The Annals of Mathematical Statistics
1970, Vol. 41, No. 2, 472–479

FUNCTIONS OF PROCESSES WITH MARKOVIAN STATES—III

BY MARTIN FOX[1] AND HERMAN RUBIN[2]

Michigan State University and Purdue University

1. Introduction. Let $\{Y_k\}$ be a stochastic process where either $k = 1, 2, \cdots$ or $k = 0, \pm 1, \pm 2, \cdots$. Suppose there exist a time n and a state ε such that $P(Y_n = \varepsilon) > 0$. In this case, the rank at time n of the state ε is defined in [3], although the notion was first considered by Gilbert in [5]. The definition is such that a state of rank 1 is Markovian.

Let $\{X_k\}$ be a second stochastic process indexed as $\{Y_k\}$. Gilbert [5] proved (but stated in far less generality) that if $\nu_n(\varepsilon)$ and $\mu_n(\delta)$ are the ranks at time n of the states ε of $\{Y_k\}$ and δ of $\{X_k\}$ respectively, and if $Y_n = f(X_n)$, then

$$\nu_n(\varepsilon) \leqq \textstyle\sum_{f_n(\delta)=\varepsilon} \mu_n(\delta), \tag{1.1}$$

Dharmadhikari [1] considered the case of stationary $\{Y_k\}$ with all states of finite rank and found an additional condition in order to guarantee $Y_k = f(X_k)$ for $\{X_k\}$ stationary and Markovian. The present authors [3] provided an example showing Dharmadhikari's result cannot be obtained without some condition beyond finiteness of rank.

In the present paper we extend the definition of rank by eliminating the condition $P(Y_n = \varepsilon) > 0$ and generalize Gilbert's, Dharmadhikari's, and our own results.

Section 2 contains the extension of the definition of rank. Sections 3 and 4 contain, respectively, the extensions of Gilbert's and Dharmadhikari's results. The extension of the Dharmadhikari-type result to the nonstationary case is discussed in Section 5. Section 6 contains an outline of the extension of the work of the present authors [4].

2. The extended definition of rank. Let U_k be the state space of $\{Y_k\}$ at time k and set $V_n = \cdots \times U_{n-2} \times U_{n-1}$ and $W_n = U_{n+1} \times U_{n+2} \times \cdots$. Let $\mathscr{A}_n$, $\mathscr{S}_n$ and $\mathscr{T}_n$ be, respectively, the σ-algebras of measurable subsets of U_n, V_n and W_n. Assume all $\mathscr{A}_i$ (and hence $\mathscr{S}_n$ and $\mathscr{T}_n$) are separable. For $A \in \mathscr{A}_n$, $S \in \mathscr{S}_n$ and $T \in \mathscr{T}_n$ let

$$P_n(S, A, T) = P((\cdots, Y_{n-2}, Y_{n-1}) \in S, \quad Y_n \in A, \quad (Y_{n+1}, Y_{n+2}, \cdots) \in T).$$

Let $Q_n(A) = P_n(V_n, A, W_n)$ and

$$p_n(S, \varepsilon, T) = \frac{dP_n(S, \cdot, T)}{dQ_n}(\varepsilon).$$

Received May 26, 1969.

[1] Sponsored by Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No. DA-31-124-ARO-D-462.

[2] Partially supported by Office of Naval Research Contract N 00014-67-A-0226-0008.

We assume p_n be a *bona fide* conditional probability on $\mathscr{S}_n \times \mathscr{T}_n$. This is guaranteed if we assume conditional expectations are regular.

We say that $C_{n\nu}$ is the *set of states of rank* ν *at time* n if there exist linearly independent functions g_i on $\mathscr{S}_n \times C_{n\nu}$ and linearly independent functions h_i on $\mathscr{T}_n \times C_{n\nu}$ $(i = 1, \cdots, \nu)$ such that

$$p_n(S, \varepsilon, T) = \sum_{i=1}^{\nu} g_i(S, \varepsilon) h_i(T, \varepsilon) \tag{2.1}$$

for all $S \in \mathscr{S}_n$, $T \in \mathscr{T}_n$ and $\varepsilon \in C_{n\nu}$ and $C_{n\nu}$ is maximal (except for a set of measure 0). Note that the $C_{n\nu}$ are only defined up to sets of measure 0.

Let $\varepsilon \in C_{n\nu}$. It is easy to see that there exist $S_{\varepsilon i} \in \mathscr{S}_n$ and $\mathscr{T}_{\varepsilon i} \in \mathscr{T}_n$ $(i = 1, \cdots, \nu)$ such that the matrix $(p_n(S_{\varepsilon i}, \varepsilon, T_{\varepsilon i}))$, is nonsingular and any larger matrix of this form is singular. Conversely the condition of this paragraph implies (2.1). This condition is the obvious parallel to the condition given in the earlier papers. Furthermore, we obtain a more explicit version of (2.1), namely,

$$p_n(S, \varepsilon, T) = \sum_{i=1}^{\nu} a_i(S, \varepsilon) p_n(S_{\varepsilon i}, \varepsilon, T) \tag{2.2}$$

which will be useful in Section 4. Note that the $a_i(\cdot, \varepsilon)$ are signed measures on $\mathscr{S}_n$.

The following lemma shows that the $S_{\varepsilon i}$ can be chosen so the functions $p_n(S_{\cdot i}, \cdot, T)$ are measurable. This implies that the $a_i(S, \cdot)$ can also be taken to be measurable.

LEMMA 2.1. *There exist measurable sets* $D_{n\nu j}$ *such that* $C_{n\nu} = \bigcup_{j=1}^{\infty} D_{n\nu j}$ *and a choice of the* $S_{\varepsilon i}$ *independent of* $\varepsilon \in D_{n\nu j}$.

PROOF. Let $\mathscr{S}_{Bn}$ and $\mathscr{T}_{Bn}$ be countable bases for $\mathscr{S}_n$ and $\mathscr{T}_n$, respectively, which are closed under finite unions and intersections and complements. Now (2.1) holds for $S \in \mathscr{S}_{Bn}$ and $T \in \mathscr{T}_{Bn}$ and no fewer terms will suffice since, in general, $p_n(S, \varepsilon, T)$ can be obtained by limits from its values on $\mathscr{S}_{Bn} \times \mathscr{T}_{Bn}$. Let $D_{n\nu k}$ be given for $k < j$. Let $D'_{n\nu j} = \bigcup_{k<j} D_{n\nu k}$. If $C_{n\nu} \sim D_{n\nu j}$ has measure 0, the decomposition has already been achieved. If not, for R a ν-element subset of $\mathscr{S}_{Bn}$ let

$$E_R = \{\varepsilon : p_n(S, \varepsilon, T) \quad \text{is of rank } \nu \text{ when} \quad S \in R, T \in \mathscr{T}_{Rn}\} \cap C_{n\nu} \sim D'_{n\nu j}.$$

Let $D_{n\nu j} = E_R$ for some R such that $Q_n(D_{n\nu j}) > \frac{1}{2} \sup_R Q(E_R)$ and let $S_{\varepsilon i}$ be the elements of a corresponding R in some order. The functions $a_i(S, \varepsilon)$ are now uniquely determined and are readily seen to have bounded Radon–Nikodym derivatives with respect to $p_n(\cdot, \varepsilon, W_n)$ and hence can be extended to all of $\mathscr{S}_n$. A similar limit argument shows that (2.2) holds for all $S \in \mathscr{S}_n$ and $T \in \mathscr{T}_n$.

3. Generalization of Gilbert's result. Let $\{X_k\}$ be another process with $Y_k = f_k(X_k)$ for measurable f_k and let P_n^*, $C^*_{n\nu}$ and Q_n^* be defined in terms of $\{X_k\}$ analogously to P_n, $C_{n\nu}$ and Q_n, respectively. For fixed $\mu = (\mu_1, \cdots, \mu_m)$ let $B_{nm\mu} \subset U_n$ be such that, for all $\varepsilon \in B_{nm\mu}$ and $j = 1, \cdots, m$, there exists $\delta_{\varepsilon j} \in C^*_{n\mu_j}$ for which $\varepsilon = f_n(\delta_{\varepsilon j})$. The desired generalization of Gilbert's result is:

THEOREM 3.1. *If* $\{X_k\}$ *satisfies the assumptions of Section* 2, *then* $B_{nm\mu} \subset \bigcup_{\nu=1}^{\lambda} C_{n\nu}$ *where* $\lambda = \sum_{j=1}^{m} \mu_j$.

PROOF. First note that, by Lemma 2.1, the $C_{n\nu}^*$ are measurable. Let $A \subset B_{nm\mu}$, $A^* = f^{-1}(A)$ and $A_\nu^* = A^* \bigcap C_{n\nu}^*$. Assume A and, hence, A^* and A_ν^* are measurable. For $C_{n\nu}^*$ let the functions analogous to the g_i and h_i, respectively, be $g_{\nu i}^*$ and $h_{\nu i}^*$. Finally, let $\tilde{\mu} = \max \mu_j$. Then, since $A^* = \bigcup_{\nu=1}^{\tilde{\mu}} A_\nu^*$ we have

$$(3.1) \qquad P_n(S, A, T) = \sum_{\nu=1}^{\tilde{\mu}} P_n^*(S^*, A_\nu^*, T^*)$$
$$= \sum_{\nu=1}^{\tilde{\mu}} \int_{A_\nu^*} \sum_{i=1}^{\nu} g_{\nu i}^*(S^*, \delta) h_{\nu i}^*(T^*, \delta)\, dQ_n^*(\delta)$$

where $(\cdots, X_{n-2}, X_{n-1}) \in S^*$ if, and only if, $(\cdots, Y_{n-2}, Y_{n-1}) \in S$ and T^* is defined similarly. Let $g_{ji}(S, \varepsilon) = g_{\nu_j i}^*(S^*, \delta_{\varepsilon j})$ and $h_{ji}(T, \varepsilon) = h_{\nu_j i}^*(T^*, \delta_{\varepsilon j})$ for $j = 1, \cdots, m$. Since $Q_n(A) = \sum_{\nu=1}^{\tilde{\mu}} Q_n^*(A_\nu^*)$ it follows that (3.1) becomes

$$P_n(S, A, T) = \int_A \sum_{\nu=1}^{\tilde{\mu}} \sum_{\nu_j=\nu} \sum_{i=1}^{\nu_j} g_{ji}(S, \varepsilon) h_{ji}(T, \varepsilon)\, dQ_n(\varepsilon)$$
$$= \int_A \sum_{j=1}^{m} \sum_{i=1}^{\nu_j} g_{ji}(S, \varepsilon) h_{ji}(T, \varepsilon)\, dQ_n(\varepsilon).$$

This completes the proof.

4. Generalization of Dharmadhikari's result. With the exception of proofs of measurability, the methods of this section are straightforward generalizations of those developed in [1]. To facilitate comparisons of steps, and whenever it is appropriate, formulae will be numbered in the form (x, y) where x will be our sequential number and y will be the sequential number in [1]. In referring to previously derived formulae only x will be used. A similar notation will be used for lemmas, etc.

We now assume $\{Y_k\}$ is a stationary process so we can drop the subscripts n and k in all notation introduced in Sections 2 and 3. We also refer to the rank of a state without reference to time. The assumptions (i) that $\mathscr{A}$ is separable and (ii) that p is a *bona fide* conditional probability are retained. Note that (2.2) is the parallel of (1.1) in [1].

Take a fixed determination of the C_ν and let $\nu(\varepsilon)$ be such that $\varepsilon \in C_{\nu(\varepsilon)}$. Let $\alpha_\varepsilon(S) = (a_1(S, \varepsilon), \cdots, a_{\nu(\varepsilon)}(S, \varepsilon))$ and $\pi_\varepsilon(T) = (p(S_{\varepsilon 1}, \varepsilon, T), \cdots, p(S_{\varepsilon\nu(\varepsilon)}, \varepsilon, T)$. Assume the $S_{\varepsilon i}$ are chosen so $\alpha_{\cdot}(S)$ and $\pi_{\cdot}(T)$ are measurable. Recall that Lemma 2.1 guarantees that this is possible. Now (2.2) can be rewritten as

$$(4.1, 1.2) \qquad p(S, \varepsilon, T) = (\alpha_\varepsilon(S), \pi_\varepsilon(T)).$$

Let $\alpha_\varepsilon = \{\alpha_\varepsilon(S) : S \in \mathscr{S}\}$ and $\pi_\varepsilon = \{\pi_\varepsilon(T) : T \in \mathscr{T}\}$. Let $\mathscr{C}(\pi_\varepsilon)$ and $\mathscr{C}(\alpha_\varepsilon)$ be the closed, convex cones generated by π_ε and α_ε, respectively.

LEMMA 4.1, 1.1. *For every ε, both $\mathscr{C}(\alpha_\varepsilon)$ and $\mathscr{C}(\pi_\varepsilon)$ have dimension $\nu(\varepsilon)$.*

PROOF. As in [1].

For any cone $\mathscr{C}$ we let $\mathscr{C}^+$ be its dual cone.

LEMMA 4.2, 1.2. *Let $b \varepsilon [\mathscr{C}(\pi_\varepsilon)]^+$, $b \neq 0$. Then,* $(b, \pi_\varepsilon(W)) > 0$.

PROOF. Now $(b, \pi_\varepsilon(\cdot))$ is a measure on $\mathscr{T}$. Hence, $(b, \pi_\varepsilon(W)) = 0$ implies $(b, \pi_\varepsilon(T)) = 0$ for all $T \in \mathscr{T}$ so that $b \perp \mathscr{C}(\pi_\varepsilon)$. But then $b = 0$ by Lemma 4.1.

We require notation for the analogy to what occurs in [1] when sequences containing two specified states are used as arguments. The functions used so far have been Radon–Nikodym derivations with respect to Q evaluated at the state at one time and are signed measures with respect to their other arguments. Hence it is natural to take another Radon–Nikodym derivative. We assume that two-dimensional derivatives exist with respect to $Q \times Q$. Let

$$a_{\mu j}^{*}(S, \varepsilon) = \frac{da_j(S \times \cdot\,, \mu)}{dQ}(\varepsilon) \qquad \text{for} \quad j = 1, \cdots, \nu(\mu);$$

$$p^{*}(S, \varepsilon, \mu, T) = \frac{dp(S, \varepsilon, \cdot \times T)}{dQ}(\mu);$$

$$\pi_{\varepsilon}^{*}(\mu, T) = (p^{*}(S_{\varepsilon i}, \varepsilon, \mu, T), \cdots, p^{*}(S_{\varepsilon\nu(\varepsilon)}, \varepsilon, \mu, T));$$

$$\alpha_{\mu}^{*}(S, \varepsilon) = (a_{\mu i}^{*}(S, \varepsilon), \cdots, a_{\mu\nu(\mu)}^{*}(S, \varepsilon));$$

and let $A_{\varepsilon\mu}$ be the matrix with component $a_{\mu j}^{*}(S_{\varepsilon i}, \varepsilon)$ in its ith row and jth column. We have the alternative expression

$$p^{*}(S, \varepsilon, \mu, T) = \frac{dP(S, \cdot, \cdot \times T)}{d(Q \times Q)}(\varepsilon, \mu)$$

so that $p^{*}(S, \cdot, \cdot\, T)$ is measurable. Furthermore, the choice of the $S_{\varepsilon i}$ in Lemma 2.1 guarantees measurability of $p^{*}(S_{\cdot\, i}, \cdot, \cdot, T)$. From stationarity

$$p^{*}(S, \varepsilon, \mu, T) = \frac{dp(S \times \cdot\,, \mu, T)}{dQ}(\varepsilon)$$

the right side of which, by (2.2), is equal to $\sum_{j=1}^{\nu(\mu)} a_{\mu j}^{*}(S, \varepsilon)p(S_{\mu j}, \mu, T)$. Setting $S = S_{\varepsilon i}$ this reduces to

$$\pi_{\varepsilon}^{*\prime}(\mu, T) = A_{\varepsilon\mu}\pi_{\mu}{}'(T). \tag{4.2, 1.4}$$

Hence there exists a measurable determination of $A_{\cdot\,\cdot}$.

Taking inner product with respect to b on both sides of (4.2) we obtain

$$(b, \pi_{\varepsilon}^{*}(\mu, T)) = (bA_{\varepsilon\mu}, \pi_{\mu}(T)) \tag{4.3, 1.5}$$

for all $b \in E^{\nu(\xi)}$ and $T \in \mathscr{T}$.

LEMMA 4.3, 1.3. *For every* $\varepsilon, \mu \in U$ *and* $S \in \mathscr{S}$ *we have* $\alpha_{\varepsilon}(S)A_{\varepsilon\mu} = \alpha_{\mu}^{*}(S, \varepsilon)$.

PROOF. From (2.2) and (4.3) we see that for every $T \in \mathscr{T}$ we have

$$\begin{aligned}(\alpha_{\varepsilon}(S)A_{\varepsilon\mu}, \pi_{\mu}(T)) &= (\alpha_{\varepsilon}(S), \pi_{\varepsilon}^{*}(\mu, T)) \\ &= p^{*}(S, \varepsilon, \mu, T) \\ &= (\alpha_{\mu}^{*}(S, \varepsilon), \pi_{\mu}(T)).\end{aligned}$$

This and the fact that $\mathscr{C}(\pi_{\mu})$ has full dimension complete the proof.

LEMMA 4.4. *For each* $\varepsilon \in U$ *let* $[\mathscr{C}(\pi_\varepsilon)]^+$ *be polyhedral and generated by* $N(\varepsilon)$ *non-zero vectors* $\beta_{\varepsilon 1}, \cdots, \beta_{\varepsilon N(\varepsilon)}$. *For* $j > N(\varepsilon)$ *set* $\beta_{\varepsilon j} = 0$. *Then, the* $\beta_{\varepsilon j}$ *can be selected so that for each* $j = 1, 2, \cdots$ *the functions* $\beta_{\cdot j}$ *on* $\mathscr{C}_\nu$ *to* E^ν *are measurable.*

PROOF. By the separability of $\mathscr{T}$ there exist $T_1, T_2, \cdots \in \mathscr{T}$ independent of ε such that the $\pi_\varepsilon(T)$ are limits of linear combinations of the $\pi_\varepsilon(T_i)$. For the remainder of this proof S with or without super- or subscripts will denote a cone generated by a rational sphere (rational center and radius). Note that a face of $\mathscr{C}(\pi_\varepsilon)$ can be identified by finding S such that (i) there exists i for which $\pi_\varepsilon(T_i) \in S$, (ii) there exists $S^* \subset S$ such that all $\pi_\varepsilon(T_i) \notin S^*$ and (iii) if $S_1, S_2 \subset S$ contain no $\pi_\varepsilon(T_i)$, then the convex hull of $S_1 \cup S_2$ contains no $\pi_\varepsilon(T_i)$. In this case we will say that S *establishes a face of* $\mathscr{C}(\pi_\varepsilon)$. Fix $q \notin R^{\nu(\varepsilon)}$ for all ε. Here $R^{\nu(\varepsilon)}$ is $\nu(\varepsilon)$-dimensional Euclidean space. If S establishes a face of $\mathscr{C}(\pi_\varepsilon)$ let $\varphi(S, \varepsilon)$ be the corner of $[\mathscr{C}(\pi_\varepsilon)]^+$ orthogonal to that face. Let $\varphi(S, \varepsilon) = q$ otherwise. With the obvious σ-field for $R^{\nu(\varepsilon)} \cup \{q\}$, we see that $\varphi(S, \varepsilon)$ is a measurable function for fixed S as ε varies over $\mathscr{C}_\nu$ for each ν. This completes the proof since there are only countably many rational spheres.

LEMMA 4.5, 2.1. *For each* $\varepsilon \in U$ *let* $[\mathscr{C}(\pi_\varepsilon)]^+$ *be polyhedral and generated by* $N(\varepsilon)$ *non-zero vectors* $\beta_{\varepsilon j}$. *Let* $r_{\varepsilon j}(T) = (\beta_{\varepsilon j}, \pi_\varepsilon(T))$ *for all* $T \in \mathscr{T}$. *Then,*

(i) *the* $\beta_{\varepsilon j}$ *can be chosen such that* $\beta_{\cdot j}$ *is measurable and for all* $\varepsilon \in U$ *and* $j = 1, \cdots, N(\varepsilon)$ *we have*

$$r_{\varepsilon j}(W) = 1; \tag{4.4, 2.1}$$

(ii) *there exists* $r^*_{\varepsilon j}(\mu, T) = (d/dQ) r_{\varepsilon j}(\cdot \times T)(\mu)$ *and* $m_{\varepsilon j, \mu k} \geqq 0$ $(j = 1, \cdots, N(\varepsilon)$; $k = 1, \cdots, N(\mu))$ *such that the* $m_{\cdot j, \cdot k}$ *are measurable and*

$$r^*_{\varepsilon j}(\mu, T) = \textstyle\sum_{j=1}^{N(\mu)} m_{\varepsilon j, \mu k} r_{\mu k}(T). \tag{4.5, 2.2}$$

Furthermore if the $r_{\varepsilon j}$ *satisfy* (4.4), *then*

$$\textstyle\int_U \sum_{k=1}^{N(\mu)} m_{\varepsilon j, \mu k} \, dQ(\mu) = 1. \tag{4.6, 2.3}$$

PROOF. (i) For fixed $\varepsilon \in U$ the $\beta_{\varepsilon j}$ are defined only to multiplicative constants and $(\beta_{\varepsilon j}, \pi_\varepsilon(W)) > 0$ by Lemma 4.2. The functions of ε which normalize so that $(\beta_{\varepsilon j}, \pi_\varepsilon(W)) = 1$ are continuous in $\pi_\varepsilon(W)$ and, hence, are measurable. Thus, measurable $\beta_{\cdot j}$ remain measurable after normalization.

(ii) Now $r_{\varepsilon j}(\cdot)$ and, hence, $r_{\varepsilon j}(\cdot \times T)$ are measures. Furthermore, $\pi_\varepsilon(\cdot \times T) \ll Q$ so $r_{\varepsilon j}(\cdot \times T) \ll Q$ and $r^*_{\varepsilon j}(\mu, T)$ exists.

From (4.3) we obtain

$$\begin{aligned}(\beta_{\varepsilon j} A_{\varepsilon\mu}, \pi_\mu(T) &= (\beta_{\varepsilon j}, \pi_\varepsilon{}^*(\mu, T)) \\ &= r^*_{\varepsilon j}(\mu, T) \geqq 0.\end{aligned} \tag{4.7, 2.4}$$

Hence, $\beta_{\varepsilon j} A_{\varepsilon\mu} \in [\mathscr{C}(\pi_\mu)]^+$ so there exist $m_{\varepsilon j, \mu k} \geqq 0$ such that

$$\beta_{\varepsilon j} A_{\varepsilon\mu} = \textstyle\sum_{k=1}^{N(\mu)} m_{\varepsilon j, \mu k} \beta_{\mu k}. \tag{4.8, 2.5}$$

We have seen that $B_{.j}$ and $A_{..}$ are measurable so that also $\beta_{.j}A_{..}$ is measurable. There is a unique representation (4.8) of $\beta_{\varepsilon j}A_{\varepsilon\mu}$ in terms of $\beta_{\mu 1}, \cdots, _{\beta l}$ for the minimal $l \leqq N(\mu)$. The $m_{\varepsilon j,\mu k}$ for this representation are continuous in $\beta_{\varepsilon j}A_{\varepsilon\mu}$ and $\beta_{\mu k}$ so that the $m_{.j,\,.k}$ so chosen are measurable.

Taking the inner product of the expressions in (4.8) with $\pi_\mu(T)$ and applying (4.7) yields (4.5). If (4.4) holds, setting $T = W$ in (4.5) we obtain

(4.9, 2.6) $$r^*_{\varepsilon j}(\mu, W) = \textstyle\sum_{k=1}^{N(\mu)} m_{\varepsilon j,\mu k}\,.$$

Hence,

$$\int_U \textstyle\sum_{k=1}^{N(\mu)} m_{\varepsilon j,\mu k}\, dQ(\mu) = \int r^*_{\varepsilon j}(\mu, W)\, dQ(\mu) = r_{\varepsilon j}(W) = 1$$

which proves (4.6).

Assume the conditions of Lemma 4.5 hold. Since $\alpha_\varepsilon(V) \in [\mathscr{C}(\pi_\varepsilon)]^+$, there exist $m^{(0)}_{\varepsilon j} \geqq 0$ $(j = 1, \cdots, N(\varepsilon))$ such that

(4.10, 2.8) $$\alpha_\varepsilon(V) = \textstyle\sum_{j=1}^{N(\varepsilon)} m^{(0)}_{\varepsilon j}\beta_{\varepsilon j}\,.$$

The existence of a measurable choice of $m^{(0)}_{.j}$ follows in the same way as for $m_{.j,\,.k}$ and we assume this choice has been made. Applying (4.1), (4.10) and (4.4) we obtain $p(V, \varepsilon, W) = \sum_{j=1}^{N(\varepsilon)} m^{(0)}_{\varepsilon j}$. Let

(4.11, 2.9) $$m^{(l+1)}_{\varepsilon j} = \int_U \textstyle\sum_{i=1}^{N(\mu)} m^{(l)}_{\mu i}\, m_{\mu i,\varepsilon j}\, dQ(\mu)$$

and $m^{(l)*}_{\varepsilon j} = m^{(l)}_{\varepsilon j}/p(V, \varepsilon, W)$ for $l = 0, 1, \cdots$. Clearly the $m^{(l)}_{.j}$ and, hence, the $m^{(l)*}_{.j}$ are measurable. Furthermore, $m^{(l)*}_{\varepsilon j} \geqq 0$ and $\sum_{j=1}^{N(\varepsilon)} m^{(l)*}_{\varepsilon j} = 1$. There exists a weak $-*$ cluster point $m^*_{\varepsilon j}$ of the $m^{(l)*}_{\varepsilon j}$ for which $m^*_{.j}$ is measurable and $\sum_{j=1}^{N(\varepsilon)} m^*_{\varepsilon j} = 1$.

Finally, $m_{\varepsilon j} = m^*_{\varepsilon j}p(V, \varepsilon, W)$ is the density with respect to $Q \times \Delta$ (Δ being counting measure on the integers) of the stationary distribution for the Markov process with transition function given by

$$g(\varepsilon, i, A) = \textstyle\sum_{j=1}^{\infty} \int_{(\mu,j)\in A} m_{\varepsilon i,\mu j}\, dQ(\mu).$$

LEMMA 4.6, 2.2. *For* $l = 0, 1, 2, \cdots$ *and every* $\varepsilon \in U$ *we have* $\sum_{j=1}^{N(\varepsilon)} m^{(l)}_{\varepsilon j}\beta_{\varepsilon j} = \alpha_\varepsilon(V)$.

PROOF. By (4.10) the result is true for $l = 0$. Assume it is true for $l = \nu$. From (4.11), (4.8) and Lemma 4.3 we obtain

(4.12) $$\begin{aligned}\textstyle\sum_{k=1}^{N(\mu)} m^{(\nu+1)}_{\mu k}\beta_{\mu k} &= \int_U \textstyle\sum_{j=1}^{N(\varepsilon)} m^{(\nu)}_{\varepsilon j} \sum_{k=1}^{N(\mu)} m_{\varepsilon j,\mu k}\beta_{\mu k}\, dQ(\varepsilon)\\ &= \int_U \alpha_\varepsilon(V)A_{\varepsilon\mu}\, dQ(\varepsilon) = \int_U \alpha_\mu{}^*(V, \varepsilon)\, dQ(\varepsilon) = \alpha_\mu(V)\end{aligned}$$

and the proof is complete by induction.

LEMMA 4.7, 2.3. *Let* $g, m_{\varepsilon j}$ *and* f *be as defined above. Let* $\{X_k\}$ *be a stationary Markov process with state space* U^*, *transition function* g *and initial distribution given by the density* $m_{\varepsilon j}$ *with respect to* $Q \times \Delta$. *Then,* $\{f(X_k)\}$ *has the same distribution as* $\{Y_k\}$.

PROOF. From Lemma 4.5 we see that $\sum_{j=1}^{N(\varepsilon)} m_{\varepsilon j}\beta_{\varepsilon j} = \alpha_\varepsilon(V)$ so that inner product with $\pi_\varepsilon(T)$ yields

(4.13, 2.11) $$\textstyle\sum_{j=1}^{N(\varepsilon)} m_{\varepsilon j}\, r_{\varepsilon j}(T) = p(V, \varepsilon, T).$$

For fixed n let $T \subset \times_{i=1}^{n} U$. It suffices to show that for every n we have

$$(4.14, 2.12) \quad P[(f(X_2), \cdots, f(X_{n+1})) \in T \mid X_1 = (\varepsilon, j)] = r_{\varepsilon j}(T \times W)$$

for, then, (4.13) implies that $\{f(X_k)\}$ has the same distribution as $\{Y_k\}$.

For $n = 1$, from (4.9) we obtain

$$\begin{aligned} r_{\varepsilon j}(T \times W) &= \int_T r_{\varepsilon j}^*(\mu, W)\, dQ(\mu) \\ &= \int_T \sum_{k=1}^{N(\mu)} m_{\varepsilon j, \mu k}\, dQ(\mu) \end{aligned}$$

so that (4.14) follows. Assume (4.14) holds for $n = \nu$. Let $T \subset \times_{i=1}^{\nu} U$ and $A \subset U$. From (4.5) and stationarity we obtain

$$\begin{aligned} &P[(f(X_2), \cdots, f(X_{\nu+2})) \in A \times T \mid X_1 = (\varepsilon, j)] \\ &\quad = \int_A \sum_{k=1}^{N(\mu)} m_{\varepsilon j, \mu k} P[(f(X_3), \cdots, f(X_{\nu+2})) \in T \mid X_2 = (\mu, k)]\, dQ(\mu) \\ &\quad = \int_A \sum_{k=1}^{N(\mu)} m_{\varepsilon j, \mu k}\, r_{\mu k}(T \times W)\, dQ(\mu) \\ &\quad = \int_A r_{\varepsilon j}^*(\mu, T \times W)\, dQ(\mu) \\ &\quad = r_{\varepsilon j}(A \times T \times W) \end{aligned}$$

which, by induction, completes the proof.

Thus, we have proved the

THEOREM 4.1, 2.1. *Let $\{Y_k\}$ ($k = 0, 1, 2 \cdots$ or $k = 0, \pm 1, \pm 2, \cdots$) be a stationary process in which the σ-field on the state space is separable, conditional probability is regular and transition probabilities are absolutely continuous with respect to marginals. For each state ε assume the rank of ε is finite and that $[\mathscr{C}(\pi_\varepsilon)]^+$ is polyhedral and generated by $N(\varepsilon)$ nonzero vectors. Then, $\{Y_k\}$ is a function of a stationary Markov process. The state ε is the image under this function of $N(\varepsilon)$ states of the Markov process.*

The analogue of Corollary 2.1 of [1] follows immediately. In fact, as in Dharmadhikari's case, we can apply the condition of rank 1 or rank 2 to some states (in addition to the full conditions of the theorem) and conclude that these states are images of 1 or 2 states, respectively, of $\{X_k\}$.

5. The nonstationary case. In Section 4 stationarity of $\{Y_k\}$ was used in only two ways: (i) to simplify notation and (ii) to obtain stationarity of $\{X_k\}$. One would expect that the result should extend to the nonstationary case through essentially the same proof. This is indeed the case through Lemma 4.5 with the modification of all notation by adding a subscript for the time parameter. In place of the stationary density $m_{\varepsilon j}$ with respect to $Q \times \Delta$ we require densities $m_{\varepsilon j; n}$ with respect to Q_n satisfying

$$(5.1) \qquad m_{\varepsilon j; n+1} = \int_{U_n} \sum_{i=1}^{N(\mu)} m_{\mu i; n}\, m_{\mu i, \varepsilon j; n}\, dQ_n(\mu)$$

where $m_{\mu i, \varepsilon j; n}$ is the transition density. In addition we must show these marginal and transition densities yield a process $\{X_k\}$ and functions f_k such that $\{f_k(X_k)\}$ has the same distribution as $\{Y_k\}$.

Define $m_{\varepsilon j\,;\,0}$ by the extension of (4.10). For $n > 0$ use (5.1) for the definition of $m_{\varepsilon j\,;\,n}$. If needed, use the backwards analogue fo (5.1) for $n < 0$. The extensions of Lemmas 4.6 and 4.7 follow. Thus we may prove the extension of Theorem 4.1.

THEOREM 5.1. *Let* $\{Y_k\}$ ($k = 0, 1, 2, \cdots$ *or* $k = 0, \pm 1, \pm 2, \cdots$) *be a process in which the σ-field on the state space at each time is separable, and conditional probability is bona fide. For each state ε and each time n assume the rank of ε at time n is finite and that* $[\mathscr{C}(\pi_{\varepsilon,n})]^+$ *is polyhedral and generated by* $N_n(\varepsilon)$ *nonzero vectors. Then, there exists a Markov process* $\{X_k\}$ *and functions* f_k *such that* $Y_n = f_n(X_n)$ *for each n. The state ε is the image under* f_n *of* $N_n(\varepsilon)$ *states of* $\{X_k\}$ *at time n.*

6. Countable images. Dharmadhikari's condition that $[\mathscr{C}(\pi_\varepsilon)]^+$ be polyhedral guarantees that each state is the image of a finite number of states. In [4] the present authors showed that without this extra condition one could split the finite rank states at a finite number of times into a countable number of states. The proof in that paper is correct even in the case of the more general definition of rank in this paper. Only measurability arguments need to be added. We will use the notation of [4] in outlining the measurability arguments.

Although the notation does not indicate this, the functions $\beta_j^{(i_1)}$ depend also on ε_1, the functions $\gamma_j^{(i_{l-1}, i_l)}$ depend also on ε_{l-1} and ε_l and the functions $\delta_j^{(i_r)}$ depend also on ε_r. These functions must be shown measurable in these arguments. But they are continuous functions of the functions $\varphi^{(i_1)}$, $\xi^{(i_{l-1}, i_l)}$ and $\psi^{(i_r)}$, respectively and the latter functions clearly have the desired measurability properties.

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. (1968). Splitting a single state of a stationary process into Markovian states. *Ann. Math. Statist.* **39** 1069-1077.

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

[4] FOX, MARTIN and RUBIN, HERMAN (1969). Functions of processes with Markovian states—II. *Ann. Math. Statist.* **40** 865–869.

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