Z. Wahrscheinlichkeitstheorie verw. Gebiete
40, 169–181 (1977)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Compactness of Stopping Times

J.R. Baxter[1] and R.V. Chacon[2]

[1] Department of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455 (USA)
[2] Department of Mathematics, University of British Columbia, Vancouver, British Columbia V6T1W5, Canada

The purpose of the present paper is to extend the results of [2] from the continuous to the quasi-left continuous case. We are indebted to P.-A. Meyer for pointing out that some of the results can be further extended to apply to regular processes. Our main goal is to show that any sequence $\{T(n)\}$ of stopping times which does not grow too rapidly admits of a subsequence converging, as far as the process in question is concerned, to a finite stopping time T. The growth condition is just that $\lim_{a\to\infty} P\{T(n)>a\}=0$ uniformly in n. In order to make the result valid it is necessary to enlarge the probability space so that there is a random variable with uniform distribution on $[0,1]$ which is independent of the process and also to properly define the notion of convergence. The proper notion of convergence is easy to state and it is that the subsequence $\{T(n(k))\}$ converges to T as far as the process is concerned provided that the distributions of $\{T(n(k))\}$ and of the process stopped at $T(n(k))$ converge in the usual way, i.e. weakly, to the distributions of T and of the process stopped at T. The enlargement of the space which is required is essential since it is easy to give examples which show that the result is false without it. This enlargement gives rise to randomized stopping times, so that the closure of the space of stopping times is the spaces of randomized stopping times.

A result of independent interest obtained in the course of the proof is that if $\{T(n)\}$ and $\{U(n)\}$ are sequences of stopping times which do not grow too rapidly in the sense already given and if $T(n)-U(n)$ converges to zero in probability then $f(X_{T(n)}, X_{U(n)})$ converges to zero in probability as well, where X_t is the process and where $f(x,y)$ is a function which can be taken to be the distance between x and y, when one exists, although in general the state space is not assumed to be a metric space.

An overall assumption made about the process is that it is standard, where the definition of the standard process is given in the same way as in the Markov case. It is not necessary, however, to assume that the process is a Markov process.

Many probabilistic constructions involve a passage from the discrete to the continuous. The main theorem in the paper gives a general method for proving convergence of any stopping times constructed in this way. The result can also be used to show that functionals of stopping times attain their extreme values.

An example which explains some of the ideas and in particular why randomization is needed is the following. Consider Brownian motion starting at the origin. Consider a sequence of non-randomized stopping times $\{T(n)\}$ satisfying the growth conditions such that the distribution of the Brownian motion stopped at $T(n)$ has half of its mass uniformly distributed on the sphere of radius $1/n$ and half of its mass uniformly distributed on the sphere of radius 1. The stopping time $T(n)$ can be constructed by stopping on the sphere of radius $1/n$ until half of the mass is deposited there and then stopping the remaining mass on the sphere of radius 1. The main theorem tells us that there exists a subsequence converging to a stopping time T in the sense described. The distribution of the Brownian motion stopped at T has half of its mass at the origin and half of its mass uniformly distributed on the sphere of radius 1. This can only happen, in the case that the dimension of the state space is two or more, if T equals zero with probability 1/2 and if T equals the first hitting time of the sphere of radius 1 with probability 1/2 so that T cannot be non-randomized even though the $T(n)$ are not, so that the main theorem cannot be sharpened to yield a non-randomized limit in the case that the sequence is made up of non-randomized stopping times.

A selection argument is used by Monroe [4] (in the proof of Theorem 11, p. 1300) to construct a time change for Brownian motion, in order to embed a right continuous martingale. The method used there to prove convergence depends on the special form of the time change involved in the construction. The methods of the present paper can be adapted to give a selection principle and convergence theorem for arbitrary sequences of time changes. This result will appear later. The authors would like to thank Professor Monroe for his helpful comments.

The plan of the paper is as follows. In Section 1 the necessary definitions are given and the main results are stated precisely. In Sections 2 and 3 the basic compactness result for sequences of stopping times is obtained. This is really a general compactness principle for random variables, and not merely for stopping times. Consider a sequence of extended-real-valued random variables. We can say (at the very least) that some subsequence converges in distribution to a random variable. This rather trivial selection principle is deficient in two ways, first because the convergence is very weak, and second because the limit is highly nonunique. However, the basic principle can be made more useful by considering a fixed countably generated σ-algebra of subsets of the sample space and then using the Cantor diagonal process to find a subsequence of the given sequence which converges in distribution on *every* set in the fixed σ-algebra (an independent example of the same idea applied to a problem in probabilistic functional analysis is given by I. Berkes and H.P. Rosenthal in a forthcoming paper [1]). Let the fixed σ-algebra be denoted by $\mathscr{G}$ and the limit by Z. Z must in general be defined on $\Omega \times [0,1]$ rather than the original sample space Ω. The random variable Z is still nonunique, but its conditional distribution on every fiber of $\mathscr{G}$ is determined. If $\mathscr{G}$ is big enough, the conditional distributions of Z on its fibers are all one needs to know about Z. Furthermore, it is shown that under general conditions many quantities associated with the convergent subsequence will be preserved in the limit. For example, it is shown in Section 2 that the limit of a sequence of stopping times is again a stopping time. Then, in Sections 4 and 5, conditions are given such that if a sequence of stopping times for a process converges in the sense just described, then the process stopped at those times converges in the same sense to the process stopped at the limit time.

1. Introduction

Let $X=(\Omega, \mathcal{M}, \mathcal{M}_t, X_t, P)$, $0 \leqq t \leqq \infty$, be a stochastic process, where

(1.1) $(\Omega, \mathcal{M}, P)$ is a probability space, $\{\mathcal{M}_t\}$ is an increasing family of σ-algebras contained in $\mathcal{M}$, and

(1.2) $X_t\colon \Omega \to E$ is $\mathcal{M}_t$-measurable for each t, where E is a topological space with the Borel σ-algebra $\mathcal{E}$. We will write X_t as $X(t)$ or $X(t,\omega)$ where convenient.

A map $T\colon \Omega \to [0,\infty]$ will be called an *$\mathcal{M}$-time* if T is $\mathcal{M}$-measurable. An $\mathcal{M}$-time for which $\{T \leqq t\} \in \mathcal{M}_t$ for all t is called an *$\{\mathcal{M}_t\}$-stopping time.* Let $\mathcal{B}$ be the Borel sets of $[0,1]$. A map $T\colon \Omega \times [0,1] \to [0,\infty]$ will be called a *randomized $\mathcal{M}$-time* if it is an $\mathcal{M} \times \mathcal{B}$-time, and will be called a *randomized $\{\mathcal{M}_t\}$-stopping time* if it is an $\{\mathcal{M}_t \times \mathcal{B}\}$-stopping time. We shall adopt the convention that *all $\mathcal{M} \times \mathcal{B}$-times are required to be nondecreasing* and *left continuous* in the second variable. *Also, any function on Ω will be considered defined on $\Omega \times [0,1]$ in the obvious way.* A probability P on $\mathcal{M}$ will induce a probability $P \times \lambda$ on $\mathcal{M} \times \mathcal{B}$, where λ is Lebesgue measure.

We wish to study the convergence of stopping times and of the associated stopped random variables. Let $\mathcal{G}$ be a fixed σ-field contained in $\mathcal{M}$. Let $T(n)$, $1 \leqq n < \infty$ and T be $\mathcal{M} \times \mathcal{B}$-times. $T(n)$ will be said to *converge weakly to T on $\mathcal{G}$*, written $T(n) \Rightarrow T(\mathcal{G})$, if the distribution of $T(n)$ on $G \times [0,1]$ converges weakly to the distribution of T on $G \times [0,1]$ for each $G \in \mathcal{G}$. That is,

(1.3) $T(n) \Rightarrow T(\mathcal{G})$ if and only if $T(n)|G \times [0,1] \Rightarrow T|G \times [0,1]$ for each $G \in \mathcal{G}$. Naturally $T(n)|G \times [0,1] \Rightarrow T|G \times [0,1]$ means

$$\int_{G \times [0,1]} f(T(n))\, d(P \times \lambda) \to \int_{G \times [0,1]} f(T)\, d(P \times \lambda) \quad \text{as} \quad n \to \infty$$

for each $f \in C([0,\infty])$, where $C([0,\infty])$ is the set of continuous functions on $[0,\infty]$.

Let $\Gamma = \Gamma(\{\mathcal{M}_t\})$ be the set of all $\{\mathcal{M}_t \times \mathcal{B}\}$-stopping times. For any $Y \in \mathcal{L}_1(\Omega, \mathcal{G}, P)$ and any $f \in C([0,\infty])$, let $\phi(Y,f)\colon \Gamma \to R$ be defined by $\phi(Y,f)(T) = E[Y, f(T)]$. Let ϕ be the set of all such ϕ. Let $\mathcal{T} = \mathcal{T}(\mathcal{G})$ be the topology on Γ generated by all ϕ in Φ. It is easy to check that $\mathcal{T}$ is also generated by those maps $\phi(\chi_G, f)$, obtained as $\mathcal{G}$ runs over an $\mathcal{L}_1$-dense subset of $\mathcal{G}$, and f runs over a sup-norm-dense subset of $C([0,\infty])$. In particular

(1.4) $T(n) \Rightarrow T(\mathcal{G})$ if and only if $T(n) \to T(\mathcal{T})$.

The following result is proved in Section (3):

(1.5) **Theorem.** *If $\{\mathcal{M}_t\}$ is right continuous then $(\Gamma, \mathcal{T})$ is compact.*

When $\mathcal{G}$ is countably generated mod P, $\mathcal{T}$ will have a countable base, and hence, if $\{\mathcal{M}_t\}$ is right continuous $(\Gamma, \mathcal{T})$ will be sequentially compact.

A stronger notion of convergence is desirable. For any $Y \in \mathcal{L}_1(\Omega, \mathcal{G}, P)$, $f \in C([0, \infty])$, and $h \in C(E)$, ($C(E)$ denotes the continuous bounded functions on E), let $\phi(Y, f, h): \Gamma \to R$ be defined by $\phi(Y, f, h) = E[Yf(T)\, h(X(T))]$. Let Φ_1 be the set of all such ϕ. Let $\mathcal{T}_1 = \mathcal{T}_1(\mathcal{G}, X)$ be the topology generated by all ϕ in Φ_1. Clearly $\mathcal{T} \subseteq \mathcal{T}_1$.

(1.6) $T(n) \Rightarrow T(\mathcal{G}, X)$ means $T(n) \to T(\mathcal{T}_1)$.

By taking $\mathcal{M}_t = \mathcal{M}$ we see that both modes of convergence, (1.4) and (1.6), are defined for arbitrary $\mathcal{M} \times \mathcal{B}$-times. A sufficient condition will be given for the two to coincide. For any $\mathcal{M} \times \mathcal{B}$-time T, let

(1.7) $T[j]$ = the smallest function $\geqq T$ of the form $k/2^j$.

(1.8) **Theorem.** *Let $X = (\Omega, \mathcal{M}, \mathcal{M}_t, X_t, P)$, $0 \leqq t \leqq \infty$ be a stochastic process, and let $\mathcal{G}$ be a σ-algebra, $\mathcal{G} \subseteq \mathcal{M}$, such that X_t is $\mathcal{G}$-measurable* mod P *for all t. Let $X.(\omega)$ be right continuous for P-almost every ω. Let D be a subset of $\Gamma(\mathcal{M})$, the set of all $\mathcal{M} \times \mathcal{B}$-times, such that for each $\phi \in \Phi_1$.*

(1.9) *$\phi(T[j]) \to \phi(T)$ as $j \to \infty$, uniformly over all $T \in D$. Then any limit point T of D under $\mathcal{T}$ which is finite* mod P *is also a limit point of D under $\mathcal{T}_1$.*

Theorem (1.8) is proved in Section 4. To give examples where it is applicable, in Section 5 we prove:

(1.10) **Theorem.** *Let $X = (\Omega, \mathcal{M}, \mathcal{M}_t, X_t, P)$ $0 \leqq t \leqq \infty$, be a stochastic process taking its values in a topological space E with a countable base. Let $\{\mathcal{M}_t\}$ be right continuous. Let $X.(\omega)$ be right continuous and have left limits for P-almost every ω. Let η be an $\{\mathcal{M}_t\}$-stopping time such that if $S(n)$, $1 \leqq n < \infty$, and S are $\{\mathcal{M}_t\}$-stopping times with $S(n) \uparrow S$ pointwise as $n \to \infty$, then*

(1.11) $X(S(n)) \to X(S)$ *pointwise* mod P *on $\{S \leqq n\}$ as $n \to \infty$.*

Now let $\bar{\eta}$ be the supremum of countably many η's of the sort just described. Let $\{T(n)\}$ be a sequence of $\{\mathcal{M}_t \times \mathcal{B}\}$-stopping times such that for any $\varepsilon > 0$ there exists $\rho: \Omega \to [0, \infty]$, $\rho\, \mathcal{M}$-measurable, such that $\rho \leqq \bar{\eta}$ mod P *and*

(1.12) $P(\{T(n) > \rho\}) < \varepsilon$ *for all n.*

Let $\{U(n)\}$ be a sequence of $\mathcal{M} \times \mathcal{B}$-times such that $T(n) \leqq U(n)$ and $U(n) - T(n) \to 0$ in probability as $n \to \infty$. Let f in $C(E \times E)$ be such that $f(x, x) = 0$ for all x in E. Then

(1.13) $f(X_{T(n)}, X_{U(n)}) \to 0$ *in probability as $n \to \infty$.*

Let X be a process with lifetime ζ. Let X be right continuous and quasi-left-continuous (for stopping times $< \zeta$). For example, X might be a standard Markov process as defined in [1], page 45. For any X that is right continuous and quasi-left-continuous, any stopping time $\eta < \zeta \bmod P$ will satisfy (1.11). If ζ is a limit of stopping times strictly less than ζ (in particular if $\zeta = \infty$) then $\bar{\eta} = \zeta$ will satisfy the assumptions of Theorem (1.10).

To apply Theorem (1.10) to Theorem (1.8), let X and $\bar{\eta}$ be as in Theorem (1.10). Let D be a collection of $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping times such that for every $\varepsilon > 0$ there exists $\rho\colon \Omega \to [0, \infty]$, $\rho\,\mathscr{M}$-measurable, such that $\rho < \bar{\eta} \bmod P$ and

(1.14) $P(\{T > \rho\}) < \varepsilon$ for all T in D.

Clearly D satisfies (1.9).

A process X with lifetime ζ is *a standard process with lifetime* ζ if the state space has a countable base, the process is right-continuous and the process is quasi-left-continuous for stopping times $< \zeta$. Sequential convergence with respect to the topology $\mathscr{T}_1$ gives us convergence in distribution of the stopping times as well as of the process, and the remarks following the statement of Theorem (1.10) yield the following corollaries.

(1.15) **Corollary.** *If X is a standard process with lifetime $\zeta = \infty$ and if $\{T(n)\}$ is a sequence of randomized stopping times such that*

$$\lim_{a \to \infty} P(\{T(n) > a\}) = 0$$

uniformly in η, then there exists a subsequence $\{T(n(k))\}$ and a randomized stopping time T such that $\operatorname{dist.} X_{T(n(k))} \Rightarrow \operatorname{dist.} X_T$ *and such that* $\operatorname{dist.} T(n(k)) \Rightarrow \operatorname{dist.} T$.

(1.16) **Corollary.** *If X is a standard process with lifetimes ζ which is the limit of stopping times strictly less than ζ, and if $\{T(n)\}$ is a sequence of randomized stopping times such that*

$$\lim_{m \to \infty} P(\{T(n) > \rho_m\}) = 0$$

uniformly in n, where $\{\rho_m\}$ is a sequence of $\mathscr{M}$-measurable functions strictly less than ζ, then there exists a subsequence $\{T(n(k))\}$ and a randomized stopping time T such that $\operatorname{dist.} X_{T(n(k))} \Rightarrow \operatorname{dist.} X_T$, *and such that* $\operatorname{dist.} T(n(k)) \Rightarrow T$.

2. Representations

Let $T\colon \Omega \times [0,1] \to [0, \infty]$ be nondecreasing and left continuous in the second variable. Define $T_v\colon \Omega \to [0, \infty]$ for each v in $[0,1]$ by

(2.1) $T_v(\omega) = T(\omega, v)$.

It is easy to check that T is an $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping time if and only if T_v is an $\{\mathscr{M}_t\}$-stopping time for each v. In particular, taking $\mathscr{M}_t = \mathscr{M}$ for a moment, T is an $\mathscr{M} \times \mathscr{B}$-time if and only if T_v is an $\mathscr{M}$-time for each v.

Define the probability $A(\omega, \cdot)$ on $[0, \infty]$ by

(2.2) $A(\omega, [0, t]) = \sup\{v \mid T(\omega, v) \leqq t\}$.

Then

(2.3) $T(\omega, v) = \inf\{t \mid A(\omega, [0,t]) \geqq v\}$.

T is an $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping time if and only if $A(\cdot, [0,t])$ is $\mathscr{M}_t$-measurable for each t, and T is an $\mathscr{M} \times \mathscr{B}$-time if and only if $A(\cdot, [0,t])$ is $\mathscr{M}$-measurable for each t.

For any probability P on $\mathscr{M}$, $A(\omega, \cdot)$ gives the conditional distribution of the $\mathscr{M} \times \mathscr{B}$-time T on $\{\omega\} \times [0,1]$. A will be called the ω-distribution of T. For any $Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$, and any bounded Borel function f on $[0, \infty]$, let

(2.4) $\alpha(Y, f) = E[Y f(T)]$.

α will be called the distribution map for T. It is easy to check that

(2.5) $\alpha(Y, f) = \int Y(\omega) \int f(t)\, A(\omega, dt)\, P(d\omega)$.

For $H \in \mathscr{M}$ write $\alpha(\chi_H, f)$ as $\alpha(H, f)$ and similarly for $K \in \mathscr{B}$ let $\alpha(Y, \chi_K) = \alpha(Y, K)$.

Clearly if T is an $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping time then $\int f(t)\, A(\cdot, dt)$ is $\mathscr{M}_s$-measurable for every bounded Borel function f on $[0, \infty]$ such that $f \equiv 0$ on (s, ∞). Conversely, if $\int f(t)\, A(\cdot, dt)$ is $\mathscr{M}_s$-measurable for every *continuous* function f with support on $[0, s]$ then T is an $\{\mathscr{M}_{t+} \times \mathscr{B}\}$-stopping time.

For any $\mathscr{M} \times \mathscr{B}$-time T, let us temporarily restrict the distribution map α to $\mathscr{L}_1(\Omega, \mathscr{M}, P) \times C([0, \infty])$. Then α has the following properties:

(2.6) α is a bilinear functional on $\mathscr{L}_1(\Omega, \mathscr{M}, P) \times C([0, \infty])$.

(2.7) $Y \geqq 0$, $f \geqq 0$ implies $\alpha(Y, f) \geqq 0$, $\alpha(1,1) = 1$, and $|\alpha(Y, f)| \leqq \|Y\|\, \|f\|$ for all Y, f. Here $\|Y\|$ means $\mathscr{L}_1$-norm Y, $\|f\|$ means sup-norm f.

If T is an $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping time then

(2.8) $\alpha(Y, f = \alpha(E[Y \mid \mathscr{M}_t], f)$ for all $Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$

and any continuous f with support on $[0, t]$.

(2.9) **Lemma.** *Let* $\{\mathscr{M}_t\}$ *be right continuous. Let* α *be any map satisfying* (2.6), (2.7), *and* (2.8). *Then there exists an* $\{\mathscr{M}_t \times \mathscr{B}\}$*-stopping time* T *with distribution map* α. T *is unique* $\operatorname{mod} P \times \lambda$.

Proof. For any $Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$ with $0 \leqq Y \leqq 1$,

$$1 = \alpha(1,1) = \alpha(Y,1) + \alpha(1-Y, 1) = \|Y\| + \|1-Y\|,$$

and $\|Y\| \geqq |\alpha(Y,1)|$, $\|1-Y\| \geqq \alpha(1-Y, 1)$. Hence $\alpha(Y,1) = \|Y\|$. By continuity, $\alpha(Y,1) = E[Y]$ for any Y in $\mathscr{L}_1(\Omega, \mathscr{M}, P)$.

For any $Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$, the map $\alpha(Y, \cdot)$ on $C([0, \infty])$ defines a bounded Borel measure on $[0, \infty]$, which we will denote again by $\alpha(Y, \cdot)$. $\alpha(Y, [0, \infty]) = E[Y]$, and if $Y \geqq 0$ then $\alpha(Y, \cdot) \geqq 0$. One may extend α to all $Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$ and all bounded Borel functions f by defining $\alpha(Y, f) = \int f(t)\, \alpha(Y, dt)$. Clearly α is bilinear, $Y \geqq 0$, $f \geqq 0$ implies $\alpha(Y, f) \geqq 0$, $\alpha(Y, 1) = E[Y]$, and $|\alpha(Y, f)| \leqq \|Y\|\, \|f\|$.

Using the right continuity of $\{\mathscr{M}_t\}$ and the fact that $\alpha(Y, \cdot)$ is continuous under bounded pointwise limits, it is easy to see that (2.8) holds for all

$Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$ and any bounded Borel function f such that $f=0$ on (t, ∞). For t rational or $t=\infty$, let $c(t)$ be a bounded $\mathscr{M}_t$-measurable function such that $\alpha(Y, [0,t]) = \int Yc(t)\,dP$ for each $Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$. We may choose $c(t)$ such that $c(\infty)=1$, $c(t) \geqq 0$. Clearly $c(s) \leqq c(t) \bmod P$ if $s \leqq t$, so we may also assume $c(s) \leqq c(t)$ everywhere, by replacing $c(t)$ by $\sup\{c(s) | s \leqq t\}$.

Let $A(\omega, \cdot)$ be the probability on $[0,\infty]$ such that $A(\omega, [0,t]) = \inf\{c(s,\omega) | t < S\}$ for all real t, $A(\omega, [0,\infty]) = 1$. Let T be the $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping time with ω-distribution A. Let α' be the distribution map for T. For any $Y \in \mathscr{L}_1(\Omega, \mathscr{M}, P)$ and any real t,

$$\begin{aligned}\alpha(Y,[0,t]) &= \inf\{\alpha(Y,[0,s]) | t<s,\ s \text{ rational}\}\\ &= \inf\{\textstyle\int Yc(s)\,dP | t<s\} = \int Y(\inf\{c(s)|t<s\})\,dP\\ &= \textstyle\int YA(\cdot,[0,t])\,dP = \alpha'(Y,[0,t]).\end{aligned}$$

Hence $\alpha = \alpha'$. T is unique by Lemma (2.14), so the lemma is proved.

Let $\Gamma = \Gamma(\{\mathscr{M}_t\})$ be the set of $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping times, and let $\Lambda = \Lambda(\{\mathscr{M}_t\})$ be the set of maps α satisfying (2.6), (2.7), and (2.8). There is a natural map Θ: $\Gamma \to \Lambda$ which takes each member of Γ to its distribution map,

(2.10) $\quad \Theta(T) = \alpha.$

Lemma (2.9) says that if $\{\mathscr{M}_t\}$ is right continuous then Θ is an onto map, and Θ is one-to-one $\bmod P \times \lambda$.

Let T be an $\mathscr{M} \times \mathscr{B}$-time with ω-distribution A. It is easy to check that $T(\omega, v)$ is the inf of all s such that $\int f(t) A(\omega, dt) \geqq v$ for some nonincreasing f in $C([0,\infty])$ with $0 \leqq f \leqq 1$ and $f=0$ on $[s,\infty]$. Let U be another $\mathscr{M} \times \mathscr{B}$-time with ω-distribution B. Clearly the following statements are equivalent:

(2.11) $\quad T(\omega, \cdot) \leqq U(\omega, \cdot),$

(2.12) $\quad A(\omega, [0,\cdot]) \geqq B(\omega, [0,\cdot]),$

(2.13) $\quad \int f(t) A(\omega, dt) \geqq \int f(t) B(\omega, dt)$ for all nonincreasing $f \in C([0,\infty])$.

(2.14) **Lemma.** *Let T and U be $\mathscr{M} \times \mathscr{B}$-times with distribution maps α and β respectively. Suppose $\alpha(Y,f) \geqq \beta(Y,f)$ for every $Y \geqq 0$ in $\mathscr{L}_1(\Omega, \mathscr{M}, P)$ and every nonincreasing f in $C([0,\infty])$. Then $T \leqq U \bmod P \times \lambda$.*

Proof. Choose a countable dense set D of nonincreasing functions in $C([0,\infty])$. Let A and B be the ω-distributions of T and U respectively. For any $f \in D$,

$$\int Y \int f(t) A(\cdot, dt)\,dP \geqq \int Y \int f(t) B(\cdot, dt)\,dP$$

for all $Y \geqq 0$ in $\mathscr{L}_1(\Omega, \mathscr{M}, P)$. Hence

$$\int f(t) A(\cdot, dt) \geqq \int f(t) B(\cdot, dt) \bmod P.$$

Let Ω_1 be a set in $\mathscr{M}$ with $P(\Omega - \Omega_1) = 0$ such that

$$\int f(t) A(\omega, dt) \geqq \int f(t) B(\omega, dt)$$

for all $f \in D$, $\omega \in \Omega_1$. Then the same inequality holds on Ω_1 for all nonincreasing f in $C([0,\infty])$. By (2.13), $T \leqq U$ on Ω_1, so the lemma is proved.

(2.15) **Lemma.** *Let T and U be $\mathcal{M} \times \mathcal{B}$-times with distribution maps α and β respectively. Let $\mathcal{G}$ be a σ-algebra, $\mathcal{G} \subseteq \mathcal{M}$. The following conditions are equivalent:*

(2.16) $\alpha(Y,f) = \beta(Y,f)$ *for all* $Y \in \mathcal{L}_1(\Omega, \mathcal{G}, P)$, $f \in \mathcal{C}([0,\infty])$,

(2.17) *distribution T on $G \times [0,1]$ = distribution U on $G \times [0,1]$ for all $G \in \mathcal{G}$.*

The proof is immediate.

(2.18) **Lemma.** *Let T be an $\mathcal{M} \times \mathcal{B}$-time. Let $\mathcal{G}$ be a σ-algebra, $\mathcal{G} \subseteq \mathcal{M}$. Then there exists a $\mathcal{G} \times \mathcal{B}$ time U for which* (2.17) *holds. U is unique* mod $P \times \lambda$.

Proof. Restrict α to $\mathcal{L}_1(\Omega, \mathcal{G}, P) \times \mathcal{C}([0,\infty])$. Apply Lemma (2.9) with $\mathcal{M} = \mathcal{M}_t = \mathcal{G}$ to obtain U. Apply Lemma (2.15) to conclude that (2.17) holds. Uniqueness follows also from Lemma (2.9) with $\mathcal{M} = \mathcal{M}_t = \mathcal{G}$.

It is easy to check that if T happens to be an $\{\mathcal{M}_t \times \mathcal{B}\}$-stopping time then U can be chosen to be an $\{\mathcal{M}_t \cap \mathcal{G} \times \mathcal{B}\}$-stopping time provided that $\{\mathcal{M}_t\}$ is right continuous and

(2.19) $E[E[\cdot|\mathcal{G}]|\mathcal{M}_t] = E[\cdot|\mathcal{M}_t \cap \mathcal{G}]$ for all t.

3. Compactness

Proof of Theorem (1.5). Let Θ be the map defined in (2.10). For each $Y \in \mathcal{L}_1(\Omega, \mathcal{M}, P)$ and each $f \in C([0,\infty])$, let $\psi(Y,f)\colon \Lambda \to R$ be defined by $\psi(Y,f)(\alpha) = \alpha(Y,f)$. Let Ψ be the set of all such ψ. Let $\mathcal{C}$ be the topology on Λ generated by all ψ in Ψ. It is easy to see that $\mathcal{T} = \Theta^{-1}(\mathcal{C})$. $\mathcal{C}$ is compact, by the same argument used to prove that unit ball in the dual of a normed linear space is compact. Since Θ is onto, $\mathcal{T}$ is compact also, so the theorem is proved.

4. Convergence

(4.1) **Lemma.** *Let $(\Omega, \mathcal{M}, X_t, P)$, $0 \leqq t \leqq \infty$ be a real-valued stochastic process. Let $\mathcal{G}$ be a σ-algebra, $\mathcal{G} \subseteq \mathcal{M}$. Let K be a finite set of extended real numbers such that X_t is bounded and $\mathcal{G}$-measurable* mod P *for each $t \in K$. Let $U(n)$, $1 \leqq n < \infty$, and U be $\mathcal{M} \times \mathcal{B}$-times such that $U(n) \in K$* mod P *for each n, and $U(n) \Rightarrow U(\mathcal{G})$. Then*

(4.2) $E[YX(U(n))] \to E[YX(U)]$ *as $n \to \infty$ for each* $Y \in \mathcal{L}_1(\Omega, \mathcal{G}, P)$.

Proof. By assumption, for each $G \in \mathcal{G}$ the distribution of $U(n)$ restricted to $G \times [0,1]$ has support in K and converges weakly to the distribution of U restricted to $G \times [0,1]$. Given $\varepsilon > 0$, choose a finite partition $\{G(1), \ldots, G(l)\}$ of G such that $G(i) \in \mathcal{G}$ for each i, and such that there exist constants $c(i,t)$, $1 \leqq i \leqq l$, $t \in K$, for which $|X(t) - c(i,t)| < \varepsilon$ mod P on $G(i)$. Let $Y = \chi_G$ and let

$$Y(n,i,t) = \chi_A \quad \text{where} \quad A = (G(i) \times [0,1]) \cap \{U(n) = t\}.$$

Then

$$E[YX(U(n))] = \sum_{i=1}^{l} \sum_{t \in K} E[Y(n,i,t)X(t)]$$
$$= \sum_{i=1}^{l} \sum_{t \in K} c(i,t) P \times \lambda((G(i) \times [0,1]) \cap \{U(n)=t\}) + v(n),$$

where

$$|v(n)| \leqq \sum_{i=1}^{l} \sum_{t \in K} E[Y(n,i,t)|X(t) - c(i,t)|] \leqq E[\varepsilon Y] = \varepsilon P(G).$$

A similar equation holds with $U(n)$ replaced by U. Since $P \times \lambda((G(i) \times [0,1]) \cap \{U(n)=t\})$ converges to $P \times \lambda((G(i) \times [0,1]) \times \{U=i\}$, (4.2) holds for $Y = \chi_G$, and the lemma follows.

(4.3) **Lemma.** *Let $X = (\Omega, \mathcal{M}, \mathcal{M}_t, X_t, P)$ be a bounded real-valued stochastic process. Let $\mathcal{G}$ be a σ-algebra, $\mathcal{G} \subseteq \mathcal{M}$. Let $x(\cdot, \omega)$ be right continuous for P-almost every ω. Let T and U be two $\mathcal{M} \times \mathcal{B}$-times such that* (2.17) *holds. Then*

(4.4) $E[YX(T)] = E[YX(U)]$ *for all* $Y \in \mathcal{L}_1(\Omega, \mathcal{G}, P)$.

Proof. For each j let $T[j]$ and $U[j]$ be defined as in (1.7). Let $T(j) = T[j] \wedge j$ on $\{T < \infty\}$, $T(j) = T$ on $\{T = \infty\}$. Define $U(j)$ similarly. Then $T(j)$ and $U(j)$ satisfy (2.17). By Lemma (4.2), with U replaced by $T(j)$ and $U(n)$ replaced by $U(j)$ for all n, (4.4) holds for $T(j)$ and $U(j)$. Letting $j \to \infty$ proves the lemma.

Proof of Theorem (1.8). Let T be a limit point of D under $\mathcal{T}$, T finite mod P. We must show that T is also a limit point of D with respect to $\mathcal{T}_1$. To do this it is enough to show that any $\phi \in \Phi_1$ is continuous at T (on $D \cup \{T\}$) with respect to $\mathcal{T}$. Fix a $\phi \in \Phi_1$, $\phi = \phi(Y, f, h)$. Let $\mathcal{G}'$ be the σ-algebra generated by Y and $\{X_t | t \text{ rational}\}$. Clearly $\mathcal{T}(\mathcal{G}') \subseteq \mathcal{T} = \mathcal{T}(\mathcal{G})$. Thus it is enough to show that ϕ is continuous at T with respect to $\mathcal{G}'$. Since $(\Omega, \mathcal{M}, X_t, P)$, $\mathcal{G}'$, and D satisfy the same hypothesis as $(\Omega, \mathcal{M}, X_t, P)$, $\mathcal{G}$, and D, we may drop the prime and assume from now on that $\mathcal{G}$ is countably generated mod P. Then $\mathcal{T}$ has a countable base, and there is a sequence $\{T(n)\} \subseteq D$ such that $T(n) \Rightarrow T(\mathcal{G})$. We must show for any such sequence that $\phi(T(n)) \to \phi(T)$ as $n \to \infty$. Since $(\Omega, \mathcal{M}, f(t)\, h(X_t), P), \mathcal{G}$, and $\{T(n)\}$ satisfy the hypothesis of Theorem (4.4) below, the proof of Theorem (1.8) has been reduced to the proof of that theorem.

(4.4) **Theorem.** *Let $(\Omega, \mathcal{M}, X_t, P)$ be a bounded real-valued stochastic process. Let $\mathcal{G}$ be a σ-algebra, $\mathcal{G} \subseteq \mathcal{M}$, such that X_t is bounded and $\mathcal{G}$-measurable* mod P *for each t. Let $X(\cdot, \omega)$ be right continuous for P-almost every ω. Let $\{T(n)\}$ be a sequence of $\mathcal{M} \times \mathcal{B}$-times such that for every Y in $\mathcal{L}_1(\Omega, \mathcal{G}, P)$,*

(4.5) $E[YX(T(n)[j])] \to E[YX(T(n))]$

as $j \to \infty$, uniformly in n. Suppose T is an $\mathcal{M} \times \mathcal{B}$-time, T finite mod P, *such that $T(n) \Rightarrow T(\mathcal{G})$. Then, for every Y in $\mathcal{L}_1(\Omega, \mathcal{G}, P)$*

(4.6) $E[YX(T(n))] \to E[YX(T)]$ *as* $n \to \infty$.

Proof. Fix Y bounded. It is enough to show that any subsequence of $\{n\}$ contains another subsequence for which the desired convergence holds. Thus it may be assumed, by using Theorem (1.5) to choose subsequences and then relabelling, that $T(n)[j] \Rightarrow T(j)(\mathscr{G})$ *as* $n \to \infty$ for each j, where $T(j)$ is some $\mathscr{M} \times \mathscr{B}$-time. By Lemma (2.18) it may be assumed that T and all the $T(j)$ are $\mathscr{G} \times \mathscr{B}$-times. For each j and each n, $T(n) \leqq T(n)[j] \leqq T(n) + 1/j$. Lemma (2.14) with $\mathscr{M} = \mathscr{G}$ shows that $T \leqq T(j) \leqq T + 1/j \bmod P$ for all j. We have

(4.7) $|E[YX(T(n))] - E[YX(T)]| \leqq a_1 + a_2 + a_3 + a_4,$

where

$a_1 = |E[YX(T(n)) - YX(T(n)[j])]|$ is small for large j, uniformly in n, by (4.5),

$a_2 = |E[YX(T(n)[j]) - YX(T(n)[j] \wedge j)]| \leqq$ const) $P(T(n)[j] > j)$ is small for large j and large n, since T is finite $\bmod P$,

$a_3 = |E[YX(T(n)[j] \wedge j) - YX(T(j) \wedge j)]|$ is small for any fixed j and sufficiently large n by Lemma (4.1),

$a_4 = |E[YX(T(j) \wedge j) - YX(T)]|$ is small for large j by right-continuity of X and finiteness of T. Thus by choosing first j and then n we see that (4.7) can be made as small as desired. This proves the theorem.

5. Approximation

(5.1) **Lemma.** *If Theorem* (1.11) *holds for* $\{\mathscr{M}_t\}$*-stopping times then it holds for* $\{\mathscr{M}_t \times \mathscr{B}\}$*-stopping times.*

Proof. We note first that if a sequence $\{T(n)\}$ of $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping times satisfies (1.12) then so does $\{T_v(n)\}$, for each $v < 1$. Indeed, if $\{T_v(n)\}$ did not satisfy (1.12), for some $v < 1$, then for some $\varepsilon > 0$ and every $\rho < \bar{\eta}$ the inequality $P(\{T_v(n) > \rho\}) \geqq \varepsilon$ would hold for some n. Then $\{T(n)\}$ could not satisfy (1.12) since $P \times \lambda(\{T(n) > \rho\}) \geqq (1 - v)\,\varepsilon$ would hold.

Now let $\{T(n)\}$ be a sequence of $\{\mathscr{M}_t \times \mathscr{B}\}$-stopping times satisfying (1.12), and let $\{U(n)\}$ be a sequence of $\mathscr{M} \times \mathscr{B}$-times such that $T(n) \leqq U(n)$, and $U(n) - T(n) \to 0$ in probability as $n \to \infty$. Let $f \in C(E \times E)$ with $f(x, x) = 0$ for all x in E. Suppose that $f(X(T(n)), X(U(n)))$ does not converge to 0 in probability. Then there exists some $\varepsilon > 0$ such that

$$P \times \lambda(\{f(X(T(n)), X(U(n))) \geqq \varepsilon\}) \geqq \varepsilon$$

for infinitely many n. By choosing a subsequence we may assume that the inequality holds for all n, and also that $U(n) - S(n) \to 0$ pointwise almost everywhere as $n \to \infty$. But then $U_v(n) - T_v(n) \to 0$ p.w.a.e. for λ-a.c. v in $[0, 1]$. Hence, by assumption,

$$P(\{f(X(T_v(n)), X(U_v(n))) \geqq \varepsilon\}) \to 0$$

as $n \to \infty$ for λ-a.e. v in $[0, 1]$. A contradiction follows, so the lemma is proved.

Let X, E, η satisfy the hypotheses of Theorem (1.10). A map $T\colon \Omega \to [0, \infty]$ will be called an $\{\mathscr{M}_t\}$*-stopping time* $\bmod P$ if

(5.2) $\{T \leqq t\} \in \mathscr{M}_t \bmod P$ for all t.

It is easy to check that T is an $\{\mathscr{M}_t\}$-stopping time mod P if and only if T is equal to an $\{\mathscr{M}_t\}$-stopping time except on a set of P-measure 0.

Let $S(n)$, $1 \leqq n < \infty$, and S be $\{\mathscr{M}_t\}$-stopping times mod P, with $S(n) \to S$ pointwise mod P as $n \to \infty$. Let η satisfy (1.11). Then it is easy to see that

(5.3) $X(S(n)) \to X(S)$ pointwise mod P on $\{S \leqq n\}$ as $n \to \infty$.

Because of Lemma (5.1) it will be enough to prove Theorem (1.10) for $\{\mathscr{M}_t\}$-stopping times. Before doing that, some definitions are needed.

Fix f in $C(E \times E)$ such that $f(x,x)=0$ for all x in E. Assume $0 \leqq f \leqq 1$. Define

(5.4) $Y(t,s) = \sup\{f(X(t), X(r)) \mid t \leqq r \leqq s\}$ for all t, s in $[0, \infty]$.

Fix a number $\varepsilon > 0$.

For each number t of the form $k/2^n$, where $k \geqq 0$ and $n > 0$ are integers, and each integer $m > 0$, let

(5.5) $H(t,m) = \{E[Y(t, t+1/m \mid \mathscr{M}_t] > \varepsilon\}$.

$H(t,m)$ is only defined mod P. Choose $H(t,m)$ in $\mathscr{M}_t$. For any choice,

$$H(t, m+1) \subseteq H(t,m) \bmod P.$$

Replacing $H(t,m)$ by $\bigcup_{j=m}^{\infty} H(t,j)$ if necessary, we can make sure that

(5.6) $H(t, m+1) \subseteq H(t,m)$.

For any ω in Ω, and any integers $n > 0$ and $m > 0$, let $R(n,m)(\omega)$ be the smallest number $k/2^n$ such that ω is in $H(k/2^n, m)$. If no such number exists, let $R(n,m)(\omega) = \infty$. Clearly.

(5.7) $R(n+1, m) \leqq R(n,m)$, $\quad R(n, m+1) \geqq R(n,m)$.

It is easy to check that each $R(n,m)$ is an $\{\mathscr{M}_t\}$-stopping time. Let $R(m)$ and R be those $\{\mathscr{M}_t\}$-stopping times such that

(5.8) $R(n,m) \downarrow R(m)$ as $n \to \infty$, $\quad R(m) \uparrow R$ as $m \to \infty$.

(5.9) **Lemma.** $P(\{R \leqq \eta\}) = 0$.

Proof. It is easy to check that

$$E[Y(R(n,m), R(n,m)+1/m) \mid \mathscr{M}(R(n,m))] > \varepsilon \bmod P \quad \text{on } \{R(n,m) < \infty\}.$$

Thus for any $\{\mathscr{M}_t\}$-stopping time ξ,

$$\int_{\{R(n,m) < \xi\}} Y(R(n,m), R(n,m)+1/m)\, dP \geqq \varepsilon P(\{R(n,m) < \xi\}).$$

By right continuity,

$$Y(R(n,m), R(n,m)+1/m) \to Y(R(m), R(m)+1/m)$$

pointwise mod$P$ as $n \to \infty$. Hence

$$\int_{\{R(m)<\xi\}} Y(R(m), R(m)+1/m)\, dP \geqq \varepsilon P(\{R(m)<\xi\}).$$

Let η satisfy (1.11).
Letting ξ approach η from above,

$$(5.10) \quad \int_{\{R(m)\leqq\eta\}} Y(R(m), R(m)+1/m)\, dP \geqq \varepsilon P(\{R(m)\leqq\eta\}).$$

Let $S(m)$ be the infimum of all r such that $R(m) \leqq r \leqq R(m)+1/m$ and $f(X(R(m)), X(r)) > \varepsilon/2$. If no such r exists, let $S(m)=R(m)+1/m$. It is easy to check that $S(m)$ is an $\{\mathscr{M}_t\}$-stopping time mod$P$. By right continuity, if

$$Y(R(m), R(m)+1/m) > \varepsilon/2 \quad \text{then } f(X(R(m)), X(S(m))) \geqq \varepsilon/2.$$

By (5.10),

$$P(\{Y(R(m), R(m)+1/m) > \varepsilon/2\} \cap \{R(m) \leqq \eta\}) + (\varepsilon/2)\, P(\{R(m) \leqq \eta\}) \geqq \varepsilon P(\{R(m) \leqq \eta\}).$$

Thus

$$P(\{f(X(R(m)), X(S(m))) \geqq \varepsilon/2\} \cap \{R(m) \leqq \eta\}) \geqq (\varepsilon/2)\, P(R(m) \leqq \eta\}).$$

Since $\{R \leqq \eta\} = \bigcap_{m=1}^{\infty} \{R(m) \leqq \eta\}$, (5.3) implies $P(\{R \leqq \eta\})=0$, which proves the lemma.

(5.11) **Lemma.** *Let U be an $\{\mathscr{M}_t\}$-stopping time. Let $\varepsilon>0$ be as in* (5.5), *$R(m)$ as in* (5.8). *Then*

$$(5.12) \quad E[Y(U, U+1/m)] \leqq P(\{R(m) \leqq U\}) + \varepsilon.$$

Proof. Let $U[j]$ be defined as in (1.7). Let $t=k/2^j$. Let $Z(j)=Y(U[j], U[j]+1/m)$. On $\{U[j]=t$,

$$E[Z(j)|\mathscr{M}(U[j])] = E[Y(t, t+1/m)|\mathscr{M}_t] \bmod P.$$

Thus

$$\{U[j]=t\} \cap \{E[Z(j)|\mathscr{M}(U[j])] > \varepsilon\} = \{U[j]=t\} \cap \{E[Y(t, t+1/m)|\mathscr{M}_t] > \varepsilon\} \bmod P.$$

Let $R(n,m)$ be the stopping time defined before (5.7). Then $R(j,m) \leqq t \bmod P$ on

$$\{U[j]=t\} \cap \{E[Z(j)|\mathscr{M}(U[j])] > \varepsilon\},$$

by definition. Thus

$$R(j,m) \leqq U[j] \bmod P \text{ on}$$

$$\{E[Y(U[j], U[j]+1/m) | \mathcal{M}(U[j])] > \varepsilon\}.$$

Hence

$$E[Y(U[j], U[j]+1/m)] \leqq P(\{R(j,m) \leqq U[j]\}) + \varepsilon,$$

so

$$E[Y(U[j], U[j]+1/m)] \leqq P(\{R(m) \leqq U[j]\}) + \varepsilon \quad \text{for all } j.$$

(5.12) follows by right continuity, so the lemma is proved.

Proof of Theorem (1.10). Let $\{T(n)\}$ and $\{U(n)\}$ be as in Theorem (1.10). Let f and ε be given, as in (5.4) and (5.5). Then

$$\begin{aligned} &E[f(X(T(n)), X(U(n)))] \\ &\quad \leqq \int_{\{U(n)-T(n) \leqq 1/m\}} f(X(T(n)), X(U(n)))\, dP + P(\{U(n)-T(n) > 1/m\}) \\ &\quad \leqq \int Y(T(n), T(n)+1/m)\, dP + P(\{U(n)-T(n) > 1/m\}) \\ &\quad \leqq P(\{R(m) \leqq T(n)\}) + \varepsilon + P(\{U(n)-T(n) > 1/m\}), \end{aligned}$$

by Lemma (5.11). By Lemma (5.9), $P(\{R < \bar{\eta}\} = 0$, so by (1.12) the first term approaches 0 for large m, uniformly in n. For fixed m the last term approaches 0 as $n \to \infty$. This proves the theorem.

Bibliography

1. Berkes, I., Rosenthal, H.P.: Almost exchangeable sequences of random variables. [To appear]
2. Chacon, R.V.: Unpublished manuscript, 1976, University of British Columbia
3. Blumenthal, R.M., Getoor, R.K.: Markov processes and potential theory. New York: Academic Press 1968
4. Monroe, I.: On embedding right continuous martingales in Brownian motion. Ann. Math. Statist. **43**, 1293–1311 (1972)

Received September 2, 1976; in revised form February 1977