Scand J Statist 5: 141–150, 1978

An Empirical Transition Matrix for Non-homogeneous Markov Chains Based on Censored Observations

ODD O. AALEN[1] and SØREN JOHANSEN

University of Tromsø and University of Copenhagen

Received August 1977, revised April 1978

ABSTRACT. A product limit estimator is suggested for the transition probabilities of a non-homogeneous Markov chain with finitely many states. The estimator is expressed as a product integral and its properties are studied by means of the theory of square integrable martingales.

Key words: Markov chains, censored observations, product limit estimator, transition probabilities

I. The model and a summary of the results

We shall consider a right continuous Markov chain $(X_t, t\in[0,1])$ on a finite state space E with intensities or forces of transition given by $Q(t)=(q_{ij}(t), i\in E, j\in E)$ where for all $i\neq j\in E$

$$q_{ij}(t)>0,\quad q_{ii}(t)<0\quad \text{and } \textstyle\sum_j q_{ij}(t)=0, \tag{1.1}$$

$q_{ij}(\cdot)$ is left continuous and has finite right hand limits, such that (1.2)

$$\int_0^1 q_{ij}(t)\,dt<\infty. \tag{1.3}$$

It is well known, see Goodman (1970), and Dobrushin (1953), that under these assumptions, the transition probabilities are given by the differential equations

$$\frac{\partial}{\partial s}P(s,t)=-Q(s)P(s,t) \tag{1.4}$$

$$\frac{\partial}{\partial t}P(s,t)=P(s,t)\,Q(t) \tag{1.5}$$

with initial condition $P(s,s)=I$. The equations hold almost surely with respect to Lebesgue-measure.

The solution to these equations $P(s,t)$ is absolutely continuous as a function of s and t and is given by the product integral

$$P(s,t)=\prod_{]s,t]}(I+Q(u)\,du)\quad 0\leqslant s\leqslant t\leqslant 1 \tag{1.6}$$

see e.g. Dobrushin (1953) or Johansen (1977).

The solution satisfies the Chapman–Kolmogorov equation

$$P(s,t)=P(s,u)P(u,t)\quad 0\leqslant s\leqslant u\leqslant t\leqslant 1. \tag{1.7}$$

Consider first the problem of estimating $P(s,t)$ on the basis of independent observations $\{X_t^{(k)}, t\in[0,1], k=1,\ldots,n\}$, where the kth process has transition probabilities $P(s,t)$ and initial distribution $p^{(k)}$.

The obvious estimator of the transition probability is

$$P_{ij}^{(1)}(s,t)=\frac{\sum_{r=1}^{n}1\{X_{s-}^{(r)}=i,\ X_t^{(r)}=j\}}{\sum_{r=1}^{n}1\{X_{s-}^{(r)}=i\}} \tag{1.8}$$

which is simply the fraction of observations, available in i at time s, which end up in j at time t.

This estimator does not satisfy equation (1.7) and does therefore not belong to the class of functions considered in the model. The idea underlying (1.8) is also not very useful when censoring is present.

The estimator can be modified as follows: We split the interval [0, 1] by a partition $\{t_m\}$ so fine that in each interval at most 1 jump occur. We then apply (1.8) to each interval and define

$$\hat{P}(s,t)=\prod_{s<t_m<t_{m+1}\leqslant t}P^{(1)}(t_m,t_{m+1}). \tag{1.9}$$

Each of these factors is either the identity, if no jump occurs, or a stochastic matrix with only 1 off diago-

[1] Supported by the Norwegian Research Council for Science and the Humanities.

nal element positive and equal to $[\sum_{r=1}^{n} 1\{X_{u-}^{(r)} = i\}]^{-1}$ placed at position (i, j) if the jump took place from i to j at time u. A different representation is given in (5.1).

This estimator does not depend on the choice of partition and can be considered a generalization of the product limit estimator discussed by Kaplan & Meier (1958) and Breslow & Crowley (1974). Notice how the estimator is constructed as a product integral, a concept which formalizes that of a product limit. It is a basic idea of this paper that Markov chain transition probabilities are constructed in exactly the same way from their intensities as the above product limit estimator is constructed from the observed jumps. It is the formalism of the product integral representation that allows us to write up the basic stochastic integral equation (3.3) which again allows the martingale theory to be applied. The estimator (1.9) therefore satisfies (1.7) and, as we shall see in a certain sense also (1.4) and (1.5). It is not however absolutely continuous and a third estimator $\hat{\hat{P}}$ can be constructed which interpolates between the steps of $\hat{P}$.

This estimator is constructed using the observation that the factors of $\hat{P}$ are simple stochastic matrices that are imbeddable in time continuous homogeneous chains, see Johansen (1973).

The methods used for analyzing the independent identically distributed observations can also be used for analyzing the situation where the processes $X^{(1)}, \ldots, X^{(n)}$ are censored. This is an important extension of the theory since in many medical and engineering applications the processes are only under observation part of the time. Therefore we present in section 2 a general model of censoring including most of those commonly considered in the literature. We will show that the multiplicative intensity model for counting processes (see Aalen 1978*b*) play a central role in the description of censored processes.

In section 3 we will show how an estimator for the integrated intensity allows us to define $\hat{P}$ as a product integral relative to the censored processes. This section also contains the exact properties of the estimator. We derive those by employing the theory of square integrable martingales as well as results on product integrals to represent $\hat{P}$ as the solution to a stochastic integral equation. From this follows certain martingale properties of $\hat{P}$.

The same technique is used in section 4 to obtain the asymptotic distribution of the process $\hat{P}$. In 5 the smoothed estimator is discussed and it is proved that it has the same asymptotic properties as $\hat{P}$. Finally, in the appendix we have collected a few items that supplement the methods used in the paper.

The estimator $\hat{P}$ was previously suggested and studied by one of the authors (Aalen, 1973 and 1978*a*) in the case when E has only one transient state. It was stated in Aalen (1973) that an extension to general Markov chains was possible. When the present paper was essentially finished, it came to our attention that Fleming (1978*a*, *b*) has suggested the same estimator $\hat{P}$ as we have. He does, however, only treat the uncensored case and gives no exact results. He also does not give the smoothed version of the estimator. Also, our methods are quite different from Flemings, especially our use of the product integral representation.

Fleming, by the way, estimates partial transition probabilities. It will be apparent from our approach that the number of processes under observation in each state may be allowed to vary quite freely. Clearly, therefore, the estimation of partial transition probabilities is really included in our treatment.

2. A general model for censoring

Define $Y_i^{(k)}(t) = 1\{X_{t-}^{(k)} = i\}$. Let $K_{ij}^{(k)}(t)$ denote the number of jumps directly from i to j that $X^{(k)}$ has performed in the time interval $[0, t]$, hence the process $K_{ij}^{(k)}$ is right-continuous. Put $K = \{K_{ij}^{(k)}, k = 1, \ldots, n, i, j \in E\}$.

Let Ω_0 be the space of possible sample paths of K for $t \in [0, 1]$. Let $(A, \mathcal{A}, P_1)$ be a probability space with P_1 not depending on the intensity Q, and let Ω be the Cartesian product of Ω_0 and A. Let $\mathfrak{F}_t$ be the product σ-algebra on Ω corresponding to the σ-algebra $\mathcal{A}$ on A and the σ-algebra on Ω_0 generated by $\{K(s), 0 \leq s \leq t\}$. The family $\{\mathfrak{F}_t, 0 \leq t \leq 1\}$ is increasing and right-continuous (see Boel et al., 1975). Let P be the product measure on $\mathfrak{F} = \mathfrak{F}_1$ generated by the measure P_1 on $\mathcal{A}$ and the measure on $\mathcal{N}_1 = \sigma\{X_t^{(k)}, 0 \leq t \leq 1, k = 1, \ldots, n\}$ given by the previously defined Markovian structure.

K is a multivariate counting process with $K_{ij}^{(k)}$ having intensity process $q_{ij} Y_i^{(k)}$ relative to $\{\mathfrak{F}_t\}$. In this paper we will exploit the recently developed martingale-based approach to counting processes, see e.g. Boel et al. (1975) or the short review in Aalen (1978*b*). A consequence of that theory is that the $M_{ij}^{(k)}$ defined by

$$M_{ij}^{(k)}(t) = K_{ij}^{(k)}(t) - \int_0^t q_{ij}(s)\, Y_i^{(k)}(s)\, ds$$

are orthogonal, square integrable martingales with variance process

$$\langle M_{ij}^{(k)}, M_{ij}^{(k)} \rangle (t) = \int_0^t q_{ij}(s)\, Y_i^{(k)}(s)\, ds.$$

(See e.g. Meyer, 1971 for these concepts.)

The *censoring process* is a stochastic process $J = (J_1, ..., J_n)$ which has piecewise constant and left-continuous sample functions taking the values 0 and 1 and with a finite number of jumps. We also assume that $J(t)$ is measurable with respect to $\{\mathfrak{F}_t\}$ for all $t \in [0, 1]$. The process $X_t^{(k)}$ is observed at those times t for which $J_k(t) = 1$.

Notice that $J(t)$ may depend in almost arbitrary ways on what has been observed in the past and on outside random variation (modeled by the space A). Hence, our censoring scheme is considerably more general than those commonly considered in the literature, see e.g. Kaplan & Meier (1958).

Define now the stochastic integrals

$$\hat{M}_{ij}^{(k)}(t) = \int_0^t J_k(s)\, dM_{ij}^{(k)}(s).$$

By the theory of stochastic integrals (see e.g. Meyer, 1971), the $\hat{M}_{ij}^{(k)}$ are orthogonal, square integrable martingales. Hence, by the above mentioned counting process theory, $\hat{K} = \{\hat{K}_{ij}^{(k)}, k = 1, ..., n, i, j \in E\}$ given by

$$\hat{K}_{ij}^{(k)} = \int_0^t J_k(s)\, dK_{ij}^{(k)}(s)$$

is a counting process with $\hat{K}_{ij}^{(k)}$ having intensity process $q_{ij} J_k Y_i^{(k)}$ relative to $\{\mathfrak{F}_t\}$. We call $\hat{K}$ the *censored process*.

Define now $N_{ij} = \sum \hat{K}_{ij}^{(k)}$, $N_i = \sum J_k Y_i^{(k)}$, $N^* = \{N_i, i \in E\}$ and $N = \{N_{ij}, i, j \in E\}$. N is a counting process with N_{ij} having intensity process $q_{ij} N_i$. The assumptions made above imply that N and N^* are observed over the time interval [0, 1].

Assume for a moment that $\mathcal{A}$ is the trivial σ-algebra. In Aalen (1978*b*) it is proved that the statistic $\{\sum_k K_{ij}^{(k)}, i, j \in E\}$ is complete for the nonparametric model defined by (1.1), (1.2) and (1.3). Since (N, N^*) is a measurable function of that statistic it follows that (N, N^*) is complete.

If (N, N^*) is sufficient, then we have a case of the multiplicative intensity model studied in Aalen (1978*b*). Obviously, the inference procedures developed for that model is applicable whether (N, N^*) is sufficient or not, but in the latter case some information will be lost by only applying those procedures.

The question of when (N, N^*) is sufficient has been treated in Aalen (1978*c*). The results indicate that (N, N^*) is in general sufficient when all J_i are decreasing processes, i.e. in the case of right-censoring, while otherwise it will generally not be sufficient. In the case of uncensored processes sufficiency is clear by a simple likelihood consideration (see Aalen (1978*b*)).

Finally, one should note that this whole problem of sufficiency only arises when one can follow each process $X^{(k)}$ *individually* over the time interval $\{t: J_k(t) = 1\}$. Sometimes this may not be the case, and one may at any time t only be able to observe the numbers of processes in each state and the jumps that occur without knowing which process that jumps. Of course, this amounts precisely to observing (N, N^*).

At any rate, in the present paper we will study estimation of $P(s, t)$ based solely on the statistic (N, N^*). Put $M_{ij}(t) = N_{ij}(t) - \int_0^t N_i(s) q_{ij}(s)\, ds$.

3. The estimator and its exact properties

It was suggested by Aalen (1978*b*) that one should use

$$\hat{B}_{ij}(t) = \int_0^t 1\{N_i(s) \geq 1\} N_i(s)^{-1} dN_{ij}(s) \quad i \neq j \qquad (3.1)$$

as an estimator for the integrated intensity $B_{ij}(t) = \int_0^t q_{ij}(s)\, ds$. This estimator has previously been suggested for life testing models by Altshuler (1970) and Nelson (1969).

From the integrated intensity (3.1) it is now possible to estimate the transition probabilities using the theory of product integrals.

The necessary theory for constructing Markov chains from integrated intensities, that are not absolutely continuous with respect to Lebesgue measure was given by Dobrushin (1953), who developed the product integral for matrix valued measures, see also Johansen (1977). A similar problem was solved by Jacobsen (1972) for countable state chains.

In terms of product integral we now define

$$\hat{P}(s, t) = \prod_{]s, t]} (I + d\hat{B}) \qquad (3.2)$$

where $\hat{B}_{ij}$ is given by (3.1) and $\hat{B}_{ii} = -\sum_{j \neq i} \hat{B}_{ij}$.

Since $\hat{B}$ is a purely discrete measure with finite support, $\hat{P}$ reduces to a finite product of stochastic matrices. For very small intervals $]s, t]$, either $\hat{B}]s, t] = 0$ in which case $\hat{P}(s, t) = I$ or there is one jump occuring in $]s, t]$, from i to j at time u, say, then $dB_{ij}(u) = N_i(u)^{-1}$ and this means that $\hat{P}(s, t)$ is a stochastic matrix with only 1 off diagonal element different from zero. Thus (3.2) is the same as the estimator (1.9).

Once the intensity is estimated one can also derive estimates of the waiting time distributions

$$G_i[0, t] = 1 - \prod_{[0, t]} (1 + dB_{ii})$$

by simply inserting $\hat{B}_{ii}$ instead of B_{ii}.

Put $F_i(t) = 1 - G_i[0, t]$. Then we get:

$$\hat{F}_i(t) = 1 - \hat{G}_i[0, t] = \prod_{[0, t]} (1 + d\hat{B}_{ii}) = \prod_{t_r \leq t} \left(1 - \frac{1}{N_i(t_r)}\right)$$

where $\{t_r\}$ denotes the times where a jump occurs from i.

This estimator generalizes, in a different sense than $\hat{P}$, the estimator suggested by Kaplan & Meier (1958).

Let

$$\tilde{B}_{ij}(t) = \int_0^t 1\{N_i(u) \geq 1\} q_{ij}(u)\, du$$

and

$$\tilde{P}(s, t) = \prod_{]s, t]} (I + d\tilde{B}).$$

Then the following result is an analogue to Theorem 6.2 of Aalen (1978*b*). See the appendix for the notation.

Theorem 3.1. *The process* $\hat{P}(0, t) = \prod_{[0, t]} (I + d\hat{B})$ *satisfies the stochastic integral equation*

$$\hat{P}(0, t)\, \tilde{P}(0, t)^{-1} - I = \int_0^t \hat{P}(0, s)\, d(\hat{B} - \tilde{B})(s)\, \tilde{P}(0, s)^{-1} \tag{3.3}$$

Further $M_t = \hat{P}(0, t)\tilde{P}(0, t)^{-1} - I$ *is a square integrable martingale and*

$$\langle M_t, M_t\rangle = \int_0^t \hat{P}(0, s) \otimes \hat{P}(0, s)\, d\langle \hat{B} - \tilde{B}, \hat{B} - \tilde{B}\rangle (s)\, \tilde{P}(0, s)^{-1} \otimes \tilde{P}(0, s)^{-1} \tag{3.4}$$

where $\otimes$ *denotes the Kronecker matrix product.*

It follows from the above relations that

$$E\hat{P}(0, t)\tilde{P}(0, t)^{-1} = I$$

and that

$$V\{\hat{P}(0, t)\, \tilde{P}(0, t)^{-1}\} = E(\hat{P}(0, t)\, \tilde{P}(0, t)^{-1} - I)\, (\hat{P}(0, t)\, \tilde{P}(0, t)^{-1} - I)' = E\int_0^t \hat{P}(0, s-) \otimes \hat{P}(0, s-)\, d\langle \hat{B} - \tilde{B}, \hat{B} - \tilde{B}\rangle (s)\, \tilde{P}(0, s)^{-} \otimes \tilde{P}(0, s)^{-1}.$$

Proof. As is shown by Johansen (1977), $\hat{P}$ satisfies the differential equation

$$\frac{d\hat{P}(0, t)}{dv_0} = \hat{P}(0, t-)\frac{d\hat{B}}{dv_0} \quad \text{a.s. } [v_0]$$

and $\tilde{P}(t, 1)$ satisfies the equation

$$\frac{d\tilde{P}(t, 1)}{dv_0} = -\frac{d\tilde{B}}{dv_0}\tilde{P}(t, 1) \quad \text{a.s. } [v_0]$$

where v_0 is a measure that dominates the measures of $\hat{B}$ and $\tilde{B}$. One can take $v_0 = -tr\hat{B} - tr\tilde{B}$.

Then

$$\frac{d}{dv_0}(\hat{P}(0, s)\, \tilde{P}(s, 1)) = \hat{P}(0, s-)\frac{d\tilde{P}(s, 1)}{dv_0} + \frac{d\hat{P}(0, s)}{dv_0}\tilde{P}(s, 1) = \hat{P}(0, s-)\left(\frac{d\hat{B}}{dv_0} - \frac{d\tilde{B}}{dv_0}\right)\tilde{P}(s, 1)$$

Integrating from 0 to t gives

$$\hat{P}(0, t)\, \tilde{P}(t, 1) - \tilde{P}(0, 1) = \int_0^t \hat{P}(0, s-)\, d(\hat{B} - \tilde{B})(s)\, \tilde{P}(s, 1) \tag{3.5}$$

Now

$$\text{Det}\, \tilde{P}(0, 1) = \exp\left\{\int_0^1 \sum_i 1\{N_i(u) \geq 1\}\, q_{ii}(u)\, du\right\} \geq \exp\left\{\sum_i \int_0^1 q_{ii}(u)\, du\right\} > 0.$$

Hence we can divide through in (3.5) by $\tilde{P}(0, 1)$ which proves the relation (3.3) as a Stieltjes integral for a given realization of the process.

In order to prove that it is a stochastic integral, and in fact a square integrable martingale we first note that each element of the matrix is the sum of a finite number of integrals of the form

$$\int_0^t \hat{p}_{ik}(0, s-)\, d(\hat{B}_{km} - \tilde{B}_{km})(s)\, \tilde{p}^{mj}(0, s)$$

The coefficients $\hat{p}_{ik}(0, s-)$ and $\tilde{p}^{mj}(0, s)$ are left continuous and measurable with respect to $\mathfrak{F}_s$. Further $\hat{p}_{ik}(0, s-) \leq 1$ and $\tilde{p}^{mj}(0, s)$ is evaluated as follows:

$$|\tilde{p}^{mj}(0, s)| = |(\text{Det}\, \tilde{P}(0, s))^{-1} \sum_\sigma (-1)^{|\sigma|} \prod_{\substack{i \neq m \\ \sigma(i) \neq j}} \tilde{p}_{i\sigma(i)}(0, s)| \leq (\text{Det}\, P(0, 1))^{-1}(k-1)!$$

where k is the number of states in E.

Thus the coefficients are bounded predictable processes. By proposition 3 of Doléans-Dadé & Meyer (1970) it now suffices to prove that

$$E\int_0^1 d|\hat{B}_{km}-\tilde{B}_{km}|(s)<\infty \tag{3.6}$$

and

$$E\int_0^1 d\langle\hat{B}_{km}-\tilde{B}_{km}, \hat{B}_{km}-\tilde{B}_{km}\rangle(s)<\infty \tag{3.7}$$

By Theorem 6.2 of Aalen (1978*b*) the process $\hat{B}_{km}-\tilde{B}_{km}$ is a square integrable martingale and

$$E|\hat{B}_{km}-\tilde{B}_{km}|(t) \leqslant E\int_0^t 1\{N_k(u)\geqslant 1\}N_k(u)^{-1}dN_{km}(u) + E\int_0^t 1\{N_k(u)\geqslant 1\}q_{km}(u)\,du$$

and

$$E\langle\hat{B}_{km}-\tilde{B}_{km}, \hat{B}_{km}-\tilde{B}_{km}\rangle(t) = E\int_0^t 1\{N_k(u)\geqslant 1\}N_k(u)^{-1}q_{km}(u)\,du. \tag{3.8}$$

All these integrals, however, are bounded by $\int_0^1|q_{kk}(u)\,du|<\infty$ which completes the proof of Theorem 3.1 (see the appendix).

Using a similar argument one can prove the following results about the generalized Kaplan–Meier estimator of the waiting time distribution.

Theorem 3.2. *The process* $\prod_{[0,t]}(I+d\hat{B}_{ii})=\hat{F}_i(t)$ *satisfies the stochastic integral equation*

$$\hat{F}_i(t)\tilde{F}_i(t)^{-1}-1=\int_0^t \hat{F}_i(s-)\,d(\hat{B}_{ii}-\tilde{B}_{ii})(s)\tilde{F}_i(s)^{-1},$$

where $\tilde{F}_i(t)=\prod_{[0,t]}(1+d\tilde{B}_{ii})$. *Hence*

$$S_t=\hat{F}_i(t)\tilde{F}_i(t)^{-1}-1,\quad i\in E,$$

are orthogonal square integrable martingales and

$$\langle S_t,S_t\rangle=\int_0^t \hat{F}_i(s-)^2 d\langle\hat{B}_{ii}-\tilde{B}_{ii},\hat{B}_{ii}-\tilde{B}_{ii}\rangle(s)\,\tilde{F}_i(s)^{-2}.$$

Remark. Notice that Theorems 3.1 and 3.2 only depend on the fact that N_{ij} is a counting process with intensity process $N_i q_{ij}$. Since the only important requirement to N_i in the counting process theory is that it be predictable, one may in principle introduce rules to control the size of N_i such that it does not become too small. This requires that one has an "infinite reservoir" of possible observations, that can be inserted into the states when needed. *If in particular* $N_i(t)\geqslant 1$ *for all i and t, then* $\tilde{P}=P$ *and so* $\hat{P}$ *is unbiased.* The same holds for the $\hat{F}_i$.

4. Asymptotic properties of the estimator

We shall first prove a general result about weak convergence of stochastic integrals, which is a modification of Theorem 2.1 in Aalen (1977). The conditions have been slightly changed so as to be easier to verify.

We start with a sequence of counting processes $N_{i,n}$, $i=1,\ldots,k$; $n=1,\ldots$ with intensities $\Lambda_{i,n}$ and we let $M_{i,n}(t)=N_{i,n}(t)-\int_0^t\Lambda_{i,n}(u)\,du$. Let $H_{i,n}$ be predictable processes satisfying $H_{i,n}\in L^2(M_{i,n})$ so that $Y_{i,n}=\int H_{i,n}\,dM_{i,n}$ is a square integrable martingale. Set $Y_n=(Y_{1,n},\ldots,Y_{k,n})$.

In order that these stochastic integrals can be evaluated as Stieltjes integrals we shall assume (Requirement A of Aalen (1977)):

$$E\int_0^1|H(s)|\,dN(s)<\infty \tag{4.1}$$

for any H and N as above.

A general result of Rebolledo (1977) about weak convergence of martingales shows that in order that the processes Y_n converge weakly to a Gaussian process the conditions of Theorem 2.1 of Aalen (1977) are sufficient:

$$\int_0^t H_{i,n}(s)^2\Lambda_{i,n}(s)\,ds \xrightarrow{P} \int_0^t g_i^2(s)\,ds \quad \forall_{i,t} \tag{4.2}$$

$$\sum_i E\int_0^1 H_{i,n}^2(s)\,1\{|H_{i,n}(s)|\geqslant\varepsilon\}\,dN_{i,n}(s)\to 0 \quad \forall\varepsilon>0 \tag{4.3}$$

Here $g_i(s)$ is some function in $L^2(0,1)$.

Theorem 4.1. *Let the following conditions be satisfied*

$$H_{i,n}(s)\xrightarrow{P}0,\quad n\to\infty\quad \forall i,s \tag{4.4}$$

$$H_{i,n}^2(s)\Lambda_{i,n}(s)\xrightarrow{P}g_i^2(s),\quad n\to\infty\quad\forall i,s \tag{4.5}$$

$$g_i^2(s)\quad \textit{bounded on } [0,1]\quad \forall i \tag{4.6}$$

$$H_{i,n}^2(s)\Lambda_{i,n}(s)\quad \textit{integrable uniformly in } (n,s,i). \tag{4.7}$$

Then if $W_1,\ldots,W_k$ *are independent Wiener processes and* $Y=(\int g_1\,dW_1,\ldots,\int g_k\,dW_k)$ *we have*

$$Y_n\overset{w}{\Rightarrow}Y.$$

Proof. We shall show that (4.2) and (4.3) are satisfied.

Since we are using uniform integrability we remind about the following standard result for a sequence of random variables $\{X_n\}$:

$$X_n \xrightarrow{P} 0, \quad X_n \text{ uniformly integrable} \Rightarrow EX_n \to 0. \quad (4.8)$$

Let now

$$K_{i,n}(s) = |H_{i,n}^2(s)\,\Lambda_{i,n}(s) - g_i^2(s)|.$$

By (4.5) $K_{i,n}(s) \xrightarrow{P} 0$ and since, by (4.6) and (4.7) $K_{i,n}(s)$ is uniformly integrable it follows from (4.8) that $EK_{i,n}(s) \to 0$. Since $EK_{i,n}(s)$ is bounded in (n, s) we get $\int_0^1 EK_{i,n}(s)\,ds \to 0$ but this easily implies (4.2).

To prove (4.3) note that

$$\int_0^t H_{i,n}^2(s)\,1\{|H_{i,n}(s)| \geq \varepsilon\}\,d\left\{N_{i,n}(s) - \int_0^s \Lambda_{i,n}(u)\,du\right\}$$

is a martingale and hence

$$E\int_0^t H_{i,n}^2(s)\,1\{|H_{i,n}(s)| \geq \varepsilon\}\,dN_{i,n}(s)$$
$$= E\int_0^t H_{i,n}^2(s)\,\Lambda_{i,n}(s)\,1\{|H_{i,n}(s)| \geq \varepsilon\}\,ds.$$

From (4.4) and (4.5) it follows that

$$Z_{i,n}(s) = H_{i,n}^2(s)\,\Lambda_{i,n}(s)\,1\{|H_{i,n}(s)| \geq \varepsilon\} \xrightarrow{P} 0,$$

but (4.7) implies that $Z_{i,n}(s)$ is uniformly integrable and hence $EZ_{i,n}(s) \to 0$ but since also $EZ_{i,n}(s)$ is bounded in (s, n) we get $\int_0^1 EZ_{i,n}(s)\,ds \to 0$, which proves (4.3).

For applications of this theorem it is worth noting that the uniform integrability of a sequence of random variables $\{X_n\}$ is implied by a condition like $E|X_n|^{1+\varepsilon} \leq c$ or, for positive variables by $X_n \overset{w}{\Rightarrow} X$ and $EX_n \to EX$.

We shall now study the asymptotic properties of $\hat{P}$ when the number n of observed processes increases to ∞. We write $J^{(n)} = (J_1^{(n)}, \ldots, J_n^{(n)})$ to indicate that each element of the censoring process may depend on all observed processes. Of course, all stochastic processes occurring below will depend on n, but we will generally suppress n from the notation. We will first give a consistency result. We define the norm of a matrix A by $|A| = \sup_i \sum_j |a_{ij}|$.

Theorem 4.2. *Make the following assumption:*

$E[1\{N_i(t) \geq 1\}N_i(t)^{-1}] + P\{N_i(t) = 0\} \to 0$ *for all i and t.*

Then:

$$\sup_t |\hat{P}(0, t) - P(0, t)| \xrightarrow{P} 0$$

Proof. We have

$$|\hat{P}(0, t) - P(0, t)| = |(\hat{P}(0, t)\tilde{P}(0, t)^{-1} - I)\tilde{P}(0, t)$$
$$+ (\tilde{P}(0, t)P(0, t)^{-1} - I)P(0, t)|$$
$$\leq |\hat{P}(0, t)\tilde{P}(0, t)^{-1} - I| + |\tilde{P}(0, t)P(0, t)^{-1} - I|$$

We will treat these two terms separately. From Theorem 3.1 we have

$$\hat{P}(0, t)\,\tilde{P}(0, t)^{-1} - I$$
$$= \int_0^t \hat{P}(0, s-)\,d(\hat{B} - \tilde{B})(s)\,\tilde{P}(0, s)^{-1}.$$

The ijth term of this integral is a sum of stochastic integrals of the form

$$\int_0^t \hat{p}_{ik}(0, s-)\,d(\hat{B}_{km} - B_{km})(s)\,\tilde{p}^{mj}(0, s)$$

and we shall therefore first study the joint limiting behaviour of the processes

$$Y_{km,n}(t) = \sqrt{n}\int_0^t \hat{p}_{ik}(0, s-)\,\tilde{p}^{mj}(0, s)\,d(\hat{B}_{km} - \tilde{B}_{km})(s)$$
$$= \int_0^t \sqrt{n}\,\hat{p}_{ik}(0, s-)\,\tilde{p}^{mj}(0, s)$$
$$\times 1\{N_k(s) \geq 1\}\,N_k(s)^{-1}\,dM_{km,n}(s) \quad (4.9)$$

where, here and in the following, i and j has been suppressed in the notation.

The stochastic integral given in (4.9) was investigated in (3.6), (3.7) and (3.8) and it was proved that

$$E\left(\frac{1}{\sqrt{n}}Y_{km,n}(1)\right)^2$$
$$= E\int_{[0,1]} \hat{p}_{ik}^2(0, s-)\,(\tilde{p}^{mj}(0, s))^2\,N_k(s)^{-1}$$
$$\times 1(N_k(s) \geq 1)\,q_{km}(s)\,ds$$
$$\leq cE\int_0^1 1\{N_k(s) \geq 1\}\,N_k(s)^{-1}q_{km}(s)\,ds$$

for some constant c. By the assumption and Lebesgues dominated convergence theorem the last expression converges to 0, and so it follows by a submartingale inequality (Doob, 1953, Theorem 3.4 and p. 354) that

$$\sup_t |\hat{P}(0, t)\,\tilde{P}(0, t)^{-1} - I| \xrightarrow{P} 0$$

Using

$$\tilde{P}(0, t)\,P(0, t)^{-1} - I = \int_0^t \tilde{P}(0, s)\,d(\tilde{B} - B)(s)\,P(0, s)^{-1}$$

we get

$$|\tilde{P}(0,t)P(0,t)^{-1}-I| \leq \int_0^t d|\tilde{B}-B|.$$

Now we have the evaluation

$$0 \leq B_{ij}(t)-\tilde{B}_{ij}(t) \leq \int_0^1 1\{N_i(s)=0\}\, q_{ij}(s)\, ds.$$

From the assumption of the theorem and Lebesgues dominated convergence theorem we have

$$E\int_0^1 1\{N_i(s)=0\}\, q_{ij}(s)\, ds = \int_0^1 P\{N_i(s)=0\}\, q_{ij}(s)\, ds \to 0$$

Hence we may conclude:

$$\sup_t |\tilde{P}(0,t)P(0,t)^{-1}-I| \xrightarrow{P} 0.$$

We will now prove a weak convergence result. In the proof we will consider the stochastic integrals $Y_{km,n}(t)$ defined by (4.9). These integrals have the form needed to apply Theorem 4.1 if we define

$$H_{km,n}(t) = \sqrt{n}\,\hat{p}_{ik}(0,t)\tilde{p}^{mj}(0,t)1\{N_k(t)\geq 1\}N_k(t)^{-1}.$$

$$\Lambda_{km,n}(t) = N_k(t)q_{km}(t),$$

$$H^2_{km,n}\Lambda_{km,n}(t) = n\hat{p}^2_{ik}(0,t-)(\tilde{p}^{mj}(0,t))^2 1\{N_k(t) \geq 1\}N_k(t)^{-1}q_{km}(t).$$

We will now make an application of Theorem 4.1. Let W_{km}, $k\neq m$, be independent Wiener processes and define

$$Y_{km}(t) = \int_0^t \left(\frac{q_{km}(s)}{p_k(s)}\right)^{\frac{1}{2}} dW_{km}(s) \quad k\neq m,$$

$$Y_{kk} = -\sum_{m\neq k} Y_{km},$$

$$U(t) = \int_0^t P(0,s)\,Y(ds)\,P(s,t).$$

Here $p_k(s)$ is a function assumed to be $\geq a>0$ for all k and s.

Theorem 4.3. *Make the following assumptions:*

(i) $\sqrt{n}\int_0^1 1\{N_i(s)=0\}q_{ii}(s)\,ds \xrightarrow{P} 0 \quad \forall i$

(ii) $\frac{1}{n}N_i(t) \xrightarrow{P} p_i(t) \quad \forall i, t,$

(iii) $n1\{N_i(t)\geq 1\}N_i(t)^{-1}$

is uniformly integrable in (n, t, i).

Then:

$$\sqrt{n}(\hat{P}-P) \Rightarrow U.$$

Proof. We first consider the stochastic integrals $Y_{km,n}$. We have to check the conditions of Theorem 4.1.

Condition (4.1) holds immediately since $H_{km,n}(t)$ is bounded and $EN_{km}(1)<\infty$.

Condition (4.4) follows from the consistency of $\hat{P}(0,t)$ together with assumption (ii). Similarly (4.5) follows and also (4.6). Condition (4.7) follows from assumption (iii) since the coefficients $\hat{p}^2_{ik}$ and $(\tilde{p}^{mj})^2$ are bounded.

Thus we have established that

$$Y_{km,n} \overset{w}{\Rightarrow} V_{km}$$

with

$$V_{km}(t) = \int_0^t p_{ik}(0,s)p^{mj}(0,s)\sqrt{\frac{q_{km}(s)}{p_k(s)}}\,dW_{km}(s)$$

$$= \int_0^t p_{ik}(0,s)p^{mj}(0,s)\,dY_{km}(s)$$

Hence

$$\sqrt{n}(\hat{P}(0,t)\tilde{P}(0,t)^{-1}-I) \overset{w}{\Rightarrow} \int_0^t P(0,s)\,dY(s)\,P(0,s)^{-1}$$

Now

$$\sqrt{n}(\hat{P}(0,t)-P(0,t)) = \sqrt{n}(\hat{P}(0,t)\tilde{P}(0,t)^{-1}-I)\tilde{P}(0,t) + \sqrt{n}(\tilde{P}(0,t)-P(0,t)).$$

Hence we have completed the proof if we can prove that

$$\sqrt{n}(\tilde{P}(0,t)-P(0,t)) \overset{w}{\Rightarrow} 0.$$

This follows, however, from

$$\tilde{P}(0,t)-P(0,t) = \int_0^t \tilde{P}(0,u)\,d(\tilde{B}-B)(u)\,P(u,t).$$

and

$$|\tilde{P}(0,t)-P(0,t)| \leq 2\sum_i \int_0^1 1\{N_i(u)=0\}\,|q_{ii}(u)|\,du$$

Hence by assumption (i)

$$\sup_t \sqrt{n}|\tilde{P}(0,t)-P(0,t)| \xrightarrow{P} 0, n\to\infty$$

We have now found the asymptotic distribution of $\hat{P}$. The covariance matrix of the limiting distribution is found as follows:

$$E\langle U, U\rangle(t) = \int_0^t P(0,s)\otimes P(0,s)\, d\langle Y, Y\rangle(s)\, P(s,t)\otimes P(s,t)$$

(see the appendix).

But

$$\langle Y_{km}, Y_{k'm'}\rangle(t) = \int_0^t \frac{q_{km}(u)}{p_k(u)}\, du \quad \text{if}$$

$(k', m') = (k, m)$ and 0 otherwise.

Let C_{ij} denote the intensity matrix with element (i, j) equal 1, element (i, i) equal -1 and the rest zero, then

$$\langle Y, Y\rangle(t) = \sum_{k\neq m} \int_0^t \frac{q_{km}(u)}{p_k(u)}\, du\, C_{km}\otimes C_{km}.$$

The following theorem gives a consistent estimator of the covariance matrix of the limiting distribution.

Theorem 4.4. *Suppose the conditions of Theorem 4.3 hold with* (iii) *substituted by*

(iii)′ $n^2 1\{N_i(t) \geq 1\} N_i(t)^{-2}$

is uniformly integrable in (n, t, i).

Define:

$$V_n(t) = \int_0^t \hat{P}(0,s)\otimes \hat{P}(0,s)\, dZ_n(t)\, \hat{P}(s,t)\otimes \hat{P}(s,t)$$

where

$$Z_n(t) = \sum_{i\neq j} \int_0^t n 1\{N_i(s)\geq 1\} N_i(s)^{-2} dN_{ij}(s)\, C_{ij}\otimes C_{ij}.$$

Then

$$\sup_t |V_n(t) - E\langle U, U\rangle(t)| \xrightarrow{P} 0.$$

Proof. Define:

$$Z_{ij,n}(t) = \int_0^t n 1\{N_i(s)\geq 1\} N_i(s)^{-2} dN_{ij}(s).$$

By the consistency of $\hat{P}$ it is enough to prove for each pair (i, j):

$$\sup_t \left| Z_{ij,n}(t) - \int_0^t \frac{q_{ij}(s)}{p_i(s)}\, ds \right| \xrightarrow{P} 0.$$

We have

$$\begin{aligned} Z_{ij,n}(t) - \int_0^t \frac{q_{ij}(s)}{p_i(s)}\, ds &= \int_0^t n 1\{N_i(s)\geq 1\} N_i(s)^{-2} dM_{ij}(s) \\ &\quad + \int_0^t [n 1\{N_i(s)\geq 1\} N_i(s)^{-1} - p_i(s)^{-1}]\, q_{ij}(s)\, ds \\ &= A_n(t) + B_n(t). \end{aligned}$$

It follows immediately from the conditions that $\sup_t |B_n(t)| \xrightarrow{P} 0$. Now, $A_n(t)$ is a square integrable martingale, and so $\sup |A_n(t)| \xrightarrow{P} 0$ if

$$E\langle A_n(1), A_n(1)\rangle = E\int_0^1 n^2 1\{N_i(s)\geq 1\} N_i(s)^{-3} q_{ij}(s)\, ds$$

converges to 0. This is, however, an immediate consequence of the assumptions, using the result (4.8).

The verification of the assumptions of the theorems in this section in the uncensored case is a straightforward exercise.

For the empirical waiting time distributions we may prove the following results in a way similar to above.

Theorem 4.5. *Make the assumption of Theorem 4.2. Then:*

$$\sup_{i,t} |\hat{F}_i(t) - F_i(t)| \xrightarrow{P} 0$$

Let k be the number of states in E.

Theorem 4.6. *Make the assumptions of Theorem 4.3. Then the vector*

$$\sqrt{n}(\hat{F}_1 - F_1, \ldots, \hat{F}_k - F_k)$$

of stochastic processes converge weakly to the vector

$$(F_1 Y_{11}, \ldots, F_k Y_{kk}).$$

An estimator of the asymptotic variance of $\sqrt{n}(\hat{F}_i - F_i)$ *is given by*

$$\hat{F}_i^2(t) \sum_{j\neq i} \int_0^t n 1\{N_i(s)\geq 1\} N_i(s)^{-2} dN_{ij}(s)$$

which is consistent under the assumptions of Theorem 4.4.

Theorems 4.5 and 4.6 generalize results of Breslow & Crowley (1974) and Aalen (1976).

5. The smooth estimator and its properties

In this section we shall discuss the smoothing of the estimator (1.9). Let us first give a different formula for the estimator $\hat{P}$. We let the jumps of the processes occur at $0 < s_1 < s_2 < \ldots$ and let the nth jump go from i_n to j_n. Then

$$\hat{P}(s,t) = \prod_{s < s_n \leqslant t} (I + N_{i_n}(s_n^-)^{-1} C_{i_n j_n}) \tag{5.1}$$

Note that a very simple algorithm exists for computing $\hat{P}(s,t)$. Assume $\hat{P}(s,u)$ has been computed and that the next jump occurs at $u_1(>u)$ and goes from i to j. Then $\hat{P}(s,u_1)$ is constructed as follows: The ith column is multiplied by $(1 - 1/N_i(u_1-))$ and the remaining is added to column j.

$\hat{P}(s,t)$ can also be computed backwards as follows: If $\hat{P}(u,t)$ has been computed and the nearest previous jump is at time $u_1(<u)$ from i to j then only the ith row is replaced by a convex combination of the ith and jth rows with weight $(1 - 1/N_i(u_1-))$ and $1/N_i(u_1-)$ respectively.

In order to smooth $\hat{P}$ we note that

$$e^{tC_{ij}} = I + C_{ij}(1 - e^{-t}). \tag{5.2}$$

This follows from $C_{ij}^2 = -C_{ij}$ together with the series expansion for the exponential function.

The matrices $e^{tC_{ij}}$ are the simplest imbeddable stochastic matrices, see Johansen (1973) and the representation (5.2) and (5.1) exhibits $\hat{P}(s,t)$ as a finite product of these elementary matrices

$$\hat{P}(s,t) = \prod_{s < s_n \leqslant t} \exp\left(\ln \frac{N_{i_n}(s_n)}{N_{i_n}(s_n) - 1} C_{i_n j_n}\right)$$

It is therefore not difficult to interpolate to make this estimator smooth. We just define

$$\hat{\hat{Q}}(t) = \frac{t - s_{n-1}}{s_n - s_{n-1}} \ln \frac{N_{i_n}(s_n)}{N_{i_n}(s_n) - 1} C_{i_n, j_n}; \qquad s_{n-1} < t \leqslant s_n$$

and

$$\hat{\hat{B}}(t) = \sum_{k=1}^{n-1} \hat{\hat{Q}}(s_k) + \hat{\hat{Q}}(t), \qquad s_{n-1} < t \leqslant s_n$$

and

$$\hat{\hat{P}}(s,t) = \prod_{]s,t]} (I + d\hat{\hat{B}}) = \hat{P}(s, s_{n-1})\, e^{\hat{\hat{Q}}(t)}, \qquad s_{n-1} < t \leqslant s_n$$

Then one easily checks that $\hat{\hat{P}}$ and $\hat{P}$ coincide at all the jump points, but whereas $\hat{P}$ is piecewise constant we have obtained that $\hat{\hat{P}}$ is absolutely continuous and satisfies the Kolmogorov equations (1.4) and (1.5) with the $\hat{\hat{B}}$ given above.

Clearly the estimates $\hat{P}$ and $\hat{\hat{P}}$ are not very different, but $\hat{\hat{P}}(0,t)$ is not measurable with respect to $\mathcal{F}_t$ since it depends on where the first jump after t is going to happen.

The difference can be evaluated as follows

$$|\hat{\hat{P}}(0,t) - \hat{P}(0,t)| \leqslant |\hat{P}(0, s_{n-1})(e^{\hat{\hat{Q}}(t)} - I)|$$
$$\leqslant |e^{\hat{\hat{Q}}(t)} - I| \leqslant |\hat{\hat{Q}}(t)|\, e^{|\hat{\hat{Q}}(t)|}, \; s_{n-1} < t \leqslant s_n.$$

Now

$$|\hat{\hat{Q}}(t)| = 2\frac{t - s_{n-1}}{s_n - s_{n-1}} \ln \frac{N_{i_n}(s_n)}{N_{i_n}(s_n) - 1} \leqslant 2 \ln \frac{N_{i_n}(s_n)}{N_{i_n}(s_n) - 1}$$

If the assumptions of Theorem 4.2 hold then $\sup_t |\hat{\hat{Q}}(t)| \xrightarrow{P} 0$ as $n \to \infty$. This implies that $\hat{\hat{P}}$ and $\hat{P}$ have the same asymptotic behaviour.

6. Appendix

In this appendix we have collected a few results which provide a background for some of the methods in the main body of the paper.

We shall discuss briefly square integrable matrix valued martingales and stochastic integrals with respect to them, and we shall prove that the process N_{ij} is in fact a counting process with intensity $N_i q_{ij}$. For the theory of square integrable martingales and stochastic integrals, see e.g. Meyer (1971).

Let now M be a matrix of square integrable martingales on $[0,1]$. We define $\langle M, M \rangle$ as the matrix we get by substituting in the Kronecker product $M \otimes M$, the element $M_{ij} M_{km}$ by $\langle M_{ij}, M_{km} \rangle$. Then $M \otimes M - \langle M, M \rangle$ is a matrix valued martingale.

Next we shall use stochastic integrals with respect to such an M. Let K and H be matrices of predictable processes, with a dimension such that the matrix product HMK has a meaning and such that

$$E \int_0^1 H_{ij}^2(u)\, K_{km}^2(u)\, d\langle M_{st}, M_{st} \rangle(u) < \infty$$

for all i, j, k, m, s, t.

Then

$$\int_0^t H(u)\, dM(u)\, K(u)$$

is defined as the matrix with elements

$$\sum_{k,m} \int_0^t H_{ik}(u)\, K_{mj}(u)\, M_{km}(du).$$

It follows that $\int_0^t HdMK$ is a square integrable martingale and some calculations show that

$$\left\langle \int HdMK, \int HdMK \right\rangle = \int H \otimes H d\langle M, M \rangle K \otimes K.$$

The only rule that is needed repeatedly is the bilinearity of $\langle\cdot, \cdot\rangle$, i.e.

$$\left\langle \sum_i \int H_i \, dM_i, \sum_j \int K_j \, dM_j \right\rangle = \Sigma_i \Sigma_j \int H_i \otimes K_j \, d\langle M_i, M_j \rangle.$$

These formulae are used in the discussion of the variance of $\hat{P}$ and $\hat{B}$.

Consider now a single uncensored process. We shall prove that the intensity of N_{ij} is in fact $q_{ij} N_i$.

Let $N = \sum_{i \neq j} N_{ij}$ and consider $\{N_{ij}\}$ a marked point process, with marks Y. The intensity of this process can be found as follows, see Brémaud & Jacod (1977). Let again the nth jump take place at time s_n and go from i_n to j_n. Let A_n denote the event $\{S_\nu = s_\nu, Y_\nu = (i_\nu, j_\nu), \nu = 1, \ldots, n-1\}$. Then if

$$F_n^{ij}(s \mid A_n) = P\{S_n > s + s_{n-1}, Y_n = (i,j) \mid A_n\} = \int_{s_{n-1}+s}^{\infty} q_{ij}(u) \exp\left(\int_u^{\infty} q_{ii}(v)\, dv\right) du 1\{j_{n-1} = i\}$$

and

$$F_n(s \mid A_n) = \sum_{i \neq j} F_n^{ij}(s \mid A_n) = \exp \int_{s_{n-1}+s}^{\infty} q_{j_{n-1}, j_{n-1}}(v)\, dv$$

we define

$$\varphi_n^{ij}(s \mid s_1, \ldots, s_{n-1}, i_1, \ldots, i_{n-1}, j_1, \ldots, j_{n-1}) = \int_0^s \frac{-F_n^{ij}(du \mid A_n)}{F_n(u- \mid A_n)} = 1\{j_{n-1} = i\} \int_{s_{n-1}}^{s_{n-1}+s} q_{ij}(u)\, du$$

Now the integrated intensity of the process N_{ij} is given by

$$\hat{N}_{ij}(t) = \sum_{\nu=0}^{N(t)} \varphi_\nu^{ij}(S_\nu - S_{\nu-1} \mid S_1, \ldots, S_{\nu-1}, Y_1, \ldots, Y_{\nu-1}) + \varphi_{N(t)+1}^{ij}(t - S_{N(t)} \mid S_1, \ldots, S_{N(t)}, Y_1, \ldots, Y_{N(t)})$$
$$= \sum_{\nu=0}^{N(t)} 1\{j_{\nu-1} = i\} \int_{S_{\nu-1}}^{S_\nu} q_{ij}(u)\, du + 1\{j_{N(t)} = i\} \times \int_{S_{N(t)}}^{t} q_{ij}(u)\, du = \int_0^t q_{ij}(u) N_i(u)\, du.$$

Note added in proof

One may easily see that the conditions of Theorem 4.4 imply those of Theorem 4.3 which further imply the condition of Theorem 4.2. This justifies our use of the consistency of $\hat{P}(0, t)$ in the proofs of Theorems 4.3 and 4.4.

References

Aalen, O. (1973). Nonparametric inference in connection with multiple decrement models. Statist. Research Report no. 6, University of Oslo.

Aalen, O. (1976). Nonparametric inference in connection with multiple decrement models. *Scand. J. Statist.* **3**, 15–27.

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

Aalen, O. (1978*a*). Nonparametric estimation of partial transition probabilities in multiple decrement models. *Ann. Statist.* **6**, 534–535.

Aalen, O. (1978*b*). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6**, no. 4.

Aalen, O. (1978*c*). The effect on the likelihood function by stopping and transformation of multivariate stochastic processes. Tech. Rep. University of Tromsø, 7 pp.

Altshuler, B. (1970). Theory for the measurement for competing risks in animal experiments. *Math. Biosciences* **6**, 1–11.

Boel, R., Varayia, P. & Wong, E. (1975). Martingales on jump processes. *SIAM J. Control* **13**, 999–1061.

Brémaud, P. & Jacod, J. (1977). Processus ponctuels et martingales : Résultat récent sur la modélisation et le filtrage. *Adv. Appl. Prob.* **9**, 362–416.

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

Dobrushin, R. L. (1953). Generalization of Kolmogorov's equations for Markov Processes, with a finite number of possible states. *Mat. Sb. N.S.* **33**, 567–596.(In Russian.)

Doléans-Dadé, C. & 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.

Doob, J. L. (1953). *Stochastic processes*. Wiley, New York.

Fleming, T. R. (1978*a*). Nonparametric estimation for non-time-homogeneous Markov processes in the problem of competing risks. *Ann. Statist.* **6**, no. 5.

Fleming, T. R. (1978*b*). Asymtotic distribution results in competing risks estimation. *Ann. Statist.* **6**, no. 5.

Goodman, G. S. (1970). An intrinsic time for non-stationary finite Markov chains. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **16**, 165–180.

Goodman, G. S. & Johansen, S. (1973). Kolmogorov's differential equations for non-stationary, countable state Markov processes with uniformly continuous transition probabilities. *Proc. Camb. Phil. Soc.* **73**, 119–138.

Jacobsen, M. (1972). A characterization of minimal Markov jump processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **23**, 32–46.

Johansen, S. (1973). The Bang-Bang problem for stochastic matrices. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **26**, 191–195.

Johansen, S. (1977). Product integrals and Markov Processes. Tech. Rep. ETH Zürich, 13 pp.

Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Statist. Ass.* **53**, 457–481.

Meyer, P. A. (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. Qual. Tech.* **1**, 27–52.

Rebolledo, R. (1977). Remarques sur la convergence en loi des martingales vers des martingales continues. *C. R. Acad. Sc. Paris* **285**, 517–520.

Odd O. Aalen
Institute of Mathematical and Physical Sciences
University of Tromsø, N-9001 Tromsø, Norway