Pacific Journal of Mathematics

BOREL SETS OF PROBABILITY MEASURES

KENNETH LAMAR LANGE

Vol. 48, No. 1 March 1973

PACIFIC JOURNAL OF MATHEMATICS
Vol. 48, No. 1, 1973

BOREL SETS OF PROBABILITY MEASURES

KENNETH LANGE

Let $M(X)$ be the collection of probability measures on the Borel sets of a Polish space X. The Borel structure of $M(X)$ generated by the weak* topology is investigated. Various collections of probability measures arising in nonparametric statistics are shown to Borel sets of $M(X)$. Attention is particularly focused on collections arising from restrictions on distribution functions, density functions, and supports of the underlying probability measures.

1. Introduction. Baynesian statisticians assume prior distributions on certain families of probability measures. This amounts to putting a probability measure on a family of probability measures. Now families of probability measures typically arising in statistics are parametrized by some Borel set of Euclidean n-space. In such cases, one has a natural Borel structure or σ-algebra of subsets with which to deal. In nonparametric situations the natural Borel structure is not so obvious. Ideally, one might desire each commonly occurring family of probability measures to be a Borel set of some properly chosen complete separable metric space. Then a prior distribution could be viewed as a probability measure on the entire space which is concentrated on the given Borel set. Our aim is to show that many, if not most, nonparametric families of probability measures are indeed Borel sets of complete separable metric spaces. This advances slightly the cause of nonparametric Baynesian statistics, but does not overcome the more difficult barrier of finding reasonable prior distributions in nonparametric situations.

In our probabilistic model we suppose X to be a complete separable metric space. Let $C(X)$ be the bounded real-valued continuous functions on X under the sup norm topology. Then the collection of probability measures $M(X)$ on the Borel sets of X can be viewed as a subset of the dual of $C(X)$ under the weak* topology. It is well known that $M(X)$ is metrizable as a complete separable metric space with this topology [16]. Our investigations will center on the Borel structure of $M(X)$.

Dubins and Freedman have done the spadework for the subsequent discussion in their basic paper [8]. Section 3 generalizes their analysis of the relationship of distribution and density functions to probability measures. Section 4 explores the connection between a probability measure and its support when the underlying space is no longer compact. Section 5 collects some further examples not considered in

[8], and §6 adds to the fund of counterexamples.

Finally, before moving on to some preliminary definitions, let us cite two other areas where the Borel structure of collections of measures can be fruitfully pursued. Much work has been done on so called ergodic decompositions of invariant measures. The original stimulus for this research came from classical statistical mechanics. The reader may consult [20] for a detailed theoretical discussion and further references. Another area of potential applications is the analysis of Poisson and point processes. See [14] for steps in this direction.

2. Preliminary definitions.

2.1. *Borel spaces.* First, let us give a compressed account of Borel spaces. The reader is advised to consult §§1-3 of Chap. 1 of [2] for a fuller treatment. A *Borel space* $(X, \mathscr{A})$ consists of a set X together with a distinguished σ-algebra of subsets $\mathscr{A}$. Quite often X itself is said to be the Borel space and $\mathscr{A}$ is tacitly understood. For example, if X is a topological space, then $\mathscr{A}$ is always taken to be the smallest σ-algebra containing the open sets. A function $f: X \to Y$ between two Borel spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ is called *Borel* if $f^{-1}(B) \in \mathscr{A}$ whenever $B \in \mathscr{B}$.

The sets in the σ-algebra $\mathscr{A}$ of a Borel space $(X, \mathscr{A})$ are also termed *Borel sets.* Every subset Z of X inherits a *relative Borel structure* $\{Z \cap A: A \in \mathscr{A}\}$. If $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ are two Borel spaces, the *product Borel structure* $\mathscr{A} \times \mathscr{B}$ is defined to be the smallest σ-algebra of subsets of $X \times Y$ containing the Borel rectangles $A \times B$, $A \in \mathscr{A}$, and $B \in \mathscr{B}$. Suppose $\sim$ is an equivalence relation on a Borel space $(X, \mathscr{A})$ and $\pi: X \to X/\sim$ is the projection taking each point into its equivalence class. The *quotient Borel structure* on $X/\sim$ is the largest σ-algebra making π a Borel map.

Certain Hausdorff topological spaces have very well behaved Borel structures. Among these are *Polish spaces.* A Polish space is a topological space which is metrizable by a complete separable metric. It is well known that any locally compact space with a countable neighborhood basis is Polish. Such spaces will be referred to as locally compact and separable.

One property of Polish spaces will be particularly useful later on: Suppose X and Y are Polish spaces. Let B be a Borel set of X equipped with the relative Borel structure. Also let $f: B \to Y$ be a one-to-one Borel map. Then $f(B)$ is a Borel set of Y. (See Cor. 3.3 of Chap. 1 of [16].) This fact will be applied to show that certain Borel maps are *Borel isomorphisms.* A map $f: X \to Y$ between two Borel spaces $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ is said to be a Borel isomorphism if it is one-to-one, onto, Borel and its inverse is Borel.

Yet another class of Borel spaces appearing in the sequel is the class of *analytic Borel spaces*. To define this notion it is necessary to mention a second notion. A Borel space $(X, \mathscr{A})$ is called *countably separated* if there exists a countable collection of Borel sets $\{A_n\}_{n=1}^{\infty}$ such that for any two points $u, v \in X$ there is some A_n with either $u \in A_n, v \notin A_n$ or $u \notin A_n, v \in A_n$. A Borel space $(X, \mathscr{A})$ is analytic if it is countably separated and the image of a Polish space under a Borel map. A subset of a Borel space is called analytic if it is an analytic Borel space with its relative Borel structure; it is called *complementary analytic* if its complement is analytic. Every Borel set of an analytic Borel space is analytic, but not every analytic set is Borel. If $f: X \to Y$ is a Borel map between two analytic Borel spaces, then the image and inverse image of every analytic set under f is analytic.

2.2. *Notation for topological spaces.* Suppose T is a topological space and Y is a subset of T. Y^- will denote the closure of Y, Y^0 the interior, and Y' the complement. For two subsets Y and Z put $Y \Delta Z = (Y \backslash Z) \cup (Z \backslash Y)$, the symmetric difference of Y and Z. R^n will mean Euclidean n-space and C^n the space of n-tuples of complex numbers. If $u = (u_1, \cdots, u_n)$ and $v = (v_1, \cdots, v_n)$ are in R^n, then $\langle u, v \rangle = \sum_{i=1}^n u_i v_i$ will denote the usual inner product and $\|u\| = (\sum_{i=1}^n u_i^2)^{1/2}$ the usual norm.

2.3. *Multi-index notation.* A multi-index $j = (j_1, \cdots, j_n)$ is a finite sequence of nonnegative integers. $|j|$ denotes the sum $\sum_{i=1}^n j_i$. If $x = (x_1, \cdots, x_n) \in R^n$, put $x^j = x_1^{j_1} \cdots x_n^{j_n}$. For a function $f: R^n \to C$ continuously differentiable of order $|j|$, set $D_i = \partial / \partial x_i$ and $D^j f = D_1^{j_1} \cdots D_n^{j_n} f$.

2.4. *Comments on probability measures.* Suppose X is a Polish space. The *support* of a probability measure $\mu \in M(X)$ can be characterized as the complement of the largest open set on which μ vanishes. If ν is a σ-finite measure on the Borel sets of X and μ is absolutely continuous with respect to ν, then the *density* of μ will mean the Radon-Nikodym derivative $d\mu/d\nu$. On R^n densities will always refer to Radon-Nikodym derivatives with respect to Lebesgue measure. The *distribution function* $F_\mu: R^n \to [0, 1]$ of $\mu \in M(R^n)$ is defined by $F_\mu(x_1, \cdots, x_n) = \mu((-\infty, x_1] \times \cdots \times (-\infty, x_n])$.

Finally, for X Polish we should mention an alternate description of the Borel structure of the collection of probability measures $M(X)$. Varadrajan proved that the Borel structure on $M(X)$ generated by the weak* topology is precisely the smallest Borel structure making each of the maps $\mu \to \mu(A)$ Borel, where $\mu \in M(X)$ and A is a Borel set of X. (See Lemma 2.3 of [20].)

3. **Borel properties of $M(R^n)$ in terms of distribution and density functions.** Let us treat the case of distribution functions first. Our opening lemma is a generalization of a well known result in the theory of stochastic processes. (See Thm. T 47 of Chap. 4 of [12].)

LEMMA 3.1. *Suppose M_1 is a Borel space and M_2 a metric space. Consider a function $f(x_1, \cdots, x_n, y)$ from $R^n \times M_1$ into M_2. If f is Borel in y for $x_1, \cdots, x_n$ fixed and right continuous in each x_i for all other variables fixed, then f is Borel in all variables jointly. The same conclusion holds if R^n is replaced by a product of intervals.*

Proof. The general case follows from the one dimensional case since one can replace M_1 by $(\prod_{i=1}^{n-1} U_i) \times M_1$, where the U_i's are intervals. So consider $f: U \times M_1 \to M_2$ satisfying the condition of the lemma, where U is an interval. For each $k = 1, 2, \cdots$, choose pairwise disjoint intervals $F_k^n (n = 1, \cdots, n_k, n_k$ finite or $\infty)$ which are closed on the right, have union U and satisfy length $(F_k^n) < 1/k$. For each k and n let x_k^n be the right endpoint of F_k^n. Now define $f_k(x, y) = f(x_k^n, y)$ whenever $x \in F_k^n$. Since f is continuous on the right in its first variable, $f_k \to f$ pointwise on $U \times M_1$. Since f is Borel in its second variable, each f_k is Borel on $U \times M_1$. Finally, because M_2 is a metric space, the limit f of the Borel functions f_k is Borel.

DEFINITIONS. For $0 \leqq k \leqq \infty$ let $C^k(R^n)$ be the topological vector space of all complex-valued functions on R^n having continuous derivatives of order k under the topology of uniform convergence on compact sets. Note that $f_m \to f$ in $C^k(R^n)$ means whenever j is a multi-index with $|j| \leqq k$, $D^j f_m \to D^j f$ uniformly on each compact set of R^n. $C^k(R^n)$ is a separable Frechet space since the polynomials with rational coefficients form a countable dense subset [19]. Lip (R^n) will denote the collection of complex-valued functions f on R^n satisfying $|f(x) - f(y)| \leqq c_f ||x - y||$ for all x and y and some constant c_f depending only on f. Finally, let $H(C^n)$ be the topological vector space of holomorphic functions in n complex variables equipped with the topology of uniform convergence on compact sets. $H(C^n)$ is also a separable Frechet space.

LEMMA 3.2. *Each of the spaces $C^k(R^n)$, $1 \leqq k \leqq \infty$, and* Lip (R^n) *is a Borel subset of $C^0(R^n)$. If $H(R^n)$ denotes the holomorphic functions in n real variables, then $H(R^n)$ is a Borel set of $C^0(R^n)$ too.*

Proof. The natural injection of $C^k(R^n)$, $1 \leqq k \leqq \infty$, into $C^0(R^n)$ is continuous. Hence its image in $C^0(R^n)$ is a Borel set. Lip (R^n) is a Borel set of $C^0(R^n)$ because the map

$$f \to \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|}$$

is lower semicontinuous in the classical sense (not to be confused with the notion of lower semicontinuity for set valued maps discussed in the next section), hence Borel. Last of all, the map $H(C^n) \to C^0(R^n)$ given by restricting a holomorphic function on C^n to R^n is certainly continuous. It is one-to-one because a holomorphic function on C^n is completely determined by its values on R^n. (See 9.4.4 of [6].)

THEOREM 3.3. *For each probability measure μ on R^n let F_μ be the distribution function of μ. Then the set of probability measure μ with F_μ satisfying any of the conditions below forms a Borel set of $M(R^n)$:*

1. *F_μ is continuous.*
2. *F_μ is Lipschitz, i.e., $|F_\mu(x) - F_\mu(y)| \leqq c\|x - y\|$ for some constant c and all x and y.*
3. *F_μ is continuously differentiable of order k, $1 \leqq k \leqq \infty$.*
4. *F_μ is holomorphic.*

Proof. Let us prove the assertion in part 1. first. Designate by $\{r_k\}_{k=1}^\infty$ the points in R^n having all components rational. Since $\mu \to \mu(A)$ is Borel from $M(R^n)$ to R for every Borel set A in R^n, the function

$$\mu \longrightarrow \inf_{\substack{n \in Z \\ n>0}} \sup_{\|r_k - r_l\| < 1/n} |F_\mu(r_k) - F_\mu(r_l)| = g(\mu)$$

is Borel. Our claim is that $\{\mu: g(\mu) = 0\}$ is the collection of all probability measures with continuous distribution functions. In fact, if $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$ are two points in R^n, the expansion

$$F_\mu(x) - F_\mu(y) = \sum_{k=1}^{n} (F_\mu(z^{k-1}) - F_\mu(z^k)),\ z_i^k = \begin{cases} x_i & \text{if } k < i \\ y_i & \text{if } k \geqq i \end{cases},$$

leads immediately to the conclusion that $g(\mu) = 0$ implies F_μ is continuous. On the other hand, if F_μ is continuous, it is actually uniformly continuous. This follows from the fact that for every $\varepsilon > 0$ there is a compact set K with $\mu(K) > 1 - \varepsilon$. The uniform continuity of F_μ then clearly entails $g(\mu) = 0$.

To prove the assertion of the theorem for parts 2.—4. it is sufficient by Lemma 3.2 to prove that the map $\mu \to F_\mu$ is Borel from the collection of probability measures having continuous distribution functions into $C^0(R^n)$. To do this it is enough according to Thm. 2 of the appendix of [13] to show $\mu \to l(F_\mu)$ Borel whenever l is a continuous linear functional on $C^0(R^n)$. Since the dual space of $C^0(R^n)$ consists of the complex measures ν with compact support (13.19.3 of [7]), the

problem reduces to showing $\mu \to \int F_\mu d\nu$ Borel for every complex measure ν with compact support. But this follows from Fubini's Theorem once one notes Lemma 3.1 says $F_\mu(x)$ is jointly Borel in μ and x.

COROLLARY 3.4. *For $|j| \leqq k$ the map $(x, \mu) \to D^j F_\mu(x)$ is jointly Borel in x and μ on the Cartesian product of R^n and the set of all $\mu \in M(R^n)$ with distribution functions in $C^k(R^n)$.*

Proof. Since the natural injection of $C^k(R^n)$ into $C^0(R^n)$ is Borel, its inverse is also. Hence the map $\mu \to F_\mu$ is Borel from $\{\mu \in M(R^n)$: $F_\mu \in C^k(R^n)\}$ into $C^k(R^n)$. Also the map $f \to D^j f$ of $C^k(R^n)$ into $C^0(R^n)$ is continuous, and the map $(x, g) \to g(x)$ of $R^n \times C^0(R^n)$ into the complex plane is Borel by Lemma 3.1. Now combine these facts and the statement of the corollary follows.

Our next aim is to prove analogues of the preceding for probability measures having densities. Some of the assertions for densities follow by noting that the density of a probability measure on R^n can be recovered by differentiating the distribution function a sufficient number of times. For the sake of completeness though, it seems preferable to proceed directly.

THEOREM 3.5. *Let S be a Polish space and μ a σ-finite measure on the Borel sets of S. The collection of probability measures absolutely continuous with respect to μ and the collection of probability measures equivalent to μ both form Borel sets of $M(S)$. Furthermore, the map taking a probability measure ν absolutely continuous with respect to μ into its density $f_\nu \in L^1(\mu)$ is a Borel isomorphism.*

Proof. Consider the separable Banach space $L^1(\mu)$ of μ-integrable functions. Take a countable collection of Borel sets $\{A_k\}_{k=1}^\infty$ in S which are dense in the measure algebra derived from μ. The set

$$B = \bigcap_{k=1}^{\infty} \left\{ f \in L^1(\mu) : \int f \chi_{A_k} d\mu \geqq 0 \right\} \cap \left\{ f \in L^1(\mu) : \int f d\mu = 1 \right\}$$

is a Borel set of $L^1(\mu)$. We claim that B consists of the function in $L^1(\mu)$ which are nonnegative μ-almost everywhere and have integral 1. Suppose $f \in B$. Then it is necessary to show $\int_A f d\mu \geqq 0$ for every Borel set A. But

$$\int_A f d\mu = \int_{A_k} f d\mu - \int_{A_k \setminus A} f d\mu + \int_{A \setminus A_k} f d\mu \quad *\ .$$

By taking $\mu(A_k \Delta A)$ small enough, we can make the second and third contributions on the right of * as small as desired.

Now map f in B into the measure $\mu_f(A) = \int_A f d\mu$. This map is one-to-one and continuous for the weak* topology on the set of probability measures. The image of B under the map is a Borel set and reduces simply to the probability measures absolutely continuous with respect to μ. Furthermore, the inverse of $f \to \mu_f$ is also a Borel map.

To prove the assertion about probability measures equivalent to μ it suffices to prove that the set of $f \in B$ with $f > 0$ μ-almost everywhere forms a Borel subset of B. Now this set is

$$\left(\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{\substack{A_k \\ \mu(A_k)>1/m}}\left\{f \in L^1(\mu): \int f\chi_{A_k} d\mu \geqq \frac{1}{n}\right\}\right) \cap \left\{f \in L^1(\mu): \int f d\mu = 1\right\} = C .$$

For, if $f \in B \backslash C$, then there is an m and a sequence $\{A_{k_i}\}_{i=1}^{\infty}$ with

$$\mu(A_{k_i}) > \frac{1}{m} \text{ and } \int_{A_{k_i}} f d\mu \to 0 .$$

Assuming $f > 0$ μ-almost everywhere, this contradicts the absolute continuity of μ with respect to μ_f. On the other hand, if $f \in C$ and $\mu(A) > 0$, then take A_k so that $\mu(A_k \Delta A)$ is small and $\mu(A_k) > 1/m$ for some m. Then our earlier representation * of $\int_A f d\mu$ shows that $\int_A f d\mu > 0$.

DEFINITIONS. $D(R^n)$ will denote the space of infinitely differentiable functions with compact support. It is well known that $D(R^n)$ is a countable union of separable Frechet spaces. [19] $S(R^n)$ will denote the space of infinitely differentiable functions rapidly decreasing at infinity together with all derivatives. With its usual topology $S(R^n)$ is a separable Frechet space. Finally, for $1 \leqq p < \infty$, $L^p(R^n)$ means the Banach space of equivalence classes of Borel functions whose pth powers are integrable.

LEMMA 3.6. *For $1 \leqq p < \infty$ each of the following spaces is a Borel set of $L^p(R^n)$:*

1. $C^k(R^n) \cap L^p(R^n)$ $\quad k = 0, 1, 2, \cdots, \infty$
2. $\text{Lip}\,(R^n) \cap L^p(R^n)$
3. $H(R^n) \cap L^p(R^n)$
4. $S(R^n)$
5. $D(R^n)$.

Proof. Since each of the spaces $C^k(R^n)$, $1 \leqq k \leqq \infty$, Lip (R^n), $H(R^n)$, $S(R^n)$, and $D(R^n)$ reside in $C^0(R^n)$ as Borel sets, it suffices to show that $C^0(R^n) \cap L^p(R^n)$ is a Borel set of $C^0(R^n)$ and that the injection $C^0(R^n) \cap L^p(R^n) \to L^p(R^n)$ is Borel. The first ingredient of the proof follows from the fact that

$$f \longrightarrow \sup_{\substack{n \in Z \\ n>0}} \int_{||x|| \leqq n} |f(x)|^p dx = \int |f(x)|^p dx$$

is a Borel map of $C^0(R^n)$ into the extended real line. As for the second, note that

$$f \longrightarrow \lim_{n \to \infty} \int_{||x|| \leqq n} f(x)g(x)dx = \int f(x)g(x)dx$$

is Borel from $C^0(R^n) \cap L^p(R^n)$ into C for every $g \in L^q(R^n)$, where $1/p + 1/q = 1$ and $L^q(R^n)$ is identified with the dual of $L^p(R^n)$. According to the criterion already cited in Theorem 3.3, this implies that the injection $C^0(R^n) \cap L^p(R^n) \to L^p(R^n)$ is Borel.

THEOREM 3.7. *The collection of probability measures on R^n whose densities lie in any of the spaces $C^k(R^n)$, Lip (R^n), $H(R^n)$, $S(R^n)$ or $D(R^n)$ form a Borel set of $M(R^n)$.*

Proof. The map taking a probability measure to its density with respect to Lebesgue measure is Borel into $L^1(R^n)$. Now apply the last lemma.

4. **Supports of probability measures.** In order to analyze the relation of a probability measure to its support, we find it convenient to introduce the Fell topology [10]. Let X be a locally compact separable space. Denote the collection of closed subsets of X by $\mathscr{C}(X)$. The Fell topology on $\mathscr{C}(X)$ can be given by specifying a basis of open sets of the form $U(C, \{V_1, \cdots, V_n\})$. C is compact in X and $\{V_1, \cdots, V_n\}$ is a finite, but possibly empty, family of open sets. $U(C; \{V_1, \cdots, V_n\}) = \{Y \in \mathscr{C}(X): Y \cap C = \phi$ and $Y \cap V_i \neq \phi, i = 1, \cdots, n\}$. Fell demonstrates that $\mathscr{C}(X)$ is a compact Hausdorff space with this topology. Furthermore, as Fell observes, $\mathscr{C}(X)$ is separable and thus a compact metric space if X is separable. Indeed, if $\mathscr{B}$ is a basis for the topology of X and each element of $\mathscr{B}$ has compact closure, then the sets $U(C; \mathscr{F})$, where C is the closure of the union of a finite subset of $\mathscr{B}$ and $\mathscr{F} \subset \mathscr{B}$ is finite, form a countable basis for $\mathscr{C}(X)$.

Next we wish to define notions of semicontinuity for maps into $\mathscr{C}(X)$. Since our definitions differ slightly from Kuratowski's [11] and Berge's [3], we feel it prudent to give a detailed discussion.

(Especially with reference to the next lemma see Chap. 6, Sec. 2 of [3].)

DEFINITION. Let $f: T \to \mathscr{C}(X)$ be a map from a topological space T into $\mathscr{C}(X)$. f is said to be *lower semicontinuous* if $\{t \in T: f(t) \cap V \neq \phi\}$ is open whenever V is open in X. f is *upper semicontinuous* if $\{t \in T: f(t) \cap C = \phi\}$ is open whenever C is compact in X. Obviously, a map $f: T \to \mathscr{C}(X)$ will be continuous if and only if it is both lower and upper semicontinuous.

LEMMA 4.1. *Suppose* $\{f_i: T \to \mathscr{C}(X)\}_{i \in \Lambda}$ *is a nonempty collection of maps from a topological space* T *into* $\mathscr{C}(X)$.

1. *If each* f_i *is lower semicontinuous, then* $t \to (\bigcup_{i \in \Lambda} f_i(t))^-$ *is lower semicontinuous.*

2. *If each* f_i *is upper semicontinuous, then* $t \to \bigcap_{i \in \Lambda} f_i(t)$ *is upper semicontinuous.*

3. *If the index set* Λ *is finite and each* f_i *is upper semicontinuous, then* $t \to \bigcup_{i \in \Lambda} f_i(t)$ *is upper semicontinuous.*

Proof. 1. Take V open in X. Then $\{t \in T: (\bigcup_{i \in \Lambda} f_i(t))^- \cap V \neq \phi\} = \{t \in T: (\bigcup_{i \in \Lambda} f_i(t)) \cap V \neq \phi\} = \bigcup_{i \in \Lambda} \{t \in T: f_i(t) \cap V \neq \phi\}$ is open in T.

2. Take C compact in X and suppose $t_0 \in T$ satisfies $\bigcap_{i \in \Lambda} f_i(t_0) \cap C = \phi$. It is enough to show the existence of some neighborhood of t_0 where $\bigcap_{i \in \Lambda} f_i(t) \cap C = \phi$ continues to hold. To this end select one function g from the collection $\{f_i: T \to \mathscr{C}(X)\}_{i \in \Lambda}$ and reduce the collection to $\{f_i: T \to \mathscr{C}(X)\}_{i \in \Sigma}$ by eliminating g. If $g(t_0) \cap C = \phi$, then our choice for the neighborhood of t_0 is obvious. Otherwise, put $K = g(t_0) \cap C$. For each $x \in K$ there is at least one index $i \in \Sigma$ with $x \notin f_i(t_0)$. Choose a neighborhood V_x of x with compact closure and satisfying $f_i(t_0) \cap V_x^- = \phi$. Since f_i is upper semicontinuous, it is possible to select a neighborhood W_x of t_0 such that $f_i(t) \cap V_x^- = \phi$ for all $t \in W_x$. Let $V_{x_1}, \cdots, V_{x_n}$ cover K. Applying the upper semicontinuity of g there also exists a neighborhood W of t_0 where $g(t) \cap (C \setminus (\bigcup_{m=1}^n V_{x_m})) = \phi$. It now follows that on $W \cap (\bigcap_{m=1}^n W_{x_m})$, $(\bigcap_{i \in \Lambda} f_i(t)) \cap C = \phi$.

3. For C compact in X, $\{t \in T: (\bigcup_{i \in \Lambda} f_i(t)) \cap C = \phi\} = \bigcap_{i \in \Lambda} \{t \in T: f_i(t) \cap C = \phi\}$.

The next lemma shows that upper and lower semicontinuous mappings are Borel. (Compare Lemma 9.4 of [17].)

LEMMA 4.2. *The following are necessary and sufficient conditions for a map* $f: T \to \mathscr{C}(X)$ *from a Borel space* T *into* $\mathscr{C}(X)$ *to be Borel.*

1. $\{t \in T: f(t) \cap V \neq \phi\}$ *is a Borel set for each open set* V.
2. $\{t \in T: f(t) \cap C = \phi\}$ *is a Borel set for each compact set* C.

Proof. Both conditions are clearly necessary for f to be Borel. To prove their sufficiency consider a basic open set $U(C; \{V_1, \cdots, V_n\})$ in $\mathscr{C}(X)$. $U(C; \{V_1, \cdots, V_n\}) = U(C; \phi) \cap U(\phi; \{V_1\}) \cap \cdots \cap U(\phi; \{V_n\})$. Let d be a metric for X and define W_m by $\{x \in X: d(x, C) = \inf_{y \in C} d(x, y) < 1/m\}$. Then $U(C; \phi) = \bigcup_{m=1}^{\infty} U(\phi; \{W_m\})'$. On the other hand, if $\{K_l\}_{l=1}^{\infty}$ is a sequence of compact sets whose union is some V_i, then $U(\phi; \{V_i\}) = \bigcup_{l=1}^{\infty} U(K_l; \phi)'$. Thus any basic open set can be expressed in terms of a countable number of open sets of type $U(K; \phi)$ or $U(\phi; \{V\})$.

LEMMA 4.3. *The map* $\mu \to$ *support* (μ) *is lower semicontinuous from* $M(X)$ *into* $\mathscr{C}(X)$.

Proof. For V open in X, $\{\mu: \text{support } (\mu) \cap V \neq \phi\} = \{\mu: \mu(V) > 0\}$ is open in $M(X)$. In fact, $\{\mu: \mu(V) = 0\}$ is closed because $\mu(V) \leqq \lim_n \inf \mu_n(V)$ holds for every converging sequence $\mu_n \to \mu$ in $M(X)$. (See Thm. 6.1 of Chap. 2 of [16]).

LEMMA 4.4. *Each of the following maps is Borel*:

1. $F: \prod_{n=1}^{\infty} \mathscr{C}(X) \to \mathscr{C}(X)$ *given by* $F(\prod_{n=1}^{\infty} Y_n) = (\bigcup_{n=1}^{\infty} Y_n)^-$. $\prod_{n=1}^{\infty} \mathscr{C}(X)$ *has the product topology.*
2. $G: \prod_{n=1}^{\infty} \mathscr{C}(X) \to \mathscr{C}(X)$ *given by* $G(\prod_{n=1}^{\infty} Y_n) = \bigcap_{n=1}^{\infty} Y_n$.
3. $H: \mathscr{C}(X) \to \mathscr{C}(X)$ *given by* $H(Y) = (Y^0)'$.
4. $J: \mathscr{C}(X) \to \mathscr{C}(X)$ *given by* $J(Y) = (Y^0)^-$.

Proof. 1. The projection $F_k: \prod_{n=1}^{\infty} \mathscr{C}(X) \to \mathscr{C}(X)$ taking $\prod_{n=1}^{\infty} Y_n$ into Y_k is continuous. Hence $F(\prod_{n=1}^{\infty} Y_n) = (\bigcup_{k=1}^{\infty} F_k(\prod_{n=1}^{\infty} Y_n))^-$ is lower semicontinuous.

2. G is upper semicontinuous.

3. It suffices to show that the set of $Y \in \mathscr{C}(X)$ satisfying $(Y^0)' \cap V = \phi$ is Borel for each open set V. Now $(Y^0)' \cap V = \phi$ iff $Y^0 \supset V$ iff $Y \supset V$ iff $Y \supset V^-$ iff $Y \cap V^- = V^-$. The collection of Y satisfying $Y \cap V^- = V^-$ is a Borel set because it is precisely the set where the Borel maps $Y \to Y \cap V^-$ and $Y \to V^-$ agree.

4. Again it is enough to prove that the collection of $Y \in \mathscr{C}(X)$ satisfying $(Y^0)^- \cap V = \phi$ is Borel for each open set V. Let $W = (V')^0$. Then $(Y^0)^- \cap V = \phi$ iff $V' \supset (Y^0)^-$ iff $V' \supset Y^0$ iff $W \supset Y^0$ iff $W' \subset (Y^0)'$ iff $W' \cap (Y^0)' = W'$. But the collection of Y satisfying $W' \cap (Y^0)' = W'$ is precisely where the Borel maps $Y \to W' \cap H(Y)$ and $Y \to W'$ agree.

THEOREM 4.5. *Each of the following collection of sets in* $\mathscr{C}(X)$

is Borel:

1. *For every positive integer k, the collection of $Z \in \mathscr{C}(X)$ having k or fewer points.*
2. *The collection of compact sets $Z \in \mathscr{C}(X)$.*
3. *The collection of compact connected sets $Z \in \mathscr{C}(X)$.*
4. *The single closed set $Y \in \mathscr{C}(X)$.*
5. *The collection of $Z \in \mathscr{C}(X)$ contained within a given closed set Y.*
6. *The collection of $Z \in \mathscr{C}(X)$ containing a given closed set Y.*
7. *The collection of $Z \in \mathscr{C}(X)$ with empty interior.*
8. *The collection on $Z \in \mathscr{C}(X)$ with no isolated points.*
9. *The collection of $Z \in \mathscr{C}(X)$ which are open as well as closed.*

Proof. 1. $\mathscr{D}^k$ is the collection of $Z \in \mathscr{C}(X)$ having k or fewer points. According to Fell [10], $\mathscr{D}^1$ is closed and therefore compact in $\mathscr{C}(X)$. Since the union map $\prod_{n=1}^{k} \mathscr{C}(X) \to \mathscr{C}(X)$ is both lower and upper semicontinuous, the image of $\prod_{n=1}^{k} \mathscr{D}^1$, which is $\mathscr{D}^k$, is compact in $\mathscr{C}(X)$.

2. Let $\{C_n\}_{n=1}^{\infty}$ be an increasing sequence of compact sets having $\bigcup_{n=1}^{\infty} C_n^0 = X$. The collection of compact sets $Z \in \mathscr{C}(X)$ is the countable union of closed sets $\bigcup_{n=1}^{\infty} U(\phi; \{C_n'\})'$.

3. The compact disconnected sets coincide with the intersection of the collection of compact sets with the union of the basic open sets $U((V_1 \cup V_2)'; \{V_1, V_2\})$, where $V_1 \cap V_2 = \phi$ and $(V_1 \cup V_2)'$ is compact.

5. $Z \in \mathscr{C}(X)$ is contained within Y iff $Z \notin U(\phi; \{Y'\})$.

6. $Z \in \mathscr{C}(X)$ contains Y iff $Z \cap Y = Y$. Hence the collection of $Z \in \mathscr{C}(X)$ containing Y is where the Borel maps $Z \to Z \cap Y$ and $Z \to Y$ agree.

7. Let H be the map of last lemma. Then Z has empty interior iff $H(Z) = X$.

8. Let J be the map of the last lemma. Then Z has no isolated points iff $J(Z) = Z$.

9. $Z \in \mathscr{C}(X)$ is open iff $H(Z) \cap Z = \phi$.

COROLLARY 4.6. *The collection of probability measures in $M(X)$ having support in any one of the families* 1.—9. *listed above is Borel.*

We now wish to introduce notions of congruence and symmetry for the space of closed sets $\mathscr{C}(X)$. To be specific, suppose G is a locally compact separable topological group which acts on the right of X. If for each fixed $g \in G$, the map $x \to xg$ is continuous, then G also acts on $\mathscr{C}(X)$. For $Y \in \mathscr{C}(X)$ and $g \in G$ define Yg to be $\{yg: y \in Y\}$.

LEMMA 4.7. *If the action $X \times G \to X$ is jointly continuous, then*

so is the action $\mathscr{C}(X) \times G \to \mathscr{C}(X)$. (Compare Prop. 2.2 of Chap. 2 of [2].)

Proof. For $Y \in \mathscr{C}(X)$ and $g \in G$ let $U(C; \{V_1, \cdots, V_n\})$ be a basic neighborhood of Yg. Since $Yg \cap V_i \neq \phi$, there exists $x_i \in Y$, a neighborhood W_i of x_i, and a neighborhood U_i of g with $W_i U_i \subset V_i$. Let $V = \bigcap_{i=1}^n U_i$ and reduce V if necessary so that V^- is compact. Also define $K = \{xh^{-1}: x \in C, h \in V^-\}$. K is compact, and by a further reduction of V it is possible to assume $Y \cap K = \phi$. Now the neighborhood $U(K; \{W_1, \cdots, W_n\}) \times V$ of (Y, g) maps into $U(C; \{V_1, \cdots, V_n\})$.

The next theorem provides some more instances of Borel sets of probability measures.

THEOREM 4.8. *If G is a locally compact separable topological group, then the collection of closed subgroups is closed in $\mathscr{C}(G)$. The collection of closed normal subgroups is also closed in $\mathscr{C}(G)$. If G acts continuously on a locally compact separable space X, the fixed points of $\mathscr{C}(X)$ under the action of G form a closed set of $\mathscr{C}(X)$. Finally, the orbit of any $Y \in \mathscr{C}(X)$ under the action of G is a Borel s et o $\mathscr{C}(X)$.*

Proof. The first statement is just Fell's observation. Indeed, the collection of closed subgroups is the complement of the union of $U(\{e\}; \phi)$ with the basic open neighborhoods of the form $U(V_1^-(V_2^-)^{-1}; \{V_1, V_2\})$, e being the identity of G and V_1^- and V_2^- being compact. The second statement is a special case of the third statement. Simply note that G acts on itself by $(x, g) \to g^{-1}xg$ and hence on $\mathscr{C}(G)$ by $(Y, g) \to Yg = \{g^{-1}xg: x \in Y\}$. The fixed points of this action form a closed subset of $\mathscr{C}(G)$. Intersecting the collection of fixed points with the collection of closed subgroups gives the closed normal subgroups. Finally, the third statement is obvious, and the fourth statement follows because an orbit in $\mathscr{C}(X)$, $\{Yg: g \in G\}$, can be written as a countable union of compact sets, $\bigcup_{n=1}^\infty \{Yg: g \in K_n\}$, if $\{K_n\}_{n=1}^\infty$ is a sequence of compact subsets of G whose union is G.

EXAMPLES. The sets in $\mathscr{C}(R^n)$ spherically symmetric about the origin are just the fixed points of $\mathscr{C}(R^n)$ under the action of the orthogonal group. Furthermore, under the action of the orthogonal group, the orbit of a subspace of dimension m, $m \leqq n$, is the collection of all subspaces of dimension m. The collection of all sets in $\mathscr{C}(R^n)$ geometrically congruent to a given closed set Y lies on the orbit of Y under the affine orthogonal group, i.e., the group generated by all orthogonal transformations and translations. As a consequence, the collection of all affine subspaces of a given dimension m forms a Borel

set of $\mathscr{C}(R^n)$. Similarly, considering the group of dilations (resp. homothetic transformation with the origin as center), it is clear that the collection of closed spheres (resp. closed spheres centered at the origin) forms a Borel set of $\mathscr{C}(R^n)$. (Consult [4] for the geometric terminology.)

The remainder of this section deals with convexity and subspaces and is inspired by [18]. Since the present proofs are different from those in [18] where results overlap, and require perhaps less background of the reader, we have furnished complete arguments.

THEOREM 4.9. *On R^n the map $Y \to clconv\,(Y)$ which takes a closed set into its closed convex hull is Borel from $\mathscr{C}(R^n)$ into $\mathscr{C}(R^n)$. Hence the collection of closed convex sets forms a Borel set of $\mathscr{C}(R^n)$, and the collection of probability measures with convex support forms a Borel set of $M(R^n)$.*

Proof. Let $\{l_i\}_{i\in A}$ be the collection of linear functionals on R^n. For each positive integer k and $i \in A$ define $L_i^k: \mathscr{C}(R^n) \to \mathscr{C}(R^n)$ as follows: $L_i^k(Y) = \{x \in R^n: l_i(x) \leqq \sup_{w\in Y\cap S_k} l_i(w)\}$, where S_k is the sphere $\{w \in R^n: \|w\| \leqq k\}$ and $L_i^k(Y)$ is taken to be ϕ if $Y \cap S_k = \phi$. Let us check that L_i^k is upper semicontinuous. For C compact $\{Y \in \mathscr{C}(R^n): L_i^k(Y) \cap C = \phi\} = \{Y \in (R^n): Y \cap S_k \cap \{t \in R^n: l_i(t) \geqq \inf_{s\in C} l_i(s)\} = \phi\}$, so L_i^k is upper semicontinuous because $S_k \cap \{t \in R^n: l_i(t) \geqq \inf_{s\in C} l_i(s)\}$ is compact. Now note that $Y \to \bigcap_{i\in A} L_i^k(Y)$ is upper semicontinuous and $\bigcap_{i\in A} L_i^k(Y)$ is the closed convex hull of $Y \cap S_k$. It follows that $Y \to \bigcup_{k=1}^{\infty} \bigcap_{i\in A} L_i^k(Y)$ is a Borel map taking Y into its closed convex hull.

The second claim of the theorem is true because the closed convex sets Y of $\mathscr{C}(R^n)$ are simply where the Borel maps $Y \to \bigcup_{k=1}^{\infty} \bigcap_{i\in A} L_i^k(Y)$ and $Y \to Y$ agree.

DEFINITION. Suppose Y is a nonempty closed convex set of R^n. For any point $x \in R^n$, prox (x, Y) is defined to be the unique point of Y closest to x. In other words, if d is the metric on R^n derived from the Euclidean norm, then prox (x, Y) is the unique point of Y where $d(x, Y) = \inf_{w\in Y} d(x, w)$ is attained.

LEMMA 4.10. *Suppose $f: T \to \mathscr{C}(R^n)$ is a Borel map with $f(t)$ a nonempty closed convex set for every $t \in T$. Then for every $x \in R^n$, $t \to$ prox $(x, f(t))$ is a Borel map of T into R^n.*

Proof. First, let us show that $t \to d(x, f(t))$ is a Borel map of T into $R^+ = \{z \in R: z \geqq 0\}$. It suffices to show that $\{t \in T: d(x, f(t)) < \varepsilon\}$

is a Borel set of T for each $\varepsilon > 0$. But $\{t \in T: d(x, f(t)) < \varepsilon\} = \{t \in T: \{y \in R^n: d(x, y) < \varepsilon\} \cap f(t) \neq \phi\}$. Next the map $s \to \{y \in R^n: d(x, y) \leqq s\}$ is Borel from R^+ into $\mathscr{C}(R^n)$ because for K compact

$$\{s \in R^+: \{y \in R^n: d(x, y) \leqq s\} \cap K = \phi\} = \begin{cases} \phi & \text{if } x \in K \\ 0 \leqq s < d(x, K) . \end{cases}$$

Finally, note that $\operatorname{prox}(x, f(t)) = f(t) \cap \{y \in R^n: d(x, y) \leqq d(x, f(t))\}$. Since $t \to \operatorname{prox}(x, f(t))$ is a Borel map into $\mathscr{C}(R^n)$, it follows at once that it is a Borel map into R^n.

LEMMA 4.11. *Suppose $f: T \to \mathscr{C}(R^n)$ has the same properties as in the last lemma. Then the map $t \to f(t)^\perp = \{y \in R^n: \forall x \in f(t)\langle x, y\rangle = 0\}$ is also Borel from T into $\mathscr{C}(R^n)$.*

Proof. Let $\{x_m\}_{m=1}^\infty$ be a dense collection of points in R^n. Since $f(t)$ equals the closure of $\bigcup_{m=1}^\infty \operatorname{prox}(x_m, f(t))$, $\{y \in R^n: \forall x \in f(t)\langle x, y\rangle \leqq 0\} = \bigcap_{m=1}^\infty \{y \in R^n: \langle \operatorname{prox}(x_m, f(t)), y\rangle \leqq 0\}$. Now $t \to \{y \in R^n: \langle \operatorname{prox}(x_m, f(t)), y\rangle \leqq 0\}$ is a Borel map into $\mathscr{C}(R^n)$ because $s \to \{y \in R^n: \langle s, y\rangle \leqq 0\}$ is Borel from R^n into $\mathscr{C}(R^n)$. In fact, for a compact set K, $s \to \inf_{y \in K} \langle s, y\rangle$ is Borel and $\{s \in R^n: \{y \in R^n: \langle s, y\rangle \leqq 0\} \cap K \neq \phi\} = \{s \in R^n: \inf_{y \in K} \langle s, y\rangle \leqq 0\}$. Similarly, $t \to \{y \in R^n: \forall x \in f(t)\langle x, y\rangle \geqq 0\}$ is Borel. To finish the lemma observe that $t \to \{y \in R^n: x \in f(t)\langle x, y\rangle = 0\}$ is the intersection of two Borel maps.

THEOREM 4.12. *The map $\mathscr{C}(R^n)\backslash\{\phi\} \to \mathscr{C}(R^n)$ taking a closed set Y into the smallest subspace containing Y is Borel. Likewise, the map $\mathscr{C}(R^n)\backslash\{\phi\} \to \mathscr{C}(R^n)$ taking Y into the smallest affine subspace containing Y is Borel.*

Proof. The smallest subspace containing Y is $[\operatorname{clconv}(Y)]^{\perp\perp}$. The smallest affine subspace containing Y is

$$[\operatorname{clconv}(Y) - \operatorname{prox}(0, \operatorname{clconv}(Y))]^{\perp\perp} + \operatorname{prox}(0, \operatorname{clconv}(Y)) .$$

The second map is Borel because the action $\mathscr{C}(R^n) \times R^n \to \mathscr{C}(R^n)$ defined by translation is jointly continuous.

COROLLARY 4.13. *For each $m \leqq n$, the collection of probability measures on R^n whose supports lie within a subspace (affine subspace) of dimension m forms a Borel set of $M(R^n)$.*

5. Further examples of Borel sets of probability measures.

EXAMPLE 5.1. Let G be a metric group acting continuously on

a Polish space X. Then G acts by translation on the set of probability measures $M(X)$. Indeed, define for $\mu \in M(X)$ and $g \in G$ μg to be the probability measure assigning measure $\mu(Ag^{-1})$ to each Borel set A of X.

THEOREM 5.2. *G acts continuously on $M(X)$.*

Proof. Suppose $\mu_n \to \mu$ in $M(X)$ and $g_n \to g$ in G as $n \to \infty$. It is sufficient to prove that

$$\int f(sg_n)d\mu_n(s) \longrightarrow \int f(sg)d\mu(s)$$

for every bounded continuous real-valued function f. For $\varepsilon > 0$ let K be a compact subset of X with $\mu_n(K) \geqq 1 - \varepsilon$ for all n. Now estimate as follows:

$$\left|\int f(sg_n)d\mu_n(s) - \int f(sg)d\mu(s)\right| \leqq \int_K |f(sg_n) - f(sg)|\,d\mu_n(s)$$
$$+ \int_{X\setminus K} |f(sg_n) - f(sg)|\,d\mu_n(s) + \left|\int f(sg)d\mu_n(s) - \int f(sg)d\mu(s)\right| . \quad **$$

The last term in ** can be made small since $\mu_n \to \mu$. The middle term on the right of ** is bounded by $2 \sup_{s\in X} |f(s)|\,\varepsilon$. The first term on the right can be made small because $sg_n \to sg$ uniformly on K, as we prove momentarily, and because f is uniformly continuous on the compact set $\{kh: k \in K, h \in \{g_n\}_{n=1}^{\infty}$ or $h = g\}$. To show that $sg_n \to sg$ uniformly on K let d be the metric on X and suppose $d(s_m g_{n_m}, s_m g) > \delta$ for some subsequence $\{g_{n_m}\}$ of $\{g_n\}$, a sequence $\{s_m\}$ of K and some $\delta > 0$. Since K is compact we may assume $s_m \to s \in K$. Then $s_m g_{n_m} \to sg$ and $s_m g \to sg$ by the joint continuity of the action of G on S. This contradicts the assumption that $d(s_m g_{n_m}, s_m g) > \delta$ for all m. Hence $sg_n \to sg$ uniformly and this completes the proof of the theorem.

COROLLARY 5.3. *The invariant measures form a closed subset of $M(X)$ since they are the fixed points for the action of G on $M(X)$. If G is a Polish group, then the orbit of any probability measure under G is Borel in $M(X)$.*

Proof. For the second assertion see Lemma 3.4 of [9].

Applications. On R^n, the orbit of any nondegenerate normal distribution under the group of invertible affine transformations is the whole collection of nondegenerate normal distributions. Also the collection of translates of any probability measure is Borel in $M(R^n)$.

Other commonly occurring groups acting on R^n are the orthogonal group and the group permuting the coordinates of any point. The latter group arises in the theory of order statistics.

REMARK. A Borel set B of a space X is called invariant under a group action if $Bg = B$ for every group element g. If μ is a probability measure on X and $\mu(B) = 0$ or 1 for every invariant Borel set B, then μ is said to be ergodic. Varadarajan has shown that the collection of invariant ergodic probability measures is Borel in $M(X)$ if X is a Polish space and the underlying group is locally compact and separable. (See Thm. 4.2 of [20].)

EXAMPLE 5.4. Suppose $\{X_n\}_{n=1}^{\infty}$ is a sequence of separable metric spaces. Consider the probability measures $M(\prod_{n=1}^{\infty} X_n)$ on the product space $\prod_{n=1}^{\infty} X_n$. We claim that the set $P = \{\mu \in M(\prod_{n=1}^{\infty} X_n)\colon \mu = \prod_{n=1}^{\infty} \mu_n, \mu_n \in M(X_n)\}$ is closed in $M(\prod_{n=1}^{\infty} X_n)$. Our reasoning goes as follows: The map $M(\prod_{n=1}^{\infty} X_n) \to \prod_{n=1}^{\infty} M(X_n)$ taking a probability measure into its sequence of marginal probability measures is continuous. Also the map $\prod_{n=1}^{\infty} M(X_n) \to M(\prod_{n=1}^{\infty} X_n)$ taking a sequence of probability measures into their product is continuous. (Modify slightly the proof of Lemma 1.1 of Chap. 3 of [16].) P is the set where the composition of these two maps agrees with the identity map on $M(\prod_{n=1}^{\infty} X_n)$.

If each X_n is the same, then the set of probability measures on $\prod_{n=1}^{\infty} X_n$ having all marginals the same is certainly closed too. Hence the set of probability measures on $\prod_{n=1}^{\infty} X_n$ which are product measures with equal components is closed.

EXAMPLE 5.5. Suppose X is locally compact and separable. According to Corollary 4.6, the collection of probability measures concentrated at k or fewer points is Borel in $M(X)$. A stronger assertion is possible.

LEMMA 5.6. *Let X be a Polish space. For each $1 \geqq \delta > 0$ the collection of probability measures having k or fewer atoms with total mass $\geqq \delta$ is closed in $M(X)$.*

Proof. An easy induction using Prohorov's Theorem. (See Thm. 6.7 of Chap. 2 of [16].)

For another application of Lemma 5.6 put $A_{k,n} = \{\mu \in M(X)\colon \mu$ has k or less atoms with total mass $\geqq 1 - 1/n\}$. Then the Borel set $\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A_{k,n}$ consists of those probability measures concentrated on a finite or countable set of points.

EXAMPLE 5.7. If one is more interested in the number or geometry of the atoms rather their total weight, one can proceed as follows in the locally the compact case: For each $\mu \in M(X)$ and $\varepsilon > 0$ let $d_\varepsilon(\mu)$ be the set of atoms of μ having individual mass of at least ε. Since $\{\mu \in M(X): d_\varepsilon(\mu) \cap C \neq \phi\}$ is closed in $M(X)$ for each compact set C of X, $d_\varepsilon: M(X) \to \mathscr{C}(X)$ is upper semicontinuous. (See Prop. I. 2. 8 of [1].) Apply part 1. of Lemma 4.4 to conclude that $\mu \to (\bigcup_{n=1}^{\infty} d_{1/n}(\mu))^-$ is a Borel map into $\mathscr{C}(X)$. $\mu \to (\bigcup_{n=1}^{\infty} d_{1/n}(\mu))^-$ can be used to keep track of the cardinality of the atoms and their positions.

EXAMPLE 5.8. The collection of probability measures on R^n with some moment (all moments) existing is a Borel set of $M(R^n)$. To prove this let $\{g_i: R^n \to [0, 1]\}_{i=1}^{\infty}$ be a collection of continuous functions satisfying

$$g_i(x) = \begin{cases} 1 & \text{if } \|x\| \leqq i \\ 0 & \text{if } \|x\| \geqq i + 1 \,. \end{cases}$$

Now note that

$$\mu \longrightarrow \sup_i \int |x^k| \, g_i(x) d\mu(x) = \int |x^k| \, d\mu(x)$$

is lower semicontinuous in the classical sense for each multi-index k. Hence the collection of probability measures having finite kth moment is a countable union of closed sets. In general, this collection is neither open nor closed. For instance, on R it is possible to show that the collection of probability measures having finite first moment is neither open nor closed. Furthermore, the collection of probability measures lacking a first moment is dense in $M(R)$.

EXAMPLE 5.9. Occasionally it is convenient to deal only with those probability measures on R^n having continuous, strictly increasing distribution functions. To characterize this family of probability measures consider for each pair of positive integers n and m the map $G_{n,m}: C^0(R^n) \to R$ defined by

$$G_{n,m}(g) = \inf_{\substack{\|z-w\|>1/n \\ \|z\| \leq m \\ \|w\| \leq m \\ z \geqq W}} g(z) - g(w) \,,$$

where $z \geqq w$ means $z_i - w_i \geqq 0$ for each component of $z - w$. It is easy to check that $G_{n,m}$ is upper semicontinuous in the classical sence. Hence $\mu \to G_{n,m}(F_\mu)$ is a Borel map into R from the collection of probability measures having continuous distribution functions. μ has strictly increasing distribution function iff $G_{n,m}(F_\mu) > 0$ for every n

and m. Similar arguments can be used to show that $\{(\mu, \nu): F_\mu, F_\nu$ continuous, $F_\mu(x) > F_\nu(x)\ \forall x\}$ is Borel in $M(R^n) \times M(R^n)$. However, removing the continuity assumptions makes both problems much more difficult.

EXAMPLE 5.10. A probability measure μ on R is said to be symmetric if the Fourier transform $\hat{\mu}(\theta)$ of μ can be written as $e^{it\theta}r(\theta)$, where $r(\theta)$ is a real-valued function of θ and t is some real constant. Now for each positive integer m, $\{e^{it\theta}r(\theta): t \in [-m, m]$ r real-valued and in $C^0(R)\}$ is closed in $C^0(R)$. Since $\mu_n \to \mu$ in $M(R)$ iff $\hat{\mu}_n \to \hat{\mu}$ in $C^0(R)$, the collection of symmetric probability measures on R is a countable union of closed sets of $M(R)$.

EXAMPLE 5.11. For μ a probability measure on R and $p \in (0, 1)$, $t \in R$ is called a pth percentile of μ if $\mu(-\infty, t) \leqq p$ and $\mu(-\infty, t] \geqq p$. When $p = 1/2$ the term median is used instead of percentile. It is easy to show that the set of pth percentiles for μ is a compact interval whose right endpoint is $\{^+(p, \mu) = \sup\{r \in Q: \mu(-\infty, r) \leqq p\}$ and whose left endpoint is $\{^-(p, \mu) = \inf\{r \in Q: \mu(-\infty, r] \geqq p\}$, where Q is the set of rationals. Moreover, $\{^+(p, \mu)$(resp. $\{^-(p, \mu)$) is right (resp. left) continuous in p for fixed μ and Borel in μ for fixed p. Hence Lemma 3.1 implies $\{^+(p, \mu)$ and $\{^-(p, \mu)$ are jointly Borel in p and μ. Using this fact one can show various hypotheses in nonparametric statistics involving the set of pth percentiles to be Borel. Perhaps it is worth pointing out that $(p, \mu) \to [\{^-(p, \mu), \{^+(p, \mu)]$ is a Borel map into $\mathscr{C}(R)$.

EXAMPLE 5.12. Let us indicate briefly now the Borel structure on $M(X)$ furnishes a natural framework for the description of several ideas in probability and statistics. For instance, in the theory of Markov processes one can define transition functions as Borel maps from X into $M(X)$. If μ is a probability measure on a Polish space X and π is a Borel map onto another Polish space Y, define $\bar{\mu}$ on Y by $\bar{\mu}(A) = \mu(\pi^{-1}(A))$ for every Borel set A of Y. Then μ has a regular conditional probability distribution given π. From our perspective this means a Borel map $y \to \mu_y$ from Y into $M(X)$ such that $\bar{\mu}$-almost all μ_y are concentrated on $\pi^{-1}(y)$ and $\mu(B) = \int \mu_y(B) d\bar{\mu}(y)$ for each Borel set B of X. Finally, we should cite empirical distribution functions. Let $\{f_i: S \to R\}_{i=1}^\infty$ be a sequence of independent and identically distributed random variables on a Borel space S with probability measure μ. For each positive integer n define a Borel map $\mu_n: S \to M(R)$ by taking $\mu_n(s)$ to be the probability measure giving equal weight to $f_1(s), \cdots, f_n(s)$. μ_n is Borel because for every Borel set

$$A \subset R,\ \mu_n(s)(A) = \frac{1}{n} \sum_{i=1}^{n} \chi_A(f_i(s)) ,$$

where χ_A is the indicator function of A.

6. **Counterexamples.** Obviously not all subsets of $M(X)$ are Borel. Here are some counter-examples.

EXAMPLE 6.1. For X a Polish space it is well known that X is homeomorphic to the collection of unit point masses, $\{\delta_w \in M(X)\colon \delta_w(\{w\}) = 1, w \in X\}$ [16]. If X is uncountable, then there exists $Y \subset X$ which is not Borel. But then $\{\delta_w \in M(X)\colon w \in Y\}$ cannot be Borel in $M(X)$ either.

EXAMPLE 6.2. Our second counterexample involves the notion of equivalence between probability measures. It is transparent that mutual absolute continuity, denoted $\sim$, is an equivalence relation on $M(X)$. By the axiom of choice it is possible to choose one representative probability measure from each equivalence class. The next theorem shows when this "transversal" can also be taken to be a Borel set of X.

THEOREM 6.3. *Suppose X is a Polish space. Then a Borel transversal exists for $\sim$ on $M(X)$ iff X is countable or finite.*

Proof. Suppose X is the set of positive integers. Give $Z_n = \{0, 1\}$ the discrete topology and consider the product space $\prod_{n=1}^{\infty} Z_n$. Subtract off from $\prod_{n=1}^{\infty} Z_n$ the countable number of sequences in which 1 appears only finitely often and call the remainder Z. Map Z into $M(X)$ by taking the sequence $\{w_n\}_{n=1}^{\infty}$ into the probability measure giving mass $w_n(1/2)^{w_1+\cdots+w_n}$ to the integer n. This map is one-to-one, Borel, and provides the desired Borel transversal. The case of X finite is even simpler.

Now assume X is uncountable. Since any two uncountable Polish spaces X and Y are Borel isomorphic, (Thm. 2.12 of Chap. 1 of [16]), it is easy to see that $M(X)$ and $M(Y)$ will be Borel isomorphic too. Hence it is enough to establish the necessary part of the theorem for the space $\prod_{n=1}^{\infty} Z_n$ above. But this is the content of Lemma 5.1 of [15]. Here it is proved that $M(\prod_{n=1}^{\infty} Z_n)/\sim$ is not countably separated. If a Borel (even analytic) transversal existed in this case, then Prop. 2.12 of Chap. 1 of [2] would be contradicted, since $M(\prod_{n=1}^{\infty} Z_n)/\sim$ cannot be analytic if it fails to be countably separated. Note that $\sim$ is Borel as a subset of $M(\prod_{n=1}^{\infty} Z_n) \times M(\prod_{n=1}^{\infty} Z_n)$ because of 2.11 of [8].

EXAMPLE 6.4. Our next two counterexamples partially justify sticking to locally compact spaces when discussing the relation of a

probability measure to its support. Suppose X is a Polish space. Define a Borel structure on the space $\mathscr{C}(X)$ of closed subsets of X by requiring every collection, $\{A \in \mathscr{C}(X): A \subset B\}$ to be Borel whenever $B \in \mathscr{C}(X)$. Christensen shows in Thm. 1 of [5] that this Borel structure is analytic and on the subspace of nonempty closed sets coincides with the Borel structure generated by the Hausdorff metric associated with any precompact metric on X. Furthermore, if X is locally compact, this is the Borel structure generated by the Fell topology.

Now it is evident that $\mu \to \text{support}\,(\mu)$ is Borel from $M(X)$ into $\mathscr{C}(X)$. If X is a real infinite dimensional separable Hilbert space, Christensen proves that the collection W of $Z \in \mathscr{C}(X)$ contained in the open unit sphere is complementary analytic but not analytic. (See Thm. 8 of [5].) Since every $Z \in W$ is the support of some $\mu \in M(X)$, the inverse image of W under $\mu \to \text{support}\,(\mu)$ fails to be Borel or even analytic.

This counterexample also illustrates that the intersection map $\mathscr{C}(X) \times \mathscr{C}(X) \to \mathscr{C}(X)$ need not be Borel when X is not locally compact. Indeed, let Y be the complement of the open unit sphere. Then $W = \{Z \in \mathscr{C}(X): Z \cap Y = \phi\}$.

EXAMPLE 6.5. The same phenomenon of Example 6.4 occurs if X is a countable Cartesian product of the positive integers. Then the collection of $Z \in \mathscr{C}(X)$ which are open as well as closed is complementary analytic but not analytic. (See Thm. 5 of [5].)

ACKNOWLEDGEMENTS. I would like to thank Richard Dudley, Eric Nordgren, Don Cohn, and M. Takesaki for assistance with some of the proofs of this paper and for pointing out valuable references.

REFERENCES

1. Alfsen, *Compact Convex Sets and Boundary Integrals*, Springer, Berlin, 1971.
2. Auslander and C. Moore, *Unitary representations of solvable Lie groups*, Memoirs Amer. Math. Soc., **62**, Amer. Math. Soc., Providence, 1966.
3. Berge, *Espaces Topologiques et Fonctions Multivoques*, Dunod, Paris, 1959.
4. Choquet, *Geometry in a Modern Setting*, Houghton Mifflin, Boston, 1964.
5. Christensen, *On some properties of Effros Borel structure on spaces of closed subsets*, Math. Ann., **195** (1971), 17-24.
6. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1960.
7. ———, *Treatise on Analysis Vol. II*, Academic Press, New York, 1970.
8. Dubins and Freedman, *Measurable sets of measures*, Pacific J. Math., **14** (1964), 1211-1222.
9. Effros, *Transformation Groups and C*-Algebras*, Ann. of Math., **81** (1965), 38-55.
10. Fell, *A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff space*, Proc. Amer. Math. Soc., **13** (1962), 472-476.
11. Kuratowski, *Topology Vol. I*, Academic Press, New York, 1966.
12. Meyer, *Probability and Potentials*, Ginn (Blaisdell), Boston, 1966.

13. R. Moore, *Measurable, continuous, and smooth vectors for semigroups and group representation*, Memoirs Amer. Math. Soc., **78**, Amer. Math. Soc., Providence, 1968.
14. Mori, *On random translations of point processes*, Yokoama Math. J., **19** (1971), 119-139.
15. Nielsen, *Borel Sets of Von Neumann Algebras*, Amer. J. Math., **95** (1973), 145-164.
16. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.
17. Ramsay, *Virtual groups and group actions*, Advances in Math., **6** (1971), 253-328.
18. Rockafellar, *Measurable dependence of convex sets and functions on parameters*, J. Math. Anal. Appl., **28** (1969), 4-25.
19. Treves, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.
20. Varadarajan, *Groups of automorphisms of Borel spaces*, Trans. Amer. Math. Soc., **109** (1963), 191-220.

Received May 11, 1972.

UNIVERSITY OF NEW HAMPSHIRE

Current address:
Department of Biomathematics
School of Medicine
University of California
Los Angeles, CA 90024.

Printed in Japan by International Academic Printing Co., Ltd., Tokyo, Japan

Pacific Journal of Mathematics
Vol. 48, No. 1 March, 1973