The Annals of Mathematical Statistics
1970, Vol. 41, No. 1, 207–213

SOME REGULAR AND NON-REGULAR FUNCTIONS OF FINITE MARKOV CHAINS

BY S. W. DHARMADHIKARI[1] AND M. G. NADKARNI

Indian Statistical Institute

1. Introduction and summary. Let J be a finite non-empty set and let $S(J)$ denote the set of all finite sequences of elements of J. If $s = (\delta_1, \cdots, \delta_m) \in S(J)$ and $t = (\mu_1, \cdots, \mu_n) \in S(J)$, then st will denote the combined sequence $(\delta_1, \cdots, \delta_m, \mu_1, \cdots, \mu_n)$. The singleton sequence (δ) will be denoted by δ. The symbol s^2 will mean the sequence ss and the symbols s^3, s^4, etc., are defined similarly.

Suppose $\{Y_n\}$ is a stationary process with state-space J. If $s \in S(J)$ and has length n, $p(s)$ denotes $P[(Y_1, \cdots, Y_n) = s]$. The *rank* $n(\delta)$ of a $\delta \in J$ is defined to be the largest integer n such that we can find $2n$ sequences $s_1, \cdots, s_n, t_1, \cdots, t_n$ in $S(J)$ such that the $n \times n$ matrix $\|p(s_i \delta t_j)\|$ is non-singular.

Suppose now that $\{Y_n\}$ is a function of a finite Markov chain (hereafter abbreviated ffMc). That is, let there exist a stationary Markov chain $\{X_n\}$ with a finite state-space I and a function f on I onto J such that $\{Y_n\}$ and $\{f(X_n)\}$ have the same distribution. Then Gilbert [5] has shown that $n(\delta) \leqq N(\delta)$ for all $\delta \in J$, where $N(\delta)$ is the number of elements in $f^{-1}[\{\delta\}]$. If we can find $\{X_n\}$ and f in such a way that $n(\delta) = N(\delta)$ for all $\delta \in J$, then $\{Y_n\}$ is said to be a *regular* ffMc. The motivation for investigating the regularity property of a ffMc has been made clear by Gilbert in the first and the last paragraphs of Section 2 of [5].

Fox and Rubin [3] have given an example of a process $\{Y_n\}$ which has $n(\delta) < \infty$ for all $\delta \in J$ but which is not a ffMc. In the first part of this paper we expand their example into a class of examples and show that some of these examples yield non-regular ffMc. These examples are of a different nature than those given in [1], Section 4. Further our method of investigation is different from that employed by Fox and Rubin.

The second part of this paper is devoted to proving that an exchangeable process which is a ffMc is a regular ffMc.

2. A class of non-regular functions of finite Markov chains. We will use without comment the notation introduced in Section 1. Throughout this section J will be a set having exactly two elements, namely, μ and δ. Before we discuss our class of examples we prove a simple lemma.

LEMMA 2.1. *Let $\{Z_n\}$ be a stationary process with state-space $J = \{\mu, \delta\}$. Suppose p_1 denotes the probability function for $\{Z_n\}$. Then*

$$p_1(\delta\mu^j) = \textstyle\sum_{k=j}^{\infty} p_1(\delta\mu^k\delta), \qquad (1)$$

and

$$p_1(\mu^j) = \beta + \textstyle\sum_{n=j}^{\infty} (n-j+1) p_1(\delta\mu^n\delta), \qquad (2)$$

where $\beta = \lim_{j\to\infty} p_1(\mu^j)$.

Received October 29, 1968; revised July 25, 1969.

[1] Work partially done at Michigan State University.

PROOF. Since $p_1(\mu^j)$ is nonincreasing in j, the limit $\beta = \lim_{j\to\infty} p_1(\mu^j)$ exists. Now it is easy to see that

$$p_1(\delta\mu^j) = p_1(\mu^j) - p_1(\mu^{j+1}), \tag{3}$$ and

$$p_1(\delta\mu^j\delta) = p_1(\delta\mu^j) - p_1(\delta\mu^{j+1}). \tag{4}$$

From (3) we see that $p_1(\delta\mu^j) \to 0$ as $j \to \infty$. Now (1) follows easily from (4). Using (1) and (3), we see that

$$\begin{aligned} p_1(\mu^j) &= \beta + \textstyle\sum_{k=j}^{\infty} p_1(\delta\mu^k) = \beta + \sum_{k=j}^{\infty}\sum_{i=k}^{\infty} p_1(\delta\mu^i\delta) \\ &= \beta + \textstyle\sum_{n=j}^{\infty}(n-j+1)p_1(\delta\mu^n\delta). \end{aligned}$$

This proves (2) and completes the proof of the lemma.

We will now state our class of examples. Let $0 < \lambda \leqq \frac{1}{2}, 0 < \alpha < 2\pi$ and $\alpha \neq \pi$. Let $c_j = \lambda^j \sin^2(j\frac{1}{2}\alpha), j = 1, 2, \cdots$; $c = \sum_{j=1}^{\infty} c_j$ and $d = [4(1+\sum_{j=1}^{\infty} jc_j)]^{-1}$. Suppose $\{X_n\}$ is a Markov chain with state-space $\{0, 1, 2, \cdots\}$, initial distribution $\{m_j, j \geqq 0\}$ and transition matrix $\|m_{jk}\|$, where $m_0 = 4d, m_{00} = 1-c$ and, for $j \geqq 1$, $m_j = 4d\sum_{k=j}^{\infty} c_k$, $m_{0j} = c_j$ and $m_{j,j-1} = 1$. It is easy to check that $\{X_n\}$ is stationary. Define f by $f(0) = \delta$ and $f(j) = \mu$ for $j \geqq 1$. Let $Y_n = f(X_n)$. Then $\{Y_n\}$ is a stationary process with state-space $J = \{\delta, \mu\}$.

The process $\{Y_n\}$ introduced in the preceding paragraph will be discussed in detail in this section. We first prove that, for this process, $n(\delta) = 1$ and $n(\mu) = 3$. The first of these assertions follows easily because δ is the image under f of exactly one state of $\{X_n\}$. To see the second assertion observe first that the result $n(\delta) = 1$ is the same as

$$p(s\delta t) = [p(s\delta)\cdot p(\delta t)]/p(\delta) \tag{5}$$

for all $s \in S(J)$ and $t \in S(J)$. Recall that $n(\mu)$ denotes the maximal rank of matrices P of the form $\|p(s_i\mu t_j)\|$ where $s_i \in S(J)$ and $t_j \in S(J)$ for all i and j. If any s can be written as $s'\delta s''$ then (5) shows that the constant $p(s'\delta)/p(\delta)$ can be factored out from the corresponding row in P. A similar argument holds for the t's. Thus we need consider only sequences s of the form $\delta\mu^i$ or μ^i and sequences t of the form $\mu^j\delta$ or μ^j. Further, since $p(\delta\mu^i t) = p(\mu^i t) - p(\mu^{i+1}t)$ for all $t \in S(J)$, the row corresponding to $s = \delta\mu^i$ is obtained by subtracting the row corresponding to $s' = \mu^{i+1}$ from the row corresponding to $s'' = \mu^i$. A similar argument again holds for the t's. It is thus seen that we need consider only the case where $s_i = t_i = \mu^i$. The matrix P thus equals $\|p(\mu^{i+j+1})\|$. Now it is easy to see from the definition of $\{Y_n\}$ that

$$\begin{aligned} p(\delta\mu^j\delta) &= 4dc_j = 4d\lambda^j \sin^2(j\tfrac{1}{2}\alpha) \\ &= d\lambda^j[2 - \sigma_1{}^j - \sigma_2{}^j], \end{aligned} \tag{6}$$

where $\sigma_1 = \exp[i\alpha]$ and $\sigma_2 = \exp[-i\alpha]$. Now $p(\mu^j) \to 0$ as $j \to \infty$ because $\{X_n\}$ is irreducible and has positive recurrent states. Therefore Lemma 2.1 and (6) show that

$$\begin{aligned} p(\mu^j) &= d\textstyle\sum_{n=j}^{\infty}(n-j+1)[2\lambda^n - (\lambda\sigma_1)^n - (\lambda\sigma_2)^n] \\ &= A_1\lambda^j + A_2(\lambda\sigma_1)^j + A_3(\lambda\sigma_2)^j, \end{aligned} \tag{7}$$

where A_1, A_2 and A_3 are non-zero constants. We are now ready to show that the rank of the matrix $P = ||p(\mu^{i+j+1})||$ is 3. If σ_0 is taken to be 1, if Λ denotes the matrix whose (i,j)th element is $(\lambda\sigma_{j-1})^i$ and if $B = \mathrm{diag}(\lambda A_0\sigma_0, \lambda A_1\sigma_1, \lambda A_2\sigma_2)$, then $P = \Lambda B \Lambda'$. This shows that the rank of P cannot exceed 3. Further, if P_3 is the leading 3×3 principal submatrix of P and if Λ_3 is the Van der Monde matrix formed by the first three rows of Λ, then $P_3 = \Lambda_3 B \Lambda_3'$. But B is clearly nonsingular and Λ_3 is also nonsingular because σ_0, σ_1 and σ_2 are all distinct. Therefore P_3 has rank 3. Thus $n(\mu) = 3$.

In the rest of this section, whenever $\{Y_n\}$ is a ffMc $\{Z_n\}$ with initial distribution $\mathbf{m}$ and transition matrix M, we will write $\mathbf{m}$ as $(\mathbf{m}_\delta, \mathbf{m}_\mu)$ and partition M into submatrices $M_{\delta\delta}$, $M_{\delta\mu}$, $M_{\mu\delta}$ and $M_{\mu\mu}$ in the natural way. We will also assume that every entry of $\mathbf{m}$ is positive. We need one more lemma.

LEMMA 2.2. *Suppose $\{Y_n\}$ is a function of a finite Markov chain $\{Z_n\}$ with transition matrix M. Let λ_1 be the nonnegative eigenvalue of $M_{\mu\mu}$ having maximal modulus. Then $\lambda_1 \leqq \lambda$.*

PROOF. The nonnegative matrix $M_{\mu\mu}$ can be written (after a permutation, if necessary) in the form ([4], page 75)

$$M_{\mu\mu} = \begin{pmatrix} A_{11}, & 0, & 0, \cdots, 0 \\ A_{21}, & A_{22}, & 0, \cdots, 0 \\ \cdot & \cdot & \cdot \ \cdots \ \cdot \\ A_{r1}, & A_{r2}, & \cdots \ A_{rr} \end{pmatrix} \tag{8}$$

where each A_{ii} is square and is either irreducible or consists of a single zero entry. We note that every eigenvalue of $M_{\mu\mu}$ is an eigenvalue of some A_{ii} and conversely.

Suppose $\lambda_1 > \lambda$ and suppose that λ_1 is an eigenvalue of A_{ii}. Since λ is positive, so is λ_1 and hence A_{ii} is irreducible. Since λ_1 is the positive eigenvalue of A_{ii} having maximal modulus, it follows ([4], page 63) that A_{ii} has at least one row sum $\geqq \lambda_1$. Therefore $M_{\mu\mu}$ also has at least one row sum $\geqq \lambda_1$. We show exactly in the same way that $M^j_{\mu\mu}$ must have at least one row sum $\geqq \lambda_1{}^j$.

Let $\mathbf{m}$ be the initial distribution of $\{Z_n\}$. Let α be the smallest entry of $\mathbf{m}_\mu$. It follows from the paragraph immediately preceding the statement of this lemma that $\alpha > 0$. Now if $\mathbf{e}$ denotes a column vector of 1's, then

$$p(\mu^j) = \mathbf{m}_\mu M^{j-1}_{\mu\mu}\mathbf{e} \geqq \alpha\lambda_1{}^{j-1} \tag{9}$$

Now $\lambda^{-j}p(\mu^j)$ is seen to be bounded because of (7) and also unbounded because of (9). This contradiction shows that $\lambda_1 \leqq \lambda$. The lemma is thus proved.

We are now ready to prove the following theorem.

THEOREM 2.1. (a) *If α is an irrational multiple of 2π, then $\{Y_n\}$ is not a function of a finite Markov chain*

(b) *If $\alpha = 2\pi v/N$, where v and N are relatively prime integers, then $\{Y_n\}$ is a function of a finite Markov chain with $(N+1)$ states. Moreover no representation as a function of a Markov chain with less than $(N+1)$ states is possible.*

PROOF. (i) Suppose $\{Y_n\}$ is a ffMc $\{Z_n\}$ with initial distribution $\mathbf{m}$ and transition matrix M. Then

$$p(\delta\mu^j\delta) = \mathbf{m}_\delta M_{\delta\mu} M_{\mu\mu}^{j-1} M_{\mu\delta}\mathbf{e}, \tag{10}$$

where $\mathbf{e}$ is again a column vector of 1's. We can write $M_{\mu\mu}$ (after a permutation, if required) in the form (8). We note that the characteristic polynomial of $M_{\mu\mu}$ is the product of the characteristic polynomials of the A_{ii}'s. Applying Frobenius' theorem ([4], page 53) to the irreducible A_{ii}'s we see that $M_{\mu\mu}$ has a nonnegative eigenvalue λ_1 such that, for every eigenvalue λ' of $M_{\mu\mu}$, we have $|\lambda'| \leqq \lambda_1$. Further, if $|\lambda'| = \lambda_1$ then $\lambda' = \lambda_1 \exp[2\pi i\xi]$, where ξ is rational.

Suppose $M_{\mu\mu}$ has s eigenvalues $\lambda_1', \cdots, \lambda_s'$ of modulus λ_1. Then $\lambda_k' = \lambda_1 \exp[2\pi i\xi_k]$, where ξ_k is rational. Let $\eta_k = \exp[2\pi i\xi_k]$. In the remainder of this proof k will always run from 1 to s. The Jordan canonical form of $M_{\mu\mu}$ can be used to show from (10) that

$$p(\delta\mu^j\delta) = \lambda_1{}^j \textstyle\sum_k \eta_k{}^j q_k(j) + Q(j), \tag{11}$$

where $q_k(j)$ is a polynomial in j of degree at most equal to the number of rows in $M_{\mu\mu}$ and Q is a sum of terms having as factors the jth powers of those eigenvalues λ' of $M_{\mu\mu}$ for which $|\lambda'| < \lambda_1$. From (6) and (11) we have

$$(\lambda_1/\lambda)^j \textstyle\sum_k \eta_k{}^j q_k(j) + \lambda^{-j} Q(j) = d(2-\sigma_1{}^j - \sigma_2{}^j). \tag{12}$$

Lemma 2.2 shows that $\lambda_1 \leqq \lambda$. If $\lambda_1 < \lambda$, then, as $j \to \infty$, the left side of (12) converges to zero, whereas the right side oscillates. Therefore $\lambda_1 = \lambda$.

Let t be the highest among the degrees of the polynomials $q_k(j)$. Write $q_k(j) = \sum_{u=0}^t a_{ku} j^u$ and let $b_u(j) = \sum_k a_{ku}\eta_k{}^j$. Then (12) together with $\lambda = \lambda_1$ shows that

$$\textstyle\sum_{u=0}^t b_u(j) j^u + \lambda^{-j} Q(j) = d(2-\sigma_1{}^j - \sigma_2{}^j). \tag{13}$$

We note that the $b_u(j)$'s are periodic in j, because the ξ_k's are rational. Therefore, if $t \geqq 1$ and $b_t(j) \neq 0$ for some j, the left side of (13) is unbounded in absolute value whereas the right side is bounded in absolute value. It follows that $b_u(j) = 0$ for all j whenever $1 \leqq u \leqq t$. Thus (13) reduces to

$$\textstyle\sum_k a_{k0}\eta_k{}^j + \lambda^{-j} Q(j) = d(2-\sigma_1{}^j - \sigma_2{}^j). \tag{14}$$

(ii) Let α be an irrational multiple of 2π. Since the ξ_k's are rational, there is an integer L such that $L\xi_k$ is an integer for every k. Put $j = nL$ in (14) and let $n \to \infty$. Then the left side does not oscillate whereas the right side oscillates. This contradiction shows that $\{Y_n\}$ cannot be a ffMc.

(iii) Let $\alpha = 2\pi v/N$, where v and N are relatively prime integers. Then (14) shows that $\lambda^{-j}Q(j)$ is periodic. But it tends to zero as $j \to \infty$. Thus $\lambda^{-j}Q(j) = 0$ for all j and (14) reduces to

$$\textstyle\sum_k a_{k0}\eta_k{}^j = d(2-\sigma_1{}^j - \sigma_2{}^j). \tag{15}$$

It follows easily from (15) that σ_1 equals one of the η_k's. This means that $\lambda\sigma_1$ is an eigenvalue of one of the irreducible A_{ii}'s. Now Frobenius' theorem shows that

$\lambda\sigma_1, \lambda\sigma_1^2, \cdots, \lambda\sigma_1^N$, are eigenvalues of $M_{\mu\mu}$. Since v and N are relatively prime, these eigenvalues are all distinct. But then $M_{\mu\mu}$ must have at least N rows and hence $\{Z_n\}$ must have at least $(N+1)$ states.

We will now exhibit an $(N+1)$-state Markov chain $\{V_n\}$ of which $\{Y_n\}$ is a function. The state-space of $\{V_n\}$ will be $\{0, 1, \cdots, N\}$. Suppose $\{v_0, \cdots, v_N\}$ and $||v_{ij}||$ denote respectively the initial distribution and the transition matrix of $\{V_n\}$. Using the quantities introduced earlier in this section, we take $v_0 = 4d$, $v_{00} = 1-c$, $v_{01} = c_1/(1-\lambda^N)$, $v_1 = 4dc/(1-\lambda^N)$, $v_{10} = (1-\lambda^N)$, $v_{1N} = \lambda^N$, and, for $2 \leqq j \leqq N$, $v_{0j} = c_j/(1-\lambda^N)$, $v_j = 4d\sum_{k=j}^{\infty} c_k/(1-\lambda^N)$ and $v_{j,j-1} = 1$. Using the fact that $c_N = 0$, we can check that $\{V_n\}$ is stationary.

Let $W_n = g(V_n)$, where $g(0) = \delta$ and $g(j) = \mu$ for $j = 1, \cdots, N$. Then we claim that $\{Y_n\}$ and $\{W_n\}$ have the same distribution. To see this denote the probability function for the W-process by p'. We want to show that $p(s) = p'(s)$ for all $s \in S(J)$, where $J = \{\mu, \delta\}$. Clearly $n(\delta) = 1$ for the W-process also. Therefore (5) and a similar equation involving p' show that it is enough to show that $p(s) = p'(s)$ whenever $s = \delta^j$, $\delta\mu^j\delta$, $\delta\mu^j$ or μ^j, where $j = 1, 2, \cdots$. For $s = \delta^j$ or $\delta\mu^j\delta$, this verification is easy. For the remaining two types of sequences, the required result follows from Lemma 2.1, if it is noted that both $p(\mu^j)$ and $p'(\mu^j)$ tend to zero as $j \to \infty$. This proves part (b) of the assertions and completes the proof of the theorem.

COROLLARY. *If $\alpha = 2\pi v/N$ where v and N are relatively prime integers and $N \geqq 4$, then $\{Y_n\}$ is a non-regular function of a finite Markov chain.*

REMARK. The example given by Fox and Rubin in Section 2 of [3] uses $\lambda = \frac{1}{2}$ and $\alpha = 2$.

3. The exchangeable case. In this section J will be an arbitrary but fixed non-empty finite set. Further $\{Y_n\}$ will be an exchangeable process with state-space J. That is, for every $k \geqq 1$, for every choice of distinct positive integers $n_1, \cdots, n_k$ and for every choice of elements $\delta_1, \cdots, \delta_k$ in J, the probability

$$P(Y_{n_1} = \delta_1, \cdots, Y_{n_k} = \delta_k)$$

depends only on k and $\delta_1, \cdots, \delta_k$. The purpose of this section is to prove that if $\{Y_n\}$ is a ffMC then it is a regular ffMc. We first introduce some notation and prove three lemmas.

Recall from Section 1 that $S(J)$ is the set of all finite sequences of elements of J. Let $H(J)$ be the set of all vectors $\xi = \{\xi(\delta), \delta \in J\}$ such that $\xi(\delta) \geqq 0$ for all δ and $\xi(\delta) > 0$ for some δ. Let $H^+(J) = \{\xi \in H(J) \mid \xi(\delta) > 0 \text{ for all } \delta \in J\}$. If $J_1 \subset J$ and $\xi \in H(J)$, then $\xi \mid J_1$ will be the vector $\{\xi(\delta) \mid \delta \in J_1\}$. Finally, if $s = (\delta_1, \cdots, \delta_m) \in S(J)$ and $\xi \in H(J)$ we use the symbol $\langle s, \xi \rangle$ to denote the product $\xi(\delta_1) \times \cdots \times \xi(\delta_m)$.

LEMMA 3.1. *Given n distinct elements in $H^+(J)$, we can find an $s \in S(J)$ such that s uses all the elements of J and the quantities $\langle s, \xi_i \rangle$, $i = 1, \cdots, n$ are all distinct.*

PROOF. Let $u =$ the number of elements in J. We use induction on u. If J has only one element δ, then $\lambda_k = \xi_k(\delta)$ are all distinct. Therefore the singleton sequence δ is an eligible s. Thus the lemma holds for $u = 1$.

Suppose the lemma holds for $u = m$. Let J have $(m+1)$ elements. Let δ_0 be a fixed element of J and let $J_1 = J - \{\delta_0\}$. Renumber $\xi_1, \cdots, \xi_n$ as $\xi_{ij}, j = 1, \cdots, n_i$, $i = 1, \cdots, k$ in such a way that $\xi_{ij} \mid J_1 = \xi_{i'j'} \mid J_1$ if, and only if, $i = i'$. Let $\eta_i = \xi_{i1} \mid J_1$. Then $\eta_1, \cdots, \eta_k$ are distinct elements of $H^+(J_1)$. By the induction hypothesis, there is a sequence $s_1 \in S(J_1)$ such that s_1 uses all the elements of J_1 and the quantities $\langle s_1, \eta_i \rangle$, $i = 1, \cdots, k$ are all distinct. Let $a_i = \langle s_1, \eta_i \rangle$ and $\lambda_{ij} = \xi_{ij}(\delta_0)$. Let $s = s_1 \delta_0{}^r$. Then $\langle s, \xi_{ij} \rangle = \langle s_1, \eta_i \rangle \cdot \lambda_{ij}^r = a_i \lambda_{ij}^r$. Since $i \neq i' \Rightarrow a_i \neq a_{i'}$ while $j \neq j' \Rightarrow \lambda_{ij} \neq \lambda_{ij'}$, there is a large enough integer r such that $\langle s, \xi_{ij} \rangle$ are all distinct. This proves that the lemma holds for $u = m+1$. The lemma is thus established.

LEMMA 3.2. *Let $\xi_1, \cdots, \xi_n$ be distinct elements of $H^+(J)$. Then there are n sequences $s_1, \cdots, s_n$ in $S(J)$ such that each s_i uses all the elements of J and the $n \times n$ matrix $\|\langle s_i, \xi_j \rangle\|$ is nonsingular.*

PROOF. Because of Lemma 3.1, there is an $s \in S(J)$ such that s uses all the elements of J and the quantities $\lambda_j = \langle s, \xi_j \rangle$ are all distinct. The λ_j's are obviously positive. Define $s_i = s^i$. Then $\langle s_i, \xi_j \rangle = \lambda_j{}^i$. The Van der Monde matrix $\|\lambda_j{}^i\|$ is nonsingular. The lemma follows.

LEMMA 3.3. *Let $\xi_1, \cdots, \xi_n$ be distinct elements of $H(J)$. Then there are n sequences $s_1, \cdots, s_n$ in $S(J)$ such that the $n \times n$ matrix $\|\langle s_i, \xi_j \rangle\|$ is nonsingular.*

PROOF. Let $G = \{\xi_1, \cdots, \xi_n\}$. Denote by $\mathscr{A}$ the class of all subsets A of J such that there is a $\xi \in G$ whose support is A. For $A \in \mathscr{A}$, let $G(A) = \{\xi \in G \mid A$ is the support of $\xi\}$. Suppose that L is the number of sets in $\mathscr{A}$ and write $\mathscr{A} = \{A_1, \cdots, A_L\}$ in such a way that $i < i' \Rightarrow$ the number of elements in $A_i \leqq$ the number of elements in $A_{i'}$. Let $G(A_i)$ be enumerated as $\{\xi_{i1}, \cdots, \xi_{in_i}\}$. Finally, let $\eta_{ij} = \xi_{ij} \mid A_i$. Observe that $\eta_{ij}, j = 1, \cdots, n_i$ are distinct elements of $H^+(A_i)$. By Lemma 3.2 there are sequences $s_{ij}, j = 1, \cdots, n_i$ in $S(A_i)$ such that each s_{ij} uses all the elements of A_i and the $n_i \times n_i$ matrix $\|\langle s_{ij}, \eta_{ij'} \rangle\|$ is nonsingular.

Let $i' < i$. Because of the order in which the sets of $\mathscr{A}$ have been written down, there is a $\delta_0 \in A_i$ such that $\delta_0 \notin A_{i'}$. Now s_{ij} uses all the elements of A_i. In particular it uses δ_0. But $\xi_{i'j'}(\delta_0) = 0$ because $\delta_0 \notin A_{i'}$. Thus $\langle s_{ij}, \xi_{i'j'} \rangle = 0$.

Let $V_{ii'}$ be the $n_i \times n_{i'}$ matrix $\|\langle s_{ij}, \xi_{i'j'} \rangle\|$ and let V be the partitioned matrix $\|V_{ii'}\|$. Then we have shown so far that each V_{ii} is nonsingular and that $i > i' \Rightarrow V_{ii'} = 0$. Therefore V is nonsingular. The sequences $s_{ij}, j = 1, \cdots, n_i, i = 1, \cdots, L$ thus satisfy the assertion of the lemma. This completes the proof of the lemma.

We are now ready to prove the following theorem.

THEOREM 3.1. *An exchangeable process which is a function of a finite Markov chain is a regular function of a finite Markov chain.*

PROOF. As stated in the first paragraph of this section, let $\{Y_n\}$ be an exchangeable process with state-space J. Let Q denote the subset of $H(J)$ corresponding to the probability distributions on J. The de Finetti theorem asserts that there is a unique probability measure μ on the Borel subsets of Q such that for every $s \in S(J)$, we have

$$p(s) = \int_Q \langle s, q \rangle \, d\mu(q). \tag{16}$$

Suppose that $\{Y_n\}$ is a ffMc. Then as stated at the end of [2], μ has finite support. Let the support of μ be $\{q_u, u = 1, \cdots, N\}$. Let $a_u = \mu(\{q_u\}) > 0$. Let J_u be the support of q_u and suppose that N_u is the number of elements in J_u. Denote $q_u \mid J_u$ by r_u. We will denote by M_u the $N_u \times N_u$ matrix all of whose rows equal r_u. Let M be the direct sum of the M_u's. The set $\{(u, \delta) \mid \delta \in J_u, u = 1, \cdots, N\}$ will be denoted by I. We set $\mathbf{m}_u = a_u r_u$ and $\mathbf{m} = (\mathbf{m}_1, \cdots, \mathbf{m}_N)$. Finally, f is the function on I onto J defined by $f[(u, \delta)] = \delta$. Then if $\{X_n\}$ is a Markov chain with state-space I, initial distribution $\mathbf{m}$ and transition matrix M, then $\{Y_n\}$ and $\{f(X_n)\}$ have the same distribution. This Markov chain $\{X_n\}$ is the same as that constructed in the proof of the theorem in [2], with the exceptions that the state-space is finite and states (u, δ) with $q_u(\delta) = 0$ have been eliminated.

We now proceed to show that the above representation of $\{Y_n\}$ as the function f of the Markov chain $\{X_n\}$ is regular. In other words we will show that $n(\delta) = N(\delta)$, where $N(\delta)$ is the number of elements in $f^{-1}[\{\delta\}]$. Because of the result of Gilbert [5] quoted in Section 1 of this paper, we always have $n(\delta) \leqq N(\delta)$. If $N(\delta) = 0$, then $q_u(\delta) = 0$ for all u and hence $p(\delta) = 0$ and $n(\delta) = 0$. So let $N(\delta) > 0$. Without loss of generality we may assume that $(u, \delta) \in I$ for $1 \leqq u \leqq N(\delta)$ and that $(u, \delta) \notin I$ for $N(\delta) < u \leqq N$. Now q_i, $1 \leqq i \leqq N(\delta)$, are distinct elements of $H(J)$. By Lemma 3.3, there are sequences $s_i, i = 1, \cdots, N(\delta)$, such that the $N(\delta) \times N(\delta)$ matrix $\|\langle s_i, q_j \rangle\|$ is nonsingular. Now (16) shows that

$$p(s_i \delta s_j) = \textstyle\sum_{u=1}^{N(\delta)} a_u q_u(\delta) \langle s_i, q_u \rangle \langle s_j, q_u \rangle,$$

because $q_u(\delta) = 0$ for $u > N(\delta)$. Define the two $N(\delta) \times N(\delta)$ matrices V and W as follows.

$$V = \|\langle s_i, q_j \rangle\| \text{ and } W = \operatorname{diag}[a_1 q_1(\delta), \cdots, a_{N(\delta)} q_{N(\delta)}(\delta)].$$

Then $\|p(s_i \delta s_j)\| = VWV'$, which is nonsingular because both V and W are nonsingular. It follows that $n(\delta) \geqq N(\delta)$. This completes the proof of the theorem.

Acknowledgment. The authors are grateful to Professors Dorian Feldman and Ashok Maitra for useful discussions.

REFERENCES

[1] DHARMADHIKARI, S. W. (1963). Sufficient conditions for a stationary process to be a function of a finite Markov chain. *Ann. Math. Statist.* **34** 1033–1041.

[2] DHARMADHIKARI, S. W. (1964). Exchangeable processes which are functions of stationary Markov chains. *Ann. Math. Statist.* **35** 429–430.

[3] FOX, MARTIN and RUBIN, HERMAN (1968). Functions of processes with Markovian states. *Ann. Math. Statist.* **39** 938–946.

[4] GANTMACHER, F. R. (1959). *The Theory of Matrices* **2**. Chelsea, New York.

[5] GILBERT, EDGAR, J. (1959). On the identifiability problem for functions of finite Markov chains. *Ann. Math. Statist.* **30** 688–697.