ILLINOIS JOURNAL OF MATHEMATICS
Volume 31, Number 2, Summer 1987

EXISTENCE OF OPTIMAL TRANSITION KERNELS

BY
KEVIN J. HASTINGS

1. Introduction

Suppose a controller watches a continuous-time stochastic process up to a fixed time t_0. The probability law P_0 of the uncontrolled process may not be completely known. Based on his knowledge of the process up to t_0, the controller must choose a new probability law on the future of the process after t_0 from a set of admissible measures dependent on the past history. The goal is to minimize the expected value of a random variable that represents total cost. It might be the case that both the expected cost under his chosen law, and the set of admissible measures depend upon his statistical estimates of some unknown parameters. For instance, in Section 3, a problem is stated in which the process is a semi-Markov jump process with uncontrollable, unknown holding parameters. The controller may choose the state transition probabilities subject to some restrictions. His choice of an "optimal" law is only as good as his statistical estimates of the holding parameters. Returning to the general case, we will answer the following question. Under what conditions does there exist a measurable function Q^* from the past history to the set of probability measures on the future (i.e., a transition probability kernel) such that for every history, the image measure is both admissible and optimal (according to the current statistical information)?

Some ways in which this work differs from much of the stochastic control literature are the implicit inclusion of statistical estimation, the lack of Markov assumptions, the idea of treating past histories as states and probability measures as actions, and the general nature of the cost variable. Of course, spaces of probability measures are hard to deal with computationally, but their topological properties are well-suited to existence theory. The paper of Kertz and Nachman (11), inspired the problem formulation. In the context of discrete-time non-stationary dynamic programming, they employed histories as states, general cost variables, and multifunctions into spaces of measures to prove existence of discrete-time "persistently optimal plans" which are optimal for the future after every fixed time t_0. Similar models using ideas of measurable selection from spaces of probability measures appeared earlier in

Received September 16, 1985.

the work of Sudderth [20], who was concerned with a discrete-time stopping rule problem, and Blackwell, Freedman, and Orkin ([3], see also the companion paper by Freedman [8]), who applied the ideas to discrete-time dynamic programming. In the problem studied here, only one intervention is made, though choice of a law on the future may be thought of as a choice of an action at each future time. So, we essentially have a single-stage measurable selection problem. An interesting, and much more difficult, problem to study is the existence of a consistent family of measures on the future, one for each time t_0, which would be the direct continuous-time analog of the discrete-time Kertz and Nachman problem. As for the measurable selection problem, the most appropriate among many useful results in this area is in Wagner [21, Theorem 9.1], which leads to the main theorem.

Schäl [12] seems to be the first to build statistical estimation directly into a discrete-time problem, connecting dynamic programming to statistical decision theory. Here the estimates are implicit in the constraints and the value function. The notion of optimality simply changes; we do not attempt, as Schäl does, to minimize maximum risk.

Some interesting references on control of jump processes are Pliska [15] and Stone [19] in which holding times are controllable as well as transition probabilities, Wan and Davis [22] and Boel and Varaiya [5], who use the idea of the past history as a state, but are more interested in martingale characterization of the optimal control. In all these works, controls are functions into an action space that give rise to probability measures. This paper takes the measures as the fundamental objects.

In the remainder of this section, we introduce notation and terminology. In Section 2, we formulate the problem rigorously and prove the main existence theorem. The theorem is applied to the jump process example alluded to above in the final section.

Let E be a Polish space (i.e., complete separable metric space). The symbols $\mathscr{B}(E)$, $\mathscr{C}(E)$, and $\mathscr{P}(E)$ respectively denote the Borel σ-algebra on E, the collection of compact subsets of E, and the set of all probability measures on $(E, \mathscr{B}(E))$. In particular $\delta_x \in \mathscr{P}(E)$ is the measure with unit mass at $x \in E$. Let F be a topological space, let $\mu \in \mathscr{P}(E)$, and let D be a multifunction D: $E \to \mathscr{C}(F)$. Then D is called *measurable* if

$$\{x \in E\colon D(x) \cap B \neq \varnothing\} \in \mathscr{B}(E) \quad \text{for all closed } B \subseteq F; \tag{1.1}$$

Let $[a, b|$ be a sub-interval of $\mathbf{R}_+$ (possibly unbounded) that is closed at the left. Define

$$\Omega_{[a,b|} = \{\omega\colon [a,b| \to E\colon \omega \text{ is right-continuous and has left limits at all } t \in [a,b|\}; \tag{1.2}$$

$$X_t(\omega) = \omega(t); \quad \mathscr{F}_t = \sigma(X_s\colon s \le t); \quad \mathscr{F} = \sigma(X_s\colon s \in [a,b|). \tag{1.3}$$

We call $\mathscr{X} = (\Omega_{[a,b|}, \mathscr{F}, \mathscr{F}_t, X_t)_{t\in[a,b|}$ the *canonical process* on $(E, \mathscr{B}(E))$ with time domain $[a, b|$. In the sequel, we write $\Omega = \Omega_{[0,\infty)}$ and $H = \Omega_{[0,t_0]}$, where t_0 is a fixed positive number. If E is a Polish space, then $\Omega_{[a,b|}$ can be given the Skorohod topology as follows. Denote by $\|\cdot\|$ the supremum norm for functions whose domain is a compact interval $[c, d]$, and let $\Lambda = \Lambda_{[c,d]}$ be the set of all strictly increasing homeomorphisms on $[c, d]$. Write e for the identity, $e \in \Lambda$. Then the following defines a metric that is consistent with the Skorohod topology on $\Omega_{[c,d]}$:

$$d(\omega, \omega') = \inf_{\lambda\in\Lambda} \{|\lambda| \vee \|\omega\circ\lambda - \omega'\|\} \tag{1.4}$$

where

$$|\lambda| = \sup_{\substack{r,s\in[c,d]\\ r\neq s}} \left|\log \frac{\lambda(s)-\lambda(r)}{s-r}\right| + \|\lambda - e\|. \tag{1.5}$$

The Skorohod topology on $\Omega_{[a,b|}$ is the natural extension of this; $(\omega_n) \to \omega$ if for each compact $[c, d]$ such that c and d are either endpoints of $[a, b|$ or points of continuity of ω, the restrictions of ω_n converge to the restriction of ω in the Skorohod topology on $[c, d]$. Then $\Omega_{[a,b|}$ is a Polish space (Whitt [23, Theorem 2.6], see also Maisonneuve [12]), and $\mathscr{B}(\Omega_{[a,b|}) = \mathscr{F}$. The latter is an important fact that enables us to appeal to Borel measurable selection results.

For $\omega, \omega' \in \Omega$, define $\omega \sim \omega' \in \Omega$ by

$$\omega \sim \omega'(s) = \begin{cases} \omega(s) & \text{if } s < t_0; \\ \omega'(s - t_0) & \text{if } s \geq t_0. \end{cases} \tag{1.6}$$

For $h \in H$, define $\hat{h} \in \Omega$ by

$$\hat{h}(s) = \begin{cases} h(s) & \text{if } s < t_0; \\ h(t_0) & \text{if } s \geq t_0. \end{cases} \tag{1.7}$$

2. Problem statement and existence theorem

Once again, let $\mathscr{X}$ be the canonical process on $(E, \mathscr{B}(E))$ with time domain $\mathbf{R}_+$. Let there be given $Z \in \mathscr{F}/\mathscr{B}(\overline{\mathbf{R}})$, called the *cost variable*. Define $Y(\omega, \omega') \equiv Z(\omega \sim \omega')$. Clearly, for each ω the mapping $\omega' \to Y(\omega, \omega') \in \mathscr{F}/\mathscr{B}(\overline{\mathbf{R}})$, and also $Y \in (\mathscr{F}_{t_0} \times \mathscr{F})/\mathscr{B}(\overline{\mathbf{R}})$. We let D be a multifunction from Ω into $\mathscr{P}(\Omega)$ such that for all $\omega \in \Omega$,

$$D(\omega) \subseteq \left\{P : \int Y^+(\omega, \omega')P(d\omega') \wedge \int Y^-(\omega, \omega')P(d\omega') < \infty\right\}$$

and

$$\omega'(s) = \omega(s) \text{ for all } s \le t_0 \quad \Rightarrow \quad D(\omega) = D(\omega'). \tag{2.1}$$

This will be called the *admissibility function*. Define the cost criterion

$$U\colon \text{graph } D \to \overline{\mathbf{R}}$$

by

$$U(\omega, P) = \int Y(\omega, \omega')P(d\omega'). \tag{2.2}$$

Then $U(\omega, P)$ is the expected total cost if ω is observed up to time t_0, and the controller chooses P as the law of the future after t_0. Denote the optimal cost variable by

$$C(\omega) = \inf_{P \in D(\omega)} U(\omega, P). \tag{2.3}$$

An optimal transition kernel Q^* is a transition probability kernel from $(\Omega, \mathscr{F}_{t_0})$ to $(\Omega, \mathscr{F})$ such that

$$Q^*(\omega, d\omega') \in D(\omega) \text{ and } U(\omega, Q^*(\omega)) = C(\omega) \quad \text{for every } \omega. \tag{2.4}$$

We will show the following existence theorem.

(2.5) THEOREM. *Suppose that E is a Polish space, Z is lower-semi-continuous in the Skorohod topology and is bounded below, and D is a non-empty, compact-valued, measurable multifunction with the property that if $\omega(s) = \omega'(s)$ for all $s \le t_0$ then $D(\omega) = D(\omega')$. Then there exists an optimal transition kernel.*

The proof of Theorem (2.5) requires the following lemma. We will use $\to$ to indicate convergence in the Skorohod topology.

(2.6) LEMMA. *Suppose that $(h_n) \to h$ in H and $(\omega_n) \to \omega$ in Ω. Then*
(i) $(\hat{h}_n) \to \hat{h}$ *in* Ω,
(ii) $(\hat{h}_n \sim \omega_n) \to \hat{h} \sim \omega$ *in* Ω.

Proof. (i) It must be shown that $(r_{[c,d]}\hat{h}_n) \to r_{[c,d]}\hat{h}$ for points of continuity c, d of $\hat{h}$, where $r_{[c,d]}$ is the restriction mapping. If $d \le t_0$, this follows immediately from the fact that $(h_n) \to h$. Suppose that $t_0 \le c$. Then

$$r_{[c,d]}h_n(s) = \hat{h}_n(t_0) \quad \text{for each } n \text{ and all } s,$$

and

$$r_{[c,d]}\hat{h}(s) = h(t_0) \quad \text{for all } s.$$

Since t_0 is a fixed point for every $\lambda \in \Lambda_{[0,t_0]}$, it is clear from (1.4) that $h_n(t_0) \to h(t_0)$, hence the restrictions converge (in fact, uniformly) in the case $t_0 \le c$. The remaining case is $c < t_0 < d$. Since c is a continuity point of h, we may choose a sequence $(\gamma_n) \subseteq \Lambda_{[0,t_0]}$ such that

$$|\gamma_n| \vee \|h_n \circ \gamma_n - h\| \to 0 \text{ as } n \to \infty, \quad \gamma_n(c) = c \text{ for all } n. \tag{2.7}$$

For each n define

$$\mu_n(s) = \begin{cases} \gamma_n(s) & \text{if } c \le s \le t_0; \\ s & \text{if } t_0 < s \le d. \end{cases}$$

Using (1.5) and the inequality

$$\left|\log \frac{x+y}{z+w}\right| \le \max\left\{\left|\log \frac{x}{z}\right|, \left|\log \frac{y}{w}\right|\right\} \quad \text{for } x, y, z, w > 0$$

it is relatively easy to show that $|\mu_n| \to 0$. Also, $\hat{h}_n \circ \mu_n(s)$ is equal to $h_n \circ \gamma_n(s)$ for $s \in [c, t_0]$, and is equal to $h_n(t_0)$ for $s \in (t_0, d]$. From this and (1.4), it follows that $(\hat{h}_n) \to \hat{h}$.

(ii) Again, if $d \le t_0$ or if $t_0 \le c$, then the hypotheses immediately imply that

$$\left(r_{[c,d]}(\hat{h}_n \sim \omega_n)\right) \to r_{[c,d]}(\hat{h} \sim \omega) \quad \text{in } \Omega_{[c,d]},$$

so we consider the case $c < t_0 < d$. Choose (γ_n) as in (2.7), and choose a similar sequence $(\lambda_n) \in \Lambda_{[0,d-t_0]}$ for (ω_n) (note that since d is a point of continuity for $\hat{h} \sim \omega$, $d - t_0$ is a point of continuity for ω). Define for each n,

$$\mu_n(s) = \begin{cases} \gamma_n(s) & \text{if } c \le s \le t_0; \\ t_0 + \lambda_n(s - t_0) & \text{if } t_0 < s \le d. \end{cases}$$

It is tedious, but straightforward to show that $|\mu_n| \to 0$ as $n \to \infty$. Now for $s \in [c, t_0)$, $\hat{h}_n \sim \omega_n(s) = h_n \circ \gamma_n(s)$, and for $s \in [t_0, d]$, $\hat{h}_n \sim \omega_n(s) = \omega_n \circ \lambda_n(s - t_0)$. Then (1.4) implies that $d(r_{[c,d]}(\hat{h}_n \sim \omega_n), r_{[c,d]}(h \sim \omega)) \to 0$ as $n \to \infty$, and (ii) follows.

Proof of Theorem 2.5. Denote by r the restriction mapping from Ω to H. Define

$$\begin{aligned} \hat{D}(h) &= D(\hat{h}), \\ \hat{Y}(h, \omega) &= Y(\hat{h}, \omega) = Z(\hat{h} \sim \omega), \\ \hat{U}(h, P) &= \int \hat{Y}(h, \omega) P(d\omega) \quad \text{for } (h, P) \in \text{graph } \hat{D}. \end{aligned}$$

We show that $\hat{D}$ is measurable. Let B be a closed subset of $\mathscr{P}(\Omega)$, and let ϕ denote the mapping from H to Ω defined by $\phi(h) = \hat{h}$. Then ϕ is 1-1, and by Lemma 2.6(i), it is also continuous, hence measurable. It is easy to check that

$$\{h \in H\colon \hat{D}(h) \cap B \neq \varnothing\} = \phi^{-1}\{\omega \in \Omega\colon D(\omega) \cap B \neq \varnothing\},$$

hence the measurability of D implies that of $\hat{D}$.

By Part (ii) of Lemma 2.6, $\hat{Y}$ is continuous (resp. l.s.c.) if Z is continuous (resp. l.s.c.). As in [16, Lemma 3.4], $\hat{U}$ is l.s.c. and bounded below. Also, it is well known that $\mathscr{P}(\Omega)$ is metrizable as a complete, separable metric space. By [21, Theorem 9.1], there is an optimal measurable selector $\hat{f}$ for $\hat{U}$ and $\hat{D}$. Define $Q^*(\omega;\cdot) = \hat{f} \circ r(\omega)$. By the equality of $\mathscr{B}(H)$ with the canonical history and the measurability of r, Q^* is a transition kernel from $(\Omega, \mathscr{F}_{t_0})$ to $(\Omega, \mathscr{F})$. Also,

$$Q^*(\omega) = \hat{f} \circ r(\omega) \in \hat{D}(r(\omega)) = D(\hat{r}(\omega)) = D(\omega).$$

We have used the second hypothesis on D in the last equality. The optimality of Q^* follows easily from the relation $U(\omega, Q^*(\omega)) = \hat{U}(r(\omega), \hat{f} \circ r(\omega))$. This finishes the proof of the theorem.

3. Control of a semi-Markov step process

Let E be a discrete state space endowed with the discrete topology, and suppose that there is a minimum distance ρ between all points of E. As usual, $\mathscr{X}$ is the canonical process on E with time domain $\mathbf{R}_+$. Let $F(t; \theta)$ be a distribution function dependent on a parameter θ that takes values in a set $\theta \subseteq \mathbf{R}^q$. We will suppose that for $x \in E$, $F(t; \theta^x)$ is the distribution of any sojourn time in state x, where F is considered to be known, but θ^x is not. For each $x \in E$, we have a random variable $\hat{\theta}^x \in \mathscr{F}_{t_0}/\mathscr{B}(\theta)$ that serves as an estimate of θ^x. Let

$$\hat{F}^x(t; \omega) = F(t; \hat{\theta}^x(\omega)),$$

which is, of course, the estimate of the sojourn time distribution for x.

Let $\overline{\Omega} = (E \times [0, \infty))^\infty$. A typical element $\overline{\omega}$ of $\overline{\Omega}$ is written

$$\overline{\omega} = (x_0, s_0, x_1, s_1, \dots).$$

For $\overline{\omega}$ in the subspace $\overline{K}_0$ of $\overline{\Omega}$ such that Σs_i diverges, define $\phi(\overline{\omega}) \in \Omega$ to be the step function that moves through states $x_0, x_1, x_2, \dots$ with jumps at times $s_0, s_0 + s_1, s_0 + s_1 + s_2, \dots$. Given $\omega \in \Omega$ and a transition matrix p on E, there is a unique probability measure $\overline{P} = \overline{P}_{\omega, p}$ on $\overline{\Omega}$ whose projections are

$$\overline{P}(dx_0, ds_0, \ldots, dx_n, ds_n) = \delta_{\omega(t_0)}(dx_0)\hat{F}^{x_0}(ds_0; \omega)\prod_{i=1}^{n} p(x_{i-1}, dx_i)\hat{F}^{x_i}(ds_i, \omega). \quad (3.1)$$

This induces a measure $P = P_{\omega,p}$ on Ω by $P = \overline{P} \circ \phi^{-1}$. If M is a set of admissible transition matrices, define

$$D(\omega) = \left\{ P_{\omega,p} = \overline{P}_{\omega,p} \circ \phi^{-1} \colon \overline{P} \text{ is as in (3.1) and } p \in M \right\}. \quad (3.2)$$

Let L be a bounded, measurable function on E, and let $\alpha > 0$. Define

$$Z(\omega) = \int_0^\infty e^{-\alpha s} L(X_s(\omega))\, ds. \quad (3.3)$$

We will show that, under some conditions, Z is Skorohod l.s.c. and D is compact-valued and measurable, hence by Theorem (2.5) there is an optimal kernel for the problem expressed by (3.1)—(3.3).

(3.4) *Remark.* By the definition of the Skorohod topology and the fact that the points of E are isolated, it is easy to see that if $(\omega_n) \to \omega$ then for any η, $T_0 > 0$ such that T_0 is a point of continuity of ω, there is $N = N(\eta, T_0) > 0$ and there is a sequence $(\lambda_n) \subseteq \Lambda_{[0,\infty)}$ such that for all $n \geq N$,

$$|\lambda_n| < \eta \quad \text{and} \quad \omega_n = \omega \circ \lambda_n \quad \text{on } [0, T_0].$$

In fact, the condition above is also sufficient for Skorohod convergence.

(3.5) LEMMA. *Z is bounded and Skorohod continuous.*

Proof. Since the proof is straightforward, we merely give a sketch. Let $(\omega_n) \to \omega$. Take T_0 large enough that

$$\int_{T_0}^\infty e^{-\alpha t} L(X_t(\omega))\, dt$$

is small, uniformly in ω, and that $\omega_n = \omega \circ \lambda_n$ on $[0, T_0]$ for n large enough, as in Remark (3.4). If the jump times of ω are $T_1, T_2, T_3, \ldots$ then the sojourn intervals for ω_n are $\lambda_n^{-1}(T_k) - \lambda_n^{-1}(T_{k-1})$, $k = 1,2,3\ldots$. For $|\lambda_n| = |\lambda_n^{-1}|$ small, these intervals are close to $T_k - T_{k-1}$, $k = 1,2,3\ldots$. Thus, in $[0, T_0]$, ω_n moves through the same states as ω, and spends nearly the same time in those states, which makes the $[0, T_0]$-discounted reward for ω_n close to that of ω.

The following lemma allows us to transfer our attention from $D(\omega)$ to

$$\overline{D}(\omega) = \left\{ \overline{P}_{\omega,p} \text{ as in (3.1) such that } p \in M \right\}.$$

(3.6) LEMMA. *ϕ is a continuous (relative to the product topology) function from $\overline{K}_0$ to Ω. Thus, if $\overline{D}(\omega)$ is both tight and closed, $D(\omega)$ is compact.*

Proof. Let $\overline{\omega}_n = (x_0^n, s_0^n, x_1^n, s_1^n, \dots) \to \overline{\omega} = (x_0, s_0, x_1, s_1, \dots)$ in the product topology on $\overline{K}_0$. Denote $\omega_n = \phi(\overline{\omega}_n)$ and $\omega = \phi(\overline{\omega})$ and let T_0 be a continuity point of ω. Denote the jump times of ω_n by $t_k^n = \sum_{i=0}^k s_i^n$ and the jump times of ω by $t_k = \sum_{i=0}^k s_i$. Suppose that $t_l < T_0 < t_{l+1}$. Given $\varepsilon > 0$ we can find N large enough that for all $n \geq N$,

$$\max_{0 \leq k \leq l} d(x_k^n, x_k) \vee \max_{0 \leq k \leq l} |s_k^n - s_k| < \rho \wedge \varepsilon, \tag{3.7}$$

which implies that for all $n \geq N$, $x_k^n = x_k$, $k = 0, \dots, l$. We lose no generality in supposing that N is large enough that $t_l^n < T_0 < t_{l+1}^n$ for all $n \geq N$. Then we can define λ_n to be the continuous, piecewise linear function such that $\lambda_n(t_k^n) = t_k$ for $k = 0, \dots, l$, $\lambda_n(T_0) = T_0$, and $\lambda_n = e$ after T_0. Then $\omega_n = \omega \circ \lambda_n$ on $[0, T_0]$. Some tedious arguments, whose details we omit, can be given to show that for any $\eta > 0$, $\varepsilon > 0$ and N can be found so that the condition $|\lambda_n| < \eta$ is satisfied. By Remark (3.4), $\omega_n \to \omega$ and ϕ is continuous.

Now a standard result (for example see [(13) Cor. 1, p. 191]) implies that if $(\overline{P}_{\omega,p_n})$ converges weakly to $\overline{P}_{\omega,p}$, then $(P_{\omega,p_n}) = (\overline{P}_{\omega,p_n} \circ \phi^{-1})$ converges weakly to $P_{\omega,p} = \overline{P}_{\omega,p} \circ \phi^{-1}$. To show $D(\omega)$ is closed, let $(P_n) = (\overline{P}_{\omega,p_n} \circ \phi^{-1})$ be a sequence in $D(\omega)$ converging to some $P \in \mathscr{P}(\Omega)$. If $\overline{D}(\omega)$ is both tight and closed, then it is compact and hence $(\overline{P}_{\omega,p_n})$ has a convergent subsequence converging to some $\overline{P}_{\omega,p}$ in $\overline{D}(\omega)$. By the observation above, (P_{ω,p_n}) has a subsequence converging to $P_{\omega,p}$, therefore $P = P_{\omega,p} \in D(\omega)$. So, $D(\omega)$ is closed. The fact that $D(\omega)$ is tight follows from the tightness of $\overline{D}(\omega)$ [13, Theorem A, p. 195]. This completes the proof.

In light of Lemma (3.6), the following lemma shows the compactness of $D(\omega)$ under some regularity conditions on M and F. Earlier, we had assumed that E is a countable set of isolated states, and in this case, any compact subset of E is actually finite. Thus, the tightness condition on M in (3.8) is weaker than it first appears, and the regularity of F in (3.9) represents no real restriction. There is value, perhaps, in stating, the conditions this way for the purpose of avoiding direct reference in the proof to the structure of E.

(3.7) LEMMA. (a) *Suppose that for all $\delta < 1$ and all compact $K \subseteq E$, there exists compact $L \subseteq E$ such that*

$$p(x, L) > \delta \quad \textit{for all} \quad x \in K,\ p \in M, \tag{3.8}$$

and for all $\omega \in \Omega$, $\delta < 1$, *compact* $K \subseteq E$, *there exists* $[c, d] \subseteq \mathbf{R}_+$ *such that*

$$F([c, d], \hat{\theta}^x(\omega)) > \delta \quad \textit{for all} \quad x \in K. \tag{3.9}$$

Then $\overline{D}(\omega)$ *is tight.*

(b) *Suppose, in addition, the following condition holds: If* $(p_n) \subseteq M$ *is a sequence of transition matrices such that for all* x *in a set* $K \subseteq E$, $p_n(x, y)$ *converges for all* $y \in E$, *then there is* $p \in M$ *such that for all* $x \in K$,

$$\lim_{n \to \infty} p_n(x, y) = p(x, y) \quad \textit{for all} \quad y \in E. \tag{3.10}$$

Then $\overline{D}(\omega)$ *is closed.*

Proof (a) Fix $\omega \in \Omega$, and let $1 > \varepsilon > 0$. By (3.8) and (3.9), we may construct a sequence of compact sets of states and compact time intervals $K_0 = \{\omega(t_0)\}$, $[c_0 d_0]$, $K_1, [c_1, d_1], \ldots$ satisfying the following:

$$F_0 \equiv F([c_0, d_0], \hat{\theta}^{\omega(t_0)}(\omega)) > 1 - \varepsilon,$$

$$P_0 \equiv \inf_{p \in M} p(\omega(t_0), K_1) > \frac{1 - \varepsilon}{F_0},$$

$$F_1 \equiv \inf_{x \in K_1} F([c_1, d_1], \hat{\theta}^x(\omega)) > \frac{1 - \varepsilon}{F_0 P_0},$$

$$P_1 \equiv \inf_{p \in M, x \in K_1} p(x, K_2) > \frac{1 - \varepsilon}{F_0 P_0 F_1}$$

$$\vdots$$

Then the set $\overline{K}_\varepsilon \equiv \prod_{i=0}^{\infty}(K_i \times [c_i d_i])$ is compact in the product topology, and the sequence of finite cylinders

$$\overline{A}_0 \equiv K_0 \times [c_0, d_0] \times \overline{\Omega}, \ \overline{A}_1 \equiv K_0 \times [c_0, d_0] \times K_1 \times [c_1, d_1] \times \overline{\Omega}, \ldots$$

decrease to $\overline{K}_\varepsilon$. It is then a straightforward matter to use the fact that

$$\prod_{i=0}^{m} F_i P_i > 1 - \varepsilon \quad \text{for all } m$$

to show that for all $p \in M$ and all m, $\overline{P}_{\omega, p}(\overline{A}_m) > 1 - \varepsilon$. Therefore,

$$\overline{P}_{\omega, p}(\overline{K}_\varepsilon) \geq 1 - \varepsilon \quad \text{for all } p \in M,$$

and hence $\overline{D}(\omega)$ is tight.

(b) Fix $\omega \in \Omega$, and let $\overline{P}_n = \overline{P}_{\omega, p_n}$ be a sequence in $\overline{D}(\omega)$ that converges weakly to some probability $\overline{P}$. Then the finite-dimensional distributions of the

$\overline{P}_n$'s converge weakly. Because E is discrete, indicator functions 1_y, $y \in E$ are continuous, hence we can define

$$\begin{aligned} p(\omega(t_0), y) &\equiv \lim_{n\to\infty} \int 1_y(x_1) p_n(\omega(t_0), dx_1), \quad && y \in E \\ &= \lim_{n\to\infty} p_n(\omega(t_0), y), && y \in E. \end{aligned}$$

Let $A_1 = \{y \in E \setminus \{\omega(t_0)\}: p(\omega(t_0), y) > 0\}$. If A_1 is empty, then let $K = \{\omega(t_0)\}$ in (3.10) and let $p \in M$ be the associated transition matrix. Then it is clear that $\overline{P}_n \to \overline{P}_{\omega, p}$, hence we now consider the case where A_1 is not empty. For each $y \in A_1$, n can be taken large enough that $p_n(\omega(t_0), y) > 0$. Now the following limit exists:

$$\begin{aligned} &\lim_{n\to\infty} \int\int 1_y(x_1) 1_z(x_2) p_n(\omega(t_0), dx_1) p_n(x_1, dx_2) \\ &\quad = \lim_{n\to\infty} p_n(\omega(t_0), y) p_n(y, z), \end{aligned}$$

and for $y \in A_1$ and $z \in E$, we can define

$$p(y, z) = \frac{\lim_n p_n(\omega(t_0), y) p_n(y, z)}{p(\omega(t_0), y)}.$$

In a similar fashion, we can define $p(z, w)$ for all $w \in E$ and all z in the set

$$A_2 = \{z \in E \setminus A_1 \setminus \{\omega(t_0)\}: p(y, z) > 0 \quad \text{for some} \quad y \in A_1\}.$$

This procedure defines $p(x, y)$ for all $x \in A \equiv \{\omega(t_0)\} \cup A_1 \cup A_2 \cup \dots$ and all $y \in E$, and clearly for such (x, y), $\lim_{n\to\infty} p_n(x, y) = p(x, y)$. By hypothesis (3.10), p can be extended to a transition matrix in M. (It is not hard to show that (3.10) implies that for $x \in A$, $\sum_{y \in E} p(x, y) = 1$.)

Now we have already seen that $\overline{D}(\omega)$ is tight, therefore since every $\overline{P}_n$ is in $\overline{D}(\omega)$, the convergence of finite-dimensional distributions is sufficient to prove that $\overline{P}_n \to \overline{P}_{\omega, p}$. This is straightforward, and the details will be omitted. It follows that $\overline{P} = \overline{P}_{\omega, p}$, hence $\overline{D}(\omega)$ is closed and the proof is finished.

It remains to find conditions under which D is measurable. Once again we may easily reduce to studying $\overline{D}$. Suppose $\overline{D}$ is measurable. Now the mapping

$$\Psi: \overline{P} \to \overline{P} \circ \phi^{-1}$$

is continuous, so that if B is a closed subset of $P(\Omega)$, then $\Psi^{-1}(B)$ is a closed subset of $P(\overline{\Omega})$. Thus,

$$\{\omega: D(\omega) \cap B \neq \varnothing\} = \{\omega: \overline{D}(\omega) \cap \Psi^{-1}(B) \neq \varnothing\} \in \mathscr{F}.$$

Therefore, measurability of $\overline{D}$ implies measurability of D.

By a theorem of Castaing and Rockefellar (Theorem 4.2 of (21)), it suffices to show that there is a countable family of measurable selectors $\alpha_i : \Omega \to P(\overline{\Omega})$ such that

$$\overline{D}(\omega) = \mathrm{cl} \bigcup_i \{\alpha_i(\omega)\} \quad \text{for all} \quad \omega. \tag{3.11}$$

For this we make the assumption that M has a separant $M_s = \{q_1, q_2, \dots\}$ in the sense that for all $p \in M$ and $\delta > 0$, there is $q \in M_s$ such that

$$\sup_x \sum_y |p(x, y) - q(x, y)| < \delta. \tag{3.12}$$

Defining $\alpha_i(\omega) = \overline{P}_{\omega, q_i}$, it then remains to show that for any $\omega \in \Omega$, $\varepsilon > 0$, $p \in M$, there exists $q \in M_s$ such that

$$d\left(\overline{P}_{\omega,p}, \overline{P}_{\omega,q}\right) < \varepsilon,$$

where

$$d\left(\overline{P}_{\omega,p}, \overline{P}_{\omega,q}\right) = \sum_{k=1}^{\infty} \min\left\{(1/2)^k, \left|\int Y_k \, d\overline{P}_{\omega,p} - \int Y_k \, d\overline{P}_{\omega,q}\right|\right\} \tag{3.13}$$

for some countable dense subset $\{Y_1, Y_2, \dots\}$ of the bounded, uniformly continuous functions on $\overline{\Omega}$. (We put the usual bounded metric on the product space $\overline{\Omega}$). Since the tail of the above sum can be made arbitrarily small, it clearly suffices to find q such that

$$\left|\int Y \, d\overline{P}_{\omega,p} - \int Y \, d\overline{P}_{\omega,q}\right| < \varepsilon \tag{3.14}$$

for each of finitely many bounded, uniformly continuous Y's. Furthermore, by setting

$$Y'(x_0, s_0, \dots, x_m, s_m) = Y(x_0, s_0, \dots, x_m, s_m, \overline{\omega}_0)$$

for some fixed ω_0 and large enough m, we may assume without loss of generality that Y is a function of only finitely many coordinates.

Let $\overline{M}$ be a bound for Y. Choose q as in (3.12) for $\delta = \varepsilon / m\overline{M}$. Let

$$f(x_0, \dots, x_m) = \int\int \cdots \int \hat{F}^{x_0}(ds_0, \omega) \cdots \hat{F}^{x_m}(ds_m, \omega) Y(x_0, s_0, \dots x_m, s_m)$$
$$\leq \overline{M}$$

Then, writing $x_0 = \omega(t_0)$, we have

$$\left|\int Y\,d\bar{P}_{\omega,p} - \int Y\,d\bar{P}_{\omega,q}\right| = \Big|\sum_{x_1} p(x_0, x_1) \cdots \sum_{x_m} p(x_{m-1}, x_m) f(x_0, \ldots, x_m)$$
$$- \sum_{x_1} q(x_0 x_1) \cdots \sum_{x_m} q(x_{m-1}, x_m) f(x_0, \ldots x_m)\Big|$$

Now,

$$\left|\sum_{x_m} (p(x_{m-1}, x_m) - q(x_{m-1}, x_m)) f(x_0, \ldots, x_m)\right| \leq \overline{M}\delta,$$

and an easy backward induction argument shows that for $k = 0, 1, \ldots, m-1$,

$$\Big|\sum_{x_{m-k}} p(x_{m-k-1}, x_{m-k}) \cdots \sum_{x_m} p(x_{m-1}, x_m) f(x_0, \ldots, x_m)$$
$$- \sum_{x_{m-k}} q(x_{m-k-1}, x_{m-k}) \cdots \sum_{x_m} q(x_{m-1}, x_m) f(x_0, \ldots, x_m)\Big|$$
$$\leq (k+1)\overline{M}\delta.$$

In particular for $k = m - 1$ we obtain (3.14). The result depended only on the bound $\overline{M}$ and on m, hence the same q will yield (3.14) for finitely many such Y, as needed. Thus, under (3.12), $\overline{D}$ is measurable.

We close by summarizing the results of the section.

(3.15) THEOREM. *Suppose that E is discrete; $F(t; \theta)$ is a distribution function dependent on a parameter $\theta \in \Theta \subseteq \mathbf{R}^q$; D and Z are defined respectively by* (3.2) *and* (3.3), *where L is bounded and $\alpha > 0$; and conditions* (3.8), (3.9), (3.10), *and* (3.12) *hold. Then there exists a transition kernel Q^* from $(\Omega, \mathscr{F}_{t_0})$ to $(\Omega, \mathscr{F})$ such that for all ω,*

$$Q^*(\omega, d\omega') \in D(\omega) \quad \text{and} \quad U(\omega, Q^*(\omega)) = \inf_{P \in D(\omega)} U(\omega, P).$$

Acknowledgment The author wishes to express his appreciation to the referee for some very helpful suggestions.

REFERENCES

1. D.P. BERTSEKAS and S.E. SHREVE, *Stochastic Optimal Control*, Academic Press, Orlando, Florida 1978.
2. P. BILLINGSLEY, *Convergence of probability measures*, *Wiley*, *New York*, 1968.
3. D. BLACKWELL, D. FREEDMAN and M. ORKIN, *The optimal reward operator in dynamic programming*, Ann. Probability, vol. 2 (1974), pp. 926–941.

4. R.M. BLUMENTHAL and R.K. GETOOR, *Markov processes and potential theory*, Academic Press, Orlando, Florida, 1968.
5. R. BOEL and P. VARAIYA, *Optimal control of jump processes*, SIAM J. Control Optimization, vol. 15 (1977), 92–119.
6. B.T. DOSHI, Continuous-time control of Markov processes on an arbitrary state space, Ann. Statist., vol. 4 (1976), 1219–1235.
7. I.V. EVSTIGNEEV, Measurable selection and dynamic programming, *Math. Operations Research*, vol. 1, (1976), 267–272.
8. D. FREEDMAN, *The optimal reward operator in special classes of dynamic programming problems*, Ann. Probability, vol. 2, (1974), pp. 942–949.
9. K.J. HASTINGS, *Control of stochastic processes in discrete and continuous time*, Ph.D. Dissertation, Northwestern Univ., 1981.
10. S.H. HOU, *On Property (Q) and Other Semicontinuity Properties of Multifunctions*, Pacific J. Math., vol. 103 (1982), pp. 39–56.
11. R. P. KERTZ and D.C. NACHMAN, *Persistently optimal plans for non-stationary dynamic programming*, Ann. Probability, vol. 7 (1979), pp. 811–826.
12. K. KURATOWSKI and C. RYLL-NARDZEWSKI, *A general theorem on selectors*, Bull. Polish Acad. Sci., vol. 13 (1965), pp. 397–411.
13. M. LOÈVE, *Probability theory I*, Springer-Verlag, New York, 1977.
14. B. MAISONNEUVE, *Topologies du Type de Skorohod*, LNM 258-Seminaire de Probabilities VI, Springer-Verlag, New York, 1970.
15. S.R. PLISKA, *Controlled jump processes*, Stochastic Processes Appl., vol. 3 (1975), pp. 259–282.
16. M. SCHÄL, *A Selection theorem for optimization problems*, Arch. Math., vol. XXV (1974), pp. 219–224.
17. ______, On dynamic programming, *Stochastic Processes Appl.*, vol. 3 (1975), pp. 345–364.
18. ______, *On dynamic programming and statistical decision theory*, Ann. Statist., vol. 7 (1979), pp. 432–445.
19. L.D. STONE, *Necessary and sufficient conditions for optimal control of semi-Markov jump processes*, SIAM J. Control, vol. 11 (1973), pp. 187–201.
20. W.D. SUDDERTH, *On the existence of good stationary strategies*, Trans. Amer. Math. Soc., vol. 135 (1969), pp. 399–414.
21. D.H. WAGNER, *Survey of measurable selection theorems*, SIAM J. Control, vol. 15 (1977), pp. 859–903.
22. C.B. WAN and M.H.A. DAVIS, *Existence of optimal controls for stochastic jump processes*, SIAM J. Control, vol. 17 (1979), pp. 511–524.
23. W. WHITT, *Some useful functions for functional limit theorems*, Math. Operations Research, vol. 5 (1980), pp. 67–85.

UNIVERSITY OF DELAWARE
NEWARK, DELAWARE