ON ORTHOGONAL PROBABILITY MEASURES[1]

AGNES BERGER

Definitions. Let X be an arbitrary set, $\mathcal{L}$ a Borel-field of some subsets B of X, and $\mathfrak{M}(\mathcal{L})$ the family of all probability measures defined on $\mathcal{L}$, i.e. the totality of all countably additive, non-negative set functions $m(B)$, $B \in \mathcal{L}$, for which $m(X) = 1$. Henceforth the word "measure" denotes an element of $\mathfrak{M}(\mathcal{L})$ and the expression "set of measures" a subset of $\mathfrak{M}(\mathcal{L})$.

Two measures m and m' have been called equivalent (notation: $m \sim m'$, as for example in [1]), if $m'(B) = 0$ for every set B in $\mathcal{L}$ for which $m(B) = 0$ and $m(B) = 0$ for every set B in $\mathcal{L}$ for which $m'(B) = 0$. A corresponding notion for sets of measures has been defined in [3] by Halmos and Savage in the following way: Two sets of measures $M = \{m\}$ and $M' = \{m'\}$ are called equivalent (notation: $M \sim M'$), if whenever $m(B) = 0$ for all m in M, B in $\mathcal{L}$, then $m'(B) = 0$ for all m' in M' and, conversely, whenever $m'(B) = 0$ for all m' in M', B in $\mathcal{L}$, then $m(B) = 0$ for all m in M. Further, Halmos and Savage call a set of measures $M = \{m\}$ dominated (notation: $M \prec \mu$), if there exists a measure μ such that $m(B) = 0$ for every m in M, whenever $\mu(B) = 0$, $B \in \mathcal{L}$. Having established this terminology, Halmos-Savage prove the following lemma, which will be used later:

LEMMA OF HALMOS-SAVAGE. *Every dominated set of measures has an equivalent countable subset.*

A complementary notion to equivalence is orthogonality. Two measures m and m' have been called orthogonal (for example in [1]) (notation: $m \perp m'$ or equivalently, $m' \perp m$), if there exists a set B in $\mathcal{L}$ such that $m(B) = 0$ and $m'(B) = 1$. As pointed out in [1], in order that $m \perp m'$ it is necessary and sufficient[2] that for any $\epsilon > 0$, there exist a set B in $\mathcal{L}$ such that $m(B) < \epsilon$ and $m'(B) > 1 - \epsilon$. To extend the notion of orthogonality to sets of measures, we shall use the following definition for the purposes of this paper: Two sets of measures $M = \{m\}$ and $M' = \{m'\}$ are orthogonal (notation: $M \perp M'$, or equivalently, $M' \perp M$), if there exist a set B in $\mathcal{L}$ such that $m(B) = 0$ for all m in M while $m'(B) = 1$ for all m' in M'. Note that even if M and M'

[2] The proof in [1] is as follows: Necessity is clear. To show sufficiency, let for each n, $B_n \in \mathcal{L}$ be a set such that $m(B_n) < 2^{-n}$ and $m'(X - B_n) < 2^{-n}$. Then $B = \lim \sup_{n \to \infty} B_n$ satisfies $m(B) = 0$ and $m'(B) = 1$.

are such that $m \perp m'$ for every pair m in M, m' in M', M and M' are not necessarily orthogonal according to this definition, since the sets B on which $m(B) = 0$ and $m'(B) = 1$ may depend on the pair m, m'.

Some specific consequences of this definition:

LEMMA 1. *Let $M = \{m\}$ and $M' = \{m'\}$ be disjoint sets of measures. If there exist a sequence M'_j, $j = 1, 2, \cdots$, of subsets of M' such that*

$$M' = \bigcup_{j=1}^{\infty} M'_j$$

and

$$M \perp M_j', \qquad\qquad i = 1, 2, \cdots,$$

then $M \perp M'$.

PROOF. By assumption, there exist B_j in $\mathcal{L}$, $j = 1, 2, \cdots$, such that $m(B_j) = 0$ for all m in M and $m'(B_j) = 1$ for all m' in M'_j, $j = 1, 2, \cdots$.

Let $B = \bigcup_{j=1}^{\infty} B_j$. Then $m(B) = 0$ for all m in M and $m'(B) = 1$ for all m' in M', since from $M' = \bigcup_{j=1}^{\infty} M'_j$ it follows that if m' is any member of M', there exists an index k such that m' is in M'_k and consequently

$$m'(B) \geqq m'(B_k) = 1.$$

LEMMA 2. *Let $M = \{m\}$ and $M' = \{m'\}$ be dominated sets of measures. A necessary and sufficient condition for M to be orthogonal to M' is that there exist a measure μ such that $M \prec \mu$ and a measure μ' such that $M' \prec \mu'$ and that μ is orthogonal to μ'.*

PROOF. Sufficiency is clear. To show necessity, recall that by the lemma of Halmos and Savage, there exist a sequence $\{m_i\}$, m_i in M, $i = 1, 2, \cdots$, such that $\{m_i\} \sim M$, and a sequence $\{m'_j\}$, m'_j in M', $j = 1, 2, \cdots$, such that $\{m'_j\} \sim M'$. Put $\mu = \sum_{i=1}^{\infty} a_i m_i / \sum_{i=1}^{\infty} a_i$, $\mu' = \sum_{j=1}^{\infty} b_j m'_j / \sum_{j=1}^{\infty} b_j$, where $\sum_{i=1}^{\infty} a_i$ and $\sum_{j=1}^{\infty} b_j$ exist, $a_i > 0$, $b_j > 0$, $i = 1, 2, \cdots$; $j = 1, 2, \cdots$. Assume that $M \perp M'$. Then there exists B in $\mathcal{L}$ such that $m(B) = 0$ for all m in M, $m'(B) = 1$ for all m' in M'. It follows that then $\mu(B) = 0$ and $\mu'(B) = 1$, q.e.d.

Infinite direct power measures. For any measure m in $\mathfrak{M}(\mathcal{L})$, let us define its infinite direct power, to be denoted by m^*, as the infinite direct product measure every factor of which is m. (For definition of infinite direct product measures, refer to Halmos [4] or Kakutani [2].

It is a known result that $m^* \perp m'^*$ whenever $m \not\equiv m'$, so that any

set of different infinite power measures consists of pairwise orthogonal elements. It follows immediately from Lemma 1 that two disjoint sequences $M^* = \{m_i^*\}$, $i = 1, 2, \cdots$, $M'^* = \{m_j'^*\}$, $j = 1, 2, \cdots$, of different infinite power measures are orthogonal sets of measures. Let now $M^* = \{m^*\}$ and $M'^* = \{m'^*\}$ be disjoint sets of different infinite power measures and assume that they consist of more than enumerably many elements. In order to find out whether $M^* \perp M'^*$ one may try to consider first the case when M^* or M'^* are dominated. That this is futile can be seen from the lemma below:

LEMMA 3. *A dominated set of pairwise orthogonal measures has at most enumerably many elements.*

PROOF. Let $M = \{m\}$ be dominated and assume that it consists of pairwise orthogonal measures. By the lemma of Halmos-Savage, there exists a sequence $\{m_i\}$, m_i in M, $i = 1, 2, \cdots$, such that $\{m_i\} \sim M$. If M has more than enumerably many elements, then there must be an m in M that does not belong to the sequence $\{m_i\}$ and, by assumption, m is orthogonal to m_i, $i = 1, 2, \cdots$. By Lemma 1, then, it follows that m is orthogonal to $\{m_i\}$, $i = 1, 2, \cdots$, contradicting the fact that $\{m_i\} \sim M$. Accordingly, if a set of measures $M = \{m\}$ has more than enumerably many elements, the corresponding set of infinite power measures $M^* = \{m^*\}$ is necessarily not a dominated set of measures.

The problem. In the following we shall formulate some conditions under which $M^* \perp M'^*$ holds in terms of the properties of M and M', applying the method of a related study [5] in which, however, only orthogonal sets of a more specialized kind were considered.

To begin, we restate a well known result (stated as Lemma 1 and accompanied by a proof in [5]).

LEMMA 4 (BERNOULLI). *Let $\{m\} = M$ and $\{m'\} = M'$ be disjoint sets of measures and assume that there exists a set A in $\mathcal{L}$ such that for some $\delta > 0$*

$$| m(A) - m'(A) | > \delta$$

for every m in M and m' in M'. Then $M^* \perp M'^*$.

DEFINITION. Let us say that the sets of measures $M = \{m\}$ and $M' = \{m'\}$ satisfy Condition A if there exists a sequence $\{M_j'\}$, $j = 1, 2, \cdots$, of subsets of M' and to every j a set A_j in $\mathcal{L}$ and a positive number δ_j such that $\bigcup_{j=1}^{\infty} M_j' = M'$ and

$$| m(A_j) - m'(A_j) | > \delta_j, \qquad m \in M, m' \in M_j', j = 1, 2, \cdots.$$

LEMMA 5. *Let* $\{m\} = M$ *and* $\{m'\} = M'$ *be disjoint sets of measures and assume that there exist a sequence of subsets of M, say* $\{M_i\}$, $i = 1, 2, \cdots$, *such that* $\bigcup_{i=1}^{\infty} M_i = M$ *and such that for every i, the pair M_i and M' satisfy Condition* A. *Then* $M^* \perp M'^*$.

PROOF. By assumption, there exist a sequence of subsets of M', say $\{M'_{ij}\}$, $i = 1, 2, \cdots$; $j = 1, 2, \cdots$, such that $\bigcup_{j=1}^{\infty} M'_{ij} = M'$ for every i and $|m(A_{ij}) - m'(A_{ij})| > \delta_{ij}$ for all m in M_i and all m' in M'_{ij} for some sequence $\{A_{ij}\}$ of sets in $\mathcal{L}$ and some sequence of positive constants δ_{ij}, $i = 1, 2, \cdots$; $j = 1, 2, \cdots$. Then $M_i^* \perp M'^*_{ij}$ for all j by Lemma 4, so that Lemma 1 implies first that $M_i^* \perp M'^*$ for all i and finally that $M^* \perp M'^*$.

Weak topology of a set of measures. The nature of Condition A suggests that by imposing suitable separability conditions on the sets M and M' it will be possible to guarantee that $M^* \perp M'^*$. In order to do this a weak topology is introduced in the subsets of $\mathfrak{M}(\mathcal{L})$ in the usual way (see von Neumann [7]), by defining a complete system of neighborhoods as follows:

Let M be any subset of $\mathfrak{M}(\mathcal{L})$. For every p in M and for every finite collection of sets $B_1, B_2, \cdots, B_k$ in $\mathcal{L}$ and $\epsilon > 0$, let $\alpha = (B_1, B_2, \cdots, B_k; \epsilon)$ and

$$V_\alpha(p) = \{m \text{ in } M \mid |m(B_i) - p(B_i)| < \epsilon, i = 1, 2, \cdots, k\},$$

i.e. the set of all those measures in M whose values assumed on the sets $B_1, B_2, \cdots, B_k$ differ less than ϵ in absolute value from the corresponding values of p. $V_\alpha(p)$ is called the neighborhood of index α of the measure p. By letting α range over all possible finite collections of sets in $\mathcal{L}$ and all positive numbers ϵ, $V_\alpha(p)$ define a complete system of neighborhoods (see Hausdorff [8]) so that M may be regarded as a topological space. Only the following elementary notions of general topology will be used:

1. A topological space is separable if it has an everywhere dense countable subset.

2. A topological space is compact if every open covering of the space has a finite subcovering.

3. A topological space is said to have the Lindelöf property if every open covering of the space has an at most countable subcovering. Examples of spaces with the Lindelöf property are unions of countably many compact spaces, separable metric spaces, etc.

Besides these we shall make use of the following result proved in [6]:

LEMMA 6. *If a set of measures M is separable with respect to the weak topology defined above, then it is dominated.*

THEOREM 1. *Let $M = \{m\}$ and $M' = \{m'\}$ be disjoint sets of measures. If, with respect to the weak topology defined above, M is the union of at most enumerably many compact sets and M' has the Lindelöf property, then $M^* \perp M'^*$.*

PROOF. In view of Lemma 1, it is sufficient to prove the theorem under the assumption that M is compact.

Let m' in M' be fixed. To every m in M, there exists a set $B = B(m)$ in $\mathcal{L}$ and a number $\delta(m) > 0$ such that

$$\left| m(B) - m'(B) \right| > \delta(m).$$

Let $\alpha = [B(m); \delta(m)/2]$ and

$$V_\alpha(m) = \left\{ p \text{ in } M \mid \left| p(B) - m(B) \right| < \delta(m)/2 \right\}.$$

Since $\bigcup_{\{m\}} V_\alpha(m)$ is an open covering of a compact space, it may be replaced by a finite subcovering, say $\bigcup_{\{m\}} V_\alpha(m) = \bigcup_{i=1}^{k} V_{\alpha_i}(m_i) = M$, $\alpha_i = [B(m_i); \delta(m_i)/2]$, $i = 1, 2, \cdots, k$. Put $\min_i \delta(m_i) = \delta$ and $B(m_i) = B_i$ and let $\beta = [B_1, B_2, \cdots, B_k; \delta]$. Consider the neighborhood of m' given by

$$V_\beta(m') = \left\{ p' \text{ in } M' \mid \left| p'(B_i) - m'(B_i) \right| < \delta/3, \ i = 1, 2, \cdots, k \right\}.$$

Let p be any measure in M. It is covered by at least one of the $V_{\alpha_i}(m_i)$, say by $V_{\alpha_l}(m_l)$, $1 \leq l \leq k$. Then if p' is any measure in $V_\beta(m')$,

$$\left| p(B_l) - p'(B_l) \right| \geq \left| m_l(B_l) - m'(B_l) \right| - \left| m_l(B_l) - p(B_l) \right|$$
$$- \left| p'(B_l) - m'(B_l) \right| > \delta(m_l) - \delta(m_l)/2$$
$$- \delta(m_l)/3 = \delta(m_l)/6 = \eta_l > 0.$$

Therefore the pair $V_\beta(m')$ and M satisfy Condition A. Because of the Lindelöf property, M' can be covered by an at most enumerable sequence of the open sets $V_\beta(m')$ and thus the theorem follows by Lemma 5.

If neither of the sets M and M' is the union of countably many compact sets, I could not prove a similarly general theorem but would have to assume that the Borel field $\mathcal{L}$ on which $\mathfrak{M}(\mathcal{L})$ is defined is separable, i.e. there exists an enumerable collection of subsets $\{B_i\}$ in $\mathcal{L}$ such that the smallest Borel-field containing the collection $\{B_i\}$ coincides with $\mathcal{L}$.

THEOREM 2. *Let $\mathcal{L}$ be a separable Borel-field and $M = \{m\}$ and $M' = \{m'\}$ two disjoint subsets of $\mathfrak{M}(\mathcal{L})$. Assume that M and M' are separable in the weak topology described above and that*

$$\underset{\substack{\text{g.l.b.} \\ m' \text{ in } M'}}{} \quad \underset{\substack{\text{l.u.b.} \\ B \text{ in } \mathcal{L}}}{} \big| m(B) - m'(B) \big| = \delta(m) > 0.$$

Then $M^* \perp M'^*$.

PROOF. For any two measures p and q in $\mathfrak{M}(\mathcal{L})$, let

$$d(p, q) = \underset{\substack{\text{l.u.b.} \\ B \text{ in } \mathcal{L}}}{} \big| p(B) - q(B) \big|.$$

It is well known that $d(p, q)$ can be considered as a metric, and that if $\mathcal{L}$ is a separable Borel-field then any dominated set of measures is separable with respect to the topology induced by d. (Proof of this can be based on Hopf's approximation-theorem, see [9; 10].)

By Lemma 6, M and M' are dominated sets, so by the assumption on $\mathcal{L}$, they become separable metric spaces with the metric d. Let m in M be fixed. Let

$$C(m) = \big\{ p \text{ in } M \,\big|\, d(p, m) < \delta(m)/6 \big\}$$

and for every m' in M' let

$$C(m') = \big\{ p' \text{ in } M' \,\big|\, d(p', m') < \delta(m)/6 \big\}.$$

Since M' is a separable metric space, it has the Lindelöf property and therefore there exists a sequence $C(m'_j)$, $j = 1, 2, \cdots$, such that the open covering $\bigcup_{m' \text{ in } M'} C(m') = \bigcup_{j=1}^{\infty} C(m'_j)$. Whatever j is, by assumption, there exists a set B_j in $\mathcal{L}$ such that

$$\big| m(B_j) - m'_j(B_j) \big| > \delta(m)/2,$$

hence for any p in $C(m)$ and any p' in $C(m'_j)$

$$\big| p(B_j) - p'(B_j) \big| \geq \big| m(B_j) - m'_j(B_j) \big| - \big| m(B_j) - p(B_j) \big|$$
$$- \big| p'(B_j) - m'_j(B_j) \big| > \delta(m)/2$$
$$- 2\delta(m)/6 = \eta(m) > 0.$$

Therefore the sets $C(m)$ and M' satisfy Condition A. Since M also satisfies the Lindelöf condition, M can be covered by an enumerable sequence of $C(m)$'s, so that the theorem follows from Lemma 5.

REFERENCES

1. S. Kakutani, *On equivalence of infinite product measures*, Ann. of Math. vol. 49 (1948) pp. 214–224.

2. ———, *Notes on infinite product measure spaces*, I, Proc. Imp. Acad. Tokyo vol. 19 (1943) pp. 148–151.

3. P. R. Halmos and L. J. Savage, *Application of the Radon-Nikodým theorem to the theory of sufficient statistics*, Ann. Math. Statist. vol. 20 (1949) p. 225.

4. P. R. Halmos, *Measure theory*, Van Nostrand, 1950.

5. A. Berger, *On uniformly consistent tests*, Ann. Math. Statist. vol. 22 (1950).

6. ———, *Remark on separable spaces of probability measures*, Ann. Math. Statist. vol. 22 (1951).

7. J. von Neumann, *Zur Algebra der Funktionaloperatoren und Theorie der normalen Operatoren*, Math. Ann. vol. 102 (1930) p. 370.

8. F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, 1914.

9. E. L. Lehmann and H. Scheffé, *Completeness, similar regions and unbiased estimation*, Sankhyā: The Indian Journal of Statistics vol. 10 (1950).

10. E. Hopf, *Ergodentheorie*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 5, no. 2, 1937, p. 4.

COLUMBIA UNIVERSITY