

AN EXTENDED MARKOV PROPERTY⁽¹⁾

BY

R. M. BLUMENTHAL⁽²⁾

Introduction. Let $\{x(t); t \geq 0\}$ be a time homogeneous Markov process. A positive random variable T defined on the same underlying probability space as $\{x(t)\}$ is called a *Markov time* for the process if the collection of functions $\{x(t+T); t \geq 0\}$ also forms a Markov process with the same transition function as the original process and if, when $x(T)$ is fixed, the new process is independent of the original one considered only up to time T . Thus a time homogeneous Markov process could be defined as a process for which T is a Markov time whenever T is identically a constant. This paper is chiefly concerned with finding hypotheses on $\{x(t)\}$ and on T which will insure that T be a Markov time. Certain related matters including the behavior of sample functions are also discussed. Whatever loose statements are made in this introduction will be made precise at the appropriate places in the body of the paper.

In probabilistic approaches to potential theory (Doob [1], Hunt [1; 2]) and in other researches which employ the familiar "first passage time relationship" the random variables T which arise are the "first passage times" or more general "*stopping times*" which will be defined below; so the question of whether these are Markov times is of especial interest. Hunt [1] has shown that if $\{x(t)\}$ is a process in n -space with stationary independent increments then any stopping time is a Markov time for the process provided $\{x(t)\}$ has right continuous sample functions (such processes can generally be normalized to have this right continuity property according to a theorem of Doob [2]). Hunt states his result in the following form: $\{x(t+T) - x(T); t \geq 0\}$ is a Markov process with the same transition function as the original process and is completely independent of the original process considered only up to time T . He also points out that this result is valid for processes taking values in more general spaces. We shall see that for a somewhat more general class of processes any stopping time is a Markov time, although the theorem for these processes can not be expressed in the same form as Hunt's result. Among these processes are the time homogeneous processes of Ito in which a drift and a spread, varying in space but not in time, are imposed upon the Brownian motion.

Received by the editors June 7, 1956.

(¹) This paper is the author's thesis, submitted to the Graduate School of Cornell University in partial fulfillment of the requirements for the degree of Ph.D. The work was directed by Professor G. A. Hunt to whom the author is deeply indebted for encouragement and guidance.

(²) This research was carried out under a contract sponsored by the Office of Naval Research.

Preliminaries. We suppose that we are given a metric space X and a probability space Ω , of points ω , with Borel field \mathfrak{B} of subsets of Ω and probability measure P on \mathfrak{B} . $\mathfrak{B}(X)$ denotes the topological Borel field of X . We assume that \mathfrak{B} and all subfields of \mathfrak{B} referred to hereafter have been completed with respect to the measure P . A stochastic process is a collection $\{x(t, \omega); t \geq 0\}$ of functions from Ω into X which are measurable relative to the fields $\mathfrak{B}(X)$ and \mathfrak{B} (that is, a collection of random variables). Usually for fixed t we denote the corresponding random variable by $x(t)$; however we occasionally denote it by $x(t, \omega)$ thereby confusing the random variable with its value at the point ω . In each case it is clear what is meant.

If $\{\mathfrak{F}(t); t \geq 0\}$ is a family of Borel subfields of \mathfrak{B} such that $\mathfrak{F}(s)$ is contained in $\mathfrak{F}(t)$ whenever s is less than t , then a *Markov process relative to* $\{\mathfrak{F}(t)\}$ is a stochastic process such that

(a) for each t , $x(t)$ is $\mathfrak{F}(t)$ measurable,

(b) for each $s < t$ and A in $\mathfrak{B}(X)$, $P\{x(t) \in A \mid \mathfrak{F}(s)\} = P\{x(t) \in A \mid x(s)\}$ ⁽³⁾.

Of course here, as elsewhere, by equality between random variables defined only up to sets of measure 0, I mean equality for almost all ω . Obviously any such process is a Markov process in the usual sense. Conversely let $\{x(t)\}$ be a process and for each t let $\mathfrak{G}(t)$ be the smallest Borel field such that $x(s)$ is $\mathfrak{G}(t)$ measurable for each s not exceeding t . Then $\{x(t)\}$ is a Markov process under the usual definition if and only if it is a Markov process relative to $\{\mathfrak{G}(t)\}$. The notation for a Markov process relative to $\{\mathfrak{F}(t)\}$ will be $\{x(t), \mathfrak{F}(t)\}$. If the process is to be considered as a Markov process in the usual sense we shall denote it by $\{x(t), \mathfrak{G}(t)\}$ or simply by $\{x(t)\}$.

Throughout this paper $\{x(t), \mathfrak{F}(t); t \geq 0\}$ will be a time homogeneous Markov process with a transition function; that is we assume the existence of a real-valued function $P(t, x, A)$ defined for t non-negative, x in X , A in $\mathfrak{B}(X)$ with the following properties:

(A) for fixed t and x , $P(t, x, \cdot)$ is a probability measure on $\mathfrak{B}(X)$,

(B) for fixed t and A , $P(t, \cdot, A)$ is Borel measurable in x ,

(C) $P(t, x, A)$ satisfies the Chapman-Kolmogorov equation and for every s less than t , and A , the random variable $P(t-s, x(s), A)$ is one version of $P\{x(t) \in A \mid x(s)\}$. At times we shall also assume that $P(t, x, A)$ has the property

(D) $\int_x \phi(y)P(t, x, dy)$ is continuous in x whenever ϕ is a bounded continuous function on X . When this assumption is being made we shall explicitly say so. Note that we use the letter P indiscriminately; at each usage its meaning should be clear.

A positive random variable (possibly taking on the value $+\infty$) is called a *stopping time* for the process if for each real a , $\{T < a\}$ is in $\mathfrak{F}(a)$. This is equivalent to the apparently stronger condition that $\{T < a\}$ is in $\mathfrak{F}(a-)$ and is also equivalent to the condition that $\{T \leq a\}$ is in $\mathfrak{F}(a+)$ where $\mathfrak{F}(a-)$ is

⁽³⁾ This specialization of the usual definition of a Markov process was suggested to me by Professor J. L. Doob.

the smallest Borel field containing $\mathfrak{F}(b)$ for every $b < a$ and $\mathfrak{F}(a+)$ is the intersection of the Borel fields $\mathfrak{F}(b)$ with $b > a$. T may be infinite with positive probability, but we shall always assume that $P\{T < \infty\} > 0$. In the sequel we shall also encounter random variables for which $\{T \leq a\}$ is in $\mathfrak{F}(a)$. They are formally at least a little less general than stopping times. As an example of a stopping time and even of this less general random time suppose $\{x(t), \mathfrak{F}(t)\}$ has continuous paths, A is a compact subset of X and $T(\omega) = \inf \{t \mid x(t, \omega) \in A\}$ or $T(\omega) = \infty$ if for every t , $x(t, \omega) \notin A$. Then $\{T \leq a\} = \bigcap_m \bigcup_r \{x(r) \in A_m\}$ where A_m is the set of points at a distance less than $1/m$ from A and r ranges over the rationals less than a , and so $\{T \leq a\}$ is in $\mathfrak{G}(a)$ and hence in $\mathfrak{F}(a)$.

Let T be a stopping time or more generally any positive random variable, let $\Omega' = \{T < \infty\}$ and suppose $P\{\Omega'\} > 0$. We construct a new Borel field \mathfrak{B}' and probability measure P' , over Ω' , by taking \mathfrak{B}' to consist of those elements of \mathfrak{B} which are subsets of Ω' and for these sets defining

$$P'\{\Delta\} = P\{\Delta\}/P\{\Omega'\}.$$

We may then consider $\{x'(t, \omega) = x(t + T(\omega), \omega); t \geq 0\}$ as a collection of functions from Ω' into X . (We have no way of knowing, at this point, whether $\{x'(t)\}$ is a stochastic process, that is, whether $x'(t)$ is \mathfrak{B}' measurable for each t .) Suppose V is a set with a Borel field $\mathfrak{B}(V)$ and g is a random variable from Ω to V , that is $g^{-1}(D)$ is in \mathfrak{B} for every D in $\mathfrak{B}(V)$. We say that g depends on the original process only up to time T if for every D in $\mathfrak{B}(V)$ and non-negative a it is true that $\{T \leq a\}$ in $\mathfrak{F}(b)$ and $b \geq a$ imply $\{g \in D, T \leq a\}$ in $\mathfrak{F}(b)$. Precisely then, T is called a *Markov time* for $\{x(t), \mathfrak{F}(t)\}$ if $\{x'(t), \mathfrak{G}'(t)\}$ is a Markov process (over $(\Omega', \mathfrak{B}', P')$) with $P(t, x, A)$ as transition function, and if g and the process $\{x'(t)\}$ are conditionally independent relative to $x'(0)$ whenever g is a random variable depending on the original process only up to time T . Here $\mathfrak{G}'(t)$ is the Borel field generated by sets of the form $\{x'(s) \in A\}$ with s not exceeding t and A in $\mathfrak{B}(X)$.

1. LEMMA 1.1. *Let T be countably valued with $\{T \leq a\}$ in $\mathfrak{F}(a)$. Then T is a Markov time for the process $\{x(t), \mathfrak{F}(t)\}$.*

Proof. For each t let $\mathfrak{H}(t)$ be the Borel field consisting of those subsets Δ of Ω' such that for each real a , $\Delta \cap \{T \leq a\}$ is in $\mathfrak{F}(t+a)$. Note that $\mathfrak{H}(t)$ actually is a Borel field and is contained in \mathfrak{B}' , and that $\mathfrak{H}(s)$ is contained in $\mathfrak{H}(t)$ if s is less than t . We shall show that $\{x'(t), \mathfrak{H}(t)\}$ is a Markov process with $P(t, x, A)$ as transition function. Let the finite values of T be a_1, a_2, \dots and let $\sum_i = \{T = a_i\}$. If t is non-negative and A is in $\mathfrak{B}(X)$ then

$$\{x'(t) \in A\} \cap \{T \leq a\} = \bigcup_{a_i \leq a} [\sum_i \cap \{x(t + a_i) \in A\}]$$

which is in $\mathfrak{F}(t+a)$. So $x'(t)$ is $\mathfrak{H}(t)$ measurable. We now must show that if

$0 \leq t \leq t_1 < \dots < t_k$, and B_1, \dots, B_k are in $\mathfrak{B}(X)$ then

$$(1.1) \quad \begin{aligned} P\{x'(t_1) \in B_1, \dots, x'(t_k) \in B_k \mid \mathfrak{F}(t)\} \\ = P\{x'(t_1) \in B_1, \dots, x'(t_k) \in B_k \mid x'(t)\} \end{aligned}$$

and that the term on the right has the appropriate expression in terms of the transition function. Let us assume that $k=2$ since this case demonstrates all the essential points of the proof. Let Δ be in $\mathfrak{F}(t)$ and let $M = \{x'(t_1) \in B_1, x'(t_2) \in B_2\}$. We want to show that

$$(1.2) \quad \int_{\Delta} \left\{ \int_{B_1} P(t_1 - t, x'(t, \omega), dy_1) P(t_2 - t_1, y_1, B_2) \right\} P'(d\omega) = P'\{\Delta \cap M\}.$$

Since the \sum_i are disjoint and their union is Ω' we are immediately reduced to showing that for each i

$$(1.3) \quad \int_{\Delta \cap \sum_i} \left\{ \int_{B_1} P(t_1 - t, x(t + a_i, \omega), dy_1) P(t_2 - t_1, y_1, B_2) \right\} P(d\omega) \\ = P\{\Delta \cap \sum_i \cap M\}.$$

Now $\Delta \cap \sum_i$ is in $\mathfrak{F}(t + a_i)$ and

$$\sum_i \cap M = \sum_i \cap \{x(t_1 + a_i) \in B_1, x(t_2 + a_i) \in B_2\}.$$

Hence the validity of (1.3) is an immediate consequence of the fact that the original process is Markovian relative to $\{\mathfrak{F}(t)\}$ with $P(t, x, A)$ as transition function. Thus $\{x'(t), \mathfrak{F}(t)\}$ is a Markov process with $P(t, x, A)$ as transition function and, since $\mathfrak{G}'(t)$ is contained in $\mathfrak{F}(t)$ for each t , the result also holds for $\{x'(t), \mathfrak{G}'(t)\}$. Taking t to be 0 in (1.1) we obtain the conditional independence relative to $x'(0)$ of $\{x'(t)\}$ and the Borel field $\mathfrak{F}(0)$ and this proves that relative to $x'(0)$, $\{x'(t)\}$ is conditionally independent of any random variable depending on the original process only up to time T .

We pause to make the observation that if A is an open subset of a metric space X then there is an increasing sequence of bounded continuous functions whose limit is the characteristic function of A for if $\rho(x)$ is the distance from x to the complement of A then $f_n(x) = n\rho(x)/1 + n\rho(x)$ is such a sequence. This fact will be useful in the next proof.

THEOREM 1.1. *Suppose that T is a stopping time, that*

$$P' \left\{ \lim_{s \downarrow 0} x'(t + s) = x'(t) \right\} = 1$$

for each t , and that $P(t, x, A)$ has the property (D) mentioned in the preliminaries. Then T is a Markov time for $\{x(t), \mathfrak{F}(t)\}$.

Proof. Let $\mathfrak{F}(t)$ consist of those subsets Δ of Ω' such that for every real a ,

$\Delta \cap \{T < a\}$ is in $\mathfrak{F}(t+a)$. $\{\mathfrak{C}(t)\}$ is then an increasing family of Borel subfields of \mathfrak{B}' . For each $n \geq 1$ define

$$T_n(\omega) = \begin{cases} \frac{k+1}{2^n} & \text{if } \frac{k}{2^n} \leq T(\omega) < \frac{k+1}{2^n}, \\ \infty & \text{if } T(\omega) = \infty \end{cases}$$

and define $\{x'(t)\}$ in terms of T and $\{x_n(t)\}$ in terms of T_n . $T_n = \infty$ if and only if $T = \infty$ so $\{x'(t)\}$ and $\{x_n(t)\}$ for each n are defined over $(\Omega', \mathfrak{B}', P')$. Each T_n is countably valued, and from the definition of T_n and the fact that T is a stopping time we have $\{T_n \leq a\}$ in $\mathfrak{F}(a)$. Thus by Lemma 1.1 each T_n is a Markov time or more specifically $\{x_n(t), \mathfrak{C}_n(t)\}$ is a Markov process where $\mathfrak{C}_n(t)$ consists of the sets Δ such that $\Delta \cap \{T_n \leq a\}$ is in $\mathfrak{F}(t+a)$. From the definition of T_n it follows that, for every t and n , $\mathfrak{C}(t)$ is a subfield of $\mathfrak{C}_n(t)$ and that $\mathfrak{C}_m(t)$ is a subfield of $\mathfrak{C}_n(t)$ if m is larger than n . T_n approaches T from above as $n \rightarrow \infty$ so, from the second hypothesis, $P'\{\lim_{n \rightarrow \infty} x_n(t) = x'(t)\} = 1$ for each t . Therefore $x'(t)$ is $\mathfrak{C}_n(t)$ measurable for every n . Now if A is in $\mathfrak{B}(X)$ and a is non-negative then

$$\{x'(t) \in A\} \cap \{T < a\} = \bigcup_n [\{x'(t) \in A\} \cap \{T < a_n\}]$$

where $\{a_n\}$ is any sequence of rationals of the form $a_n = k/2^n$ which increase to a from below. The set on the right is $\bigcup_n [\{x'(t) \in A\} \cap \{T_n \leq a_n\}]$ and since $x'(t)$ is $\mathfrak{C}_n(t)$ measurable for each n this set is in $\mathfrak{F}(t+a)$. Thus $x'(t)$ is $\mathfrak{C}(t)$ measurable. For the proof that $\{x'(t), \mathfrak{C}(t)\}$ is Markovian with the right transition function, instead of showing that (1.2) holds we shall show, equivalently, that if ϕ_1 and ϕ_2 are bounded continuous functions on X and Δ is in $\mathfrak{C}(t)$, then

$$(1.4) \quad \int_{\Delta} \left\{ \int_X \phi_1(y_1) \left\{ \int_X \phi_2(y_2) P(t_2 - t_1, y_1, dy_2) \right\} P(t_1 - t, x'(t, \omega), dy_1) \right\} P'(d\omega) \\ = \int_{\Delta} \phi_1(x'(t_1, \omega)) \phi_2(x'(t_2, \omega)) P'(d\omega).$$

(That (1.4) implies (1.2) follows for B_1 and B_2 either open or closed sets from the remark just preceding the statement of Theorem 1.1 and Lebesgue's bounded convergence theorem and hence for arbitrary B_1 and B_2 in $\mathfrak{B}(X)$ from the usual theorem on extension of measures. That (1.2) implies (1.4) is equally trivial.) Now for each n , (1.4) holds with x' replaced by x_n since $\{x_n(t), \mathfrak{C}_n(t)\}$ is a Markov process with $P(t, x, A)$ as transition function and Δ is in $\mathfrak{C}_n(t)$. So letting $n \rightarrow \infty$ and using property (D) and the fact that, for almost all ω in Ω' , $x_n(t) \rightarrow x'(t)$, $x_n(t_i) \rightarrow x'(t_i)$, $i = 1, 2$, we immediately obtain the validity of (1.4). The fact that $\{x'(t), \mathfrak{C}(t)\}$ is Markovian implies that relative to $x'(0)$, the process $\{x'(t)\}$ is conditionally independent of any

random variable depending on the original process only up to time T . Thus the theorem is proved.

We note that the hypothesis immediately preceding (D) in the statement of Theorem 1.1 will be fulfilled whenever almost all sample functions of $\{x(t)\}$ are continuous on the right.

As a trivial application of some of the foregoing results we have a

ZERO-ONE LAW. *Suppose $\{x(t), \mathfrak{F}(t)\}$ has right continuous sample functions, that $x(0)$ is constant with probability 1, and that condition (D) holds. Then if Δ is in $\mathfrak{G}(\delta)$ for every $\delta > 0$ we have either $P\{\Delta\} = 1$ or $P\{\Delta\} = 0$.*

Proof. Let $T \equiv 0$. All the hypotheses of Theorem 1.1 are fulfilled and so relative to $x'(0)$, $\mathfrak{K}(0)$ and the process $\{x'(t)\}$ are conditionally independent where $\mathfrak{K}(0)$ consists of those subsets Γ of Ω such that $\Gamma \cap \{T < a\}$ is in $\mathfrak{F}(a)$. But $x'(0) = x(0)$ which is a constant so the conditional independence becomes full independence and $\{x(t)\} = \{x'(t)\}$ so Δ , which is $\mathfrak{K}(0)$ measurable, is also measurable on the sample space of $\{x'(t)\}$. Hence Δ is independent of itself so $P\{\Delta \cap \Delta\} = P\{\Delta\}P\{\Delta\}$ and $P\{\Delta\}$ must then be 0 or 1.

EXAMPLES RELEVANT TO THEOREM 1.1. Let $\{x(t)\}$ be the Poisson process with the sample functions normalized to be continuous on the left and with the usual definition of the transition function, and let $T(\omega)$ be the infimum of those t for which $x(t, \omega) = 1$. Then condition (D) is satisfied and T is a stopping time, but T is obviously not a Markov time for the process. Therefore the right continuity requirement can not be eliminated from the hypotheses of Theorem 1.1. In §3 we shall consider replacing this right continuity requirement with more stringent requirements on the stopping time and transition function. A less reasonable example is obtained by letting $\{x(t)\}$ be Brownian motion with $x(0) = 0$ and defining $P(t, x, A)$ as usual unless $x = 1$ in which case we set $P(t, 1, A) = 1$ or 0 according to whether 1 is or is not in A . $P\{x(t) = 1\} = 0$ for each t , so $P(t, x, A)$ serves perfectly well as a transition function for this process. If T is defined as in the first example then once again T is not a Markov time. In this example (D) is not satisfied.

REMARK 1.1. If, to be specific, $X = R^n$ and $P(t, x, A)$ is such that for any probability distribution Q on X one can construct a process with Q as initial distribution, with $P(t, x, A)$ as transition function, and with independent increments then, as one easily verifies, $P(t, x, A)$ must satisfy condition (D). Moreover under slight additional conditions any separable process arising from this construction can be normalized in such a way that its sample functions are right continuous; (see Doob [2, p. 388]). The familiar additive processes with stationary increments arise in this way and hence it follows from Theorem 1.1 that for any such process (with sample functions taken to be right continuous) any stopping time is a Markov time. As we commented earlier this result was first obtained by Hunt [1] (and, as he points out, holds for spaces more general than R^n). We state it here merely for completeness.

REMARK 1.2. One should note that Lemma 1.1 holds without any assumption about X or about $P(t, x, A)$ (other than the basic one that X has a Borel field $\mathfrak{B}(X)$ in terms of which the process and its transition function are defined). Let us assume for the moment merely that X is topological with $\mathfrak{B}(X)$ its topological Borel field. Then the passage from Lemma 1.1 to Theorem 1.1 can be carried out with essentially no change in the proof given above provided X has the following property: for any n , a finite measure π on $\mathfrak{B}(X^n)$ is determined by a knowledge of $\int_{x^n} \phi d\pi$ for every bounded continuous function ϕ . Any metric space has this property.

REMARK 1.3. Suppose $\{x(t), \mathfrak{G}(t)\}$ is a Markov process and T' is the first passage time to a compact set (see the preliminaries). Occasionally one is interested (see Hunt [2]) in knowing whether $T = \min(T', K)$ is a Markov time where K is a positive random variable which is independent of the process $\{x(t)\}$. If \mathfrak{K} denotes the smallest Borel field with respect to which K is measurable and we let $\mathfrak{F}(t)$ be the Borel field generated by $\mathfrak{G}(t)$ and \mathfrak{K} , then we see easily that $\{x(t), \mathfrak{F}(t)\}$ is a Markov process and that T is a stopping time for this process. Thus the considerations of Lemma 1.1 and Theorem 1.1 apply to this stopping time. If $\{x(t), \mathfrak{F}(t)\}$ is any Markov process for which Q is a stopping time and $\{y(t)\}$ is obtained as a function of the original process, say $y(t) = g(x(t))$, then the considerations of this section apply to Q and the new process provided that $\{y(t), \mathfrak{F}(t)\}$ is also a Markov process. These examples indicate the usefulness of the specialization of the usual Markov process definition which we have adopted.

2. Before proceeding further we must derive certain facts about the behavior of Markov process sample functions from hypotheses on the transition function. For our discussion of sample function behavior, and from now on in this paper, we shall assume that X is a separable, locally compact Hausdorff space. X is then metrizable so the considerations of §1 apply. We select once and for all any one from among the admissible metrics on X and denote it by ρ . For a sequence x_n from X we say that $x_n \rightarrow \infty$ if given any compact subset C of X there is an $N(C)$ such that $x_n \notin C$ if $n > N(C)$. A subset A of X is said to be bounded if it is contained in a compact subset of X . We note that there is an increasing sequence $\{D_n\}$ of compact subsets of X such that any compact subset of X is contained in all but a finite number of the D_n .

All that is important throughout this section is that $\{x(t)\}$ be a Markov process relative to $\{\mathfrak{G}(t)\}$. From now on we shall assume that the Markov process $\{x(t)\}$ is separable relative to the compact subsets of X (which we shall denote by saying merely " $\{x(t)\}$ is separable.") We select once and for all a fixed, but otherwise arbitrary, t set satisfying the conditions of the separability definition and denote it by S . The assumption of separability imposes no restriction on the transition function of the Markov process in the sense that given any process $\{x(t)\}$ in X there is a separable process $\{y(t)\}$ such that $P\{x(t) = y(t)\} = 1$ for each t . ($y(t)$ may take on values in the compactification of X although of course $P\{y(t) \in X\} = 1$ for each t .) If

$\{x(t)\}$ is Markovian with $P(t, x, A)$ as transition function it follows that $\{y(t)\}$ is also. (To see that such a process $\{y(t)\}$ exists one has merely to observe that the proof of this fact given by Doob [2, Theorem 2.4] for the case in which X is the real line carries over essentially word for word to the case in which X is merely assumed separable and locally compact.) The results of the first part of §2 are to a large extent due to Kinney [1]. In particular Theorems 2.1 and 2.2 below correspond to Kinney's Theorems IV (ii) and VI (i) although our hypotheses are formally at least a little different than his. Our proofs of these two theorems are essentially those given by Kinney although they are couched in slightly different terminology. Dynkin [1] also obtains the conclusion of Theorem 2.2 under hypotheses somewhat stronger than Kinney's.

We shall now list various conditions, on the transition function and sample functions, to which we shall later refer:

(E) For every $\delta > 0$ and compact subset C of X there is a $K(\delta, C)$ such that $P(t, x, I_x^\delta) \geq 1 - Kt$ for every $t \geq 0$ and x in C (where $I_x^\delta = \{y \in X \mid \rho(x, y) < \delta\}$).

(E') For every $\delta > 0$ and compact subset C of X , $P(t, x, I_x^\delta) \rightarrow 1$ uniformly in x in C as $t \rightarrow 0$.

(E'') For every $\delta > 0$ and x in X , $P(t, x, I_x^\delta) \rightarrow 1$ as $t \rightarrow 0$.

(F) There is an $r > 0$ such that $\sup_{t \leq r, x \in D_n} P(t, x, C) \rightarrow 0$ as $n \rightarrow \infty$ whenever C is a compact subset of X .

(G) With probability 1 the sample functions have left-hand and right-hand limits at every value of t ; that is, there is a subset Δ of Ω with $P\{\Delta\} = 1$ such that if ω is in Δ and t is any positive number then $\lim_{s \uparrow t} x(s, \omega)$ exists and $\lim_{s \downarrow t} x(s, \omega)$ exists.

(G') For each t , $P\{\lim_{s \uparrow t} x(s) \text{ exists}\} = P\{\lim_{s \downarrow t} x(s) \text{ exists}\} = 1$. In (G) and (G') by "limits" we mean limits in X . One should note that given a transition function $P(t, x, A)$, the conditions (E), (E') and (E'') do not depend on ρ , nor does (F) depend on the particular choice of the increasing sequence $\{D_n\}$ of compact sets.

THEOREM 2.1. *Under condition (F) almost all sample functions are bounded on any finite t interval.*

Proof. If X is compact then all sample functions are, trivially, bounded so the theorem holds in this case. Otherwise suppose the conclusion is false. Then there is a $t' > 0$ such that with positive probability the sample functions are unbounded in $[0, t']$. Hence we can find a $\delta > 0$ and an interval $[s_1, s_2]$ in $[0, t']$ with $s_2 - s_1$ not exceeding the constant r appearing in (F) such that with probability greater than δ the sample functions are unbounded in (s_1, s_2) . Let C be a fixed compact subset of X and let D_k be a member of our increasing sequence of compact sets. Insert in (s_1, s_2) a finite number of points $t_1 < \dots < t_n$ of S and define

$$T(\omega) = \begin{cases} s_2 & \text{if } x(t_i, \omega) \in D_k \text{ for all } i \leq n, \\ \text{smallest value of } t_i & \text{for which } x(t_i, \omega) \notin D_k, \text{ otherwise.} \end{cases}$$

Then T is finitely valued with $\{T \leq a\}$ in $\mathcal{G}(a)$ and it follows easily (or, if one likes, is a consequence of Lemma 1.1) that

$$(2.1) \quad P\{x(s_2) \in C\} = \int_{\Omega} P(s_2 - T(\omega), x(T(\omega), \omega), C)P(d\omega).$$

The right side of (2.1) is

$$(2.2) \quad \int_{T < s_2} P(s_2 - T, x(T), C)P(d\omega) + \int_{T = s_2} P(s_2 - T, x(T), C)P(d\omega) \\ \leq \delta_k P\{x(t_i) \notin D_k \text{ for some } t_i\} + P\{x(t_i) \in D_k \text{ for all } t_i\}$$

where $\delta_k = \sup_{t \leq r, x \notin D_k} P(t, x, C)$. This estimate does not depend on the number or position of the points t_i in (s_1, s_2) , so letting n increase and t_1, \dots, t_n increase to include all of $S \cap (s_1, s_2)$ we have

$$(2.3) \quad P\{x(s_2) \in C\} \leq P\{x(t) \in D_k \text{ for every } t \text{ in } S \cap (s_1, s_2)\} + \delta_k$$

and by the separability the right side is

$$(2.4) \quad P\{x(t) \in D_k \text{ for every } t \text{ in } (s_1, s_2)\} + \delta_k \leq 1 - \delta + \delta_k.$$

By (F), $\lim_{k \rightarrow \infty} \delta_k = 0$, so $P\{x(s_2) \in C\} \leq 1 - \delta$ regardless of the choice of compact subset C of X . This is impossible and thus the theorem is proved.

Let us assume for Lemma 2.1 and Theorem 2.2, that X is not compact. If X is compact we delete the now absurd condition (F) and, with obvious modifications in the proofs, all the results remain valid.

LEMMA 2.1. *Suppose (E') and (F) hold, that C is a compact subset of X , and that $P\{x(0) \notin C\} < \delta$. Then given any $\epsilon > 0$ and $\eta > 0$ there is a $t'(\epsilon, \eta, C) > 0$ (not depending on the distribution of $x(0)$) such that $P\{\sup_{t < t'} \rho(x(0), x(t)) > \eta\} < \epsilon + 2\delta$.*

Proof. We can find a strictly positive η' less than η such that the set of points whose distance from C is less than or equal to η' is compact. Call this set D . Let $\delta_k = \sup_{t \leq r, x \notin D_k} P(t, x, D)$. By (F) we can pick k large enough that D_k contains D and $\delta_k < \epsilon/3$. By (E') we can then select a t' strictly positive but less than r with the property that $P(t, x, I_x^{\eta'/2}) > 1 - \epsilon/3$ for every $t \leq t'$ and x in D_k . Now insert in the interval $(0, t')$ a finite number of points $t_1 < \dots < t_n$ of S and define

$$T(\omega) = \begin{cases} t' & \text{if } \rho(x(0, \omega), x(t_i, \omega)) \leq \eta' \text{ for all } i \leq n, \\ \text{smallest value of } t_i & \text{for which } \rho(x(0, \omega), x(t_i, \omega)) > \eta', \text{ otherwise.} \end{cases}$$

As in the proof of Theorem 2.1 we have

$$\begin{aligned}
 & P \left\{ \rho(x(0), x(t')) < \frac{\eta'}{2} \right\} \\
 &= \int_{\Omega} P(t' - T(\omega), x(T(\omega), \omega), I_{x(0, \omega)}^{\eta'/2}) P(d\omega) \\
 (2.5) \quad &= \int_{x(0) \notin C} + \int_{\substack{x(0) \in C, T < t' \\ x(T) \notin D_k}} + \int_{\substack{x(0) \in C, T < t' \\ x(T) \in D_k}} + \int_{x(0) \in C, T = t'} \\
 &\leq \delta + \frac{\epsilon}{3} + \frac{\epsilon}{3} + P\{\rho(x(0), x(t_i)) \leq \eta' \text{ for all } t_i\}.
 \end{aligned}$$

Letting (t_1, \dots, t_n) increase to include all of $S \cap (0, t')$, using the separability and the fact that the left side of (2.5) is greater than $(1 - \delta)(1 - \epsilon/3)$ we obtain

$$(2.6) \quad P \left\{ \sup_{t < t'} \rho(x(0), x(t)) > \eta' \right\} < \epsilon + 2\delta.$$

In obtaining (2.6) we have used no knowledge of the distribution of $x(0)$ other than the fact that $P\{x(0) \notin C\} < \delta$. Since $\eta' \leq \eta$ the proof is complete.

LEMMA 2.2. *If in Lemma 2.1 conditions (E') and (F) are replaced by (E) the conclusion remains valid.*

Proof. Let D and η' be as in the proof of Lemma 2.1. By (E) we can select a $K(\eta'/4, D)$ such that $P(t, x, I_x^{\eta'/4}) \geq 1 - Kt$ for every t and x in D . Pick t' so small that $Kt' < \epsilon/3$, insert in $(0, t')$ a finite number of points $t_1 < \dots < t_n$ of S , and define

$$T(\omega) = \begin{cases} t' & \text{if } \rho(x(0, \omega), x(t_i, \omega)) \leq \eta'/2 \text{ for all } i \leq n, \\ \text{smallest value of } t_i \text{ for which } \rho(x(0, \omega), x(t_i, \omega)) > \eta'/2, & \text{otherwise.} \end{cases}$$

As in the two previous proofs we then have

$$\begin{aligned}
 & P \left\{ \rho(x(0), x(t')) \leq \frac{\eta'}{4} \right\} \\
 &\leq P\{x(0) \notin C\} \\
 (2.7) \quad &+ P \left\{ \rho(x(0), x(t_i)) \leq \frac{\eta'}{2} \text{ for all } t_i, x(0) \in C \right\} \\
 &+ P \left\{ x(T) \in D, \rho(x(T), x(t')) > \frac{\eta'}{4} \right\} + P\{x(0) \in C, x(T) \notin D\}.
 \end{aligned}$$

The first summand is by hypothesis less than δ and, since $t' - T \leq t'$, the third is less than $\epsilon/3$. The fourth summand is equal to

$$\begin{aligned}
 (2.8) \quad & \sum_{i=1}^n P\{x(0) \in C, x(T) \notin D, T = t_i\} \\
 & \cong \sum_{i=1}^n P\left\{\rho(x(t_{i-1}), x(t_i)) > \frac{\eta'}{4}, x(t_{i-1}) \in D\right\} \\
 & \cong \sum_{i=1}^n K(t_i - t_{i-1}) < Kt' < \frac{\epsilon}{3}
 \end{aligned}$$

where $t_0=0$. The left side of (2.7) is greater than $(1-\delta)(1-\epsilon/3)$ so letting (t_1, \dots, t_n) increase to include all of $S \cap (0, t')$ and using the separability we obtain

$$(2.9) \quad P\left\{\sup_{t < t'} \rho(x(0), x(t)) > \frac{\eta'}{2}\right\} \cong \epsilon + 2\delta.$$

Since $\eta'/2 < \eta$ and no knowledge of the distribution of $x(0)$ has been used the proof is complete.

THEOREM 2.2. *If conditions (E') and (F) hold then the process satisfies condition (G).*

Proof. Suppose it is not true that almost all sample functions have right and left hand limits at every value of t . By Theorem 2.1 almost all sample functions are bounded on finite intervals so the nonexistence of limits is due to oscillatory behavior. Therefore we can find a $t' > 0$ and $\epsilon > 0$ such that for every ω in a set Γ of strictly positive probability there is a $\tau(\omega) < t'$ such that $x(t, \omega)$ oscillates infinitely often (i.o.) by more than ϵ either as t decreases to $\tau(\omega)$ or as t increases to $\tau(\omega)$. Let us select a compact C so large that $\Gamma \cap \{x(t) \in C \text{ for every } t \text{ in } [0, t']\}$ has strictly positive probability and such that the set of points whose distance from C does not exceed ϵ is compact. (To achieve this it may be necessary to decrease ϵ .) Now select a strictly positive δ small enough that if $\{y(t)\}$ is any separable Markov process with $P(t, x, A)$ as transition function and $y(0)$ confined to C then $P\{\sup_{t < \delta} \rho(y(0), y(t)) \geq \epsilon/2\} \leq 1/2$. This is possible by Lemma 2.1. From what we have done so far it then follows that we can find an interval $(\sigma, \sigma + \delta)$ contained in $[0, t']$ and a set Δ of strictly positive probability such that $x(t, \omega)$ is in C for every ω in Δ and t in $[0, t']$ and such that for every ω in Δ there is a sequence $\{s_n\}$ from $S \cap (\sigma, \sigma + \delta)$ which is either increasing or decreasing and for which $\rho(x(s_n, \omega), x(s_{n+1}, \omega)) > \epsilon$ for every n . For each n select a subset $S_n = (s_{n1} < \dots < s_{nn})$ of $S \cap (\sigma, \sigma + \delta)$ such that $S_n \subset S_{n+1}$ and $\bigcup_n S_n = S \cap (\sigma, \sigma + \delta)$. For each k and n with $k < n$ define $\Lambda_{k,n}$ as follows: ω is in $\Lambda_{k,n}$ if there are points $t_1 < t_{\mu_1} \leq t_2 < t_{\mu_2} \leq \dots \leq t_k < t_{\mu_k}$ from S_n such that $\rho(x(t_i, \omega), x(t_{\mu_i}, \omega)) > \epsilon$ for every $i \leq k$. Obviously $\Lambda_{k,n}$ is a subset of $\Lambda_{k,n+1}$ and if $\Lambda_k = \bigcup_n \Lambda_{k,n}$ then Δ is a subset of Λ_k . Now fix k and n , set $T_0 = s_{n1}$ and for $j=0, \dots, k-1$ define recursively the random variables

$$T_{j+1} = \begin{cases} s_{nn} & \text{if } \rho(x(T_j), x(s)) < \epsilon/2 \text{ for every } s \in S_n \text{ such that } s > T_j, \\ \text{smallest } s \in S_n \text{ such that } s > T_j \text{ and } \rho(x(T_j), x(s)) \geq \epsilon/2, & \text{otherwise.} \end{cases}$$

Each T_j is a finitely valued stopping time and $T_0(\omega) < \dots < T_k(\omega)$ if ω is in $\Lambda_{k,n}$. Let $x_{k-1}(t) = x(t + T_{k-1})$. Then

$$\begin{aligned} P\left\{\Delta \cap \bigwedge_{k,n}\right\} &\leq P\{T_0 < \dots < T_k, x(T_i) \in C \text{ for } i \leq k-1\} \\ &\leq \int_{\left\{T_0 < \dots < T_{k-1}, x(T_i) \in C \text{ for } i \leq k-1\right\}} P\left\{\sup_{t \leq \delta} \rho(x_{k-1}(0), x_{k-1}(t))\right. \\ (2.10) \quad &\geq \frac{\epsilon}{2} \mid x_{k-1}(0)\left.\right\} P(d\omega) \\ &\leq \frac{1}{2} P\{T_0 < \dots < T_{k-1}, x(T_i) \in C \text{ for } i \leq k-2\} \end{aligned}$$

and proceeding in this way we find that

$$P\left\{\Delta \cap \bigwedge_{k,n}\right\} \leq \left(\frac{1}{2}\right)^k.$$

Letting $n \rightarrow \infty$ we obtain $P\{\Delta\} = P\{\Delta \cap \bigwedge_k\} \leq (1/2)^k$. Since k is arbitrary we have $P\{\Delta\} = 0$ and this contradiction proves the theorem.

REMARK 2.1. It should be noted that in the hypotheses of Theorem 2.2 we can replace condition (F) by the weaker condition that "almost all sample functions of $\{x(t)\}$ are bounded on finite intervals" for, with this replacement, by an adaptation of the proof of Lemma 2.1 we can still obtain the validity of equation (2.10) and this is the critical point in the proof of Theorem 2.2.

THEOREM 2.3. *If conditions (D), (E'') and (G') hold then $\{x(t)\}$ has no stationary discontinuities.*

Proof. It follows immediately from (E'') and (G') that at every value of t the sample functions are right continuous with probability 1. So suppose there is a stationary discontinuity at $t' > 0$. Let $a(\omega) = \lim_{t \uparrow t'} x(t, \omega)$ and $b(\omega) = \lim_{t \downarrow t'} x(t, \omega)$. Of course $b(\omega) = x(t', \omega)$ for almost all ω . Since there is a stationary discontinuity there are strictly positive ϵ and δ such that $P\{\rho(a(\omega), b(\omega)) > 5\epsilon\} > 3\delta$. Also there is a strictly positive τ and a set Δ with $P\{\Delta\} < \delta$ such that $\rho(x(t, \omega), a(\omega)) < \epsilon$ and $\rho(x(s, \omega), b(\omega)) < \epsilon$ for every t in $[t' - \tau, t')$, s in $[t', t' + \tau]$ and ω not in Δ . Then $P\{\rho(x(t), x(s)) > 3\epsilon\} > 2\delta$ if t and s are as in the preceding sentence. Thus we have

$$(2.11) \quad 1 - 2\delta > P\{\rho(x(t), x(s)) \leq 3\epsilon\} \geq \int_{\Omega} P(s - t, x(t, \omega), I_{a(\omega)}^{2\epsilon}) P(d\omega) - \delta.$$

Let c be any strictly positive number less than τ and take sequences of points s_k from $[t', t' + \tau]$ and t_k from $[t' - \tau, t']$ such that $s_k - t_k = c$ and such that t_k increases to t' as $k \rightarrow \infty$. Then

$$(2.12) \quad \begin{aligned} 1 - \delta &\geq \liminf_{k \rightarrow \infty} \int_{\Omega} P(s_k - t_k, x(t_k, \omega), I_{a(\omega)}^{2\epsilon}) P(d\omega) \\ &\geq \int_{\Omega} \liminf_{k \rightarrow \infty} P(c, x(t_k, \omega), I_{a(\omega)}^{2\epsilon}) P(d\omega). \end{aligned}$$

Now I_y^ϵ is for any ϵ and y an open set and hence by (D) $P(t, x, I_y^\epsilon)$ is lower semi-continuous in x . Therefore

$$(2.13) \quad 1 - \delta \geq \int_{\Omega} P(c, a(\omega), I_{a(\omega)}^{2\epsilon}) P(d\omega).$$

This holds regardless of the choice of c and so letting $c \rightarrow 0$ we obtain a contradiction. Thus the theorem is proved.

THEOREM 2.4. *If in Theorem 2.3 conditions (D) and (E'') are replaced by (E') the conclusion remains valid.*

Proof. From (E') and (G') it follows that at every value of t the sample functions are right continuous with probability 1. So suppose there is a stationary discontinuity at $t' > 0$. It follows then, using the notation of the previous proof, that $P\{\rho(x(t), x(s)) \geq 3\epsilon\} > 3\delta$ for every t in $[t' - \tau, t']$ and s in $[t', t' + \tau]$. Select a compact set C so large that $P\{x(t) \in C\} > 1 - \delta$ for every t in $[t' - \tau, t']$; (by (G') this can be done although a readjustment of τ might first be necessary). Then by (E') we can select a $\tau' > 0$ such that $P(t, x, I_x^{3\epsilon}) > 1 - \delta$ for every $t < \tau'$ and x in C . Then for every t in $[t' - \tau, t']$ and s such that $s - t < \tau'$ we have

$$(2.14) \quad \begin{aligned} P\{\rho(x(t), x(s)) < 3\epsilon\} &= \int_{\Omega} P(s - t, x(t, \omega), I_{x(t, \omega)}^{3\epsilon}) P(d\omega) \\ &> (1 - \delta)^2 > 1 - 2\delta \end{aligned}$$

which contradicts the third sentence of the proof.

REMARK 2.2. One can easily verify that if a process satisfies condition (G) and has no stationary discontinuities it can then be normalized in such a way that almost all its sample functions are continuous on the right or in such a way that almost all its sample functions are continuous on the left. Thus the theorems of this section provide some information about when such normalization is possible.

3. To simplify the discussion in this section we shall cease, as of now, to consider the question of whether the new process $\{x(t+T)\}$ is conditionally independent of the original one considered up to time T and will concern

ourselves only with whether the new process is Markovian relative to $\{\mathcal{G}'(t)\}$ with the same transition function as the original process.

The example given in §1 of the Poisson process with left-continuous sample functions imposes some restriction on the weakening of the hypotheses of Theorem 1.1. For example, it is not enough to assume merely that $\{x(t)\}$ is separable. To obtain the conclusion of Theorem 1.1 for separable processes we must assume more about the stopping time and the transition function. This is the content of Theorem 3.1. Before stating Theorem 3.1 I make a simple observation which will be useful in its proof: if π is a finite measure on $\mathfrak{B}(X)$ then an open set A is called a continuity set for π if $\pi(A) = \pi(\bar{A})$. (\bar{A} means the closure of A .) Suppose B is a subset of X and B_r is the set of points in X whose distance from B is strictly less than r . If r is not equal to q then $\bar{B}_r - B_r$ and $\bar{B}_q - B_q$ are disjoint and so $\pi(\bar{B}_r - B_r) = 0$ for all but a countable number of values of r . Therefore if B is a closed set we can find a sequence $r_n \rightarrow 0$ such that B_{r_n} is a continuity set for every n , so B is the intersection of a decreasing sequence of continuity sets. Also in a metric space any open set is the union of a sequence of closed sets, and the sets which are simultaneously countable unions of closed sets and countable intersections of open sets form a field which generates $\mathfrak{B}(X)$. Therefore if μ is any finite measure on $\mathfrak{B}(X)$ it is determined by its values on the continuity sets of π .

THEOREM 3.1. *Suppose $\{x(t), \mathcal{F}(t)\}$ is separable and (D), (E') and (F) hold. Suppose T is an almost everywhere finite positive random variable such that $T = \lim_{n \rightarrow \infty} T_n$ where each T_n is a stopping time with $P\{T_n < T\} = 1$. Then $\{x'(t)\} (= \{x(t+T)\})$ is a Markov process with $P(t, x, A)$ as transition function.*

Proof. That T is actually a stopping time follows trivially from the fact that $T = \lim_{n \rightarrow \infty} T_n$, but this is of no consequence. For each positive n and k define

$$Q_{k,n}(\omega) = \frac{j+1}{k} \text{ if } \frac{j}{k} \leq T_n(\omega) < \frac{j+1}{k}.$$

Then $Q_{k,n}$ is for each k, n countably valued and almost everywhere finite with $\{Q_{k,n} \leq a\}$ in $\mathcal{F}(a)$. Now $P\{T_n < T\} = 1$ and $|T_n - Q_{k,n}| \leq 1/k$ so there is an integer $k(n)$ such that $P\{Q_{k,n} \geq T\} < 1/2^n$ for all $k \geq k(n)$. Set $Q_{k(n),n} = Q_n$. Then the Q_n are countably valued and almost everywhere finite with $\{Q_n \leq a\}$ in $\mathcal{F}(a)$ and $\lim_{n \rightarrow \infty} Q_n = T$ with probability 1. Also by the Borel-Cantelli lemma there is, for almost all ω , an $N(\omega)$ such that $Q_n(\omega) < T(\omega)$ if $n > N(\omega)$. Now for each n define $\{x_n(t)\}$ as usual by $x_n(t) = x(t+Q_n)$. Each of these processes is separable and, by Lemma 1.1, Markovian with $P(t, x, A)$ as transition function. Given an $\epsilon > 0$ we select R so large that $P\{T > R\} < \epsilon/2$. By Theorem 2.1 we can find a compact subset C of X such that $P\{x(t) \notin C \text{ for some } t \leq R\} < \epsilon/2$. Then for each n , $P\{x_n(0) \notin C\} < \epsilon$. It is therefore an immediate consequence of Lemma 2.1 that $P\{\lim_{n \rightarrow \infty} x_n(0)$

$=x'(0)\} > 1 - \epsilon$ and since ϵ was arbitrary this probability is 1. In exactly the same way for any finite set s_1, \dots, s_k of parameter values we consider the stopping times $Q_n + s_i$ and show that $P\{\lim_{n \rightarrow \infty} x_n(s_i) = x'(s_i) \text{ for all } i \leq k\} = 1$. This shows at least that each $x'(t)$ is a random variable. Now the argument proceeds almost exactly as in the proof of Theorem 1.1. Each side of (1.4) in §1 defines a finite completely additive set function of Δ and so by the remark just preceding the statement of Theorem 3.1 it suffices to show that (1.4) holds when Δ is of the form $\{x'(s_1) \in B_1, \dots, x'(s_k) \in B_k\}$ where B_1, \dots, B_k are continuity sets for the distributions in X of $x'(s_1), \dots, x'(s_k)$ respectively and $0 \leq s_1 < \dots < s_k \leq t$. Let $\Delta_n = \{x_n(s_1) \in B_1, \dots, x_n(s_k) \in B_k\}$. By Lemma 1.1, (1.4) holds with x' replaced by x_n and Δ replaced by Δ_n . Now the characteristic function of Δ_n tends to that of Δ with probability 1 as $n \rightarrow \infty$ so passing to the limit as in the proof of Theorem 1.1 we obtain the validity of (1.4). This completes the proof of Theorem 3.1.

Once again in Theorem 3.1 as in Theorem 2.2 the hypothesis that (F) holds can be replaced by "almost all sample functions are bounded on finite intervals."

The assumption in Theorem 3.1 that T is almost everywhere finite is not essential. If $T_n \leq T_{n+1}$, $T = \lim_{n \rightarrow \infty} T_n$, and $T_n < T$ on Ω' , the sets where T is finite, then one can show that $\{x'(t)\}$ is a Markov process on $(\Omega', \mathfrak{B}', P')$ with $P(t, x, A)$ as transition function. In this case the above proof is complicated by the fact that $\{T_n < \infty\} \neq \{T < \infty\}$.

4. Suppose $\{T_n\}$ is a sequence of stopping times and T is a stopping time such that $\lim_{n \rightarrow \infty} T_n = T$ for almost all ω in Ω' , the set where T is finite. It is useful in some applications of probability to potential theory to be able to assert that in this case $x(T_n) \rightarrow x(T)$ for almost all ω in Ω' ; (see Hunt [2]). Therefore one would like to have conditions which guarantee that this assertion is valid. Of course it is always valid if the sample functions are continuous. We shall concern ourselves chiefly with increasing sequences of stopping times.

THEOREM 4.1. *Suppose almost all sample functions of $\{x(t), \mathfrak{F}(t)\}$ are right continuous and (D) and (E) hold. Suppose that the sequence $\{T_n\}$ is increasing and that for every $\delta > 0$ there is a compact set $C(\delta)$ such that $P\{T_n < \infty, x(T_n) \in C\} < \delta$ for every n . Then $\lim_{n \rightarrow \infty} x(T_n) = x(T)$ for almost all ω in Ω' .*

The last hypothesis will be satisfied if, for example, T_n is the smallest value of t for which $x(t)$ is distant from a compact set A by less than or equal to $1/n$. This kind of sequence arises in applications.

Proof. Let $\Omega_n = \{T_n < \infty\}$ and define $\{x_n(t)\}$ over $(\Omega_n, \mathfrak{B}_n, P_n)$ in terms of T_n . By the right continuity of the sample functions and (D) it follows from Theorem 1.1 that $\{x_n(t)\}$ is for each n a Markov process with $P(t, x, A)$ as transition function. Now $P_n\{x_n(0) \in C\} < M\delta$ where $M = 1/P\{\Omega'\}$ and so by Lemma 2.2 given any $\epsilon > 0$ and $\eta > 0$ there is a $t' > 0$ such that

$P_n \{ \sup_{t < t'} \rho(x_n(0), x_n(t)) > \eta \} < \epsilon + 2M\delta$ for every n . And so, since Ω' is a subset of Ω_n we have

$$P' \left\{ \Omega' \cap \left\{ \sup_{t < t'} \rho(x_n(0), x_n(t)) > \eta \right\} \right\} < M\epsilon + M^2\delta.$$

Since T_n increases to T on Ω' and ϵ and δ are arbitrary it follows immediately that $P' \{ \Omega' \cap \{ \lim_{n \rightarrow \infty} x(T_n) = x(T) \} \} = 1$.

The condition that (D) holds can be eliminated from the hypotheses of Theorem 4.1. This is done by replacing the sequence $\{T_n\}$ by a suitably constructed sequence $\{Q_n\}$ of countably valued stopping times and working with the processes $\{x(t + Q_n)\}$, to which Lemma 1.1 applies. We shall not carry this out.

In Theorem 4.2 the hypotheses concerning $P(t, x, A)$ are weakened while those concerning sample function behavior are for all practical purposes strengthened. The method of proof differs somewhat from that employed in Theorem 4.1.

THEOREM 4.2. *If (D), (E'') and (G) hold and the sequence $\{T_n\}$ is increasing then $\lim_{n \rightarrow \infty} x(T_n) = x(T)$ for almost all ω in Ω' .*

Proof. We first need some manipulations with the stopping times: $\{T_n\}$ is increasing so it has a limit T' (possibly infinite) and $T' = T$ for almost all ω in Ω' . We redefine T to be T' . This changes essentially nothing in the problem at hand. Now let $D_n = \{T_n = T\}$. Obviously D_n is a subset of D_{n+1} . Let R be some fixed positive number and $\Delta_n = \{T_n \geq R\}$. Then Δ_n is a subset of Δ_{n+1} . Let $D = \bigcup_n D_n$ and $\Delta = \bigcup_n \Delta_n$ and define

$$Q(\omega) = \begin{cases} R + 1 & \text{if } \omega \in (D \cup \Delta), \\ T(\omega) & \text{if } \omega \in (D \cup \Delta)^* \end{cases}$$

(where $*$ denotes complement in Ω). For each pair k, n of positive integers define

$$Q_{k,n}(\omega) = \begin{cases} R + 1 & \text{if } \omega \in (D_n \cup \Delta_n), \\ \frac{j + 1}{k} & \text{if } \frac{j}{k} \leq T_n(\omega) < \frac{j + 1}{k} \text{ and } \omega \in (D_n \cup \Delta_n)^*. \end{cases}$$

Then each $Q_{k,n}$ is countably valued with $\{Q_{k,n} \leq a\}$ in $\mathfrak{F}(a)$, although the verification of this is not quite as simple as in previous cases. This verification is unenlightening and hence is deferred to the end of the proof of Theorem 4.2. Proceeding now exactly as in the proof of Theorem 3.1 we construct a sequence $\{Q_n\}$ of countably valued stopping times such that for almost all ω , $Q_n \rightarrow Q$ and $Q_n < Q$ for all n large enough (how large depending on ω). By (G) and Theorem 2.3 we can normalize $\{x(t)\}$ so that it has left-continuous sample functions; that is, we can construct a process $\{y(t)\}$ almost all of

whose sample functions are left-continuous and such that $P\{x(t)=y(t)\}=1$ for every t . $\{y(t), \mathfrak{F}(t)\}$ is then a Markov process with the same transition function as $\{x(t)\}$. Now $\{Q_n \leq a\}$ is in $\mathfrak{F}(a)$ so $\{y_n(t)\} (= \{y(t+Q_n)\})$ is by Lemma 1.1 Markovian with $P(t, x, A)$ as transition function. From the left-continuity of the sample functions and the fact that $Q_n \rightarrow Q$ essentially from below we conclude, exactly as in the proof of Theorem 3.1 that $\{y(t+Q)\}$ is Markovian with $P(t, x, A)$ as transition function. Having set up all this machinery we are now ready to complete the proof: suppose on a subset of Ω' of positive measure $x(T_n) \rightarrow x(T)$. The set of nonconvergence must clearly be a subset of D^* . Select R positive and large enough that $\{x(T_n) \rightarrow x(T), T < R\}$ has positive measure. Define Δ_n, Q_n, Q etc., as above in terms of this R . By the assumption of nonconvergence there is a subset A of $(\Delta \cup D)^* \cap \Omega'$ of positive measure such that $x(T_n(\omega), \omega) \rightarrow x(Q(\omega), \omega)$ if ω is in A . By (G) it then follows that for almost all ω in A , $x(Q_n(\omega), \omega) \rightarrow x(Q(\omega), \omega)$. (G) also implies that $\lim_{t \downarrow 0} x(t+Q)$ exists for almost all ω . Hence for almost all ω in A , $\lim_{n \rightarrow \infty} x(Q_n) \neq \lim_{t \downarrow 0} x(t+Q)$, for if the equality held on a subset of A of positive measure we would have a set of positive measure for which the above two limits exist and are equal, but where this limit is not $x(Q)$ and this would violate the separability of the $\{x(t)\}$ process (unless Q itself takes on values only in S in which case it is countably valued and the whole theorem becomes trivial). The result of all this is that we can find an $\epsilon > 0, \delta > 0$ and $t' > 0$ such that for every ω in a subset B of A with $P\{B\} > \delta$ we have

$$\rho\left(\lim_{n \rightarrow \infty} x(Q_n(\omega), \omega), x(t + Q(\omega), \omega)\right) > \epsilon,$$

for every t in $(0, t')$. Now $P\{x(Q_n)=y(Q_n)\}=1$ for each n so $\lim_{n \rightarrow \infty} x(Q_n) = \lim_{n \rightarrow \infty} y(Q_n) = y(Q)$ with probability 1. Also from the way $\{y(t)\}$ is constructed it follows that $\rho(y(Q), y(t+Q)) > \epsilon$ for every ω in B and t in $(0, t')$. Hence for any t in $(0, t')$ we have

$$(4.1) \quad \delta < P\{\rho(y(Q), y(Q+t)) \geq \epsilon\} = 1 - \int_{\Omega} P(t, y(Q), I_{y(Q)}) P(d\omega)$$

and this clearly violates (E''). All that remains now is the verification concerning $Q_{k,n}$: to proceed with this fix k, n and consider first

$$\{Q_{k,n} = R + 1\} = \{T_n \geq R\} \cup \{T_n = T, T_n < R\}.$$

The first set on the right is obviously in $\mathfrak{F}(R+1)$. Concerning the second set on the right, since $\{T_n < R\}$ is in $\mathfrak{F}(R+1)$ it suffices to show that $\{T_n < T, T_n < R\}$ is in $\mathfrak{F}(R+1)$. To this end let $\Gamma_r = \{T_n < r < T\}$. Then $\{T_n < T, T_n < R\} = \cup_r \Gamma_r$, the union being over all rational $r < R$. Each Γ_r is in $\mathfrak{F}(R+1)$ since T_n and T are stopping times so $\{T_n < T, T_n < R\}$ is in $\mathfrak{F}(R+1)$. Next consider $\{Q_{k,n} = (j+1)/k\}$ where of course $(j+1)/k < R+1/k$. As above

$$\left\{ Q_{k,n} = \frac{j+1}{K} \right\} = \begin{cases} \left\{ \frac{j}{k} \leq T_n < \frac{j+1}{k}, T_n < T \right\} & \text{if } \frac{j+1}{k} \leq R, \\ \left\{ \frac{j}{k} \leq T_n < R, T_n < T \right\} & \text{if } \frac{j}{k} < R < \frac{j+1}{K} \end{cases}$$

and in either case an argument similar to the one above shows that each set on the right is in $\mathfrak{F}(j+1/k)$. Thus the proof of Theorem 4.2 is complete.

REMARK 4.1. If we assume that almost all sample functions of $\{x(t)\}$ are right-continuous then the question of whether $x(T_n) \rightarrow x(T)$ for $\{T_n\}$ an arbitrary sequence of stopping times can be reduced to a question concerning an increasing sequence as follows: suppose $T_n \rightarrow T$ on Ω' . Define $T' = \liminf_{n \rightarrow \infty} T_n$. Then T' is a stopping time and $T' = T$ on Ω' . Let $K_n = \min(T_n, T')$. K_n is obviously a stopping time. Now define recursively

$$Q_1 = \inf_n K_n = \min_n K_n,$$

$$Q_n(\omega) = \begin{cases} T'(\omega) & \text{if } K_p(\omega) \leq Q_{n-1}(\omega) \text{ for every } p, \\ \inf_p K_p(\omega) = \min_p K_p(\omega) & \text{where for fixed } \omega \text{ the infimum is taken over} \\ & \text{all } p \text{ such that } K_p(\omega) > Q_{n-1}(\omega), \text{ otherwise.} \end{cases}$$

From the definition of T' and K_n it follows that $K_n \rightarrow T'$ from below and hence the infima in the above definitions are always attained. It can be shown, with some labor, that for each n , Q_n is a stopping time. Also $Q_n \leq Q_{n+1}$ and $Q_n \rightarrow T'$. From the right-continuity of the sample functions it follows that if for some ω in Ω' , $x(T_n(\omega), \omega) \rightarrow x(T(\omega), \omega)$ then $x(Q_n(\omega), \omega) \rightarrow x(T(\omega), \omega)$ and hence if $\lim_{n \rightarrow \infty} x(Q_n) = x(T)$ for almost ω in Ω' then $\lim_{n \rightarrow \infty} x(T_n) = x(T)$ for almost all ω in Ω' .

5. In this section we consider the question of whether a Markov process stopped at time T yields a Markov process. That is, if $\{x(t)\}$ is a Markov process, T is a stopping time and $\{y(t)\}$ is defined by

$$y(t, \omega) = \begin{cases} x(t, \omega) & \text{if } t < T(\omega), \\ x(T(\omega), \omega) & \text{if } t \geq T(\omega) \end{cases}$$

we ask whether $\{y(t)\}$ is also a Markov process. Of course it is not reasonable to expect the stopped process to have the same transition function as the original process. For T an arbitrary stopping time it is not true in general that $\{y(t)\}$ is Markovian as the following example indicates: let the space consist of 2 points a and b , let $\{x_n\}$ be a Markov chain with transition $P(1, a, a) = P(1, a, b) = P(1, b, a) = P(1, b, b) = 1/2$ and with $P\{x_0 = a\} = P\{x_0 = b\} = 1/2$, and let T be the second time that $x_n = b$. If $\{y_n\}$ is the stopped process then $P\{y_2 = a | y_1 = b\} = 1/4$ while $P\{y_2 = a | y_1 = b, y_0 = b\} = 0$ so the stopped process is not Markovian. In view of this we shall restrict our attention to the case in which T is the first passage time to a set E , that is

$T(\omega) = \inf \{t \mid x(t, \omega) \in E\}$ or $T(\omega) = \infty$ if for all t $x(t, \omega) \notin E$.

For Theorem 5.1 we make the following assumptions: $P(t, x, A)$ is a transition function satisfying (D) and such that for any probability distribution μ on $\mathfrak{B}(X)$ there is a Markov process $\{x_\mu(t)\}$ with right continuous paths, initial distribution μ and $P(t, x, A)$ as transition function. r will denote the measure which puts mass 1 at a point r in X . $\mathfrak{G}_\mu(a)$ denotes the smallest Borel field with respect to which $x_\mu(t)$ is measurable for each t not exceeding a . Let E be a fixed subset of X and define T_μ as the first passage time to E of $\{x_\mu(t)\}$. We also assume that for each μ , $\{T_\mu < a\}$ is in $\mathfrak{G}_\mu(a)$. Let p_r be the probability that $T_r = 0$. By the zero-one law p_r is either 0 or 1. For any μ the set of points r where p_r is 1 differs from a set in $\mathfrak{B}(X)$ by a set of μ measure 0. Now let $\{x(t)\}$ be a Markov process in the usual sense (that is relative to $\{\mathfrak{G}(t)\}$) with right continuous paths and $P(t, x, A)$ as transition function, let T be the first passage time to E and let $\{y(t)\}$ be the stopped process.

THEOREM 5.1. $\{y(t), \mathfrak{G}(t+)\}$ is a Markov process.

Proof. We first note that since (D) and the right continuity hold it follows from the proof of Theorem 1.1 that $\{x(t), \mathfrak{G}(t+)\}$ is a Markov process. For each n define T_n as in the proof of Theorem 1.1 and let $\{y_n(t)\}$ be $\{x(t)\}$ stopped at T_n . One verifies easily that $y_n(t)$ is $\mathfrak{G}(t)$ measurable. By the right continuity of the sample functions $y_n(t) \rightarrow y(t)$ as $n \rightarrow \infty$ so $y(t)$ is $\mathfrak{G}(t)$ measurable. Suppose now B is in $\mathfrak{B}(X)$ and s is less than t . We want to show that there is a version of $P\{y(t) \in B \mid \mathfrak{G}(s+)\}$ which is a function of $y(s)$ only. Let $C = \{r \mid p_r = 1\}$, let Q be s plus the first passage time to E of the process $\{x'(q) = x(q+s)\}$ and define this conditional probability as follows:

$$(5.1) \quad f(\omega) = \chi_{y(s) \in B \cap C} + \chi_{y(s) \notin C} \{P\{x(t) \in B, Q > t \mid x(s)\} + P\{x(Q) \in B, Q \leq t \mid x(s)\}\}$$

where χ_Γ denotes the characteristic function of a set Γ . From the definition of C it follows that $\{y(s) \in C\}$ and $\{T \leq s\}$ differ by at most a set of probability 0. We also note that the right side of (5.1) is measurable with respect to the completion of the Borel field induced by $y(s)$ and that the sets $\{x(t) \in B, Q > t\}$ and $\{x(Q) \in B, Q \leq t\}$ are measurable on the sample space of the process $\{x'(q)\}$. We want to show that if Δ is in $\mathfrak{G}(s+)$ then

$$(5.2) \quad \int_\Delta f(\omega) P(d\omega) = P\{\Delta \cap \{y(t) \in B\}\}.$$

We break up the left and right sides into

$$(5.3) \quad \int_{\Delta \cap \{T \leq s\}} f(\omega) P(d\omega) + \int_{\Delta \cap \{T > s\}} f(\omega) P(d\omega) = P\{\Delta \cap \{T \leq s\} \cap \{y(t) \in B\}\} + P\{\Delta \cap \{T > s\} \cap \{y(t) \in B\}\}.$$

The first terms on each side of (5.3) are equal so it remains only to show that the second terms are equal. We write the second term of the left side as

$$(5.4) \int_{\Delta \cap \{T > s\}} [P\{x(t) \in B, Q > t \mid x(s)\} + P\{x(Q) \in B, Q \leq t \mid x(s)\}] P(d\omega).$$

The range of integration is a set in $\mathfrak{G}(s+)$ so (5.4) becomes

$$(5.5) P\{x(t) \in B, Q > t, T > s, \Delta\} + P\{x(Q) \in B, Q \leq t, T > s, \Delta\}.$$

Q is equal to T if T exceeds s so (5.5) is $P\{\Delta \cap \{T > s\} \cap \{y(t) \in B\}\}$ which is the second term on the right of (5.3). So (5.3) holds and the theorem is proved. If E is an open set then $\{T_\mu < a\} = \bigcup_t \{x_\mu(t) \in E\}$, the union being over the rational values of t strictly less than a , and thus $\{T_\mu < a\}$ is in $\mathfrak{G}_\mu(a)$. Hunt [2] gives a thorough discussion of the measurability assumptions about the first passage time to an arbitrary analytic set.

When E is a closed set the situation is simpler. In this case we assume merely that $\{x(t)\}$ is a Markov process (no time-homogeneity, no transition function) with right continuous paths and with the additional property that if $\{T_n\}$ is an increasing sequence of stopping times whose limit is T then $x(T_n) \rightarrow x(T)$ for almost all ω in the set where T is finite. Now let E_n be the set of points whose distance from E is strictly less than $1/n$, let T_n be the first passage time to E_n and $T = \lim_{n \rightarrow \infty} T_n$. Since $\{T_n < a\}$ is in $\mathfrak{G}(a)$ so is $\{T < a\}$, and from the fact that $x(T_n) \rightarrow x(T)$ it follows that T is the first passage time to E . By the right continuity of the paths and the fact that E is closed we have $\{T \leq a\} = \{T < a\} \cup \{x(a) \in E\}$ so $\{T \leq a\}$ is in $\mathfrak{G}(a)$ and $y(s)$ is in E if and only if $s \leq T$. Now in the right side of (5.1) replace C by E and in (5.2) assume that Δ is in $\mathfrak{G}(s)$. Then $\Delta \cap \{T > s\}$ is in $\mathfrak{G}(s)$ and the validity of (5.3) follows from that fact that $\{x(t)\}$ is Markovian and that $\{x(t) \in B, Q > t\}$ and $\{x(Q) \in B, Q \leq t\}$ are measurable on the sample space of the $x(q)$ for $q \geq s$. Thus we have shown that *if E is closed then $\{y(t), \mathfrak{G}(t)\}$ is a Markov process*. Even in the case where E is an open set the simpler assumptions of this paragraph are not sufficient to imply that the stopped process is Markovian, for one can easily construct a Markov process with continuous paths and a T which is the first passage time to an open set such that the process stopped at time T is not a Markov process.

BIBLIOGRAPHY

J. L. DOOB

1. *Semi-martingales and subharmonic functions*, Trans. Amer. Math. Soc. vol. 77 (1954) pp. 86-121.
2. *Stochastic processes*, New York, 1953.

E. B. DYNKIN

1. *Criteria of continuity and of absence of discontinuities of the second kind for trajectories of a Markov random process*, Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya vol. 16 (1952) pp. 563-572.

G. A. HUNT

1. *Some theorems concerning Brownian motion*, Trans. Amer. Math. Soc. vol. 81 (1956) pp. 294–319.
2. *Probability and potentials*, to be published.

J. R. KINNEY

1. *Continuity properties of sample functions of Markov processes*, Trans. Amer. Math. Soc. vol. 74 (1953) pp. 280–302.

CORNELL UNIVERSITY,
ITHACA, N. Y.