MIXTURES OF MARKOV PROCESSES[1]

BY DAVID A. FREEDMAN[2]

Princeton University

1. Introduction. Kriloff and Bogoliouboff proved in [4] that stationary stochastic processes may be represented as mixtures of metrically transitive processes. It follows easily (as in [3]) from the strong law of large numbers that this representation is unique. A more illuminating proof is presented here for mixtures of Markov processes (even without assuming stationarity). All processes under discussion are natural number valued, and their time parameter ranges over the natural numbers.

The proof is based on the cycles of a process, which are the (finite) sequences of states beginning and ending at the same state. There is a function on the space of cycles which is associated with each process; namely, assign to a cycle the conditional probability of observing it given its initial state. In Section 2 it will be shown that by and large this association is $1 - 1$ for Markov processes, so that the transition probabilities may be determined from the cycle probabilities. Hence a probability distribution over a family of Markov processes may be thought of as a distribution in the space of cycle probabilities. If the mixture of these processes with respect to some distribution is known, so are the probabilities of repeating given cycles a given number of times. But these are the moments of the distribution of cycle probabilities, and determine that distribution. Thus the mixing distribution can be recovered from the mixture, and if a process can be represented as a mixture of Markov processes, there is essentially only one way to do this. Section 2 below discusses the cycles, and Section 3 gives the uniqueness theorem. The terminology and theory of [2, Chapter 15] is used.

2. Cycles. The symbol $\mathbf{c}$ is reserved for specific cycles, and c for a variable whose domain is some set of cycles. If $\mathbf{c}$ is the sequence $(i_j : 1 \leqq j \leqq n + 1)$, where $i_1 = i_{n+1}$, it is said to be of length n, and its cycle probability is defined as $k_{\mathbf{c}}(p) = p_{i_1 i_2} \cdots p_{i_n i_{n+1}}$, where $p \ \varepsilon\ [0, 1]^{Z \times Z}$ and Z is the set of natural numbers. Here and throughout the paper, subscripts on a point in a product space indicate its coordinates. Usually, p is a matrix of transition probabilities associated with a Markov process, with the understanding that if i or j is not in the state space of the process, $p_{ij} = 0$. The domains of c are C_i, the set of cycles beginning and ending at i; C_i^n, the subset of C_i of cycles of length n; and C_{ij}^n, the subset of C_i^n of cycles whose first transition is (i, j); ${}_jC_{ij}^n$ is the subset of C_{ij}^n of cycles which pass through j only once. Then

THEOREM 1. *The following features of a stochastic matrix are determined by its*

Received May 3, 1961.

[1] The author wishes to acknowledge the very generous financial support provided by McGill University and the RAND Corporation.

[2] Now at the University of California, Berkeley.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org