The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 865–869

FUNCTIONS OF PROCESSES WITH MARKOVIAN STATES—II

By Martin Fox[1] and Herman Rubin[2]

Michigan State University and Purdue University

1. Summary. In [4] the rank of a state of a stochastic process was defined, although the notion, without the name, is originally due to Gilbert [5]. Let $I = \{1, 2, \cdots\}$. The purpose of the present paper is to prove the

THEOREM. *Let $\{Y_k\}$ have state space U_k at time k ($k = 1, 2, \cdots$ or $k = 0, \pm 1, \pm 2, \cdots$). Let N be a finite subset of the index set and, for each $n \varepsilon N$, let $V_n \subset U_n$ be a set of states of finite rank at time n. Without loss of generality, assume $(U_n \sim V_n) \cap (V_n \times I) = \emptyset$. Then, there exists a process $\{X_k\}$ such that*

(i) *$\{X_k\}$ has state space U_k at time $k \not\varepsilon N$ and state space $(U_n \sim V_n) \cup (V_n \times I)$ at time $n \varepsilon N$;*

(ii) *The states (ϵ, i) for $\epsilon \varepsilon V_n$, $i \varepsilon I$ and $n \varepsilon N$ are Markovian; and*

(iii) *$Y_k = F_k(X_k)$ where $F_k(\delta) = \delta$ if $\delta \varepsilon U_k \sim V_k$ (take $V_k = \emptyset$ for $k \not\varepsilon N$) and $F_n(\epsilon, i) = \epsilon$ if $\epsilon \varepsilon V_k$.*

This theorem is a generalization of Theorem 1 of [4]. Its proof is in Section 2. Section 3 contains corollaries which are the analogues of the corollaries to Theorems 1 and 2 of [4]. These show that if, in addition to the $\epsilon \varepsilon V_n$ for $n \varepsilon N$, there are states of rank 1 (Markovian states) or 2, then $\{X_k\}$ can be constructed so as to preserve the ranks of these states.

Section 4 contains a third corollary giving conditions under which N may be infinite. In particular, under these conditions N may be the whole index set and, for each $n \varepsilon N$, we may let V_n be the set of all states of finite rank at time n. Corollary 4, also in Section 4, states that, under the conditions of Corollary 3, stationarity in $\{Y_k\}$ may be preserved in $\{X_k\}$.

Dharmadhikari [1], [2], [3] has given conditions under which $\{X_k\}$ can be constructed to be a stationary, finite Markov chain. In [2] he requires the condition, among other, that each state of $\{Y_k\}$ be of finite rank. We have weakened our conditions by not insisting on stationarity or finiteness of the state space of $\{Y_k\}$ and by imposing finiteness of rank only on some states. We have completely dropped his condition that certain cones be polyhedral. This last is the reason that we have countably many Markovian states mapping into a single state of $\{Y_k\}$ instead of finitely many.

2. Proof of the Theorem. Let the rank of ϵ at time n be $\nu_n(\epsilon)$. Let $\mathcal{S}_n$ and $\mathfrak{I}_n$ be the families of all measurable sets of sequences of states prior to and after,

Received 30 August 1968.

[1] Research partly supported by Contract No. DA-31-124-ARO-D-462 at Mathematics Research Center, U. S. Army, Madison, Wisconsin. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] Research partly supported by Project No. NR-042-216 at Michigan State University and by Project No. N00014-67-A-0226-0008 at Purdue University.

respectively, time n. Then, there exist functions ${\Phi_\lambda}^{(n)}$ on $\mathcal{S}_n \times V_n$ and ${S_\lambda}^{(n)}(\epsilon) \varepsilon \mathcal{S}_n$ $(\lambda = 1, \cdots, \nu_n(\epsilon))$ such that

$$
\begin{aligned}
&P((\cdots, Y_{n-2}, Y_{n-1}) \varepsilon S, Y_n = \epsilon, (Y_{n+1}, Y_{n+2}, \cdots) \varepsilon T) \\
(2.1) \qquad &= \sum_{\lambda=1}^{\nu_n(\epsilon)} {\Phi_\lambda}^{(n)}(S, \epsilon) P((\cdots, Y_{n-2}, Y_{n-1}) \varepsilon {S_\lambda}^{(n)}(\epsilon), Y_n = \epsilon, \\
&\qquad (Y_{n+1}, Y_{n+2}, \cdots) \varepsilon T)
\end{aligned}
$$

for all $S \varepsilon \mathcal{S}_n$, $T \varepsilon \mathfrak{I}_n$.

Let $i_1 < \cdots < i_r$ be elements of N and set $i_0 = -\infty$ and $i_{r+1} = \infty$. For $l = 1, 2, \cdots, r + 1$, let $\mathcal{R}_l$ be the family of all measurable sets of sequences of states between times i_{l-1} and i_l not inclusive. Let

$$
\begin{aligned}
&p_{i_1, \cdots, i_r}(R_1, \cdots, R_{r+1}; \epsilon_1, \cdots, \epsilon_r) \\
&\qquad = P((\cdots, Y_{i_1-2}, Y_{i_1-1}) \varepsilon R_1, Y_{i_1} = \epsilon_1, (Y_{i_1+1}, \cdots, Y_{i_2-1}) \varepsilon R_2 \\
&\qquad Y_{i_2} = \epsilon_2, \cdots, Y_{i_r} = \epsilon_r, (Y_{i_r+1}, Y_{i_r+2}, \cdots) \varepsilon R_{r+1}).
\end{aligned}
$$

For $R_l \varepsilon \mathcal{R}_l$ $(l = 1, \cdots, r + 1)$ and $\epsilon_j \varepsilon V_j$, repeated application ($r$ times) of (2.1) yields

$$
\begin{aligned}
&p_{i_1, \cdots, i_r}(R_1, \cdots, R_{r+1}; \epsilon_1, \cdots, \epsilon_r) \\
(2.2) \qquad &= \sum_{\lambda_1=1}^{\nu_{i_1}(\epsilon)} \cdots \sum_{\lambda_r=1}^{\nu_{i_r}(\epsilon)} \Phi_{\lambda_1}^{(i_1)}(R_1, \epsilon_1) \Phi_{\lambda_2}^{(i_2)}({R_2}^{(\lambda_1)}, \epsilon_2) \\
&\qquad \cdots \Phi_{\lambda_r}^{(i_r)}({R_r}^{(\lambda_{r-1})}, \epsilon_r) \Psi_{\lambda_r}^{(i_r)}(R_{r+1}, \epsilon_r)
\end{aligned}
$$

where $(\cdots, Y_{i_l-2}, Y_{i_l-1}) \varepsilon {R_l}^{(\lambda_{l-1})}$ if, and only if, $(\cdots, Y_{i_{l-1}-2}, Y_{i_{l-1}-1}) \varepsilon S_{\lambda_{l-1}}^{(i_{l-1})}, Y_{i_{l-1}} \varepsilon \epsilon_{i_{l-1}}, (Y_{i_{l-1}+1}, \cdots, Y_{i_l-1}) \varepsilon R_l$ and ${\Psi_\lambda}^{(j)}(T, \epsilon) = p_j({S_\lambda}^{(j)}, T; \epsilon)$.

Let $\Xi_{\alpha\beta}^{(l)}(R_l) = {\Phi_\alpha}^{(i_l)}({R_l}^{(\beta)}, \epsilon_l)$ and let $\Xi^{(l)}(R_l)$ be the matrix whose element in the αth row and βth column is $\Xi_{\alpha\beta}^{(l)}(R_l)$. Similarly define column vectors $\Phi(R)$ and $\Psi(R)$. Note that we have deleted the superscripts on Φ and Ψ. Since Φ only occurs for time i_1 and Ψ only for i_r, this will cause no confusion. In this notation the dependence of $\Phi(R_1)$ on ϵ_1, of $\Xi^{(l)}$ on ϵ_{l-1} and ϵ_l and $\Phi(R_{r+1})$ on ϵ_r has been suppressed. This will cause no confusion since throughout this argument $\epsilon_1, \cdots, \epsilon_r$ are fixed. Then (2.2) may be written as

$$
\begin{aligned}
(2.3) \quad &p_{i_1, \cdots, i_r}(R_1, \cdots, R_{r+1}; \epsilon_1, \cdots, \epsilon_r) \\
&\qquad = \Phi(R_1)' \Xi^{(2)}(R_2) \cdots \Xi^{(r)}(R_r) \Psi(R_{r+1}).
\end{aligned}
$$

For $l = 1, \cdots, r + 1$, let $\mathfrak{X}_l = U_{i_{l-1}+1} \times \cdots \times U_{i_l-1}$ (recall that $i_1 = -\infty$ and $i_{r+1} = \infty$.

Let $Q_l(R) = p_{i_1, \cdots, i_r}(\mathfrak{X}_1, \cdots, \mathfrak{X}_{l-1}, R, \mathfrak{X}_{l+1}, \cdots, \mathfrak{X}_{r+1}; \epsilon_1, \cdots, \epsilon_r)$. By the proof used in the analogous step in [4] we see that $\Phi \ll Q_1$, $\Psi \ll Q_{r+1}$ and $\Xi^{(l)} \ll Q_l$ for $l = 2, \cdots, r$. Since these set functions on the $\mathcal{R}_l$ are matrix valued measures, we have the existence of $\varphi = d\Phi/dQ_1$, of $\psi = d\Psi/dQ_{r+1}$ and

$\xi^{(l)} = d\Xi^{(l)}/dQ_l$ for $l = 2, \cdots, r$. Hence, (2.3) becomes

$$\begin{aligned} &p_{i_1,\cdots,i_r}(R_1, \cdots, R_{r+1}; \epsilon_1, \cdots, \epsilon_r) \\ (2.4) \qquad &= \int \cdots \int_{R_1\times\cdots\times R_{r+1}} \varphi(x_1)'\xi^{(2)}(x_2) \cdots \xi^{(r)}(x_r)\psi(x_{r+1}) \\ &\qquad \cdot dQ_1(x_1) \cdots dQ_{r+1}(x_{r+1}). \end{aligned}$$

In the next part of this argument any norm will do. We adopt the convention $0/0 = 0$ since this will yield the proper results in the computations which follow. Let $\varphi^*(x) = \varphi(x)/\|\varphi(x)\|$, $\psi^*(x) = \psi(x)/\|\psi(x)\|$ and $\xi^{(l)*}(x) = \xi^{(l)}(x)/\|\xi^{(l)}(x)\|$ for $l = 2, \cdots, r$. Since the integrand in (2.4) is nonnegative a.e. $(Q_1 \times \cdots \times Q_{r+1})$ we obtain

$$\begin{aligned} &p_{i_1,\cdots,i_r}(R_1, \cdots, R_{r+1}; \epsilon_1, \cdots, \epsilon_r) \\ (2.5) \qquad &= \int \cdots \int_{R_1\cdots R_{r+1}} \|\varphi(x_1)\| \, \|\xi^{(2)}(x_2)\| \cdots \|\xi^{(r)}(x_2)\| \, \|\psi(x_{r+1})\| \\ &\qquad \cdot |\varphi^*(x_1)'\xi^{(2)*}(x_2) \cdots \xi^{(r)*}(x_r)\psi^*(x_{r+1})| \, dQ_1(x_1) \cdots dQ_{r+1}(x_r). \end{aligned}$$

We now consider the matrix function $|W_1'W_2 \cdots W_{r+1}|$ where W_1 and W_{r+1} are vectors. For W_i of norm 1 this is a nonnegative continuous function on a compact product set. By a theorem of Rubin [6] we may write

$$(2.6) \qquad |W_1'W_2 \cdots W_{r+1}| = \sum_{j=1}^{\infty} \prod_{k=1}^{r+1} \alpha_{kj}(W_k)$$

where the functions on the right hand side are nonnegative and continuous. Set $\beta_j(x) = \|\varphi(x)\| \, \alpha_{1j}(\varphi^*(x))$; $\gamma_j^{(l)}(x) = \|\xi^{(l)}(x)\| \, \alpha_{lj}(\xi^{(l)*}(x))$ for $l = 2, \cdots, r$ and $\delta_j(x) = \|\psi(x)\| \, \alpha_{r+1,j}(\psi^*(x))$ and apply (2.6). Then (2.5) yields

$$\begin{aligned} &p_{i_1,\cdots,i_r}(R_1, \cdots, R_{r+1}; \epsilon_1, \cdots, \epsilon_r) \\ (2.7) \qquad &= \sum_{j=1}^{n} \int \cdots \int_{R_1\times\cdots\times R_{r+1}} \beta_j(x_1)\gamma_j^{(2)}(x_2) \cdots \gamma_j^{(r)}(x_r)\delta_j(x_{r+1}) \\ &\qquad \cdot dQ_1(x_1) \cdots dQ_{r+1}(x_{r+1}) \end{aligned}$$

where all factors of the integrands are nonnegative. For each $j = 1, 2, \cdots$ these functions can be normalized so that

$$\int_{\mathfrak{X}_{r+1}\sim\mathfrak{Y}_{r+1}} \delta_j \, dQ_{r+1} = 1 \qquad \text{and} \qquad \int_{\mathfrak{X}_l\sim\mathfrak{Y}_l} \gamma_j^{(l)} \, dQ_l = 1$$

for $l = 2, \cdots, r$ where $\mathfrak{Y}_l$ is the set of elements of $U_{i_{l-1}+1} \times \cdots \times U_{i_l-1}$ which visit some $\epsilon \, \varepsilon \, V_n$ for some n.

We now define a process $\{Z_k\}$. For $n \not\varepsilon N$ let the state space of $\{Z_k\}$ be U_n. For $n \, \varepsilon \, N$ the state space will be clear by the following. For $l = 1, \cdots, r + 1$ assume no element of R_l visits a state $\epsilon \, \varepsilon \, V_n$ for $n \, \varepsilon \, N$. Let

$$(2.8) \quad P((\cdots, Z_{i_1-2}, Z_{i_1-1}) \, \varepsilon \, R_1, Z_{i_1} = (\epsilon_1, i_1, \cdots, i_r, j)) = \int_{R_1} \beta_j \, dQ_1 .$$

For $l = 2, \cdots, r$ let

$$P((Z_{i_{l-1}+1}, \cdots, Z_{i_l-1}) \varepsilon R_l, Z_{i_l} = (\epsilon_l, i_1, \cdots, i_r, j) | Z_{i_{l-1}} = (\epsilon_{l-1}, i_1, \cdots, i_r, k)) \tag{2.9}$$
$$= \int_{R_l} \gamma_j^{(l)} \, dQ_l \quad \text{if} \quad k = j$$
$$= 0 \quad \text{if} \quad k \neq j.$$

Let

$$P((Z_{i_r+1}, Z_{i_r+2}, \cdots) \varepsilon R_{r+1} \mid Z_{i_r} = (\epsilon_r, i_1, \cdots, i_r, j)) = \int_{R_{r+1}} \delta_j \, dQ_{r+1}. \tag{2.10}$$

Finally, for R which consists only of sequences never visiting an $\epsilon \varepsilon V_k$ for any $n \varepsilon N$, let

$$P((\cdots, Z_{-1}, Z_0, Z_1, \cdots) \varepsilon R) = P((\cdots, Y_{-1}, Y_0, Y_1, \cdots) \varepsilon R). \tag{2.11}$$

Let $X_k = f_k(Z_k)$ where the f_k are one-to-one functions chosen so that the state space of $\{X_k\}$ will be as in (i) of the theorem.

Consistency of the probabilities in (2.8) through (2.11) follows from the normalization. Also (2.8) through (2.11) imply Markovianness of the elements of $V_n \times I$ for $n \varepsilon N$. By (2.7) we see that $\{Y_k\}$ and $\{f_k(X_k)\}$ have the same distribution.

3. Corollaries concerning ranks 1 and 2. Corollaries 1 and 2 are analogous to the corollaries to Theorems 1 and 2 of [4], respectively. Their proofs are so similar that only outlines are given here.

COROLLARY 1. *Under the conditions of the theorem, there exists $\{X_k\}$ satisfying the conclusions but such that, for all n, every state $\delta \not\varepsilon V_n$ which is Markovian at time n in $\{Y_k\}$ is Markovian at time n in $\{X_k\}$.*

We restrict the construction in Section 2 to the case in which no sequence in $R_2, \cdots, R_r$ visits a Markovian state and if a sequence in $R_1(R_{r+1})$ visits a Markovian state all sequences which differ from it only at earlier (later) times are in $R_1(R_{r+1})$. We can then complete the construction by piecing together the various parts using the Markovian property.

COROLLARY 2. *Under the conditions of the theorem it is possible to construct $\{X_k\}$ satisfying the conclusions but such that, for all n, every state $\delta \not\varepsilon V_n$ which is of rank at most 2 has its rank preserved in $\{X_k\}$.*

We first apply Theorem 2 of [4] to split all states of rank 2 which are not elements of V_n. Then we apply Corollary 1. Let $\{Z_k\}$ be the resulting process and set $X_k = g_k(Z_k)$ where g_k reconstitutes the states of rank 2 not elements of V_k and is the identity map otherwise. The ranks of other states are preserved while the ranks in $\{X_k\}$ of these states of rank 2 in $\{Y_k\}$ is 2.

4. Countably many times. Corollary 3 gives conditions under which N may be any subset of the index set. Corollary 4 covers the stationary case. Only outlines of the proofs of these corollaries are given.

COROLLARY 3. *Under the conditions of the theorem, if the set of states of ranks* 1 *and* 2 *in* $\{Y_k\}$ *is recurrent, then* N *need not be a finite set.*

The proof is the proof of the theorem with the following modifications:

(i) Apply Theorem 2 of [4] to all states of rank 2. This can be accomplished since the set of all such states is countable. This yields a process $\{Z_k\}$ in which the states of rank not 2 in $\{Y_k\}$ are preserved with their ranks. Furthermore the set of Markovian states in $\{Z_k\}$ is recurrent. For the remainder of the construction use $\{Z_k\}$ in place of $\{Y_k\}$;

(ii) Let $R_1 (R_{r+1})$ contain only sequences with a Markovian state at some fixed time $m < i_1$ $(m > i_r)$ and no Markovian states between m and i_1 (i_r and m) with the further restriction that if we modify on elements of R_1 (R_{r+1}) before (after) time m_1 then the result is also an element of R_1 (R_{r+1}).

(iii) For $l = 2, \cdots, r$, let R_l consist only of sequences not visiting a Markovian state; and

(iv) Apply Corollary 1 to Markovian states of $\{Z_k\}$.

Rank 2 states of $\{Y_k\}$ which are not elements of any V_n may be reconstituted as in Corollary 2.

COROLLARY 4. *Under the conditions of Corollary* 3, *if* $\{Y_k\}$ *is stationary and the* V_n *are identical, then a stationary process* $\{X_k\}$ *can be constructed satisfying the conclusions of the theorem.*

In this case clearly the probabilities given by (2.8) to (2.10) are stationary. In particular, from (2.8) and the choice of R_1 in Corollary 3, any sequence beginning and ending in states of the form (ϵ, j) for $\epsilon \varepsilon V_n$ has stationary probability. Sequences containing no such states have the stationary probabilities given in $\{Y_k\}$. In the case $k = 1, 2, \cdots$ we guarantee stationarity of sequences starting at time 1 by extending the process in a stationary manner to $k = 0, \pm 1, \pm 2, \cdots$ and then using the distribution for $(X_1, \cdots, X_n)$ which is so obtained.

REFERENCES

[1] DHARMADHIKARI, S. W. (1963). Functions of finite Markov chains. *Ann. Math. Statist.* **34** 1022–1032.

[2] 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.

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

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

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

[6] RUBIN, HERMAN (1968). Representations of nonnegative continuous functions on product spaces. *Bull. Amer. Math. Soc.* **74** 332–333.