ON STATIONARY MARKOV PROCESSES[1]

By Richard Isaac

Hunter College

1. Introduction. Consider Markov processes $(X_n, n \geqq 0)$ with given stationary transition probabilities and (σ-finite) stationary measure α. The state space Ω is arbitrary; Σ is a σ-field of measurable subsets of Ω. First, we prove that the strictly stationary process $(X_n, n \geqq 0)$ is embeddable in a strictly stationary Markov process $(X_n, -\infty < n < \infty)$ which we call the *extended process* (see [5]). This was a fact assumed true in [5], but no proof was given. We also examine the invariant random variables for these processes in Theorem 2. Also briefly discussed is the reversed Markov process. In the event that Ω is the real or complex field, Theorem 1 is known ([1], p. 456) and if α is finite Theorem 2 is known ([1], pp. 458–460). However, counterexamples are offered illustrating the difficulties arising when α is infinite.

This note is a sequel to [5]. Besides the gap there mentioned above, the language of [5] suggested that Theorem 2 is true in general, i.e., without condition (A). Section 4 of this note will set matters straight.

2. Main results.

Lemma. *Let Σ be separable, that is, Σ is generated by a countable family of sets. Then the strictly stationary Markov process $(X_n, n \geqq 0)$ may be embedded in an extended process $(X_n, -\infty < n < \infty)$.*

Proof. Consider bilateral sequence space Ω_1 with elements $\omega = (\cdots \omega_{-1}, \omega_0, \omega_1, \cdots)$. Let Λ_0 and ${}_0\Lambda$ be the σ-fields generated by cylinders in Ω_1 with non-negative coordinates and non-positive coordinates respectively. Using the transition probabilities, for each x a conditional probability measure $P(\cdot \mid X_0 = x)$ may be constructed on Λ_0 according to [1], p. 614. With α as initial measure on X_0-space, it is easily seen that a shift-invariant measure α_0 may be defined on Λ_0 by putting $\alpha_0(U) = \int P(U \mid X_0 = x)\alpha(dx)$ for $U \varepsilon \Lambda_0$ (see Lemma 1 of [5]). Proceed as in [1], p. 456, to assign a mass α_1 to cylinder sets in Ω_1 by setting $\alpha_1(C) = \alpha_0(T^{-j}C)$ where $T^{-j}C \varepsilon \Lambda_0$, T is the shift, and C is a cylinder of Ω_1. To prove that α_1 determines a measure on the σ-field Σ_1 of Ω_1 determined by the cylinder sets (and hence that $(X_n, n \geqq 0)$ is embedded in $(X_n, -\infty < n < \infty)$) it is necessary to prove α_1 countably additive on the cylinders.

Kolmogorov's extension theorem fails because Ω here is arbitrary. It is already known that α_1 restricted to Λ_0 is countably additive and equal to α_0. Now we check α_1 restricted to ${}_0\Lambda$ is countably additive. To see this, observe that since $X_0, X_1, \cdots$ is a Markov process with initial distribution α, the process $\cdots X_n, X_{n-1}, \cdots, X_0$ is also Markovian (see [1], p. 83; the restriction to real

Received 21 January 1966; revised 26 August 1966.

[1] Supported in part by NSF Grant GP-3819.

or complex-valued processes there is not essential). Let $Q_{n,n-1}(E, x) = P(X_{n-1} \varepsilon E \mid X_n = x)$. Since Σ is separable, there is a version of this conditional probability a.e. (α); the proof follows the real case, [1], p. 30, by using the selection principle on a countable field of generators for Σ. Moreover, the stationarity of α easily shows $Q = Q_{n,n-1}$ is independent of n, and α is stationary for Q, that is, the reversed Markov process has stationary transition probabilities and is strictly stationary with stationary measure α. Therefore the reversed Markov process has, for each x, a conditional probability $Q(\cdot \mid X_0 = x)$ defined on ${}_0\Lambda$. The proof of this fact is identical with [1], p. 614, except that we consider non-positive rather than non-negative coordinates. The formula

$$ {}_0\alpha(U) = \int Q(U \mid X_0 = x)\alpha(dx), $$

as in the case of α_0, defines a measure, this time on ${}_0\Lambda$, and ${}_0\alpha(U) = \alpha_1(U)$ for $U \varepsilon {}_0\Lambda$. This proves α_1 countably additive on ${}_0\Lambda$. Consider now a countable collection of arbitrary disjoint cylinders $\{C_i\}$ whose union C is a cylinder. We must show $\sum \alpha_1(C_i) = \alpha_1(C)$. α is σ-finite so that there exists an expanding sequence of sets $A_k = [X_0 \varepsilon \tilde{A}_k]$ where $\tilde{A}_k \uparrow \Omega$ and $\alpha(\tilde{A}_k) < \infty$. Let $C_i \cap A_k = C_i^{(k)}$ and $C \cap A_k = C^{(k)}$. Putting $C^{(k)} - \bigcup_{i \leq n} C_i^{(k)} = D_n^{(k)}$, one has $D_n^{(k)} \downarrow \varphi$ for each fixed k. Since $D_n^{(k)} = N_n^{(k)} \cap P_n^{(k)}$ where $N_n^{(k)} \varepsilon {}_0\Lambda$ and $P_n^{(k)} \varepsilon \Lambda_0$, it follows that, for fixed k, either $N_n^{(k)} \downarrow \varphi$ or $P_n^{(k)} \downarrow \varphi$. To fix ideas, suppose $k = 1$ and $N_n^{(1)} \downarrow \varphi$. α_1 is countably additive on ${}_0\Lambda$ and $\alpha_1(N_n^{(1)}) \leqq \alpha_1(A_1) = \alpha(\tilde{A}_1) < \infty$, so that $\alpha_1(D_n^{(1)}) \leqq \alpha_1(N_n^{(1)}) \to 0$, and so $\sum \alpha_1(C_i^{(1)}) = \alpha_1(C^{(1)})$ or, more generally, $\sum_i \alpha_1(C_i^{(k)}) = \alpha_1(C^{(k)})$ for each k. $C_i^{(k)} \uparrow C_i$ and $C^{(k)} \uparrow C$, so the monotone convergence theorem yields $\alpha_1(C) = \lim_k \alpha_1(C^{(k)}) = \lim_k \sum_i \alpha_1(C_i^{(k)}) = \sum_i \alpha_1(C_i)$, and the proof is concluded.

THEOREM 1. *The process* $(X_n, n \geqq 0)$ *may be embedded in an extended process* $(X_n, -\infty < n < \infty)$.

PROOF. The theorem merely asserts the truth of the lemma even if Σ is not supposed separable. α_1 is still finitely additive on cylinders, and again we wish to show α_1 countably additive. That α_1 may be defined on cylinders at all is a consequence of the existence of *given* transition probabilities; if these are not given the theorem is not necessarily true. See [1], p. 614. Suppose α_1 not countably additive; then there exists a sequence of disjoint cylinders $\{A_i\}$ with $\bigcup A_i = A$, a cylinder, and $\sum \alpha_1(A_i) \neq \alpha_1(A)$. Now observe that there is a sequence of sets B_n in Σ such that each set A_i is defined only in terms of the sets B_n. For, each set A_i is a union of "rectangles" (sets of form $[X_{i_1} \varepsilon C_1, \cdots, X_{i_j} \varepsilon C_j]$), and since A_i is defined in terms of at most a countable collection of Σ sets and there are a countable number of sets A_i, the result follows. There is then an admissible subfield $\tilde{\Sigma} \subseteq \Sigma$, i.e., $\tilde{\Sigma}$ is separable and $P(\cdot, E)$ is measurable with respect to $\tilde{\Sigma}$ for each $E \varepsilon \tilde{\Sigma}$ and $\{B_n\} \subset \tilde{\Sigma}$ (see [1], p. 209 and [8]). The process may now be restricted to $\tilde{\Sigma}$, and one may check that α restricted to $\tilde{\Sigma}$ is stationary for the restricted process. Thus the mass α_1 is countably additive on the cylinders generated by sets in $\tilde{\Sigma}$ by the lemma; since the A_i are cylinders of this type, a contradiction results. Hence α_1 is countably additive on the cylinders, and the proof is concluded.

The lemma has the following immediate

COROLLARY. *If Σ is separable, the stationary Markov process $(X_n, n \geqq 0)$ (or equivalently $(X_n, -\infty < n < \infty)$) has associated with it a Markov process $(Y_n, -\infty < n < \infty)$ where $Y_n = X_{-n}$. This process, the* reversed X_n-process, *has stationary transition probabilities and is strictly stationary with stationary measure α.*

"Reversing the chain" is a useful device in the case of discrete state spaces (cf. [2], p. 373). Most recently the concept has proved valuable in the potential theory of Markov chains and the analysis of the Martin boundary. The corollary is a generalization of this standard reversal procedure and may be of interest in the potential theory of general Markov processes.

α is said to satisfy condition (A) if, for every $E \varepsilon \Sigma$, $P(X_n \varepsilon E$ infinitely often $| X_0 = x) = 1$ a.e. (α) for $x \varepsilon E$ (see [6] and [7]).

THEOREM 2. *Let α satisfy condition* (A). *Then the process $(X_n, n \geqq 0)$ and the extended process have the same invariant random variables and any invariant random variable is measurable with respect to X_0.*

This generalizes results in [1], pp. 458–460, proved for finite measures α. To see this, observe that condition (A) is automatically valid for α finite; this is merely the Poincaré recurrence theorem [3], p. 10, in a probability setting.

As a tool in the proof of Theorem 2 we employ the process on A (see [4]). Since possibly $P_A(x, A) < 1$ for some points $x \varepsilon A$, the process on A cannot be defined exactly as Harris does. But because condition (A) holds, by excluding an α-null set, we may define the process on B, $B \subseteq A$, where $P_B(x, B) = 1$ for all $x \varepsilon B$. Thus, without loss of generality, assume $P_A(x, A) = 1$ for all $x \varepsilon A$.

LEMMA. *Let α be a stationary measure for $(X_n, n \geqq 0)$ satisfying condition* (A) *and let A be a set with $0 < \alpha(A) < \infty$. Then α restricted to A is a finite stationary measure for the process on A.*

PROOF. In [4] a finite stationary measure α on A was extended to a σ-finite measure on the entire space. The argument there works in the opposite direction as well, as is easily checked (see [4], p. 116, (4.4)), and this proves the lemma.

PROOF OF THEOREM 2. To prove the first assertion, let

$$\omega = (\cdots, \omega_{-1}, \omega_0, \omega_1, \cdots)$$

with $\omega_0 \varepsilon A$, and define the transformation $T_A\omega = T^r\omega$ where $r > 0$ is the first index to satisfy $\omega_r \varepsilon A$ and $\omega_i \not\varepsilon A$ for $0 < i < r$. Here T is the usual shift for the original process, and T_A corresponds to the shift for the process on A. T_A is measurable and invertible. We are now ready to adapt Doob's proof [1], p. 458. It is only required to prove that a function y invariant with respect to the extended process $(X_n, -\infty < n < \infty)$ is measurable with respect to the X_n's with n $\geqq 0$. Let $A \subseteq \Omega$ be chosen with $0 < \alpha(A) < \infty$. To every positive integer k there is a random variable y_k, measurable with respect to the σ-field determined by a finite number of X_n's, such that

$$\text{(a)} \qquad \alpha_1(\tilde{A} \cap [|y(\omega) - y_k(\omega)| > k^{-1}]) < 2^{-k}$$

where $\tilde{A} \subseteq \Omega_1$ and $\tilde{A} = (\omega: X_0(\omega) \in A)$. There exists a positive integer j such that $T^j y_k$ is measurable on the space of X_n, $n \geqq 0$; certainly then $T_A{}^j y_k(\omega) = y_k(T_A{}^j \omega)$ is also measurable relative to this sample space. Now employ the facts that $T_A{}^j \tilde{A} = \tilde{A}$ and $T_A{}^j(M \cap N) = T_A{}^j M \cap T_A{}^j N$ for any positive integer j and sets M and N in the domain of T_A. The second relation follows from the invertibility of T_A. Thus, (a) yields, using the lemma

$$\text{(b)} \qquad \alpha_1(\tilde{A} \cap [|y(T_A{}^j\omega) - y_k(T_A{}^j\omega)| > k^{-1}]) < 2^{-k}.$$

But $y(T_A{}^j\omega) = y(T^{k(\omega)}\omega) = y(\omega)$ a.e. (α_1) by the invariance of y, so that y is invariant under $T_A{}^j$. Therefore (b) holds where $y(\omega)$ may be substituted for $y(T_A{}^j\omega)$. Thus (b) says

$$\lim_{k\to\infty} T_A{}^j y_k = y \quad \text{a.e.} \quad (\alpha_1) \quad \text{on} \quad \tilde{A}$$

so that y on $\tilde{A}$ is the limit of functions measurable with respect to X_n, $n \geqq 0$. Since the entire space is the union of such sets $\tilde{A}$, the proof is concluded by piecing together a countable number of functions corresponding to sets $\tilde{A}$. This concludes the proof of the first assertion. The second follows by adapting the proof of Theorem 1.1 [1], p. 460, using the preceding method and similar arguments.

α is called ergodic for a process if the only invariant sets are trivial up to sets of measure zero.

Corollary. *α is ergodic for the process $(X_n, n \geqq 0)$ if and only if α is ergodic for the extended process.*

3. Some examples.

Example 1. Theorem 2 is not generally true without condition (A). Consider the state space $\Omega = (a_1, a_2, \cdots; b_1, b_2, \cdots; 0, 1, 2, \cdots)$ with $p(a_n, a_{n-1}) = p(b_n, b_{n-1}) = p(a_1, 0) = p(b_1, 0) = 1$ for $n > 1$; $p(n, n+1) = 1$ for $n \geqq 0$.

Let α be the stationary measure assigning mass 1 to each of the points a_n and b_n and mass 2 to each "number" point. The process $(X_n, n \geqq 0)$ has trivial invariant field since the only bounded regular functions f (i.e., $Pf = f$) are the constants (see [6]). On the other hand, the sets

$$A = \{\omega: \omega = T^k(\cdots, a_2, a_1, 0, 1, 2, \cdots) \quad \text{for some integer} \quad k\}$$

and

$$B = \{\omega: \omega = T^k(\cdots, b_2, b_1, 0, 1, 2, \cdots) \quad \text{for some integer} \quad k\}$$

are invariant sets, each with α_1 measure ∞.

Example 2. An invariant function may not be X_0 measurable if the process has an infinite stationary measure α for which condition (A) fails to hold. Consider the same state space as in Example 1, but set $p(a_n, a_{n+1}) = p(b_n, b_{n+1}) = 1$, $n \geqq 1$; $p(n, n-1) = 1, n \geqq 1$; $p(0, a_1) = p(0, b_1) = \frac{1}{2}$. Let α assign mass as in Example 1. Then α is stationary. If $z(\omega) = 1$ for $\omega = (\omega_0, \omega_1, \cdots)$ containing an infinite number of a_n's and $z(\omega) = 0$ if ω contains an infinite number of b_n's, z is defined a.e. and is invariant. But z is not a function of X_0, for if $z(\omega) = h(X_0(\omega))$

a.e. (α_1) for some function h, then $\omega_1 = (0, a_1, a_2, \cdots)$ and $\omega_2 = (0, b_1, b_2, \cdots)$ each have α_0 measure 1, and $z(\omega_1) = z(\omega_2)$, which is false by definition of z.

4. Conclusion. All page references to follow apply to [5]. On page 1782, top, it was stated that the $(X_n, n \geqq 0)$ process and the extended process are simultaneously ergodic. As we have seen (Example 1) this is not necessarily true unless condition (A) is valid. Thus, for the validity of Theorem 1 on p. 1782, it is necessary to restrict attention to processes such that the original and extended processes have the same invariant random variables. To see that Theorem 1[5] may fail otherwise, consider Example 1 where β assigns mass 1 to each point a_n and each "number" point and mass 0 to the points b_n. β is stationary, yet β is not a constant multiple of α, the stationary measure described in Example 1.

Call α *strongly ergodic* for $(X_n, n \geqq 0)$ if α is ergodic for the extended process. Then strong ergodicity implies ergodicity, but not conversely. Theorem 2 asserts the equivalence of strong ergodicity and ergodicity under condition (A). In the example on p. 1783 the non-constant bounded function k_i was erroneously asserted to be a solution to a certain equation of regularity, whereas it only satisfies the equation for $i \geqq 1$. The only bounded regular functions for this process are the constants, hence ergodicity follows for every stationary measure. Thus $\alpha + \beta$ is ergodic but not strongly ergodic, whereas both α and β are ergodic, contrary to the statement there.

A final comment: On p. 1784, for $P(V/X_0 = t) = 1$ for *every* $t \varepsilon \Omega$ to hold (line 6), V must be strictly invariant, i.e., $T^{-1}V = V$. However, for the other conclusions there it suffices for V to be α_0-invariant, i.e., $T^{-1}V$ and V differ by an α_0-null set.

REFERENCES

[1] Doob, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[2] Feller, W. (1960). *An Introduction to Probability Theory and Its Applications.* 1 (2nd ed). Wiley, New York.

[3] Halmos, P. R. (1956). *Lectures on Ergodic Theory.* Math. Soc. Jap.

[4] Harris, T. E. (1956). The existence of stationary measures for certain Markov processes. *Proc. Third Berkeley Symp. Math. Statist. Prob.* 113–124. Univ. of California Press.

[5] Isaac, R. (1964). A uniqueness theorem for stationary measures of ergodic Markov processes. *Ann. Math. Statist.* **35** 1781–1786.

[6] Isaac, R. (1966). On regular functions for certain Markov processes. (To appear in *Proc. Amer. Math. Soc.*)

[7] Isaac, R. (1967). On the ratio-limit theorem for Markov processes recurrent in the sense of Harris (To appear in *Illinois J. Math.*)

[8] Jamison, B. and Orey, S. (1966). *Markov Chains Recurrent in the Sense of Harris.* Univ. of Minnesota, mimeographed.