Theorem 5.1 of Billingsley, and Lemma 5.1, we have the following theorem.

THEOREM 5.1. $\operatorname{vec} N^{\frac{1}{2}}(\hat{\xi}(t, t_1) - \xi(t, t_1)) \Rightarrow \operatorname{vec} \int_t^{t_1} \Psi(t, s)\, ds$.

If we define $Z_{ij}^{(1)}(t, t_1) = \int_t^{t_1} \psi_{ij}(t, s)\, ds$, we have $EZ_{ij}^{(1)}(t, t_1) = 0$ for any $1 \leqq i, j \leqq m$ and for any $t \in [t_0, t_1]$. Furthermore, for any $t_0 \leqq x \leqq y \leqq t_1$, and $1 \leqq i, j, k, l \leqq m$, we have

$$\begin{aligned}
&\operatorname{Cov}[Z_{ij}^{(1)}(x, t_1), Z_{kl}^{(1)}(y, t_1)] \\
&\quad = E \int_x^{t_1} \int_y^{t_1} \psi_{ij}(x, s)\psi_{kl}(y, u)\, du\, ds \\
&\quad = \int_x^{t_1} \int_y^{t_1} \operatorname{Cov}(\psi_{ij}(x, s), \psi_{kl}(y, u))\, du\, ds \qquad (5.1) \\
&\quad = \sum_{\alpha=1}^m \sum_{\alpha=1;\beta\neq\alpha}^m \int_y^{t_1} \int_y^{t_1} \int_y^{(u\wedge s)} P_{i\alpha}(x, v)[P_{\beta j}(v, s) - P_{\alpha j}(v, s)] \\
&\qquad \times P_{k\alpha}(y, v)[P_{\beta l}(v, u) - P_{\alpha l}(v, u)]\nu_{\alpha\beta}(v)\pi_\alpha^{-1}(v)\, dv\, du\, ds\,,
\end{aligned}$$

where the integral is a stochastic integral in the quadratic mean; hence the second equality in equation (5.1) follows from equation (5.3.8) of Cramér and Leadbetter. This last equality follows from equations (4.3) and (4.4).

Note that equation (5.1) implies that

$$\begin{aligned}
\operatorname{Var} Z_{ij}^{(1)}(x, t_1) = 2 \sum_{\alpha=1}^m \sum_{\beta=1;\beta\neq\alpha}^m &\int_x^{t_1} \int_s^{t_1} \int_x^s P_{i\alpha}^2(x, v)[P_{\beta j}(v, s) - P_{\alpha j}(v, s)] \\
&\times [P_{\beta j}(v, u) - P_{\alpha j}(v, u)]\nu_{\alpha\beta}(v)\pi_\alpha^{-1}(v)\, dv\, du\, ds\,.
\end{aligned}$$

5.2. $e_i(t, t_1)$.

$$\begin{aligned}
e_i(t, t_1) &= \int_t^{t_1} P(V_A(s) \in A_l \cap A \mid V_A(t) = i)\, ds \\
&= \int_t^{t_1} \sum_{j \in A_l \cap A} P_{ij}(t, s)\, ds \qquad \text{for any} \quad 1 \leqq i \leqq m\,.
\end{aligned}$$

For any $i \in A$, define

$$\begin{aligned}
\zeta_i &= 1 \qquad \text{if} \quad i \in A_l \\
&= 0 \qquad \text{if} \quad i \notin A_l\,.
\end{aligned}$$

If $\mathbf{e}(t, t_1)$ is defined by

$$\mathbf{e}(t, t_1) = (e_1(t, t_1), e_2(t, t_1), \cdots, e_m(t, t_1))'$$

and $\boldsymbol{\zeta}$ by $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \cdots, \zeta_m)'$ then

$$\begin{aligned}
N^{\frac{1}{2}}(\hat{\mathbf{e}}(t, t_1) - \mathbf{e}(t, t_1)) &= \int_t^{t_1} N^{\frac{1}{2}}(\hat{\mathscr{P}}(t, s) - \mathscr{P}(t, s))\boldsymbol{\zeta}\, ds \\
&= \int_t^{t_1} \mathscr{Y}(t, s)\boldsymbol{\zeta}\, ds\,.
\end{aligned}$$

If we define $\int_t^{t_1} \Psi(t, s)\boldsymbol{\zeta}\, ds = \mathbf{Z}^{(2)}(t, t_1)$ and $\int_t^{t_1} (\Psi(t, s)\boldsymbol{\zeta})_i\, ds = Z_i^{(2)}(t, t_1)$, then by the same token that we arrived at Theorem 5.1, we have the following:

THEOREM 5.2. $N^{\frac{1}{2}}(\hat{\mathbf{e}}(t, t_1) - \mathbf{e}(t, t_1)) \Rightarrow \mathbf{Z}^{(2)}(t, t_1)$.

$\mathbf{Z}^{(2)}(t, t_1)$ is a Gaussian vector process such that for any $1 \leqq i \leqq m$ and $t \in [t_0, t_1]$, $EZ_i^{(2)}(t, t_1) = 0$.

Furthermore for any $t_0 \leqq x \leqq y \leqq t_1$, and $1 \leqq i, j \leqq m$, we have

$$\begin{aligned}
&\operatorname{Cov}[Z_i^{(2)}(x, t_1), Z_j^{(2)}(y, t_1)] \\
&\quad = E \int_x^{t_1} \int_y^{t_1} \sum_{k \in A_l \cap A} \psi_{ik}(x, s) \sum_{l \in A_l \cap A} \psi_{jl}(y, t)\, dt\, ds \\
&\quad = \int_x^{t_1} \int_y^{t_1} \operatorname{Cov}(\sum_{k \in A_l \cap A} \psi_{ik}(x, s), \sum_{l \in A_l \cap A} \psi_{jl}(y, t))\, dt\, ds \\
&\quad = \sum_{\alpha=1}^m \sum_{\beta=1}^m \int_y^{t_1} \int_y^{t_1} \int_y^{u\wedge s} \sum_{k \in A_l \cap A} \{P_{i\alpha}(x, v)[P_{\beta k}(v, s) - P_{\alpha k}(v, s)]\} \\
&\qquad \times \sum_{l \in A_l \cap A} \{P_{j\alpha}(y, v)[P_{\beta l}(v, u) - P_{\alpha l}(v, u)]\}\nu_{\alpha\beta}(v)\pi_\alpha^{-1}(v)\, dv\, du\, ds
\end{aligned}$$

where the last equality follows from equation (5.1) and equations (4.3) and (4.4).

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 inference for a family of counting processes. *Ann. Statist.* **6** 701–726.

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

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.

CRAMÉR, H. and LEADBETTER, M. (1967). *Stationary and Related Stochastic Processes.* Wiley, New York.

DOLÉANS-DADÉ, C. and MEYER, P. A. (1970). Intégrales stochastiques par rapport aux martingales locales. *Seminaire de probabilités IV. Lecture Notes in Mathematics* **124** 77–107. Springer-Verlag, Berlin.

FIX, E. and NEYMAN, J. (1951). A simple stochastic model of recovery, relapse, death and loss of patients. *Hum. Biol.* **23** 205–241.

FLEMING, T. R. (1978). Nonparametric estimation for nonhomogeneous Markov processes in the problem of competing risks. *Ann. Statist.* **6** 1057–1070.

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

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

MEYER, P. (1971). Square integrable martingales, a survey. *Lecture Notes in Mathematics* **190** 32–37. Springer-Verlag, Berlin.

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

YANG, G. (1976). Estimation of a biometric function. *Ann. Statist* **6** 112–116.

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