The Annals of Statistics
1976, Vol. 4, No. 2, 370–374

THE STOCHASTIC PROCESSES OF BOREL GAMBLING AND DYNAMIC PROGRAMMING[1]

By David Blackwell

University of California, London

Associated with any Borel gambling model G or dynamic programming model D is a corresponding class of stochastic processes $M(G)$ or $M(D)$. Say that $G(D)$ is *regular* if there is a $D(G)$ with $M(D) = M(G)$. Necessary and sufficient conditions for regularity are given, and it is shown how to modify any model slightly to achieve regularity.

1. Introduction and summary. Bellman's (1957) dynamic progamming and Dubins and Savage's (1965) gambling are close relatives, as noted in Dubins and Savage (1965) and in Blackwell (1965). We here explore one aspect of the relationship: the classes of stochastic processes associated with the two theories. Our formulation is in a Borel, countably additive setting: the sets and functions defining the models are assumed Borel, and the probability measures in the models are countably additive.

A *dynamic programming model* is a triple $D = (X, A, q)$, where X and A are nonempty Borel sets (i.e., Borel subsets of polish spaces) and q is a Borel map from $X \times A$ to the set $P(X)$ of probability distribution on (the Borel subsets of) X. The points of X are called *states* and the points of A are called *acts*. The interpretation is that if you are in state x and choose act a, you move to state x' selected according to $q(x, a)$.

A *gambling model* is a pair $G = (X, B)$, where X is a nonempty Borel set and B is a Borel subset of $X \times P(X)$ such that, for each $x \in X$, the x-section B_x of B, i.e., the set of all $m \in P(X)$ for which $(x, m) \in B$, is nonempty. The points of X are called *fortunes* and the points of B_x are called *gambles available at x*. The interpretation is that if you have fortune x and choose gamble $m \in B_x$, you move to fortune x' selected according to m.

Associated with each D or G is a class of stochastic processes, as follows. An initial state or fortune x_1 is selected according to some specified distribution, and shown to you. You then choose an act a_1 or available gamble m_1 and move to x_2. You observe x_2, choose a_2 or m_2 (your choice may depend on x_1 as well as x_2), move to x_3, etc. More formally, denote by Ω the space of infinite sequences of points of X and by x_n, $n \geqq 1$, the nth coordinate function on Ω. For any $p \in P(\Omega)$, denote by $d_n = d_n(p)$ a version of the conditional distribution of x_{n+1} given $x_1, \cdots, x_n$ with respect to p, so that d_n is a Borel, $(x_1, \cdots, x_n)$-measurable

Received November 1974; revised June 1975.

[1] This research was prepared with the support of National Science Foundation Grant No. GP-43085.

AMS 1970 *subject classifications.* Primary 49C15; Secondary 28A05.

Key words and phrases. Borel, gambling, dynamic programming.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.* www.jstor.org

function from Ω into $P(X)$. We shall say that p is *producible in* D if there is a sequence $\{f_n, n \geqq 1\}$, where f_n is a Borel, $(x_1, \cdots, x_n)$-measurable function from Ω into A, such that, with p-probability 1 for each n,

$$d_n = q(x_n, f_n) .$$

We shall say that p is *producible in* G if for each n, with p-probability 1,

$$(x_n, d_n) \in B .$$

Denote by $M(D)$ and $M(G)$ the class of all p that are producible in D and G respectively.

Call a model of either type *regular* if there is a model of the other type with the same M class: D is *regular* if there is a G with $M(G) = M(D)$; G is *regular* if there is a D with $M(D) = M(G)$.

THEOREM 1. (a) *D is regular if and only if the range $R(D)$ of the function (π, q) from $X \times A$ into $X \times P(X)$ is Borel, where π denotes projection onto X: $\pi(x, a) = x$. If $R(D)$ is Borel, then, with $G = (X, R(D))$, we have $M(G) = M(D)$.*

(b) *G is regular if and only if B contains a Borel graph*, i.e., *there is a Borel f from X into $P(X)$ whose graph is a subset of B. If there is such an f, then, with $D = (X, P(X), q)$, we have $M(D) = M(G)$, where*

$$\begin{aligned} q(x, a) &= a \qquad \text{if} \quad (x, a) \in B \\ &= f(x) \qquad \text{if} \quad (x, a) \notin B . \end{aligned}$$

Theorem 1 describes how to go from a regular model of one type to an equivalent (necessarily regular) model of the other type. The next two theorems describe how to go from a general model of either type to a regular model of the same type.

First, if we are given a $D = (X, A, q)$, it is a small conceptual change to regard yesterday's act as a part of today's state. If we do this, we obtain the *derived model* $D' = (A \times X, A, q')$, where q' describes the joint distribution of today's a and tomorrow's x as a function of yesterday's a, today's x, and today's a:

$$q'((a, x), a') = \delta(a') \times q(x, a') ,$$

where $\delta(a')$ denotes the probability measure concentrated at a'.

THEOREM 2. *For any D, the derived D' is regular.*

Next, if we are given a $G = (X, B)$, it is a small conceptual change to introduce a single new fortune x^*, to which we can always move, and from which we cannot escape. If we do this, we obtain the derived model $G' = (X \cup \{x^*\}, B^*)$, where $B^* = B \cup \{(x, \delta(x^*)) : x \in X \cup \{x^*\}\}$.

THEOREM 3. *For any G, the derived G' is regular.*

Theorems 1, 2, and 3, in giving a general method for going from either type to (a regular) one of the other are a partial explanation of the fact that writers find it largely a matter of taste which model to use, for either developing general theory or studying particular problems. The rest of the explanation presumably lies in relating the kinds of utility or reward functions used, which is outside the scope of this paper.

Finally we give a result whose content can be described, by straining the meaning of words a bit, as follows: (a) in a D-model, you may not be able to determine whether a given process can be produced but you can always, for each initial state, find a producible process with that initial state, while (b) in a G-model, you can always determine whether a given process is producible, but you may not be able, for each initial state, to find a producible process with that initial state.

THEOREM 4. (a) *There is a D for which $M(D)$ is not Borel. For any D there is a Borel f from X into $P(\Omega)$ such that, for each x, $f(x) \in M(D)$ and assigns probability 1 to $\{x_1 = x\}$.*

(b) *For any G, $M(G)$ is Borel. There is a G for which there is no Borel f from X into $P(\Omega)$ such that, for each x, $f(x) \in M(G)$ and assigns probability 1 to $\{x_1 = x\}$.*

2. Proofs and intermediate results. We shall need the following result of Mackey (1957, Theorem 6.3).

MACKEY SELECTION THEOREM. *Let X and Y be Borel sets, let m be a probability measure on X and let B be a Borel subset of $X \times Y$ such that there is an X-set of m-measure 1 on which B_x is nonempty. Then there is a Borel g from X into Y whose graph is in B almost everywhere with respect to m:*

$$m\{x: (x, g(x)) \in B\} = 1 .$$

LEMMA 1. *For any G, any $m \in P(X)$ and any Borel g from X into $P(X)$ whose graph is in B almost everywhere with respect to m, there is a $p \in M(G)$ giving x_1 distribution m and with $d_1 = g(x_1)$ with p-probability 1.*

PROOF. Since m is the distribution of x_1, p will be determined by the specification of d_n, $n \geqq 1$. Of course $d_1 = g(x_1)$. Inductively, suppose $d_1, \cdots, d_n$ defined, and let m_n be the resulting distribution of x_{n+1}. Apply the Mackey selection theorem to X, $P(X)$, m_n, B to obtain g_n, and define $d_{n+1} = g_n(x_{n+1})$. The resulting p has the required properties.

LEMMA 2. *For any G and any $m \in P(X)$, there is a $p \in M(G)$ giving x_1 distribution m.*

PROOF. The Mackey selection theorem applied to X, $P(X)$, m gives a g that satisfies the hypothesis of Lemma 1.

We remark that the Mackey selection theorem is in turn an easy consequence of Lemma 2, or indeed of the fact that every $M(G)$ is nonempty.

LEMMA 3. *For any D and G with the same X*

(a) *$M(D) \subset M(G)$ if and only if $R(D) \subset B$.*
(b) *$M(D) \supset M(G)$ if and only if $R(D) \supset B$.*

PROOF. (a) Suppose $R(D) \subset B$ and let $p \in M(D)$. Say $d_n = q(x_n, f_n)$. Then $(x_n, d_n) \in R(D)$, so $(x_n, d_n) \in B$ and $p \in M(G)$.

Now suppose $M(D) \subset M(G)$ and let $(x, m) \in R(D)$. Say $m = q(x, a_0)$. Let p concentrate x_1 at x and have $d_n = q(x_n, a_0)$. Then $p \in M(D)$, so $p \in M(B)$, so with p-probability 1, $(x_1, d_1) \in B$, i.e. $(x, m) \in B$.

(b) Suppose $R(D) \supset B$ and let $p \in M(G)$, so that $p\{(x_n, d_n) \in B\} = 1$. Apply the Mackey selection theorem to Ω, A, p, and $J_n = \{(w, a) : q(x_n, a) = d_n\}$, obtaining g_n from Ω into A with $q(x_n, g_n) = d_n$ with p-probability 1. Since the a-sections of J_n are $(x_1, \cdots, x_n)$-measurable, there is a g_n that is also $(x_1, \cdots, x_n)$-measurable. Thus $p \in M(D)$.

Now suppose $M(D) \supset M(G)$ and let $(x, \mu) \in B$. Apply Lemma 1 to obtain $p \in M(G)$ concentrating x_1 at x and having $d_1 = \mu$. Since $p \in M(D)$, there is an a_0 with $\mu = q(x, a_0)$ i.e. $(x, \mu) \in R(D)$.

COROLLARY. *$M(D) = M(G)$ if and only if $R(D) = B$.*

Theorem 1(a) follows immediately from the corollary. To prove Theorem 1(b), suppose first that G is regular, say $M(D) = M(G)$. Then for any a, the graph of $q(\cdot, a)$ is in B, since $B = R(D)$. Conversely, if B contains the graph of the Borel f, the D defined in Theorem 1(b) clearly has $R(D) = B$, so that $M(D) = M(G)$.

To prove Theorem 2 it suffices, from Theorem 1(a), to see that $R(D')$ is Borel. This is shown by the description

$$R(D') = \{(x, m) : m_A \in \Delta \quad \text{and} \quad m = \phi(m_A) \times q(x_1, \phi(m_A))\} ,$$

where m_A is the A-marginal distribution of m (the map $m \to m_A$ is Borel), Δ is the (Borel) subset of $P(A)$ consisting of distributions that are concentrated at single points, and ϕ is the (Borel) function on Δ that associates with each μ in Δ the point on which it concentrates: $\mu = \delta(\phi(\mu))$. (See Dubins and Freedman (1968).)

To prove Theorem 3, apply Theorem 1(b) to G', with f the constant function with value $\delta(x^*)$.

For the example of Theorem 4(a), take any X, A, h, where h is a Borel function from A into X whose range is not Borel, and take $D = (X, A, q)$, where $q(x, a) = \delta(h(a))$. The set of x for which the measure that concentrates on $(x, x, x, \cdots)$ is in $M(D)$ is just the range of h, which is not Borel, so $M(D)$ is not Borel. But for any D and any $a_0 \in A$, the function f such that $f(x)$ concentrates x_1 at x and has $d_n = q(x_n, a_0)$ has the property asserted in Theorem 4(a).

Any (X, B) for which B contains no Borel graph gives an example for Theorem 4(b) since, if there is an f with the properties of 4(b), the function $d_1(f)$ is a Borel function with graph in B.

Finally, we must show that any $M(G)$ is Borel. According to a nice result of Sudderth (1969, Lemma 2, page 402), for any two Borel functions u and v on a Borel set Z, there is a Borel ϕ from $P(Z)$ into $P(P(V))$, where V is the range space of v, such that $\phi(p)$ is the distribution of the conditional distribution of v given u with respect to p. We apply Sudderth's result with $Z = \Omega$, $u = (x_1, \cdots, x_n)$, $v = (x_n, x_{n+1})$ to obtain ϕ_n. The set H of measures in $P(V)$ of the form $\delta(x) \times m$ for some $(x, m) \in B$ is Borel, and $M(G) = \{p: \phi_n(p) \in H$ for all $n\}$.

REFERENCES

[1] BELLMAN, RICHARD (1957). *Dynamic Programming.* Princeton Univ. Press.

[2] BLACKWELL, D. (1965). Discounted dynamic programming. *Ann. Math. Statist.* **36** 226–235.

[3] DUBINS, LESTER and FREEDMAN, DAVID (1965). Measurable sets of measures. *Pacific J. Math.* **14** 1211–1222.

[4] DUBINS, L. E. and SAVAGE, L. J. (1965). *How to Gamble If You Must.* McGraw-Hill, New York.

[5] MACKEY, G. W. (1957). Borel structures in groups and their duals. *Trans. Amer. Math. Soc.* **85** 134–165.

[6] SUDDERTH, W. (1969). On the existence of good stationary strategies. *Trans. Amer. Math. Soc.* **135** 399–414.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720