NON-MARKOVIAN PROCESSES WITH THE SEMIGROUP PROPERTY

By William Feller[1]

Princeton University

1. Introduction. Every $N \times N$ stochastic matrix P defines the transition probabilities of a Markovian process with positive discrete time parameter. Its n-step transition probabilities satisfy the Chapman-Kolmogorov, or semigroup, relation $P^{n+m} = P^n P^m$. We shall show that *for* $N \geqq 3$ *there exist non-Markovian processes with* N *states whose transition probabilities satisfy the same equation.*[2] All elements of P will equal N^{-1}. The process may be chosen *strictly stationary*. A simple modification leads to non-Markovian processes with *continuous time* parameter and the semigroup property with N states or a continuum of states.

The triviality of the following example should not obscure the interest of the problem concerning the existence of non-Markovian processes satisfying the Chapman-Kolmogorov equation. As so many other basic problems in probability theory, it has been formulated by P. Lévy who with his usual ingenuity gave the first counter-example to the obvious conjecture.

2. Let $\mathfrak{P}$ be the sample space whose points $(x^{(1)}, \cdots, x^{(N)})$ are the random permutations of $(1, 2, \cdots, N)$ each carrying probability $1/N!$ Let $\mathfrak{R}$ be the set of the N points $(x^{(1)}, \cdots, x^{(N)})$ such that $x^{(i)} = \nu$ for all $1 \leqq i \leqq N$ where ν is a fixed integer $1 \leqq \nu \leqq N$; each point of $\mathfrak{R}$ carries probability $1/N$. Finally, Let $\mathfrak{S}$ be the mixture of $\mathfrak{P}$ and $\mathfrak{R}$ with $\mathfrak{P}$ carrying weight $1 - N^{-1}$ and $\mathfrak{R}$ weight N^{-1}.

More formally, $\mathfrak{S}$ contains the $N! + N$ arrangements $(x^{(1)}, x^{(2)}, \cdots, x^{(N)})$ which represent either a premutation of $(1, 2, \cdots, N)$ or the N-fold repetition of an integer ν, $1 \leqq \nu \leqq N$. To each point of the first class we attribute probability $(1 - N^{-1})(N!)^{-1}$, to each point of the second class probability N^{-2}.

Then clearly

$$(1) \qquad P\{x^{(i)} = \nu\} = N^{-1}, \qquad P\{x^{(i)} = \nu, \qquad x^{(j)} = \mu\} = N^{-2}$$

for all $i \neq j$. Thus all transition probabilities are equal:

$$(2) \qquad P\{x^{(i)} = \nu \mid x^{(j)} = \mu\} = N^{-1}.$$

Given, say, that $x^{(1)} = 1$, $x^{(2)} = 1$ the probability that $x^{(3)} \neq 1$ is zero, and hence the process is non-Markovian.

3. To extend the process to all integral values of the time parameter consider, in the usual manner, a double infinity of independent repetitions of the described

Received February 4, 1959.

[1] Research supported by the Office of Ordnance Research U. S. Army Contract DA-36-034-ORD-2001.

[2] [Added in proof.] D. Blackwell has pointed out to me that the variables of our process represent a sequence of *random variables which are pairwise independent without being mutually independent.*

sample space. In other words, we consider the product space $\cdots \mathfrak{S} \times \mathfrak{S} \times \mathfrak{S} \cdots$ with product measure; its points are the doubly infinite sequences $x = \{x^{(i)}\}$ such that for each integer r the N-dimensional block $(x^{(rN+1)}, x^{(rN+2)}, \cdots, x^{(r+1)N})$ represents the projection of x onto a coordinate space $\mathfrak{S}$. This represents a non-Markovian process with the stationary transition probabilities (2). However, the process itself is not stationary in view of the periodicity modulo N.

To obtain a *stationary* process of the same type it suffices to introduce N replicas of our process with time shifts 0, 1, 2, $\cdots$, $N - 1$ and define a new process as their mixture with equal weights.

4. To construct a process of a *similar character defined for all* $t \geqq 0$ consider the above discrete process and a Poisson process $\{N(t)\}$ with mean t independent of it. Define a new process by

$$x(t) = x^{(N(t))}. \tag{3}$$

Its absolute probabilities are given by

$$P\{x(t) = \nu\} = \sum_{i=0}^{\infty} P\{N(t) = i\} \cdot P\{x^{(i)} = \nu\} = N^{-1}. \tag{4}$$

The joint probabilities for $0 \leqq s < t$ are calculated in like manner, but the possibility that $N(t) = N(s)$ makes it necessary to consider separately the cases $\nu = \mu$ and $\nu \neq \mu$. Clearly

$$\begin{aligned} P_{\nu\mu}(t) &= P\{x(s+t) = \mu \mid x(s) = \nu\} \\ &= N^{-1}(1 - e^{-t}), \qquad \text{if } \nu \neq \mu, \\ P_{\nu\nu}(t) &= e^{-t} + N^{-1}(1 - e^{-t}) \end{aligned} \tag{5}$$

and a simple calculation shows that

$$P_{\nu\mu}(s + t) = \sum_{\lambda=1}^{N} P_{\nu\lambda}(s) P_{\lambda\mu}(t) \tag{6}$$

even though the process is clearly non-Markovian.

Finally, one might replace the N states by N intervals and an appropriate motion within them.

REFERENCE

[1] P. Lévy, "Exemples de processus pseudo-Markoviens," *Comptes Rendus Académie Sciences*, Paris, Vol. 228, (1949), pp. 2004–2006.