ISRAEL JOURNAL OF MATHEMATICS, Vol. 46, No. 4, 1983

DISTRIBUTIONS OF RANDOM SETS AND RANDOM SELECTIONS

BY
ZVI ARTSTEIN

ABSTRACT
Distributions of selections of a random set are characterized in terms of inequalities, similar to the marriage problem. A consequence is that the ensemble of such distributions is convex compact and depends continuously on the distribution of the random set.

1. Introduction

Set valued mappings and their measurable selections arise in several applications. These include pattern analysis, stochastic geometry, optimization and mathematical economics. In the latter applications one is interested primarily in the structure and the distributions of the measurable selections. In the former fields the set-valued map is often viewed as a random set, and the distributions of these random variables are of interest. General references to such applications are Grenander [3], Kendall [8], Matheron [9] and Hildenbrand [7].

In this paper we analyze the connections between the distribution of a random set and the ensemble of distributions induced by possible selections. In Section 2 we characterize the distributions of possible selections in terms similar to the well known marriage problem. Indeed, a continuous version of this problem serves as the main tool in the proof. This generalization of the marriage problem is given in the self-contained Section 3. A by-product of the proof is a construction of a canonical measure space in which all the possible distributions of selections can be realized. This result is also stated in Section 2. The proofs of these theorems are given in Section 4. We apply the results and show in Section 5 that the ensemble of distributions of possible selections is compact convex and varies continuously. Further comments are listed in the closing section.

2. Characterization of selectionable distributions

We first set the framework and collect a few known facts. Let (X, d) be a

Received May 31, 1983

complete separable metric space. (Euclidean spaces and Banach spaces appear in most of the applications.) Let $\mathcal{K}$ denote the family of compact subsets of X. With the Hausdorff distance, given by

$$h(K,L)=\max\left(\max_{x\in K}\min_{y\in L}d(x,y),\ \max_{y\in L}\min_{x\in K}d(x,y)\right),$$

the space $\mathcal{K}$ becomes a complete separable metric space.

Let (Ω,ν) be a probability measure space. A *random set* is a measurable mapping $F:\Omega\to\mathcal{K}$, the latter considered with its Borel structure. The distribution DF of F is, as usual, the probability measure on $\mathcal{K}$ given by

$$DF(A)=\nu(\{\omega : F(\omega)\in A\}).$$

A *selection* of F is a measurable mapping $f:\Omega\to X$ with $f(\omega)\in F(\omega)$ for ν-almost every ω. The distribution Df of f is a probability measure on X. Given a probability distribution σ on $\mathcal{K}$ we wish to know which distributions ρ on X are induced by selections of random sets F with $DF=\sigma$. Notice that checking one such F might not be enough, since the distribution of a random set does not determine the collection of distributions of its selections. (See more on this in the closing section.) We therefore say that a distribution ρ on X is *selectionable* with respect to a distribution σ on $\mathcal{K}$ if there is a measure space (Ω,ν) and a random set F on it with $DF=\sigma$, and a selection f of F with $Df=\rho$. (Notice that for C closed in X the set $\{K : K\cap C\neq\emptyset\}$ is closed, hence measurable, in $\mathcal{K}$.)

THEOREM 2.1. *A distribution ρ on X is selectionable with respect to a distribution σ on $\mathcal{K}$ if and only if*

$$(\#)\qquad \rho(C)\leqq\sigma(\{K : K\cap C\neq\emptyset\})\qquad \textit{for all } C\subset X \textit{ closed.}$$

Denote by $\mathcal{K}_{\text{ex}}$ the product space $[0,1]\times\mathcal{K}$. For a probability distribution σ on $\mathcal{K}$ let σ_{ex} be defined on $\mathcal{K}_{\text{ex}}$ as the product of the Lebesgue measure on $[0,1]$ and σ. Let F_{ex} be the set-valued function defined on $\mathcal{K}_{\text{ex}}$ by $F_{\text{ex}}(t,K)=K$. It is clear that the distribution of F_{ex} with respect to $(\mathcal{K}_{\text{ex}},\sigma_{\text{ex}})$ is equal to σ.

THEOREM 2.2. *Let σ be a distribution on $\mathcal{K}$. Then any distribution ρ on X which is selectionable with respect to σ is the distribution Df of a selection of the random set F_{ex} on $(\mathcal{K}_{\text{ex}},\sigma_{\text{ex}})$.*

3. A result on matchings

Let G and B (girls and boys) be separable complete metric spaces and let σ and ρ be probability distributions defined on G and B respectively. Let F be a closed subset (friendships) of the product space $G \times B$. We write $F(g)$ for the set $\{b : (g, b) \in F\}$. A *proper matching* is a measurable function $f : G \to B$ such that $f(g) \in F(g)$ for almost all g, and which preserves densities, namely $\sigma(f^{-1}(A)) = \rho(A)$ for every $A \subset B$ closed.

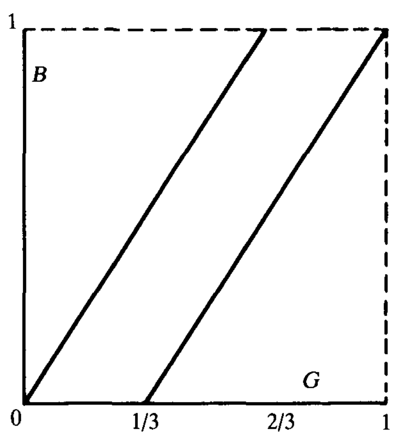
It is clear that a necessary condition for the existence of a proper matching is

$$(*) \qquad \sigma(F^{-1}(A)) \geqq \rho(A) \qquad \text{for all } A \subset B \text{ closed.}$$

(Here $F^{-1}(A) = \{g : F(g) \cap A \neq \emptyset\}$; it is clearly a measurable set.)

When G and B are finite, same number, and with the counting measure then the setting fits the well known marriage problem, and Hall's theorem says that (*) is both necessary and sufficient for the existence of a proper matching. The same is true when σ is atomless and ρ is purely atomic; see Bollobás and Varopoulos [2]. In general, condition (*) is not sufficient, even if the matching f is allowed to be multivalued. Here is an example. The sets G and B are copies of the unit interval with the Lebesgue measure. Each $b \in B$ is friendly with exactly two elements in G, which are $2b/3$ and $(2b+1)/3$. Hence F consists of the two lines of slope $\frac{3}{2}$ in Fig. 1. It is not hard to check that condition (*) is valid. If f is a proper matching then necessarily $f(g) = (3g-1)/2$ for $g \in (\frac{2}{3}, 1]$. Hence in order to satisfy the matching condition for an interval $A = [\alpha, \beta]$ which is a subinterval of $(\frac{1}{2}, 1]$ in B the value $3g/2$ must equal (or belong to) $f(g)$ for one half of the points in $(2\alpha/3, 2\beta/3) \subset G$. This should hold for every choice (α, β). Therefore $f(g)$ cannot be measurable.

We shall prove two positive results. The second establishes the existence of a proper matching when each type of the elements in G is large. The proof relies on the first result which is concerned with general G and B but applies only to



matchmakers with a sense to probability, as follows. A *mixed matching* is a measurable function $\mu(g)$ which associates with each $g \in G$ a probability measure on B with support included in $F(g)$, and which preserves densities in the sense that $\rho(A) = \int_G \mu(g)(A) d\sigma$ for all $A \subset B$ closed.

Notice that if the support of $\mu(g)$ is singleton, say $f(g)$, then the mixed matching is a proper matching. Notice also that in the example described in Fig. 1 it is quite easy to find a mixed matching, namely with $\mu(g)$ equally distributed among $F(g)$. Our general result is as follows.

THEOREM 3.1. *Suppose that for every g the set $F(g)$ is compact. Then condition* (*) *is necessary and sufficient for the existence of a mixed matching.*

The second positive result is:

THEOREM 3.2. *Suppose G is of the form $[0,1] \times G_0$ and σ is the product of the Lebesgue measure on $[0,1]$ and a distribution σ_0 on G_0. Suppose also that for every $g = (t, g_0)$ in G the set $F(g)$ is compact and depends only on the G_0-coordinate g_0. Then condition* (*) *is necessary and sufficient for the existence of a proper matching.*

The proofs involve several steps which we list as lemmas.

LEMMA 3.3. *There is a set B_1 in B which is a countable union of compact sets and such that $F(g) \subset B_1$ for almost every g in G.*

PROOF. Let $b_1, b_2, \cdots$ be a dense sequence in B and let $A_{n,i}$ denote the closed ball with radius $1/i$ around b_i. For a given $\varepsilon > 0$, and for each n there is an index $i(n)$ such that $\sigma(\{g : F(g) \subset C_n\}) \geqq 1 - \varepsilon 2^{-n}$, when $C_n = A_{n,1} \cup \cdots \cup A_{n,i(n)}$. The set $C = \bigcap C_n$ is then totally bounded, hence compact, and $\sigma(\{g : F(g) \subset C\}) \geqq 1 - \varepsilon$. The union of such sets C determined by a sequence of ε which converge to 0 is the desired set B_1.

Let $\mathcal{P}$ denote the collection of probability distributions on B, considered with the weak convergence of measures. The latter convergence is metrizable; see Billingsley [1], chapter 1 and appendix III. Let $\mathcal{M}$ denote the family of measurable mappings $\mu : G \to \mathcal{P}$. With each $\mu \in \mathcal{M}$ we associate the probability measure on $G \times B$ obtained as the integration of $\mu(g)$ with respect to the marginal σ (i.e. the measure of $C \times A$ is $\int_C \mu(g)(A) d\sigma$). We use the same notation μ to denote the associate probability measure on $G \times B$, and the weak convergence of these measures is served as a convergence notion on $\mathcal{M}$. With each $\mu \in \mathcal{M}$ we also associate a distribution $p(\mu)$ on B which is the marginal distribution on B of the probability measure μ on $G \times B$.

LEMMA 3.4. *$p(\mu): \mathcal{M} \to \mathcal{P}$ is continuous.*

PROOF. Well known; see Billingsley [1], theorem 3.1.

In $\mathcal{M}$ we identify a subfamily $\mathcal{M}_0$ which consists of all the elements μ in $\mathcal{M}$ such that the support of $\mu(g)$ is contained, for almost every g, in $F(g)$. Notice that in Theorem 3.1 we promised an element μ in $\mathcal{M}_0$ with $p(\mu) = \rho$. Recall that a collection of probability measures on a metric space S is tight if for every $\varepsilon > 0$ there is a compact set $K \subset S$ such that $r(K) \geqq 1 - \varepsilon$ for every r in the collection.

LEMMA 3.5. *The collection of distributions on $G \times B$ which are associated with elements in $\mathcal{M}_0$ is tight.*

PROOF. Let B_1 be given by Lemma 3.3 and let $A_0 \subset B_1$ be compact and such that $\sigma(\{g : F(g) \subset A_0\}) > 1 - \varepsilon/2$. The measure σ itself is tight (Billingsley [1], theorem 1.4) and let $C \subset G$ be compact with $\sigma(C) > 1 - \varepsilon/2$. Let C_0 be the closure of $C \cap \{g : F(g) \subset A_0\}$. Then $\sigma(C_0) > 1 - \varepsilon$. The condition that $\mu(g)$ is supported in A_0 for $g \in C_0$ implies that $\mu(C_0 \times A_0) > 1 - \varepsilon$ for $\mu \in \mathcal{M}_0$. This completes the proof.

LEMMA 3.6. *Every sequence in $\mathcal{M}_0$ has a converging sequence with limit in $\mathcal{M}_0$.*

PROOF. Prohorov's theorem (Billingsley [1], theorem 6.1) implies that every sequence μ_k has a converging subsequence. Let ν be the weak limit. We have to check that ν is in $\mathcal{M}_0$, namely that it can be disintegrated with respect to σ, and that in the disintegration, $\nu(g)$ is a probability measure supported in $F(g)$. The measure σ serves as the marginal measure for every μ_k, hence $\nu(g)$ is well defined by the disintegration. It is supported in $F(g)$ since every μ_k is supported in F and the support can only shrink with passage to a weak limit. Since $\mu_k(C \times B) = \sigma(C)$ for every $C \subset G$, and since $\mu_k(g)$ is almost surely a probability measure, it follows that $\nu(g)$ is a probability measure. This completes the proof of the lemma.

PROOF OF THEOREM 3.1. Necessity of the condition is clear. We first check the sufficiency of the condition in the case that σ has no atoms. Let $A_1, A_2, \cdots$ be disjoint subsets of B_1, with union equal to B_1 and each with diameter less than a prescribed $\varepsilon > 0$. Suppose also that each A_i is a difference of two closed sets in B_1. Such a sequence can be constructed since B is separable. Then, it is an easy exercise, the sets $\{g : F(g) \cap A_i \neq \emptyset\}$ are measurable. Let $U = \{1, 2, \cdots\}$ with the probability measure $\rho_0(\{j\}) = \rho(A_j)$. Consider the matching problem with respect to the spaces (U, ρ_0) and (G, σ), and when the friendship relation contains

(j, g) if $F(g) \cap A_j \neq \emptyset$. It is clear that for this matching problem condition $(*)$ holds. By the extension of Hall's theorem to the case with purely atomic versus atomless measures we have a mapping $f_0 : G \to U$ which is a proper matching. (For this extension of Hall's theorem see Bollobás and Varopoulos [2].) We create now a mapping $f : G \to B$ by choosing $f(g) \in A_i$ if $f_0(g) = i$. The function f can be chosen measurable by standard selection arguments (see e.g. Wagner [11]). The mapping f, viewed as an element in $\mathcal{M}_0$, has the property that $p(f)$ (the marginal on B) is close to ρ if ε is small. This follows easily from the definition of weak convergence, and the properties that f_0 is a proper matching between σ and ρ_0 and that the diameters of A_i are small. The continuity of $p : \mathcal{M}_0 \to \mathcal{P}$ with the compactness of $\mathcal{M}_0$ (Lemmas 3.4 and 3.6) supply a limit μ_0 in $\mathcal{M}_0$ of the choices f when $\varepsilon \to 0$ such that $p(\mu_0) = \rho$. This completes the case when σ is atomless.

When σ has atoms, say $g_1, g_2, \cdots$, we change G by attaching to each atom g_i an interval $J_i = [0, \alpha_i]$. Denote by G_1 the extended space. Then G_1 is metrizable. We extend the relation F to G_1 by letting $F(t) = F(g_i)$ for $t \in J_i$. With a proper choice of α_i the graph of the extended relation is still closed. Finally, σ is extended to G_1 by distributing the value $\sigma(\{g_i\})$ uniformly on $[0, \alpha_i]$. It is clear that condition $(*)$ holds for the extended σ as well. By our proof for the atomless case we have a $\mu(g)$ defined on G_1 with $p(\mu) = \rho$. We modify this μ and define it on the atoms of σ by $\mu(g_i)(A) = \int_0^{\alpha_i} \mu(t)(A)dt$. It is easy to verify that this generates a mixed matching on G. This completes the proof.

Proof of Theorem 3.2. Let $\mu : G_0 \to \mathcal{P}$ be the mixed matching guaranteed by Theorem 3.1. The idea is to define $f(t, g_0)$ such that for g_0 fixed the distribution of $f(\cdot, g_0)$ is equal to $\mu(g)$. Once this is done, it is clear that f is the desired proper matching. But recall that f should be measurable in the two variables. Here is a possible construction. Let $A_{n,i}$ be measurable subsets of B_1 such that: For a fixed n the sets $A_{n,i}$ are disjoint, with union equal to B_1 and the diameter less than $1/n$. The sequence $A_{n+1,j}$ is required to be a refinement of $A_{n,i}$ in the sense that an increasing sequence of integers $j_1 < j_2 < \cdots$ exists such that $A_{n,i}$ is the union of $A_{n+1,j}$ for $j_i < j \leqq j_{i+1}$. This can be arranged since B_1 is countably compact, hence $A_{1,i}$ can be chosen with compact closure. Let $b_{n,i}$ be a point in $A_{n,i}$, this for each (n, i). Define $\theta_n(g_0, i)$ to be $\mu(g_0)(A_{n,1} \cup \cdots \cup A_{n,i})$. Then, since $\mu(g)$ is measurable, the functions θ_n are measurable. We define $f_n(t, g_0) = b_{n,i}$ if $\theta_n(g_0, i-1) < t \leqq \theta_n(g_0, i)$, and here $\theta_n(g_0, 0) = 0$. The functions f_n are, clearly, measurable, and for fixed g_0 the distribution of $f_n(\cdot, g_0)$ is close in the weak convergence to $\mu(g_0)$. The choice of $A_{n+1,j}$ as a refinement of $A_{n,i}$

implies that for fixed (t, g_0) the values $f_n(t, g_0)$ form a Cauchy sequence. The completeness of B implies that the limit, as $n \to \infty$, exists. This limit, $f(t, g_0)$, is the desired proper matching.

4. Proof of Theorems 2.1 and 2.2

The necessity of condition (#) is clear.

Suppose now that σ on $\mathcal{K}$ and ρ on X satisfy condition (#). Consider a matching situation with $B = X$ and $G = \mathcal{K}_{ex}$ and with the probability distributions ρ and σ_{ex} on X and $\mathcal{K}_{ex}$ respectively. Let the friendship relation be determined by F_{ex}, namely (K, x) in F if $x \in K$. The graph of F_{ex} is indeed closed and the values of F_{ex} are, by definition, compact. Condition (*) for the matching problem is then identical with condition (#). By Theorem 3.2 there is a selection f on F_{ex} with distribution equal to ρ. This proves that ρ is selectionable and that it can be realized as a distribution of a selection of F_{ex}. This completes the proof of the two theorems.

5. The ensemble of selections

For a distribution σ on $\mathcal{K}$ we denote by $S(\sigma)$ the family of distributions on X which are selectionable with respect to σ. We use the construction of the previous sections to prove that $S(\sigma)$ is convex, compact and depends continuously on σ.

We use the Prohorov metric $P(r_1, r_2)$ to describe the weak convergence of probability measures, namely $P(r_1, r_2)$ is the minimal $\varepsilon > 0$ such that $r_1(A) \leqq r_2(N_\varepsilon(A)) + \varepsilon$ and $r_2(A) \leqq r_1(N_\varepsilon(A)) + \varepsilon$ where $N_\varepsilon(A)$ is the ε-neighborhood of the measurable set A in the metric space on which r_1 and r_2 are defined. We consider probability distributions on X and on $\mathcal{K}$; both are separable, hence convergence in the Prohorov metric is equivalent to weak convergence of measures; see Billingsley [1], page 239.

PROPOSITION 5.1. *Let σ be a distribution on $\mathcal{K}$, then $S(\sigma)$ is a convex, compact set.*

PROOF. Condition (#) can be written with C open instead of closed. Hence, by condition (iv) in Billingsley [1], page 24, condition (#) determines a closed set under weak convergence of measures. The distribution σ is tight, therefore a compact subset $\mathcal{K}_0$ of $\mathcal{K}$ exists with $\sigma(\mathcal{K}_0) \geqq 1 - \varepsilon$. The union of the sets K in $\mathcal{K}_0$ is a compact subset of X; denote it by X_0. If ρ is a selectionable distribution of σ then $\rho(X_0) \geqq 1 - \varepsilon$, say by condition (#). Hence the family $S(\sigma)$ is tight, and therefore precompact (Billingsley [1], page 37). Convexity of $S(\sigma)$ is implied by (#). This completes the proof.

The continuity that we wish to establish for $S(\sigma)$ is that σ_1 close to σ_2 in the Prohorov metric implies that the compact sets $S(\sigma_1)$ and $S(\sigma_2)$ are close in the Hausdorff distance; the latter is defined with respect to the Prohorov distance between elements $\rho \in S(\sigma)$. Compare with the definition in Section 2. The statement and the proof will follow two lemmas.

In what follows we consider $\mathcal{K}_{\mathrm{ex}} = [0,1] \times \mathcal{K}$ as a metric space, where the metric is the sum of the line distance on [0, 1] and the Hausdorff distance on $\mathcal{K}$. Recall that $F_{\mathrm{ex}}(t, K) = K$ is defined on $\mathcal{K}_{\mathrm{ex}}$ with values in $\mathcal{K}$.

LEMMA 5.2. *Let $A_1, \cdots, A_q$ be disjoint measurable subsets of $\mathcal{K}_{\mathrm{ex}}$, and $B_1, \cdots, B_q$ be also disjoint measurable subsets of $\mathcal{K}_{\mathrm{ex}}$. Suppose each A_i and each B_i has diameter less than θ, and that $h(A_i, B_i) < \theta$ for each i. Let τ_1 and τ_2 be atomless distributions on $\mathcal{K}_{\mathrm{ex}}$ such that $\tau_1(\bigcup A_i) \geqq 1 - \theta$, $\tau_2(\bigcup B_i) \geqq 1 - \theta$ and $\Sigma |\tau_1(A_i) - \tau_2(B_i)| \leqq \theta$. Then, for every selection f_1 of F_{ex} there is a selection f_2 of F_{ex} such that $P(\rho_1, \rho_2) < 4\theta$ when ρ_i is the distribution of f_i with respect to τ_i, $i = 1, 2$.*

PROOF. We first augment A_i and B_i so that $\tau_1(A_i) = \tau_2(B_i)$. This can be done, since τ_i are atomless, with only, possibly, changing the estimates to $\tau_1(\bigcup A_i) \geqq 1 - 2\theta$, $\tau_2(\bigcup B_i) \geqq 1 - 2\theta$ and $h(A_i, B_i) < 2\theta$. Let $H_1, H_2, \cdots$ be disjoint measurable subsets of X, with union equal to X and such that each H_i has diameter less than θ. These can be found since X is separable. For $i = 1, \cdots, q$ and for all $j = 1, 2, \cdots$ let $\alpha_{i,j} = \tau_1(f_1^{-1}(H_j) \cap A_i)$. Since τ_2 is atomless, each B_i can be partitioned into measurable subsets $B_{i,j}$ such that $\tau_2(B_{i,j}) = \alpha_{i,j}$. If $\alpha_{i,j} > 0$ we define $f_2(t, K)$ for $(t, K) \in B_{i,j}$ to be a point x in K with distance from H_j less than 3θ. Such an x can be found since the distance between (t, K) and the elements in A_i is less than 3θ, and there is an element (s, L) in A_i with $f_1(s, L) \in H_j$. Furthermore, since $(t, K) \to K \cap N_{3\theta}(H_j)$ determines a measurable set valued mapping on $B_{i,j}$, the function f_2 can be chosen measurable by standard selection procedures; see Wagner [11]. This defines f_2 on $\bigcup B_i$. On the complement we define f_2 arbitrarily.

We claim that with this definition $P(\rho_1, \rho_2) < 4\theta$. Indeed, let $H \subset X$. Let $J = \{j : H_j \cap H \neq \emptyset\}$ and let H_0 be the union of H_j, $j \in J$. Then $\rho_1(H) \leqq \rho_1(H_0)$ and the latter is equal to $\tau_1(f_1^{-1}(H_0))$. The latter number does not exceed 2θ plus the sum of $\alpha_{i,j}$ for $i = 1, \cdots, q$ and $j \in J$. The latter sum is dominated by $\tau_2(f_2^{-1}(N_{3\theta}(H_0)))$. Since $N_{3\theta}(H_0)$ is contained in $N_{4\theta}(H)$ we get that $\rho_1(H) \leqq \rho_2(N_{4\theta}(H)) + 2\theta$. This is one of the desired inequalities. To check the second one let $J = \{j : H \cap N_{3\theta}(H_j) \neq \emptyset\}$, and H_0 be the union of $N_{3\theta}(H_j)$ for $j \in J$. Then $\rho_2(H) \leqq \rho_2(H_0)$. If $B_{i,j}$ belongs to $f_2^{-1}(H_0)$ then $A_{i,j}$ belongs to $f_1^{-1}(N_{3\theta}(H_0))$. Since $\tau_2(X \setminus \bigcup A_{i,j})$ is less than 2θ and since $\tau_1(A_{i,j}) = \tau_2(B_{i,j})$ it follows that $\rho_2(H_0) \leqq \rho_1(N_{3\theta}(H_0)) + 2\theta$. This yields the second inequality needed for the claim. This completes the proof.

LEMMA 5.3. *Let M_0 be a compact subset of a complete metric space M. Let $\theta > 0$ and let $A_1, \cdots, A_q$ be a partition of M_0. Then a $\delta > 0$ exists such that whenever τ_1 and τ_2 are atomless distributions on M with $\tau_1(M_0) \geqq 1 - \theta$ and $P(\tau_1, \tau_2) < \delta$ then disjoint sets $E_1, \cdots, E_q$ of M can be found such that $\tau_2(E_i) = \tau_1(A_i)$ and $\Sigma_{i=1}^{q} |\tau_1(A_i) - \tau_2(E_i \cap N_\delta(A_i))| \leqq \theta$.*

PROOF. Let $\delta = \theta 2^{-(q+1)}$. Denote $A_0 = M \setminus M_0$. We check now the $22^q - 1$ nonempty subsets of $\{0, 1, \cdots, q\}$. For each subset J let A_J be the union of A_i for $i \in J$. Let $a_J = \max(0, \tau_1(A_J) - \tau_2(N_\delta(A_J)))$. For each such J we choose, arbitrarily, a set C_J in the complement of $N_\delta(A_J)$ with $\tau_2(C_J) = a_J$. Let C_0 denote the union of the sets C_J. The choice of δ and the inequality $P(\tau_1, \tau_2) < \delta$ imply that $\tau_2(C_0) \leqq \theta$. To each A_i we associate now the set $F(A_i) = C_0 \cup N_\delta(A_i)$. The choice of C_0 implies that $\tau_1(A_J) \leqq \tau_2(F(A_J))$ for all J, where $F(A_J)$ is the union of $F(A_i)$ for $i \in J$. By the extension of the marriage lemma given in Bollobás and Varopoulos [2], it follows that disjoint subsets $E_0, E_1, \cdots, E_q$ exist with $E_i \subset F(A_i)$ and $\tau_2(E_i) = \tau_1(A_i)$. For $i \geqq 1$ the differences $E_i \setminus N_\delta(A_i)$ are disjoint, and are contained in C_0. Hence $E_1, \cdots, E_q$ are the desired sets. This completes the proof.

THEOREM 5.4. *$S(\sigma)$ is continuous in σ.*

PROOF. Given σ_1 and $\varepsilon > 0$ we ought to produce a $\delta > 0$ such that $P(\sigma_1, \sigma_2) < \delta$ implies $h(S(\sigma_1), S(\sigma_2)) < \varepsilon$. Let $\theta = \varepsilon/8$. Since σ_1 is tight, a compact subset $\mathcal{K}_0$ of $\mathcal{K}$ exists such that $\sigma_1(\mathcal{K}_0) \geqq 1 - \theta$. Let $M_0 = [0, 1] \times \mathcal{K}_0$ and let $A_1, \cdots, A_q$ be a partition of M_0 into disjoint sets, each with diameter less than θ. Let $\delta > 0$ be determined by Lemma 5.3. If $P(\sigma_1, \sigma_2) < \delta$ then it is easy to see that $P(\tau_1, \tau_2) < \delta$ where τ_i is the extension of σ_i to $\mathcal{K}_{ex}$ (by cross product with the Lebesgue measure). In particular $\tau_1(M_0) \geqq 1 - \theta$ and let $E_1, \cdots, E_q$ be the sets guaranteed for τ_2 by Lemma 5.3. Define $B_1, \cdots, B_q$ by $B_i = E_i \cap N_\delta(A_i)$. The sets $A_1, \cdots, A_q$, and $B_1, \cdots, B_q$ and the atomless distributions τ_1 and τ_2 satisfy the conditions of Lemma 5.2 with respect to 2θ. The conclusion of Lemma 5.2 implies, therefore, that $h(R_1, R_2) < 8\theta$ when R_i is the family of distributions of selections of F_{ex} with respect to $(\mathcal{K}_{ex}, \tau_i)$. Theorem 2.2, however, implies that the family R_i is identical with the selectionable distributions of σ_i, hence the desired inequality $h(S(\sigma_1), S(\sigma_2)) < \varepsilon$.

6. Comments

The compactness established in Proposition 5.1 can be used for the following trivial generalization.

COROLLARY 6.1. *Let Q be a property of distributions on X which prevails under passage to weak limits. Let σ be a distribution on $\mathcal{K}$. Then the collection of distributions selectionable with respect to σ and which satisfy property Q is closed and compact.*

An example for such results can be found in Hart, Hildenbrand and Kohlberg [6], where the distributions which determine the Walras equilibrium form a closed set. The authors there also use a standard, or canonical, space, similar to our space $\mathcal{K}_{ex}$, in which all equilibria can be realized. Indeed, in view of Theorem 2.2, Proposition 5.1 and Corollary 6.1 can be stated with reference to selections of F_{ex} on $(\mathcal{K}_{ex}, \sigma_{ex})$. The passage to a large space is needed since the distribution of the random set does not determine the ensemble of distributions of selections. A trivial example is a distribution σ concentrated on one point in $\mathcal{K}$, say K. Then σ is the distribution of any $F(\omega) \equiv K$. But such a mapping F defined on an interval $\omega \in [0,1]$ produces quite different selections than the one defined on a singleton $\{\omega\}$. Even if the underlying measure space is required to be atomless would not help. Here is the example due to Debreu, and discussed in detail by Hart and Kohlberg [5]: $\Omega = [0,1]$ with the Lebesgue measure and X is the real line. $F_1(t) = \{t, -t\}$ and $F_2(t) = \{s(t), -s(t)\}$ with $s(t) = 2t (\mathrm{mod}\, 1)$. The selection $f_2(t) = s(t)$ on $[0, \frac{1}{2}]$ and $f_2(t) = -s(t)$ on $[\frac{1}{2}, 1]$, of F_2, has a distribution which is not shared by any selection of F_1. This example also shows that the distributions induced by selections of one random set might not form a convex set. (Notice that $[0,1]$ in this example can be regarded as a parametrization of the family $\{t, -t\}$ in $\mathcal{K}$, hence to consider selections of distributions σ on $\mathcal{K}$ with the identity map $K \to K$ does not yield a closed set of selectionable distributions even if σ is atomless.) Hart and Kohlberg [5] proved that for two identically distributed random sets, each on an atomless measure space, the closure of the family of distributions are the same. (Hart and Kohlberg [5] state this result for subsets of a euclidean space, but the proof works, with only minor modifications, for any separable metric space.) Since σ_{ex} on $\mathcal{K}_{ex}$ is atomless, and in view of Theorems 2.1 and 2.2, this result can be phrased as follows.

PROPOSITION 6.2. *Let $F:\Omega\to\mathcal{K}$ be measurable, defined on the atomless measure space (Ω,μ). Then the closure (in the weak convergence) of $\{Df: f$ a selection of $F\}$ coincides with the selectionable distributions of DF.*

The previous result combined with Theorem 5.3 produce the following result. (Naturally, F_k converges in distribution to F_0 if DF_k converges weakly to DF_0.) A partial result along this line was obtained by Salinetti and Wets [10].

THEOREM 6.3. *Let $F_i:\Omega_i\to\mathcal{K}$ be measurable, defined respectively on the measure spaces (Ω_i,μ_i) and assume that each μ_i is atomless, this for $i=0,1,\cdots$. If F_k converges in distribution to F_0 then the closures of $\{Df: f$ a selection of $F_k\}$ converge in the Hausdorff metric to the closure of $\{Df: f$ a selection of $F_0\}$.*

Our final comment is that if X is locally compact then Theorems 2.1 and 2.2 hold with random sets having values in the family of closed (rather than compact) sets. The metric on the family of closed sets is then determined by the closed convergence; see Hildenbrand [7]. The same applies to Theorems 3.1 and 3.2 where the condition that $F(g)$ is compact can be dropped. The reason for this is that we can apply our results for the one-point compactification of X. Closed convergence of closed sets in X is equivalent to the convergence in the Hausdorff distance in the compactification. The distributions Df guaranteed by Theorems 2.1 and 2.2 for the compactification are then actually supported at X, since ρ is supported at X. Proposition 5.1, Theorem 5.4 and the results of the present section can also be extended to the closed sets setting if X is locally compact, provided we allow the distributions to assign positive weight to the point at infinity.

REFERENCES

1. P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

2. B. Bollobás and N. Th. Varopoulos, *Representation of systems of measurable sets*, Math. Proc. Cambridge Phil. Soc. **78** (1974), 323–325.

3. U. Grenander, *Pattern Analysis — Lectures in Pattern Theory — I*, Springer-Verlag, New York, 1978.

4. P. R. Halmos and H. E. Vaughan, *The marriage problem*, Am. J. Math. **72** (1950), 214–215.

5. S. Hart and E. Kohlberg, *Equally distributed correspondences*, J. Math. Econ. **1** (1974), 167–174.

6. S. Hart, W. Hildenbrand and E. Kohlberg, *On equilibrium allocations as distributions on the commodity space*, J. Math. Econ. **1** (1974), 159–166.

7. W. Hildenbrand, *Core and Equilibria of a Large Economy*, Princeton University Press, Princeton, 1974.

8. D. G. Kendall, *Foundation of a theory of random sets*, in *Stochastic Geometry* (E. F. Harding and D. G. Kendall, eds.), Wiley, London, 1974, pp. 322–376.

9. G. Matheron, *Random Sets and Integral Geometry*, Wiley, London, 1975.

10. G. Salinetti and R. J.-B. Wets, *On the convergence in distribution of measurable multifunctions, normal integrands, processes and stochastic infima*, preprint, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1982.

11. W. Wagner, *Survey of measurable selection theorems*, SIAM J. Control Optim. **15** (1977), 859–901.

Department of Theoretical Mathematics
The Weizmann Institute of Science
Rehovot 76100, Israel