Functions of a Markov Process that are Markovian

M. ROSENBLATT*

Summary. In this paper we are primarily concerned with discrete time parameter Markov processes $\{X(n)\}$, $n = 0, 1, 2, \dots$, with stationary transition mechanism. The processes $\{Y(n)\} = \{f(X(n))\}$ generated by a given many-one function f and the processes $\{X(n)\}$ with a fixed stationary transition mechanism are constructed. The processes $\{Y(n)\}$ are in a one-to-one correspondence with the possible initial distributions of $\{X(n)\}$. The object of the paper is to determine conditions under which $\{Y(n)\}$ is Markovian, whatever the initial distribution of $\{X(n)\}$. Necessary and sufficient conditions for the new processes $\{Y(n)\}$ to be Markovian are obtained under the assumption that the family of measures corresponding to the fixed transition mechanism (of $\{X(n)\}$) is dominated [4]. The conditions are expressed, of course, in terms of the function $f(\cdot)$ and the transition mechanism. Generally the processes $\{Y(n)\}$ do not have stationary transition mechanism. The conditions simplify in the case of a continuous time parameter Markov chain. Some of the discussions may at times have the flavor of those used in considering the concept of sufficiency [4]. Some aspects of the problem discussed in the paper are touched on in [2].

Basic Concepts. Let Λ be a space of points x and $\mathfrak{B}$ a Borel field of Λ sets. The sets consisting of single points $\{x\}$, $x \in \Lambda$, are assumed to be in $\mathfrak{B}$. A function $P(\cdot) = P_0(\cdot)$ of $B \in \mathfrak{B}$ and a family of functions $P_n(\cdot, \cdot)$ ($n = 0, 1, \dots$) of $x \in \Lambda$, $B \in \mathfrak{B}$, satisfying the following recursive conditions are assumed given:

- (i) $P(B) = P_0(B)$ is a probability measure in $\mathfrak{B}$.
- (ii) $P_0(x, B)$ for almost any fixed x (P_0 measure) is a probability measure in $\mathfrak{B}$ and for every $B \in \mathfrak{B}$ is a measurable function with respect to $\mathfrak{B}$ in x .
- (iii) Let $P_n(\cdot)$, $P_n(\cdot, \cdot)$ be given where $P_n(B)$ is probability measure in $\mathfrak{B}$ and where $P_n(x, B)$ is a probability measure in $\mathfrak{B}$ for almost every fixed x (P_n measure) and for every $B \in \mathfrak{B}$ is a measurable function with respect to $\mathfrak{B}$ in x . Then

$$(1) \quad P_{n+1}(B) = \int_{\Lambda} P_n(dx) P_n(x, B);$$

*This research was supported by the Office of Naval Research under Contract Onr 908 (10). Reproduction in whole or in part is permitted for any purpose of the United States Government.

and $P_{n+1}(x, B)$ is a probability measure in $\mathfrak{B}$ for almost every x (P_{n+1} measure) and for every $B \in \mathfrak{B}$ is a measurable function with respect to $\mathfrak{B}$ in x .

Define a probability measure in the space Ω of sequences $w = (x_0, x_1, \dots)$, $x_i \in \Lambda$, as follows. Let $X_n(w)$, $n = 0, 1, 2, \dots$, be the n^{th} coordinate function, that is, $X_n(w) = x_n$ if w is the point $(x_0, x_1, \dots)$. Set

$$(2) \quad P(X_0(w) \in B_0) = P_0(B_0), \quad B_0 \in \mathfrak{B},$$

and, more generally,

$$(3) \quad \begin{aligned} &P(X_0(w) \in B_0, \dots, X_n(w) \in B_n) \\ &= \int_{B_0} P_0(dx_0) \int_{B_1} P_0(x_0, dx_1) \cdots \int_{B_{n-1}} P_{n-1}(x_{n-1}, B_n), \quad B_i \in \mathfrak{B}. \end{aligned}$$

The sequence of functions $\{X_n(w)\}$ is called the Markov process with initial distribution $P_0(\cdot)$ and stochastic transition functions (or, more loosely, transition probabilities) $P_n(\cdot, \cdot)$, $n = 0, 1, \dots$.

Notice that a Markov process, as introduced above, may have a non-stationary stochastic transition function (not independent of the time n) and our interest is in functions $Y(n) = f(X(n))$ of Markov processes $X(n)$ with stationary transition functions. Since we may be led to processes with non-stationary transition mechanism, the more general definition above has been given.

As already remarked, the Markov process is said to have a *stationary transition function* if $P_n(x, B) = P(x, B)$ (independent of n). Consider the family of Markov processes with stationary transition function generated by the fixed transition function $P(x, B)$ and any initial distribution $P(B)$. Since we speak about any initial distribution $P(\cdot)$, a more stringent condition has to be imposed on $P(x, B)$. $P(x, B)$ must be assumed to be a probability measure in $\mathfrak{B}$ for every fixed $x \in \Lambda$. The family of Markov processes introduced will be referred to as the *family of Markov processes generated by $P(x, B)$* .

Let Λ' be a space of points y with a Borel field $\mathfrak{B}'$ of Λ' sets. Let $f(\cdot)$ be a many-one function from Λ onto Λ' measurable (with respect to $\mathfrak{B}$). The field Λ' is also assumed to contain the sets consisting of single points $\{y\}$, $y \in \Lambda'$. Given a function $f(\cdot)$ and a stationary transition function $P(x, B)$, we are interested in finding conditions under which all the processes $Y(n) = f(X(n))$, derived from $f(\cdot)$ and the family of processes generated by $P(x, B)$, are Markovian. Notice that the sub-Borel field of $\mathfrak{B}$ induced by f ,

$$f^{-1}f\mathfrak{B} = \{f^{-1}fB \mid B \in \mathfrak{B}\} = f^{-1}\mathfrak{B}',$$

is of importance. The results may be phrased in terms of $f^{-1}f\mathfrak{B}$ or $\mathfrak{B}'$, depending on which is more convenient.

We shall want to obtain stochastic transition functions (in the sense given above) for the derived processes $Y(n)$. This will mean that the conditional probabilities computed have to be consistent in such a way that conditional probability measures (the stochastic transition functions) can be manufactured

from them. This is possible if there is a one-to-one mapping of the range space Λ' (of f) into the real line taking the sets $\mathcal{B}'$ into Borel sets (of the real line) with the inverse images of Borel sets of the family $\mathcal{B}'$ (see [1, 3, 5]). This assumption is made throughout this paper. This includes the case where Λ' is a Borel set of countably dimensional Euclidean space with $\mathcal{B}'$ the Borel field of Borel subsets of Λ' .

A Dominated Stochastic Transition Function. Consider the family of Markov processes generated by the stochastic transition function $P(x, B)$, $x \in \Lambda$, $B \in \mathcal{B}$. The transition function is a family of probability measures indexed by the points $x \in \Lambda$. The family of measures $P(x, f^{-1}B')$, $x \in \Lambda$, $B' \in \mathcal{B}'$, on $f^{-1}\mathcal{B}' = f^{-1}f\mathcal{B}'$ is induced by the transition function and the function f . Necessary and sufficient conditions for the processes $Y(n) = f(X(n))$, derived from $f(\cdot)$ and the Markov processes $X(n)$ generated by $P(x, B)$, to be Markovian will be obtained under the assumption that the measures $P(x, f^{-1}B')$ are dominated by a sigma-finite measure (see [4] for a discussion of this terminology).

If the family of measures $P(x, f^{-1}B')$ is dominated, it is well known that one can find an equivalent countable subset of measures $P(x_i, f^{-1}B')$, $x_i \in \Lambda$, $i = 1, 2, \dots$ [4]. The dominating (in fact, equivalent) measure can be taken, and shall be taken, as the probability measure

$$\mu(f^{-1}B') = \sum_{i=1}^{\infty} \alpha_i P(x_i, f^{-1}B'), \quad \alpha_i > 0, \quad i = 1, 2, \dots$$

Call a set $S \in f^{-1}\mathcal{B}'$ a single entry set if $\mu(S) > 0$ and there is a point $y \in \Lambda'$ such that $P(x, S) = 0$ if $x \notin f^{-1}\{y\}$. Note that since $\mu(S) > 0$ (μ equivalent to the family of measures $P(x, f^{-1}B')$) there are points $x \in f^{-1}\{y\}$ such that $P(x, S) > 0$. S is a maximal single entry set if it is a single entry set and is not contained in a single entry set of larger μ measure. We shall identify a single entry set with all sets differing from it by at most a set of μ measure zero. The point $y \in \mathcal{B}'$ from which one can enter a single entry set in one step with positive probability is called a single entry point.

Lemma 1. *If the measures $P(x, f^{-1}B')$, $x \in \Lambda$, $B' \in \mathcal{B}'$, are dominated, there are at most a countable number of disjoint maximal single entry sets and corresponding single entry points.*

This follows immediately from the fact that μ is a finite measure. Lemma 1 ensures the existence of a unique (up to sets of μ measure zero) countable maximal collection of maximal single entry sets. Let the single entry sets of this maximal collection be $S_1, S_2, \dots$. Let $M = \Lambda - \bigcup S_i$.

Lemma 2. *Under the hypothesis of Lemma 1, if $\mu(M) = 0$ then the processes $Y(n) = f(X(n))$ are Markov processes. The space Λ' is then the countable set of single entry points.*

Let y_i be the single entry point leading into the maximal single entry set

S_i . Since $\mu(M) = 0$ any point $x \in \Lambda$ must fall in some set $f^{-1}\{y_i\}$, $j = 1, 2, \dots$. The space Λ' is therefore the countable set of points y_i . Every set $f^{-1}\{y_i\}$ for which $\mu(f^{-1}\{y_i\}) > 0$ is a single entry set. In verifying that the processes $Y(n)$ are Markovian it is enough to look at the probabilities

$$(4) \quad P(Y(j) = y_{m(j)}; j = 0, \dots, n) = \int_{f^{-1}\{y_{m(0)}\}} P(dx_0) \\ \cdot \int_{f^{-1}\{y_{m(1)}\}} P(x_0, dx_1) \cdots \int_{f^{-1}\{y_{m(n-1)}\}} P(x_{n-1}, f^{-1}\{y_{m(n)}\})$$

where $P(\cdot)$ is the initial probability distribution. Let

$$(5) \quad P^{(1)}(x, B) = P(x, B), \\ P^{(n+1)}(x, B) = \int_{\Lambda} P^{(n)}(x, dx') P(x', B), \quad n = 1, 2, \dots$$

It is apparent that under the above assumptions

$$(6) \quad P(Y(j) = y_{m(j)}; j = 0, \dots, n) = P(Y(n) = y_{m(n)}) \\ = \int_{\Lambda} P(dx_0) P^{(n)}(x_0, f^{-1}\{y_{m(n)}\})$$

if (4) is positive. Notice that we need only consider points y_i for which $P(f^{-1}\{y_i\}) > 0$. Let $Q(\cdot)$ be the initial distribution (on $\mathcal{B}'$) and $Q_n(\cdot, \cdot)$ the stochastic transition functions (of $y \in \Lambda'$ and $B' \in \mathcal{B}'$) of $Y(n)$ that are to be constructed. Clearly $Q(\{y_i\}) = P(f^{-1}\{y_i\})$. If $P_n(f^{-1}\{y_i\}) > 0$ let

$$(7) \quad Q_n(y_i, \{y_k\}) = P(Y(n+1) = y_k | Y(n) = y_i) \\ = \begin{cases} P_{n+1}(f^{-1}\{y_k\})/P_n(f^{-1}\{y_i\}) & \text{if } y_i \text{ is the single entry point for the set of } f^{-1}\{y_k\} \\ 0 & \text{otherwise.} \end{cases}$$

It is then clear that $Y(n)$ is a Markov process with initial distribution $Q(\cdot)$ and transition functions $Q_n(\cdot, \cdot)$ since (3) is satisfied with $Q(\cdot)$, $Q_n(\cdot, \cdot)$, Λ' , $\mathcal{B}'$ in place of $P(\cdot)$, $P_n(\cdot, \cdot)$, Λ , $\mathcal{B}$. Since this is true for any initial distribution $P(\cdot)$ the lemma has been proved.

The Markovian property of $Y(n)$ was trivially verified when $\mu(M) = 0$. The interesting case is clearly that in which $\mu(M) > 0$. There are then single entry sets if and only if $\mu(M) < 1$ (assuming μ to be a probability measure). If S is a single entry set, we may be able to enter S from a point $x \in M$ outside the single entry sets in a finite number of steps with positive probability. If this is possible, there is an integer n and there are points $y_{m(1)}, \dots, y_{m(n)} \in \mathcal{B}'$ with the following properties. The sets $f^{-1}\{y_{m(j)}\}$, $j = 2, \dots, n$, are single entry sets while $f^{-1}\{y_{m(1)}\}$ is not. Further, the point $y_{m(j)}$ is the single entry point for $f^{-1}\{y_{m(j+1)}\}$, $j = 1, \dots, n-1$, and the point $y_{m(n)}$ is the single entry point for S . Given any single entry set S into which entry is possible from points $x \in M$ with positive probability

(that is, there is an integer $n' \geq 1$ such that $P^{(n')}(x, S)$ for some $x \in M$), the integer n and the points $y_{m(i)}$ as specified above are uniquely determined. We shall then say that a thread of length n consisting of the points $y_{m(1)}, \dots, y_{m(n)} \in \mathcal{B}'$ enters S .

Theorem 1. Let the measures $P(x, f^{-1}B')$, $x \in \Lambda$, $B' \in \mathcal{B}'$, be dominated. A necessary and sufficient condition for the processes $Y(n) = f(X(n))$, derived from $f(\cdot)$ and the family of Markov processes $X(n)$ generated by $P(x, B)$, to be Markovian is given as follows:

(a) There is a function $R(x, B)$ of $x \in M$, $B \in f^{-1}\mathcal{B}'$ that is measurable in x with respect to $f^{-1}\mathcal{B}'$ for each $B \in f^{-1}\mathcal{B}'$ and satisfies

$$(8) \quad \int_A P(x_0, dx_1)P(x_1, B) = \int_A P(x_0, dx_1)R(x_1, B)$$

for every $x_0 \in \Lambda$ and $A, B \in f^{-1}\mathcal{B}'$ where $A \subset M$.

(b) If S is a maximal single entry set with a thread of length n entering it, there is a function ${}_nR(x, B)$ of $x \in S$, $B \in f^{-1}\mathcal{B}'$ that is measurable in x with respect to $f^{-1}\mathcal{B}'$ for each $B \in f^{-1}\mathcal{B}'$ and satisfies

$$(9) \quad \int_A P^{(n+1)}(x_0, dx_{n+1})P(x_{n+1}, B) = \int_A P^{(n+1)}(x_0, dx_{n+1}){}_nR(x_{n+1}, B)$$

for all $x_0 \in \Lambda$, $A, B \in f^{-1}\mathcal{B}'$, $A \subset S$.

First consider the necessity of the conditions. If $\mu(M) = 0$ the conditions are vacuously satisfied. Assume therefore that $\mu(M) > 0$ and let the necessity of condition (a) be our concern. Take any two points $x, x' \in \Lambda$ and the corresponding measures $P(x, B)$, $P(x', B)$ on the sets $B \in f^{-1}\mathcal{B}'$, $B \subset M$. Let $M(x, x') \subset M$ be the set on which these two measures are mutually absolutely continuous in M . Set

$$(10) \quad P(x, A, B) = \int_A P(x, dx')P(x', B), \quad A, B \in f^{-1}\mathcal{B}', \quad A \subset M,$$

and let $P(x, df^{-1}y, B)/P(x, df^{-1}y)$ denote the Radon-Nikodym derivative of (10) with respect to $P(x, A)$ on the Borel field $f^{-1}\mathcal{B}'$ restricted to subsets of M . We want to show that

$$(11) \quad \frac{P(x, df^{-1}y, B)}{P(x, df^{-1}y)} = \frac{P(x', df^{-1}y, B)}{P(x', df^{-1}y)}$$

almost everywhere with respect to the dominating μ measure on $M(x, x')$ for each $B \in f^{-1}\mathcal{B}'$. Let us first consider points $x, x' \in \Lambda$ with $x' \notin f^{-1}f\{x\}$. Let $P(\cdot)$ be the initial distribution with mass α , $0 < \alpha < 1$, at x and mass $\beta = 1 - \alpha$ at x' . Then

$$(12) \quad \begin{aligned} P(Y(2) \in fB \mid Y(1) = y) &= \frac{\alpha P(x, df^{-1}y, B) + \beta P(x', df^{-1}y, B)}{\alpha P(x, df^{-1}y) + \beta P(x', df^{-1}y)} \\ &= \frac{\alpha g(x, f^{-1}y, B) + \beta g(x', f^{-1}y, B)}{\alpha g(x, f^{-1}y) + \beta g(x', f^{-1}y)} \end{aligned}$$

where $g(x, f^{-1}y, B)$, $g(x', f^{-1}y)$ are the Randon-Nikodym derivatives of $P(x, A, B)$, $P(x, A)$ respectively with respect to $\mu(A)$, $A, B \in f^{-1}\mathfrak{B}$, $A \subset M$. Since the $Y(n)$ process is Markovian, whatever the initial distribution of $X(n)$, we must, in particular, have

$$(13) \quad \begin{aligned} P(Y(2) \in fB, Y(1) \in fA \mid Y(0) = fx) &= \int_A P(x, dx')P(x', B) \\ &= \int_{fA} P(x, df^{-1}y)P(Y(2) \in fB \mid Y(1) = y) \end{aligned}$$

for $A, B \in f^{-1}\mathfrak{B}'$, $A \subset M(x, x')$. But this cannot happen for every α unless (11) holds almost everywhere with respect to μ measure on $M(x, x')$. Thus (11) is valid for pairs of points $x, x' \in \Lambda$ with $x' \notin f^{-1}f\{x\}$. Now consider a pair $x, x' \in \Lambda$ with $x' \in f^{-1}f\{x\}$ and the set $M(x, x')$ as defined above. Assume that $\mu(M(x, x')) > 0$ since otherwise (11) is obviously valid. Consider all points $x'' \in \Lambda$, $x'' \notin f^{-1}f\{x\}$ and the corresponding sets $M(x'', x)$. There is an integer $n > 0$ such that for some such x'' , $\mu(M(x'', x)M(x', x)) > (\frac{1}{2})^n$. Choose the smallest such integer n and a corresponding point $x'' \notin f^{-1}f\{x\}$ and call them n_1, x_1 . Assume that one has points $x_i \notin f^{-1}f\{x\}$ and corresponding integers $n_i, i = 1, \dots, m, n_1 \leq n_2 \leq \dots \leq n_m$, with

$$(14) \quad \mu\left(M(x, x')M(x, x_i) - \bigcup_{i=1}^{i-1} M(x, x_i)\right) > \frac{1}{2^{n_i}}, \quad i = 1, \dots, m.$$

If

$$(15) \quad \mu\left(M(x, x') - \bigcup_{i=1}^m M(x, x_i)\right) > 0$$

there is an integer $n \geq n_m$ and a corresponding point $x'' \notin f^{-1}f\{x\}$ with

$$(16) \quad \mu\left(M(x, x')M(x, x'') - \bigcup_{i=1}^m M(x, x_i)\right) > \frac{1}{2^n}.$$

Choose the smallest such n and a corresponding point x'' and call them n_{m+1}, x_{m+1} . If this construction stops in a finite number of steps,

$$(17) \quad \mu\left(M(x, x') - \bigcup_{i=1}^m M(x, x_i)\right) = 0$$

where m is the number of steps. Otherwise we have an infinite sequence of x_i 's with corresponding $n_i \rightarrow \infty$ as $j \rightarrow \infty$. Moreover,

$$(18) \quad \mu\left(M(x, x') - \bigcup_{i=1}^{\infty} M(x_i, x)\right) = 0$$

for otherwise we would have a contradiction. Now $P(x, df^{-1}y, B)/P(x, df^{-1}y)$, $P(x', df^{-1}y, B)/P(x', df^{-1}y)$ are both equal to $P(x_i, df^{-1}y, B)/P(x_i, df^{-1}y)$, $B \in f^{-1}\mathfrak{B}'$, on $M(x, x')M(x_i, x)$ and hence equal to each other there with respect

to μ measure. Because of (17) or (18) they are therefore equal almost everywhere on $M(x, x')$ with respect to μ measure. Assume that we have taken the equivalent measure

$$\mu(B) = \sum_{i=1}^{\infty} \alpha_i P(x'_i, B), \quad \alpha_i > 0, \quad \sum \alpha_i = 1, \quad B \in f^{-1}\mathfrak{B}',$$

where $\{P(x'_i, B)\}$ is an equivalent countable subset of the collection of measures $\{P(x, B), x \in \Lambda\}$, $B \in f^{-1}\mathfrak{B}'$. Now consider the Radon-Nikodym derivative of

$$\int_A \sum \alpha_i P(x'_i, dx) P(x, B)$$

with respect to $\mu(A)$, $A, B \in f^{-1}\mathfrak{B}'$, and call it $R(x, B)$. Now it is clear that $R(x, B)$ is measurable in x with respect to $f^{-1}\mathfrak{B}'$ for each $B \in f^{-1}\mathfrak{B}'$. Further,

$$(19) \quad \frac{P(x_1, df^{-1}y, B)}{P(x_1, df^{-1}y)} = R(x_2, B), \quad x_2 \in f^{-1}\{y\},$$

almost everywhere with respect to $P(x, A)$ measure, $A \in f^{-1}\mathfrak{B}'$, $A \subset M$. But this is condition (a).

An argument paralleling the one given above for the necessity of condition (a) suffices for the necessity of condition (b) if S is a maximal single entry set with a thread of length n entering it. The function ${}_nR(x, B)$ can be taken as the Radon-Nikodym derivative of

$$(20) \quad \int_A \sum_i \alpha_i P^{(n+1)}(x'_i, dx) P(x, B)$$

with respect to $\sum_i \alpha_i P^{(n+1)}(x'_i, A)$, $A, B \in f^{-1}\mathfrak{B}'$, $A \subset S$.

We now consider the sufficiency of conditions (a), (b). If $\mu(M) = 0$ conditions (a), (b) are vacuously satisfied and Lemma 2 tells us that the new process $Y(n)$ is Markovian. Let us therefore consider $\mu(M) > 0$. Let $P(\cdot)$ be the initial distribution of $X(n)$. The initial distribution of $Y(n)$ is then $Q(B') = P(f^{-1}B')$, $B' \in \mathfrak{B}'$. Now consider how to construct the stochastic transition functions $Q_n(\cdot, \cdot)$ of the $Y(n)$ process. If $y \in fM$ let $Q_n(y, B') = R(x, f^{-1}B')$, $n \geq 1$, where $x \in f^{-1}\{y\}$ and B' is any set in $\mathfrak{B}'$. Set $Q_0(y, B')$ equal to the Radon-Nikodym derivative of $P(Y(0) \in A', Y(1) \in B')$ with respect to $P(Y(0) \in A')$, $A', B' \in \mathfrak{B}'$, $y \in \Lambda'$. If $y \notin fM$, $f^{-1}\{y\}$ is contained in some single entry set S . If there is no thread of length $m < n$ entering S let

$$(21) \quad Q_n(y, B') = \int_{f^{-1}\{y\}} P_n(dx) P(x, f^{-1}B') / P_n(f^{-1}\{y\}), \quad B' \in \mathfrak{B}',$$

if $P_n(f^{-1}\{y\}) > 0$. If there is a thread of length $m < n$ entering S , let

$$(22) \quad Q_n(y, B') = {}_mR(x, f^{-1}B'), \quad B' \in \mathfrak{B}',$$

where $x \in f^{-1}\{y\}$. The $Q_n(\cdot, \cdot)$'s have been set equal to conditional probabilities. Since we have assumed that there is a one-to-one mapping of the range space

Λ' into the real line taking the sets of $\mathcal{B}'$ into Borel sets of the real line and conversely, the $Q_n(\cdot, \cdot)$ can be chosen so that they are conditional probability measures. The process $Y(n)$ must be shown to be Markovian with the specification of the stochastic transition functions $Q_n(\cdot, \cdot)$ as given above, namely

$$(23) \quad P(Y(0) \in B'_0, \dots, Y(n) \in B'_n) = \int_{B'_0} Q(dy_0) \\ \cdot \int_{B'_1} Q_0(y_0, dy_1) \cdots \int_{B'_{n-1}} Q_{n-1}(y_{n-1}, B'_n), \quad B'_i \in \mathcal{B}'.$$

It is enough to show that relation (23) holds for sets of the form

$$(24) \quad (Y(j) \in fB_j; j = 0, \dots, n), \quad B_j \in \mathcal{B},$$

where for each j separately either $B_j \subset M$ or else B_j is a single entry set, since every set $(Y(j) \in fB_j; j = 0, \dots, n)$, $B_j \in \mathcal{B}$, is the sum of a countable number of sets of the form (24). However, it is apparent from the construction of the stochastic transition functions $Q_n(\cdot, \cdot)$ that (23) holds for sets of the form (24). The proof of the sufficiency is complete.

Probably the most interesting context covered in Theorem 1 is that in which there are no single entry sets. Now the original process $X(n)$ has stationary transition mechanism. Suppose we insist that the new process $Y(n)$ (assuming it to be Markovian) also have stationary transition mechanism (which might conceivably depend on the initial distribution $P(\cdot)$ of the $X(n)$ process). Condition (8) simplifies since stationarity of the transition mechanism implies that the Radon-Nikodym derivative of

$$(25) \quad \int_{f^{-1}A'} P(dx)P(x, f^{-1}B')$$

with respect to $P(f^{-1}A')$, $A', B' \in \mathcal{B}'$, whatever the initial distribution $P(\cdot)$, must be equal to $Q(y, B')$ almost everywhere with respect to the μ measure. Corollary 1 follows.

Corollary 1. *Assume that the hypotheses of Theorem 1 are satisfied and that there are no single entry sets. A necessary and sufficient condition for the processes $Y(n)$, derived from $f(\cdot)$ and the processes $X(n)$, to be Markovian and have stationary transition mechanism ($Q_n(\cdot, \cdot) = Q(\cdot, \cdot)$) is given by*

$$(26) \quad P(x, f^{-1}B') \quad \text{measurable in } x$$

with respect to $\mathcal{B}'$ for every $B' \in \mathcal{B}'$.

However, condition (26) of Corollary 1 is not generally valid if one allows single entry sets. The following example illustrates this. Consider a Markov chain $X(n)$ (Λ a countable set which is, for convenience, labeled by the integers) with transition probability matrix $P = (p_{ij})$. The function f collapses the states (elements of Λ) into three distinct points y_1, y_2, y_3 ($\Lambda' = \{y_1, y_2, y_3\}$)

in such a way that at least two $X(n)$ states are mapped into each y point. Let

$$(27) \quad p_{ii} = \begin{cases} c \delta_{ii} & \text{if } i, j \in f^{-1}\{y_1\}, \quad 0 < c < 1, \\ 0 & \text{if } i \notin f^{-1}\{y_1\}, \quad j \in f^{-1}\{y_1\} \end{cases}$$

while

$$(28) \quad P(i, f^{-1}\{y_\beta\}) = \sum_{i \in f^{-1}\{y_\beta\}} p_{ii} = Q(y_\alpha, \{y_\beta\})$$

for $i \in f^{-1}\{y_\alpha\}$, $\alpha, \beta = 2, 3$. The remaining p_{ij} ($i \in f^{-1}\{y_1\}$, $j \notin f^{-1}\{y_1\}$) are arbitrary aside from the non-negativity condition and

$$(29) \quad P(i, \Lambda - f^{-1}\{y_1\}) = 1 - c, \quad i \in f^{-1}\{y_1\}.$$

The processes $Y(n) = f(X(n))$ are Markovian whatever the initial distribution $P(\{i\}) = p_i$ of $X(n)$. Further, the transition probability matrix of $Y(n)$ is stationary, whatever the initial distribution p_i of $X(n)$. Of course, it does depend on the initial distribution of $X(n)$ since

$$(30) \quad P(X(n+1) \in f^{-1}\{y_2\} \mid X(n) \in f^{-1}\{y_1\}) \\ = \sum_{i \in f^{-1}\{y_1\}} p_i P(i, f^{-1}\{y_2\}) / \sum_{i \in f^{-1}\{y_1\}} p_i.$$

If we insist that the transition mechanism of $Y(n)$ not only be stationary, but also independent of the initial distribution $P(\cdot)$ of $X(n)$, we are again led to (26) as a necessary and sufficient condition. Condition (26) arises in some of the results of [2, 6, 7, 8, 9].

Notice that condition (8) implies that the transition mechanism of $Y(n)$ is almost stationary. It is, in fact, then stationary for $n \geq 1$. There are simple conditions that imply (8), for example, condition (26). The assumption that the measures $P(x, B)$ are of the form

$$(31) \quad P(x, B) = \int_B f(x, z) \nu(dz),$$

where ν is a probability measure on $\mathcal{B}$ and $f(x, z)$ is measurable in (x, z) on the product space $\Lambda \times \Lambda$ with respect to the Borel field $\mathcal{B} \times f^{-1}\mathcal{B}'$, is enough to ensure (8). It is reasonable to ask whether (26) or (31) or trivial variants of them are the only conditions that imply (8). A simple discussion indicates that this is not the case generally.

Suppose that the function $P(x, B)$ is given as a function in $x \in \Lambda$ and $B \in f^{-1}\mathcal{B}'$. It is measurable for each fixed $B \in f^{-1}\mathcal{B}'$ in x with respect to $\mathcal{B}$ and a probability measure for each fixed $x \in \Lambda$ on $f^{-1}\mathcal{B}'$. We wish to find the family of stochastic transition functions $P(\cdot, \cdot)$ of $x \in \Lambda$ and $B \in \mathcal{B}$ which are consistent with $P(x, B)$, $B \in f^{-1}\mathcal{B}'$, that is, they agree with the latter on the restricted field $f^{-1}\mathcal{B}'$ and, further, condition (8) is satisfied. Assume that the given family of measures $P(x, B)$, $x \in \Lambda$, $B \in f^{-1}\mathcal{B}'$, is equivalent to the probability measure $\bar{\mu}$ on $f^{-1}\mathcal{B}'$. Call μ a refinement of $\bar{\mu}$ if it is a probability measure on $\mathcal{B}$ and agrees with $\bar{\mu}$

on $f^{-1}\mathfrak{B}'$. The following corollary characterizes the family of stochastic transition functions with corresponding sets of measures dominated by a probability measure μ and consistent with $P(x, B)$, $B \in f^{-1}\mathfrak{B}'$.

Corollary 2. *Let $P(x, B)$, $x \in \Lambda$, $B \in f^{-1}\mathfrak{B}'$, be given and be dominated by the probability measure $\bar{\mu}$ on $f^{-1}\mathfrak{B}'$. Then $P(\cdot, \cdot)$ is a dominated (by a probability measure) stochastic transition function consistent (so that (8) is satisfied) with $P(x, B)$, $B \in f^{-1}\mathfrak{B}'$, if and only if it is a transition function that can be expressed in the form*

$$(32) \quad P(x, B) = \int_B (1 + a(x, z))g(x, z)\mu(dz), \quad x \in \Lambda, \quad B \in \mathfrak{B},$$

where μ is a refinement of $\bar{\mu}$, where $g(x, z) \geq 0$ is a function of $\Lambda \times \Lambda$ measurable with respect to $\mathfrak{B} \times f^{-1}\mathfrak{B}'$ and $a(x, z)$ is a function on $\Lambda \times \Lambda$ measurable with respect to $\mathfrak{B} \times \mathfrak{B}$ and satisfying

$$(33) \quad a(x, z) \geq -1, \quad x, z \in \Lambda,$$

$$(34) \quad \int_{f^{-1}A'} a(x, z)g(x, z)\mu(dz) = 0,$$

$$(35) \quad \int_{f^{-1}A'} a(x, z)g(x, z)\mu(dz)P(z, f^{-1}B') = 0$$

for all $x \in \Lambda$ and $A', B' \in \mathfrak{B}'$.

Suppose $P(\cdot, \cdot)$ is a dominated stochastic transition function consistent with $P(x, B)$, $B \in f^{-1}\mathfrak{B}'$. Let μ be a probability measure on $\mathfrak{B}$ equivalent to the family of measures generated by $P(\cdot, \cdot)$. Since $P(\cdot, \cdot)$ agrees with $P(x, B)$, $B \in f^{-1}\mathfrak{B}'$, on $B \in f^{-1}\mathfrak{B}'$, μ can be taken to be a refinement of $\bar{\mu}$. Let $g(x, z)$ be the Radon-Nikodym derivative of $P(x, B)$, $B \in f^{-1}\mathfrak{B}'$, with respect to $\bar{\mu}$. The function $g(x, z)$ is defined on $\Lambda \times \Lambda$ and measurable with respect to $\mathfrak{B} \times f^{-1}\mathfrak{B}'$. Let $h(x, z)$ be the Radon-Nikodym derivative of $P(x, B)$, $B \in \mathfrak{B}$, with respect to

$$(36) \quad \int_B g(x, z)\mu(dz), \quad B \in \mathfrak{B}.$$

Set $a(x, z) = h(x, z) - 1$. Then $P(\cdot, \cdot)$ has the representation (32) and $g(x, z) \geq 0$. We only have to verify conditions (33–35). But $h(x, z) = 1 + a(x, z) \geq 0$. Condition (34) follows from

$$(37) \quad \int_B h(x, z)g(x, z)\mu(dz) = \int_B g(x, z)\mu(dz), \quad B \in f^{-1}\mathfrak{B}',$$

and (35) follows from (8). On the other hand, if $P(\cdot, \cdot)$ has a representation of the form (32) it is clearly a stochastic transition function consistent with $P(x, B)$, $x \in \Lambda$, $B \in f^{-1}\mathfrak{B}'$. Actually (34) is just the special case of (35) obtained by taking $B' = \Lambda'$. It was written out because we made use of it to get (35).

Corollary 2 provides us with a method of constructing transition functions

consistent with a specification of $P(x, B)$, $x \in \Lambda$, $B \in \mathcal{B}'$. The specific transition function $P(\cdot, \cdot)$ obtained will depend on the choice of the refinement μ of $\bar{\mu}$ and the function $a(x, z)$. There is always one possible choice of $a(x, z)$, namely $a(x, z) \equiv 0$. Generally, there will be many other nontrivial choices of $a(x, z)$ but this will, of course, depend on how strong the restriction (35) is (see the example in section 3 of [2]). The condition $a(x, z) \equiv 0$ amounts to the representation (31).

The statement of Theorem 1 in the special case of a countable state space is given in [10]. It is clear that the notion of a maximal single entry set does not make sense if the condition that the measures $P(x, f^{-1}B')$, $x \in \Lambda$, $B' \in \mathcal{B}'$, be dominated is relaxed. The question of determining when a function $Y(n) = f(X(n))$ of a stationary Markov process $X(n)$ (not any initial distribution) is Markovian was brought up in [2]. It is clear that (8) provides a sufficient condition broader than (26).

Continuous Parameter Markov Chains. Now consider what happens in the case of a continuous parameter Markov chain $X(t)$ with stationary transition mechanism, $0 \leq t < \infty$. Let the state space Λ be labeled by the positive integers. The chain is assumed to have a decent transition probability function

$$P(t) = (p_{ij}(t); i, j = 1, 2, \dots) \\ = (P(X(t + \tau) = j \mid X(\tau) = i)), \quad 0 \leq t, \quad \tau < \infty,$$

continuous in t , $0 \leq t < \infty$, with

$$(38) \quad \lim_{t \downarrow 0} P(t) = I.$$

Let $S_\alpha = f^{-1}\{y_\alpha\}$, $\alpha = 1, 2, \dots$, be the disjoint sets of states generated by the function f . Necessary and sufficient conditions for $Y(t) = f(X(t))$ to be Markovian, whatever the initial distribution $p_i = P(X(0) = i)$, $i = 1, 2, 3, \dots$, of $X(t)$, are obtained. We shall say that S_α is (is not) a single entry set at time $t > 0$ according as to whether it is (is not) a single entry set with respect to the matrix $P(t)$ and its powers (in the sense of the previous section). If S_α is not a single entry set at time $t > 0$, it is not a single entry set for all $\tau \geq t$. Suppose S_α is a single entry set at time $t > 0$. Then S_α can only be entered from itself in time $\tau \leq t$ because of (38). It must then be a single entry set for all $\tau \geq t$. For if it were not, there would be a maximal $T < \infty$ such that for $t < T$ S_α is a single entry set and for $t > T$ S_α is not a single entry set. Then for every $\epsilon > 0$ there is an $i \notin S_\alpha$ such that

$$(39) \quad p_{i, S_\alpha}(T + \epsilon) = \sum_{j \in S_\alpha} p_{ij}(T + \epsilon) > 0.$$

But

$$(40) \quad p_{i, S_\alpha}(T + \epsilon) = \sum_j p_{ij}(T - \epsilon)p_{j, S_\alpha}(2\epsilon) = \sum_{j \in S_\alpha} p_{ij}(T - \epsilon)p_{j, S_\alpha}(2\epsilon) = 0$$

if $\epsilon < \frac{1}{2}T$, a contradiction. Thus there are only two possibilities. Either S_α is a single entry set for all t or else it is not a single entry set for all t , $\alpha = 1, 2, \dots$.

If S_α is a single entry set for all t

$$(41) \quad p_{i, s_\alpha}(t) \equiv 0$$

for all $i \notin S_\alpha$. Suppose S_α is not a single entry set for all t . Then

$$(42) \quad \sum_{i \in S_\beta} p_{ii}(t) p_{i, s_\beta}(\tau) / p_{i, s_\alpha}(t) = Q_{s_\alpha, s_\beta}(t, \tau),$$

$t, \tau > 0$, for all $i \in S_\alpha$ and all β . As $t \downarrow 0$ the left side of (42) approaches $p_{i, s_\beta}(\tau)$. But the right side of (42) depends on i only through S_α . Thus

$$(43) \quad p_{i, s_\beta}(\tau) = Q_{s_\alpha, s_\beta}(\tau)$$

for all $i \in S_\alpha$ and we have the following result.

Theorem 2. Let $X(t)$, $0 \leq t < \infty$, be a continuous parameter Markov chain with stationary transition probability function $P(t) = (p_{ij}(t); i, j = 1, 2, \dots)$. Let $P(t)$ be continuous in t , $0 \leq t < \infty$, with

$$(44) \quad \lim_{t \downarrow 0} P(t) = I.$$

The process $Y(t) = f(X(t))$ will be Markovian, whatever the initial distribution of $X(t)$, if and only if each set S_α of states generated by f separately satisfies one or the other of the following two conditions for all t :

$$(45) \quad (i) \quad p_{i, s_\alpha}(t) \equiv 0 \quad \text{for all } i \notin S_\alpha,$$

$$(46) \quad (ii) \quad p_{i, s_\beta}(t) = Q_{s_\alpha, s_\beta}(t) \quad \text{for all } i \in S_\alpha \text{ and all } \beta.$$

It would be interesting to find out under what conditions the results of Theorem 2 hold in the case of a more general state space for the processes $X(t)$.

REFERENCES

- [1] R. R. BAHADUR, Sufficiency and statistical decision functions, *Ann. Math. Stat.*, **25** (1954), pp. 423-462.
- [2] C. J. BURKE & M. ROSENBLATT, A Markovian function of a Markov chain, *Ann. Math. Stat.*, **29** (1958), pp. 1112-1122.
- [3] J. L. DOOB, *Stochastic Processes*, New York (1952).
- [4] P. R. HALMOS & L. J. SAVAGE, Application of the Radon-Nikodym theorem to the theory of sufficient statistics, *Ann. Math. Stat.*, **20** (1949), pp. 225-241.
- [5] M. LOEVE, *Probability Theory*, New York (1955).
- [6] B. RANKIN, *The concept of enchainment—a relation between stochastic processes*, unpublished (1955).
- [7] D. ROSENBLATT, On aggregation and consolidation in linear systems, to be published in the *Naval Research Logistics Quarterly*.
- [8] D. ROSENBLATT, *On aggregation and consolidation in linear systems*, Technical Report C, prepared under Contract Nonr-1180 (00) (NR-047-012) 1956.
- [9] D. ROSENBLATT, Abstracts of "On aggregation and consolidation in finite substochastic systems I, II, III, IV," *Ann. Math. Stat.*, **28** (1957), pp. 1060-1061.
- [10] M. ROSENBLATT, Abstract of "Functions of Markov Chains," *Ann. Math. Stat.*, **29** (1958), p. 1291.

Indiana University
Bloomington, Indiana