The Annals of Mathematical Statistics
1969, Vol. 40, No. 1, 97–115

PROBABILISTIC FUNCTIONS OF FINITE STATE MARKOV CHAINS

BY T. PETRIE

The Institute for Defense Analysis

0. Introduction. This paper is statistically motivated; the content mathematical. The motivation is this: Given is an $s \times s$ ergodic stochastic matrix $A = ((a_{ij}))$ and an $s \times r$ stochastic matrix $B = ((b_{jk}))$. A generates a stationary Markov process $\{W_t\}$ according to $a_{ij} = P[W_{t+1} = j \mid W_t = i]$ and B generates a process $\{Y_t\}$ described by $P[Y_t = k \mid Wt = j] = b_{jk}$. If R is the set of integers $1, 2 \cdots r$, $R_t = R$, and $R^\infty = \prod_{t=1}^{\infty} R_t$ (a point $Y \varepsilon R^\infty$ has coordinates Y_t), then the matrices A and B define a measure $P_{(A,B)}$ on R^∞ by

$$(0.1) \quad P_{(A,B)}\{Y_1 = k_1, Y_2 = k_2 \cdots Y_n = k_n\} = \textstyle\sum_{i_0,\cdots,i_n \varepsilon S} a_{i_0} a_{i_0 i_1} b_{i_1 k_1} a_{i_1 i_2} b_{i_2 k_2} \cdots a_{i_{n-1} i_n} b_{i_n k_n}$$

where $S = \{1, \cdots, S\}$ and $\{a_{i_0}\}$ is the stationary absolute distribution for A, and $k_i \varepsilon R$. The resulting process $\{Y_t\}$ is called a probabilistic function of the Markov process $\{W_t\}$. Let Λ_1 be the space of $s \times s$ ergodic stochastic matrices, $\tilde{\Lambda}_1$ be the space of $s \times s$ stochastic matrices, Λ_2 the space of $s \times r$ stochastic matrices, $\Pi = \Lambda_1 \times \Lambda_2$ and $\tilde{\Pi} = \tilde{\Lambda}_1 \times \Lambda_2$. The above associates to $\pi = (A, B) \varepsilon \Pi$ and the stationary vector **a** for A a stationary measure P_π on R^∞. We write $P_\pi(Y_1, Y_2 \cdots Y_n)$ for that function on R^∞ whose value at $Y_1 = k_1, \cdots Y_n = k_n$ is given by (0.1) if $\pi = (A, B)$.

We also find it necessary to introduce $R_{-\infty} = \Pi_{i=0}^{-\infty} R_i$ and define the measure $P_{(A,B)}$ on $R_{-\infty}$ by $P_{(A,B)}$ $(Y_{-n} = k_{-n}, Y_{-n+1} = k_{-n+1}, \cdots Y_0 = k_0) = P_{(A,B)}$ $(Y_1 = k_{-n}, Y_2 = k_{-n+1}, \cdots Y_{n+1} = k_0)$.

The *problem*: Fix $\pi_0 \varepsilon \Pi$ and let a sample $Y_1, Y_2 \cdots Y_n$ be generated according to the distribution P_{π_0}. From the sample $Y_1, Y_2 \cdots Y_n$ obtain estimators $\theta_n(Y)$ of π_0 so that $\theta_n(Y) \to \pi_0$ a.e. P_{π_0}. Throughout this paper π_0 is fixed and π varies in Π.

The *mathematics*: Chapter I (Classification of Equivalent Processes) demonstrates that the problem has a solution in the following sense: Let $M[\pi_0] = \{\pi \varepsilon \Pi \mid P_\pi = P_{\pi_0}$ as measures on $R^\infty\}$. Clearly the points of $M[\pi_0]$ can't be distinguished by any finite or infinite sample. The description of $M[\pi_0]$ is crucial in our study. Let $\mathfrak{S}_S$ be the group of permutations of the integers 1 through s. $\mathfrak{S}_S$ acts on Π by $\sigma(A, B) = (\sigma A, \sigma B)$, $(\sigma A)_{ij} = a_{\sigma(i),\sigma(j)}$, $(\sigma B)_{jk} = b_{\sigma(j)k}$ for $\sigma \varepsilon \mathfrak{S}_S$. Observe that $P_{\sigma\pi} = P_\pi$ as measures on R^∞.

The subset $\sum_{j=1}^{S} a_{ij} = 1$, $a_{ij} \geqq 0$, is part of an $s - 1$ dimensional hyperplane in Euclidean s space and has finite non zero $(s - 1)$-dimensional Lebesgue measure $\delta_{(s-1)}$; similarly the set $\sum_{k=1}^{r} b_{jk} = 1$, $b_{jk} \geqq 0$, has $(r - 1)$-dimensional Lebesgue measure $\lambda_{(r-1)} \neq 0$. It follows that Π with the product measure has

Received 16 February 1967; revised 23 August 1968.