PACIFIC JOURNAL OF MATHEMATICS
Vol. 94, No. 2, 1981

NEARLY STRATEGIC MEASURES

THOMAS E. ARMSTRONG AND WILLIAM D. SUDDERTH

Every finitely additive probability measure α defined on all subsets of a product space $X\times Y$ can be written as a unique convex combination $\alpha=p\mu+(1-p)\nu$ where μ is approximable in variation norm by strategic measures and ν is singular with respect to every strategic measure.

1. Introduction. For each nonempty set X, let $P(X)$ be the collection of finitely additive probability measures defined on all subsets of X. A *conditional probability* on a set Y given X is a mapping from X to $P(Y)$. *A strategy* σ on $X\times Y$ is a pair (σ_0, σ_1) where σ_0 is in $P(X)$ and σ_1 is a conditional probability on Y given X. Each strategy σ on $X\times Y$ determines a *strategic measure*, also denoted σ, in $P = P(X\times Y)$ by the formula

$$\sigma g = \iint g(x, y)d\sigma_1(y|x)d\sigma_0(x) ,$$

where g is a bounded, real-valued function on $X\times Y$. The collection Σ of all strategic measures was studied by Lester Dubins [3], who proved that, if X or Y is finite, then every member of P is *nearly strategic* in the sense that it can be approximated arbitrarily well in the sense of total variation by a strategic measure. However, Dubins also showed that if X and Y are infinite, then the collection $\bar{\Sigma}$ of all nearly strategic measures is a proper subset of P and, moreover, there exist elements in $\Sigma^{\perp}(=\bar{\Sigma}^{\perp})$, the set of measures in P singular with respect to every measure in Σ. (As usual, the finitely additive probability measures μ and ν are mutually singular if, for every positive ε, there is a set A such that $\mu(A) < \varepsilon$ and $\nu(A) > 1 - \varepsilon$.)

Here is our main result.

THEOREM 1. $\Sigma^{\perp\perp} = \bar{\Sigma}$.

This answers a question posed by Dubins in [3]. As Dubins pointed out, the following corollary is a consequence of Theorem 1 together with results of Bochner and Phillips [1].

COROLLARY 1. *Every μ in P can be written in the form*

$$\mu = p\sigma + (1 - p)\tau$$

with $\sigma\in\bar{\Sigma}$, $\tau\in\Sigma^{\perp}$, and $0\leqq p\leqq 1$ where $p\sigma$, $(1-p)\tau$, and p are unique.

The next section presents a proof of Theorem 1. The final section gives a generalization.

2. The proof of Theorem 1. Let $\mathscr{B}$ be the algebra of all subsets of $X \times Y$ and let $P = P(X \times Y)$ be the set of all finitely additive probability measures on $\mathscr{B}$. Equip P with the topology induced by the total variation norm which is defined, for $\mu, \nu \in P$, by

$$\|\mu - \nu\| = \sup\{|\mu(B) - \nu(B)| : B \in \mathscr{B}\} . \tag{1}$$

Recall that ν is absolutely continuous with respect to μ, written $\nu \ll \mu$, if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that, for all $B \in \mathscr{B}$, $\mu(B) < \delta$ implies $\nu(B) < \varepsilon$. By a *simple function* f is meant a real-valued function defined on $X \times Y$ which assumes only a finite number of values. A *μ-density* is a bounded nonnegative function on $X \times Y$ whose *μ-integral* is equal to one. The measure whose value at $B \in \mathscr{B}$ is $\int_B f d\mu$ is denoted $fd\mu$.

LEMMA 1. *The following three conditions on a closed subset S of P are equivalent.*

(a) $\mu \in S, \nu \ll \mu \Rightarrow \nu \in S$.

(b) $\mu \in S, k > 0, \nu \leqq k\mu \Rightarrow \nu \in S$.

(c) $\mu \in S, f$ *a simple μ-density* $\Rightarrow fd\mu \in S$.

Proof. That (a) $\Rightarrow$ (b) $\Rightarrow$ (c) is trivial. That (c) $\Rightarrow$ (a) follows from Bochner's finitely additive Radon-Nikodym theorem [2] and the assumption that S is closed. □

PROPOSITION 1. *For a closed, convex subset S of P to satisfy $S = S^{\perp\perp}$, it suffices that any (all) of the conditions of Lemma 1 be satisfied.*

Proof. Let M be the linear space spanned by S in the space L of all finite, finitely additive, signed measures on $\mathscr{B}$. The major part of the proof consists of the verification that M is a closed vector lattice which satisfies (4) below. Several properties of M will be established. For the first, make the harmless assumption that S is not empty.

(2) For every $\mu \in M$, there exist $\lambda \in S$ and $k > 0$ such that $|\mu| \leqq k\lambda$.

To see this, write $\mu = a_1\mu_1 - a_2\mu_2$ where $a_i \geqq 0$ and $\mu_i \in S$. Let $k = a_1 + a_2$. If $k = 0$, then $\mu = 0$ and (2) is trivial. If $k > 0$, set

$\lambda = k^{-1}(a_1\mu_1 + a_2\mu_2)$. By the convexity of S, $\lambda \in S$. Clearly, $|\mu| \leqq k\lambda$.

The following partial converse to (2) is an easy consequence of condition (b) of Lemma 1.

(3) If μ is a nonnegative, nonzero element of L and if $\mu \leqq k\lambda$ for some $\lambda \in S$ and $k > 0$, then $\|\mu\|^{-1}\mu \in S$ and, hence, $\mu \in M$.

It is now possible to check the following.

$$(4) \qquad \mu \in M, \nu \in L, |\nu| \leqq |\mu| \Longrightarrow \nu \in M .$$

For by (2), $\nu^+ \leqq |\nu| \leqq |\mu| \leqq k\lambda$ for some $k > 0$ and $\lambda \in S$. By (3), $\nu^+ \in M$. Similarly, $\nu^- \in M$. Hence, $\nu = \nu^+ - \nu^- \in M$.

To see that M is a lattice, use (2) and the convexity of S to see that the supremum of two elements of M is dominated in absolute value by a scalar multiple of an element of S. Then use (4).

To check that M is closed in the total variation norm topology of L, let $\mu_n \in M$ and suppose μ_n converges to μ, a nonzero element of L. Assume first that the μ_n are nonnegative. Then, for n large, $\|\mu_n\| \geqq 2^{-1}\|\mu\| > 0$. By (2), each μ_n is dominated by a scalar multiple of some element of S and so, by (3) the measures $\nu_n = \|\mu_n\|^{-1}\mu_n$ belong to S. Clearly, ν_n converges to $\nu = \|\mu\|^{-1}\mu$. Since, by hypothesis, S is closed, $\nu \in S$. Hence, $\mu \in M$. The general case follows by taking positive and negative parts. So M is indeed a closed vector lattice which satisfies (4). This implies that $M = M^{\perp\perp}$, which is the content of Theorem 2 of Bochner and Phillips [1]. Consequently,

$$S^{\perp\perp} \subset P \cap M^{\perp\perp} = P \cap M \subset S .$$

The first inclusion and the equality are obvious. The final inclusion follows from properties (2) and (3). □

COROLLARY 2. *For a subset S of P to satisfy $\bar{S} = S^{\perp\perp}$, it suffices that these two conditions hold*: (i) $\mu, \nu \in S \Rightarrow (\mu + \nu)/2 \in \bar{S}$, (ii) $\mu \in S, f$ *a simple μ-density* $\Rightarrow fd\mu \in \bar{S}$.

Proof. Condition (i) implies that $\bar{S}$ contains the convex hull of S and, hence, is the closure of the convex hull of S and, in particular, a convex set. From condition (ii) it easily follows that condition (c) of Lemma 1 holds when S is replaced there by $\bar{S}$. Proposition 1 now applies. □

The conditions of Proposition 1 and Corollary 2 are not only sufficient, but as can be shown, necessary. In addition, the arguments presented show that these results hold for a general Boolean algebra of sets and not only for the algebra $\mathscr{B}$ of special interest here.

The rest of this section is devoted to the verification of conditions (i) and (ii) of Corollary 2 when S is the set Σ of strategic measures $X\times Y$. The argument is given in three lemmas. To state the first, associate to each $\alpha \in P(X\times Y)$ its marginal $\alpha_0 \in P(X)$ where $\alpha_0(E) = \alpha(E\times Y)$ for all $E\subset Y$.

LEMMA 2. *Suppose Z is a finite set, $\alpha \in P(X\times Z)$, and $\varepsilon > 0$. Then there is a strategy β on $X\times Z$ such that $\beta_0 = \alpha_0$ and $\|\alpha - \beta\| < \varepsilon$.*

Proof. This is a special case of Dubins [3, Proposition 1]. □

LEMMA 3. *If $\sigma, \tau \in \Sigma$, then $(\sigma + \tau)/2 \in \bar{\Sigma}$.*

Proof. Let $\varepsilon > 0$ and set $\mu = (\sigma + \tau)/2$. It suffices to find $\nu \in \Sigma$ such that

$$\|\mu - \nu\| \leqq \varepsilon \ . \tag{5}$$

Define $\nu_0 = \mu_0$; that is, $\nu_0 = (\sigma_0 + \tau_0)/2$. To define ν_1, first let $Z = \{0, 1\}$ and consider the strategy λ on $Z\times X$ which has $\lambda_0 = (\delta(0) + \delta(1))/2$, $\lambda_1(0) = \sigma_0$, and $\lambda_1(1) = \tau_0$. (Here $\delta(\mathrm{i})$ denotes the measure which assigns mass 1 to the singleton $\{\mathrm{i}\}$.) Next consider the measure α on $X\times Z$ obtained from λ by reversing the cordinates; in other terms, for each bounded, real-valued function g on $X\times Z$, $\alpha g = \lambda\tilde{g}$ where $\tilde{g}(z, x) = g(x, z)$. Notice that

$$\alpha_0 = (\sigma_0 + \tau_0)/2 = \nu_0 \ .$$

Apply Lemma 2 to obtain a strategy β on $X\times Z$ with

$$\beta_0 = \alpha_0 \ , \quad \|\alpha - \beta\| < \varepsilon \ . \tag{6}$$

Now define

$$\nu_1(x) = \beta_1(x)(\{0\})\sigma_1(x) + \beta_1(x)(\{1\})\tau_1(x)$$

for each $x \in X$. It remains to verify (5).

To that end, let $A\subset X\times Y$ and define $g\colon X\times Z \to [0, 1]$ by

$$g(x, 0) = \sigma_1(x)(Ax) \ , \quad g(x, 1) = \tau_1(x)(Ax) \ ,$$

where

$$Ax = \{y\colon (x, y) \in A\} \ .$$

It follows from (6) that

$$(7) \qquad |\alpha g - \beta g| \leqq \varepsilon .$$

However,

$$(8) \qquad \begin{aligned} \alpha g = \lambda \tilde{g} &= \iint g(x, z) d\lambda_1(x|z) d\lambda_0(z) \\ &= \frac{1}{2}\int \sigma_1(x)(Ax) d\sigma_0(x) + \frac{1}{2}\int \tau_1(x)(Ax) d\tau_0(x) \\ &= (\sigma(A) + \tau(A))/2 \\ &= \mu(A) , \end{aligned}$$

and

$$(9) \qquad \begin{aligned} \beta g &= \iint g(x, z) d\beta_1(z|x) d\beta_0(x) \\ &= \int [\beta_1(x)(\{0\}) g(x, 0) + \beta_1(x)(\{1\}) g(x, 1)] d\beta_0(x) \\ &= \int \nu_1(x)(Ax) d\nu_0(x) \\ &= \nu(A) . \end{aligned}$$

Because A is an arbitrary subset of $X \times Y$, the desired inequality (5) now follows from (7), (8), and (9). □

The next lemma can be viewed as a variant of Bayes formula and its proof is hardly different from the proof in the countably additive case as given, for example, by Renyi [4, Example 5.1.1].

LEMMA 4. *If $\sigma \in \Sigma$ and f is a σ-density, then $\nu = f d\sigma \in \Sigma$. Indeed, if $g(x) = \int f(x, y) d\sigma_1(y|x)$, then ν is the strategy (ν_0, ν_1) where $\nu_0 = g d\sigma_0$,*

$$\nu_1(x) = \frac{f(x, .)}{g(x)} d\sigma_1(\cdot|x) \quad \textit{if} \quad g(x) > 0 ,$$

and $\nu_1(x)$ is an arbitrary probability measure on Y if $g(x) = 0$.

Proof. Let $B = \{x \in X: g(x) > 0\}$. It is easy to verify that $\nu_0(B) = 1$. Now let φ be a bounded function on $X \times Y$ and calculate as follows:

$$\begin{aligned} \nu\varphi &= \int (\varphi \cdot f) d\sigma \\ &= \int_B \int \varphi(x, y) \frac{f(x, y)}{g(x)} d\sigma_1(y|x) g(x) d\sigma_0(x) \\ &= \iint \varphi(x, y) d\nu_1(y|x) d\nu_0(x) . \end{aligned}$$ □

Theorem 1 now follows from Corollary 2, Lemma 3, and Lemma 4.

3. Nearly disintegrable measures. Let T be a mapping which assigns to each $x \in X$ a nonempty subset T_x of Y. A measure $\mu \in P(Y)$ is *T-disintegrable* if there is a strategy σ on $X \times Y$ such that $\sigma_1(x)(T_x) = 1$ for all x and

$$\mu(A) = \int \sigma_1(x)(A \cap T_x) d\sigma_0(x)$$

for all $A \subset Y$. Let D be the collection of all such T-disintegrable measures.

THEOREM 2. $D^{\perp\perp} = \bar{D}$.

COROLLARY 3. *Every $\alpha \in P(Y)$ can be written in the form*

$$\alpha = p\mu + (1 - p)\nu$$

with $\mu \in \bar{D}, \nu \in D^{\perp}$, and $0 \leqq p \leqq 1$ where $p\mu$, $(1 - p)\nu$, and p are unique.

In the special case when $Y = X \times Z$ and $T_x = \{x\} \times Z$ for all x, Theorem 2 easily reduces to Theorem 1 for the product space $X \times Z$.

The proof of Theorem 2, like that of Theorem 1, is based on Corollary 2. Let E be that subset of $X \times Y$ given by $E = \{(x, y): y \in T_x\}$ and let P_E be the set of μ in $P(X \times Y)$ such that $\mu(E) = 1$. That properties (i) and (ii) of Corollary 2 hold for D follows from the fact that they hold for Σ together with the fact that D is the image of $\Sigma \cap P_E$ under the affine mapping which sends a measure on $X \times Y$ to its marginal on Y.

It should be remarked that the notion of disintegrability used here is slightly more general than the usual one which is that a measure μ in $P(Y)$ is disintegrable under the mapping φ of Y onto X if there is a $\sigma_0 \in P(X)$ and, for each $x \in X$, there is a $\sigma_1(x) \in P(\varphi^{-1}(x))$, such that

$$\mu(A) = \int \sigma_1(x)(A \cap \varphi^{-1}(x)) d\sigma_0(x)$$

for all $A \subset Y$. The main difference is that the definition here does not require that the sets $\{T_x\}$ form a partition of Y as do the sets $\{\varphi^{-1}(x)\}$.

ACKNOWLEDGMENT. After we had written this paper, Lester Dubins sent us a copy of unpublished notes of Annibal Sant'anna. These notes, written several years ago when Sant'anna was a

graduate student at U. C. Berkeley, contain results which represent a genuine contribution towards an affirmative answer to the question raised by Dubins whether $\Sigma^{\perp\perp}$ is $\bar{\Sigma}$.

References

1. S. Bochner and R. S. Phillips, *Additive set functions and vector lattices*, Annals of Math., **42** (1941), 316–324.
2. L. E. Dubins, *An elementary proof of Bochner's finitely additive Radon-Nikodym theorem*, Amer. Math. Monthly, **76** (1969), 520–523.
3. ———, *Finitely additive conditional probabilities, conglomerability, and disintegrations*, Annals of Probability, **3** (1975), 89–99.
4. Alfred Renyi, *Foundations of Prabability*, Holden-Day, San Francisco, (1970.

Received February 26, 1979. Research supported by National Science Foundation Grants MCS74-05786-A02 (for Armstrong) and MCS77-28424 (for Sudderth).

UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455