Product Integrals and Markov Processes

Søren Johansen
Institute of Mathematical Statistics, University of Copenhagen
Universiteitsparken 5, DK-2100 Copenhagen Ø, Denmark

The solution of the Kolmogorov differential equations for the transition probabilities of a finite state-space Markov process can be represented by a product-integral. For statistical applications it is useful to extend this representation to the general case when the transition probabilities are not absolutely continuous, indeed not even continuous. The correct definition of product integral for this purpose is given here, and some of its properties derived.

1. Introduction

One of the most widely used mathematical models for describing stochastic processes with dependence is that of a Markov chain.

An example where such models are applied is in the study of diseases, where a person can be in one of three states: healthy, sick or dead. The history of a patient is then described as a sequence of time points where transitions or jumps between the states healthy and sick occurs, possibly terminating with a transition to the state death.

The Markov assumption is that once the patient is healthy, say, the waiting time in that state is independent of the past history of the patient and only depends on the state he is in.

These models and some of their generalisations form the mathematical substance of the topic *Survival Analysis* which, briefly described, is the study of the life history of patients with respect to the occurence of certain events, see Andersen and Borgan [2] for a comprehensive review.

For another example from a totally different area take the following typical chemical reaction

$$A + B \rightleftarrows C.$$

If $A(t)$ denotes the number of molecules of type A in a solution, then if $A(t) = n$, a transition to $n-1$ means that an A molecule and a B molecule have collided in the solution and formed a C molecule. Similarly the transition to $n+1$ means that a C molecule has split into an A molecule and a B molecule. For a survey of the theory of Markov chains in physics and chemistry see van Kampen [11].

A finite state Markov chain is then a model for a stochastic process $X(t)$, $t \geq 0$ with values in the finite state space E. The probability measure $\mathbb{P}$,

which describes the probabilistic properties of X, satisfies the Markov property

$$\mathbb{P}(X(t_{n+1})=i_{n+1}|X(t_1)=i_1,\ldots,X(t_n)=i_n) =$$
$$\mathbb{P}(X(t_{n+1})=i_{n+1}|X(t_n)=i_n)$$

for any states $i_1, \ldots, i_n, i_{n+1}$ in E and time points $t_1<\ldots<t_n<t_{n+1}$. This property is the mathematical formulation of the intuitive property that the process only remembers its immediate past.

This condition ensures that the probability measure is completely described by the initial distribution $p(i) = \mathbb{P}(X(0) = i)$ and the transition probabilities

$$p_{ij}(s,t) = \mathbb{P}(X(t) = j|X(s) = i),\ i\in E,\ j\in E,\ s<t.$$

It is not difficult to see that the matrix $\mathbf{P}(s,t)$ with elements $p_{ij}(s,t)$ satisfies the Chapman-Kolmogorov equation

$$\mathbf{P}(s,t) = \mathbf{P}(s,u)\mathbf{P}(u,t) \quad 0\leqslant s\leqslant u\leqslant t<\infty$$

which under smoothness conditions gives rise to the Kolmogorov equations

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

and

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

which together with the initial condition $\mathbf{P}(s,s) = \mathbf{I}$ (the identity matrix) determine the function $\mathbf{P}(s,t)$ uniquely. The function $\mathbf{Q}(t)$ is determined as

$$\mathbf{Q}(t) = \frac{\partial}{\partial t}\mathbf{P}(s,t)|_{s=t} = -\frac{\partial}{\partial s}\mathbf{P}(s,t)|_{s=t}$$

From the relation

$$1-p_{ii}(t,t+h)\approx h\cdot(-q_{ii}(t))$$

we see that $h\cdot(-q_{ii}(t))$ is the probability that in a short time interval a transition from state i will take place, thus $-q_{ii}(t)$ is the intensity with which the process leaves the state i at time t.

Similarly $q_{ij}(t)/(-q_{ii}(t))$ is the probability that the transition from i which takes place at time t goes to state j.

Note that whereas $\mathbf{P}(s,t)$ is a stochastic matrix, i.e. satisfies

$$p_{ij}(s,t)\geqslant 0,\ \sum_j p_{ij}(s,t) = 1,$$

$\mathbf{Q}(t)$ is an intensity matrix, i.e.

$$q_{ij}(t)\geqslant 0,\ i\neq j,\ q_{ii}(t)\leqslant 0,\ \sum_j q_{ij}(t) = 0.$$

Thus the study of the stochastic process $\mathbf{X}(t)$ is reduced to the study of the bilinear differential equation for the matrix valued function $\mathbf{P}(s,t)$. The solution to the Kolmogorov equations can be given by the subject of this note, the

product integral

$$\mathbf{P}(s,t) = \prod_s^t (\mathbf{I}+\mathbf{Q}(u)du).$$

For reference to this, see SCHLESINGER [13] and DOBRUŠIN [3] or the monograph by DOLLARD and FRIEDMAN [4].

In the study of the estimation problem for a nonhomogeneous Markov Process (AALEN and JOHANSEN [1]) it turned out that we needed the product-integral representation of a transition probability which is piecewise constant between discrete jumps. This note contains a definition of the product integral, where the measure $\int_A \mathbf{Q}(u)du$ is replaced by an arbitrary matrix valued measure ν on [0,1].

The monograph by DOLLARD and FRIEDMAN [4] gives a survey of various definitions and applications of product integrals but does not contain the definition that allows the Kolmogorov equations to describe the transition probabilities when these are only assumed to be right continuous and of bounded variation. In order to get a simple exposition of the basic properties of the product integral we shall define it by the Peano series, see Definition 2.1. This exploits the usual measure theory for matrix valued measures, in particular Fubini's theorem, and thus gives the existence and basic properties easily.

Section 2 contains the product integral and some of its properties and Section 3 then applies this integral to the representation of the transition probabilities in terms of its integrated intensities. We thus obtain a different approach to some of the results of JACOBSEN [9].

The present note was finished in 1977 as a technical report. The recent interest in the application of product integration in survival analysis and other areas of statistics justifies its publication now. A more comprehensive review of the application of product integrals in the statistical analysis of counting processes is forthcoming, see GILL and JOHANSEN [7].

2. THE PRODUCT INTEGRAL

Let $\mathcal{B}$ denote the Borel sets of [0,1] and let ν be a σ-additive finite signed measure with values in the set of $n \times n$ matrices, i.e. a matrix of n^2 real signed measures. We let $\nu^{(n)}$ denote product measure on $[0,1]^n$ defined by

$$\nu^{(n)}(B_1 \times \ldots \times B_n) = \nu(B_1) \cdot \ldots \cdot \nu(B_n), \; B_i \in \mathcal{B}.$$

We shall use the notation $\|\cdot\|$ to denote the matrix norm $\|\nu(B)\| = \sup_i \sum_j |\nu_{ij}(B)|$, and introduce the real positive measure

$$\nu_0 = \sum_i \sum_j |\nu_{ij}|$$

where $|\nu_{ij}| = \nu_{ij}^+ + \nu_{ij}^-$.

One easily checks that $\|\nu(B)\| \leq \nu_0(B)$ and that

$$\|\nu^{(n)}(B_1 \times \ldots \times B_n)\| \leq \prod_{i=1}^n \nu_0(B_i).$$

We shall use the shorthand notation

$$\{u_1<u_2<\ldots<u_n\}$$

for the set

$$\{(u_1,u_2,\ldots,u_n):u_1<u_2<\ldots<u_n\}.$$

We now give the basic definition:

DEFINITION 2.1. *For $B \in \mathfrak{B}$ we define the product integral*

$$\prod_B(\mathbf{I}+d\nu)=\mathbf{I}+\sum_{n=1}^{\infty}\nu^{(n)}(B\times\ldots\times B\cap\{u_1<\ldots<u_n\}).$$

Notice that the convergence of the series follows from the inequality

$$\|\nu^{(n)}(B\times\ldots\times B\cap\{u_1<\ldots<u_n\})\|\leqslant$$

$$\nu_0^{(n)}(B\times\ldots\times B\cap\{u_1<\ldots<u_n\})\leqslant\frac{1}{n!}\nu_0(B)^n.$$

The following inequalities follow easily from the definition.

PROPOSITION 2.1. *For $B \in \mathfrak{B}$ we have*

$$\|\mathbf{I}+\nu(B)\|\leqslant e^{\nu_0(B)}$$

$$\|\prod_B(\mathbf{I}+d\nu)\|\leqslant e^{\nu_0(B)}$$

$$\|\prod_B(\mathbf{I}+d\nu)-\mathbf{I}\|\leqslant\nu_0(B)e^{\nu_0(B)}$$

$$\|\prod_B(\mathbf{I}+d\nu)-\mathbf{I}-\nu(B)\|\leqslant\frac{1}{2}\nu_0(B)^2e^{\nu_0(B)}$$

Now we can immediately prove the basic multiplicativity property:

THEOREM 2.1. *For any $t\in[0,1]$ we have*

$$\prod_B(\mathbf{I}+d\nu)=\prod_{B\cap[0,t]}(\mathbf{I}+d\nu)\prod_{B\cap]t,1]}(\mathbf{I}+d\nu).$$

Thus $\prod_B(\mathbf{I}+d\nu)$ is multiplicative over disjoint intervals, which is the reason for its name.

PROOF. For $i=1,\cdots,n-1$ let

$$A(B,i,n)=B\times\ldots\times B\cap\{u_1<\ldots<u_i\leqslant t<u_{i+1}<\ldots<u_n\}$$

with the obvious modification for $i=0$ and n. We let

$$A(B,n)=B\times\ldots\times B\cap\{u_1<\ldots<u_n\}=\bigcup_{i=0}^{n}A(B,i,n).$$

Now

$$\nu^{(n)}(A(B,n)) = \sum_{i=0}^{n} \nu^{(n)}(A(B,i,n))$$

$$= \sum_{i=0}^{n} \nu^{(i)}(A(B \cap [0,t],i))\nu^{(n-i)}(A(B \cap]t,1],n-i)).$$

Summing over n gives the result. □

EXAMPLES 2.4. If $d\nu = \mathbf{Q}dt$, where $\mathbf{Q}$ is a fixed matrix then

$$\nu^{(n)}(0 \leqslant u_1 < \ldots < u_n \leqslant t) = \frac{t^n \mathbf{Q}^n}{n!}$$

and hence

$$\prod_{[0,t]} (\mathbf{I} + d\nu) = e^{t\mathbf{Q}}.$$

If $d\nu = \mathbf{Q}_1 dt$ for $0 \leqslant t \leqslant t_1$ and $d\nu = \mathbf{Q}_2 dt$ for $t_1 < t \leqslant 1$, then using Theorem 2.1 and the previous example, we get

$$\Pi_{[0,t]}(\mathbf{I} + d\nu) = \begin{cases} e^{\mathbf{Q}_1 t} & 0 \leqslant t \leqslant t_1 \\ e^{\mathbf{Q}_1 t_1} e^{\mathbf{Q}_2 (t - t_1)}, & t_1 < t \leqslant 1. \end{cases}$$

As a final example we let $\nu = \mathbf{Q}\epsilon_{t_1}$ where $\mathbf{Q}$ is fixed and ϵ_{t_1} is a one point measure at t_1, then

$$\prod_{[0,t]} (\mathbf{I} + d\nu) = \begin{cases} \mathbf{I} & , 0 \leqslant t < t_1 , \\ \mathbf{I} + \mathbf{Q} & , t_1 \leqslant t \leqslant 1 . \end{cases}$$

Thus we get a piecewise constant function of t.

The following results give a different and perhaps more intuitive definition of the product integral. The definition we have chosen seems to give the basic results in a very efficient manner, since we can use existing integration and measure theory.

THEOREM 2.2. *Let* $0 = t_0 < t_1 < \ldots < t_n = 1$, *then*

$$\| \prod_{]0,1]} (\mathbf{I} + d\nu) - \prod_{i=0}^{n-1} (\mathbf{I} + \nu]t_i, t_{i+1}]) \| \leqslant c \max_i \nu_0]t_i, t_{i+1}[.$$

PROOF. We split the product integral into the corresponding factors and define

$$\mathbf{M}_i = \prod_{]t_i, t_{i+1}]} (\mathbf{I} + d\nu), \quad \mathbf{N}_i = \mathbf{I} + \nu]t_i, t_{i+1}]$$

then by Proposition 2.1, we get

$$\|\mathbf{M}_i\| \leqslant e^{\nu_0]t_i, t_{i+1}]}, \quad \|\mathbf{N}_i\| \leqslant e^{\nu_0]t_i, t_{i+1}]}.$$

We also get

$$\|\mathbf{M}_i-\mathbf{N}_i\| = \|\prod_{]t_i,t_{i+1}[}(\mathbf{I}+d\nu)(\mathbf{I}+\nu[t_{i+1}])-\mathbf{I}-\nu]t_i,t_{i+1}[-\nu[t_{i+1}]\|$$

$$\leqslant\|\prod_{]t_i,t_{i+1}[}(\mathbf{I}+d\nu)-\mathbf{I}-\nu]t_i,t_{i+1}[\| + \|\prod_{]t_i,t_{i+1}[}(\mathbf{I}+d\nu)-\mathbf{I}\|\|\nu[t_{i+1}]\|$$

$$\leqslant\frac{1}{2}(\nu_0]t_i,t_{i+1}[)^2 e^{\nu_0]t_i,t_{i+1}[}+\nu_0]t_i,t_{i+1}[e^{\nu_0]t_i,t_{i+1}[}.\nu_0[t_{i+1}]$$

$$\leqslant e^{\nu_0]t_i,t_{i+1}[}\ (\max_i \nu_0]t_i,t_{i+1}[)\nu_0]t_i,t_{i+1}[.$$

Using these evaluations we get

$$\|\prod_{i=0}^{n-1}\mathbf{M}_i-\prod_{i=0}^{n-1}\mathbf{N}_i\|\leqslant\sum_{i=0}^{n-1}\|\mathbf{M}_0\ldots\mathbf{M}_{i-1}(\mathbf{M}_i-\mathbf{N}_i)\mathbf{N}_{i+1}\ldots\mathbf{N}_{n-1}\|$$

$$\leqslant e^{\nu_0[0,1]}(\sum_i \nu_0]t_i,t_{i+1}])\max_i \nu_0]t_i,t_{i+1}[$$

which gives the result we wanted to prove. □

COROLLARY 2.1. *Let t_{in} satisfy the conditions*

$$a)\ 0 = t_{0n}<t_{1n}<\ldots<t_{nn} = 1$$

and

$$b)\ \lim_{n\to\infty}\max_i \nu_0]t_{in},t_{(i+1)n}[= 0,$$

then the product integral can be computed as

$$\prod_{]0,1]}(\mathbf{I}+d\nu) = \lim_{n\to\infty}\prod_{i=0}^{n-1}(\mathbf{I}+\nu]t_{in},t_{(i+1)n}]).$$

Notice that condition b) can always be satisfied, since we can make sure that the atoms of ν(*i.e. of* ν_0) eventually are among the division points.

The next results are about differentiability of the product integral.

THEOREM 2.3. *For $B\in\mathfrak{B}$ we have*

$$\prod_B(\mathbf{I}+d\nu)-\mathbf{I} = \int_B \prod_{B\cap[0,u[}(\mathbf{I}+d\nu)\nu(du)\ .$$

PROOF. Using Fubini's theorem on the $(n+1)$-th term we get

$$\nu^{(n+1)}(B\times\ldots\times B,\ 0\leqslant u_1<\ldots<u_{n+1}\leqslant 1)$$

$$= \int_B \nu^{(n+1)}(B\times\ldots\times B,\ 0\leqslant u_1<\ldots<u_{n+1}\leqslant 1|u_{n+1} = u)\nu(du)$$

$$= \int_B \nu^{(n)}(B\times\ldots\times B,\ 0\leqslant u_1<\ldots<u_n<u)\nu(du).$$

Summing over n gives the result. □

THEOREM 2.4. *For $B \in \mathcal{B}$ we have*

$$\prod_B(\mathbf{I}+d\boldsymbol{\nu})-\mathbf{I} = \int_B \boldsymbol{\nu}(du) \prod_{]u,1]\cap B}(\mathbf{I}+d\boldsymbol{\nu})$$

PROOF. Similar to that of Theorem 2.3.

If we define the function $\mathbf{F}$ by

$$t \to \prod_{[0,t]}(\mathbf{I}+d\boldsymbol{\nu})$$

then $\mathbf{F}$ is right continuous by Theorem 2.3 and it is of bounded variation. It thus determines a matrix valued measure, which by Theorem 2.3 is absolutely continuous with respect to ν_0. The integral relation can thus be reformulated as

$$\frac{d\mathbf{F}}{d\nu_0}(t) = \mathbf{F}(t-)\frac{d\boldsymbol{\nu}}{d\nu_0}(t), \text{ a.e.}[\nu_0].$$

If $\mathbf{H}:t \to \prod_{]t,1]}(\mathbf{I}+d\boldsymbol{\nu})$, then Theorem 2.4 can be reformulated as

$$\frac{d\mathbf{H}}{d\nu_0}(t) = -\frac{d\boldsymbol{\nu}}{d\nu_0}(t)\mathbf{H}(t), \text{ a.e.}[\nu_0].$$

3. MARKOV PROCESSES

In constructing a Markov Process one can start with the transition probabilities, satisfying the Chapman-Kolmogorov equations, then construct the process, i.e. the measure on a suitable function space, by the Kolmogorov consistency theorem, see DOOB [5], or via a general theorem of extension of continuous linear functionals, see NELSON [12], or GOODMAN and JOHANSEN [8]. In this case the discussion of the differential equations for the transition probabilities becomes a discussion of when a process is determined by its infinitesimal properties.

One can also start out with the waiting time distributions and the jump intensities and then construct the measure very directly and then prove that certain variables form a Markov Process and define the transition probabilities in terms of these. The differential equations can now be viewed as a convenient and different way of obtaining the transition probabilities, see JACOBSEN [9].

We shall here start with the intensities or rather the integrated intensities $\boldsymbol{\nu}$, i.e. we assume that

$$\nu_{ii} \leqslant 0,\ \nu_{ij} \geqslant 0,\ i \neq j \text{ and } \sum_j \nu_{ij} = 0.$$

From this measure we construct the transition probabilities by a product integral and this also gives us the differential equations for $\mathbf{P}$. Thus the approach is highly non-probabilistic as opposed to that of JACOBSEN [9]. The solution, however, is the same, as we shall show.

Thus, we let ν be a finite signed measure on $[0,1]$ with values in the set of intensity matrices, then the following holds:

THEOREM 3.1. *If* $\nu[t]+\mathbf{I}$ *is a stochastic matrix, i.e. if* $\nu_{ii}[t] \geqslant -1$, *then*

$$\mathbf{P}(B) = \prod_B (\mathbf{I}+d\nu)$$

is a stochastic matrix.

PROOF. Assume first that $\nu_{ii}[t] > -1$, $i = 1, \ldots, n$, i.e. no atoms of ν are as large as -1. Then let us choose the partition t_{in} of $[0,1]$ so fine that $\mathbf{I}+\nu]t_{in}, t_{(i+1)n}]$ is a stochastic matrix. By Corollary 2.1 $P(B)$ is the limit of stochastic matrices, hence stochastic.

In general there can only be a finite number of points $t_1, \ldots, t_k$, such that some $\nu_{ii}[t_r] = -1$. By writing

$$\mathbf{P}(B) = (\mathbf{I}+\nu(B \cap [0])) \cdot \prod_{i=1}^{k-1} \prod_{B \cap]t_i, t_{i+1}[} (\mathbf{I}+d\nu)(\mathbf{I}+\nu(B \cap [t_{i+1}]))$$

we see that **P(B)** is stochastic.

For a given ν we now define

$$\mathbf{P}(s,t) = \mathbf{P}(]s,t]) = \prod_{]s,t]} (\mathbf{I}+d\nu),\ 0<s<t\leqslant 1,$$

$$\mathbf{P}(0,t) = \mathbf{P}([0,t]) = \prod_{[0,t]} (\mathbf{I}+d\nu),\ 0\leqslant t\leqslant 1,$$

then it is seen that $\mathbf{P}(s,t)$ is right continuous in s and t (except for $s = 0$, $t \downarrow 0$), and that

$$\mathbf{P}(t-,t) = \mathbf{I}+\nu[t],\ 0<t\leqslant \mathrm{I}$$

$$\mathbf{P}(t,t+) = \mathbf{I},\ 0<t<1$$

$$\mathbf{P}(0,0+) = \mathbf{I}+\nu[0].$$

The multiplicativity of the product integral now immediately gives that $\mathbf{P}(s,t)$ satisfies the Chapman-Kolmogorov equations

$$\mathbf{P}(s,t) = \mathbf{P}(s,u)\mathbf{P}(u,t)\ \ 0\leqslant s<u<t\leqslant 1,$$

and in this formulation, Theorem 2.3 gives the forward differential equation

$$\frac{\partial \mathbf{P}(s,t)}{\partial \nu_0(t)} = \mathbf{P}(s,t-)\frac{d\nu}{d\nu_0}(t) \quad \text{a.e.}[\nu_0]$$

whereas Theorem 2.4 gives the backward equation

$$\frac{\partial \mathbf{P}(s,t)}{\partial \nu_0(s)} = -\frac{d\nu}{d\nu_0}(s)\mathbf{P}(s,t), \quad \text{a.e.}[\nu_0]$$

which shows that $\mathbf{P}(s,t)$ does in fact have ν as integrated intensity measure.

Using a result similar to Theorem 2.2 one can prove that for $s<t_{0n}<\ldots<t_{nn} = t$ such that $\max_j \nu_0]t_{jn}, t_{(j+1)n}[\to 0$ we have

$$\prod_{]s,t]}(1+d\nu_{ii}) = \lim_{n\to\infty} \prod_{j=0}^{n-1} p_{ii}(t_{jn}, t_{(j+1)n})$$

which is nothing but the waiting time distribution in state i, i.e.

$$\mathbb{P}\{X_u = i,\ s<u\leqslant t | X_s = i\} = \prod_{]s,t[}(1+d\nu_{ii}).$$

With the notation

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

and

$$\pi_{ij}(t) = -\frac{d\nu_{ij}}{d\nu_{ii}}(t),\ i\neq j,\ \pi_{ii}(t) = 0$$

we can now prove that the solution provided by JACOBSEN [9], starting with $\mathbf{G}$ and $\boldsymbol{\pi}$ is in fact the same as the solution provided here starting from $\boldsymbol{\nu}$.

We shall now assume that $\boldsymbol{\nu}$ satisfies the following extra conditions

1) $\boldsymbol{\nu}[0] = \mathbf{0}$

2) $\nu_i[t]>-1$.

i.e. $\mathbf{P}(s,t)$ becomes right continuous, also at 0, and no atom is as large as -1.

Notice that $\boldsymbol{\nu}$ can be recovered from $\mathbf{G}$ and $\boldsymbol{\pi}$, by the relations

$$\nu_{ii}(A) = -\int_A \frac{G_i(du)}{1-G_i[0,u[}$$

and

$$\nu_{ij}(A) = -\int_A \pi_{ij}(u)\nu_{ii}(du).$$

It is then seen that G_i is continuous at zero, and that $G_i[0,t] = 0$ if $\nu_{ii}[0,t] = 0$ and that $G_i[0,t]<1$, since $\nu_{ii}[0,1]$ is finite. In order to see the last result, where condition 2) is needed, we argue as follows: The largest atom of $|\nu_{ii}|$ is $1-\epsilon$ say. Now choose $0 = t_0<\ldots<t_n = 1$ such that $|\nu_{ii}]t_j,t_{j+1}[|<\frac{\epsilon}{2}$ then $1+\nu_{ii}]t_j,t_{j+1}]>\frac{\epsilon}{2}$ and

$$\log(1+\nu_{ii}]t_j,t_{j+1}]) \geqslant \frac{\log\frac{\epsilon}{2}}{-1+\frac{\epsilon}{2}}\nu_{ii}]t_j,t_{j+1}].$$

Summing over j gives

$$\prod_j(1+\nu_{ii}]t_j,t_{j+1}]) \geqslant c>0$$

which again implies that $\prod_B(1+d\nu_{ii})\geqslant c>0$.

Thus the functions $\mathbf{G}$ and $\boldsymbol{\pi}$ satisfy the conditions of Jacobsen and his

solution $\tilde{\mathbf{P}}(s,t)$ is constructed to satisfy the integral equation:

$$\tilde{p}_{ij}(s,t) = \delta_{ij}1-\delta_{ij}\frac{1-G_i[0,t]}{1-G_i[0,s]}+\sum_{k\neq i}\int_{]s,t]}\pi_{ik}(u)\tilde{p}_{kj}(u,t)\frac{G_i(du)}{1-G_i[0,s]}.$$

Using the definition of $(\mathbf{G},\pi)$ in terms of ν this is

$$\tilde{p}_{ij}(s,t) = \delta_{ij}\prod_{]s,t[}(1+d\nu_{ii})-\sum_{k\neq i}\int\tilde{p}_{kj}(u,t)\prod_{]s,u]}(1+d\nu_{ii})\nu_{ik}(du) \qquad (3.1)$$

which is known to have a unique solution, see FELLER [6].

The function $\mathbf{P}(s,t) = \prod_{]s,t]}(\mathbf{I}+d\boldsymbol{\nu})$ is known to satisfy the equation

$$p_{ij}(s,t)-\delta_{ij} = \sum_{k}\int_{]s,t]}\nu_{ik}(du)p_{kj}(u,t). \qquad (3.2.)$$

In this equation we multiply by $\prod_{]a,s[}(1+d\nu_{ii})$ and integrate with respect to $\nu_{ii}(ds)$. If we then use the results:

$$\int_{]a,t]}\prod_{]a,s[}(1+d\nu_{ii})\nu_{ii}(ds) = \prod_{]a,t]}(1+d\nu_{ii})-1$$

and

$$\int_{]a,u]}\prod_{]a,s]}(1+d\nu_{ii})\nu_{ii}(ds) = \prod_{]a,u[}(1+d\nu_{ii})-1.$$

then we get after some reduction that $\mathbf{P}$ also satisfies equation (3.1.), hence $\mathbf{P} = \tilde{\mathbf{P}}$.

In fact the equations (3.1) and (3.2) are equivalent. If (3.1) is integrated with respect to $\nu_{ii}(ds)$ on $]a,t]$, we arrive at (3.2).

It should of course be pointed out that we are only dealing with a finite number of states, whereas Jacobsen treats the more general situation of a countable number of states.

Example

The results above are considerably simplified in case we want to describe a time homogeneous chain. The intuition behind the time homogenity is that the intensities $q_{ij}(t)$ do not depend on t. Thus in this case the intensity measure satisfies

$$\nu(A) = \mathbf{Q}_0\lambda(A)$$

where λ is Lebesgue measure and $\mathbf{Q}_0$ is an intensity matrix. The transition probabilities are now given by

$$\mathbf{P}(s,t) = \prod_{s}^{t}(\mathbf{I}+\mathbf{Q}_0du) = \exp((t-s)\mathbf{Q}_0)$$

The construction of the process from $\mathbf{G}$ and π is as before with the simplification that

$$G_i(0,t) = 1-\prod_{[0,t]}(1+q_{ii}du) = 1-\exp(tq_{ii})$$

showing that the waiting time distribution in each state is exponential with a state dependent parameter $-q_{ii}$.

REFERENCES

1. O.O. AALEN, S. JOHANSEN (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scand. J. Statist. 5,* 141-150.
2. P.K. ANDERSEN, Ø. BORGAN (1985). Counting process models for life history data. A review. *Scand. J. Statist. 5,* 97-158.
3. R.L. DOBRUŠIN (1953). Generalisation of Kolmogorov's equations for Markov processes with a finite number of possible states. *Mat. Sb. (N.S.) 33,* 567-596.
4. J.D. DOLLARD, C.N. FRIEDMAN (1979). Product integration. *Encyclopedia of Mathematics and its Applications,* Addison-Wesley, London.
5. J.L. DOOB (1953). *Stochastic Processes,* Wiley, New York.
6. W. FELLER (1940). On the integro-differential equations of purely discontinuous Markoff processes. *Trans. Amer. math. Soc. 48,* 488-515. Errata *Ibid. 58,* 474(1945).
7. R.D. GILL, S. JOHANSEN (1987). Product integrals and counting processes (in preparation).
8. G.S. GOODMAN, S. JOHANSEN (1973). Kolmogorov's differential equations for non-stationary countable state Markov processes with uniformly continuous transition probabilities. *Proc. Camb. Phil. Soc. 73,* 119-138.
9. M. JACOBSEN (1972). A characterisation of minimal Markov jump processes. *Z. Wahrsch. verw. Gebiete 23,* 32-46.
10. S. JOHANSEN (1973). A central limit theorem for finite semigroups and its application to the imbedding problem for finite state Markov chains. *Z. Wahrsch. verw. Gebiete 26,* 171-190.
11. N.G. VAN KAMPEN (1981). *Stochastic Processes in Physics and Chemistry,* North-Holland, Amsterdam.
12. E. NELSON (1959). Regular probability measures on function space. *Ann. Math. 69,* 630-643.
13. L. SCHLESINGER (1931). Neue Grundlagen für einen Infinitesimalkalkül der Matrizen. *Math. Z. 33,* 33-61.