ON INVARIANT PROBABILITY MEASURES I

J. R. BLUM[1] AND D. L. HANSON[2]

1. Introduction. Let Ω be a set and let $\mathscr{A}$ be a σ-algebra of subsets of Ω. Let T be a one-to-one bimeasurable transformation mapping Ω onto itself. T then induces the group of transformations $\{T^i, i = 0, \pm 1, \cdots\}$ defined in the usual way. If $A \in \mathscr{A}$, T^iA is defined to be the set of images of the elements of A under the transformation T^i.

Let $\mathscr{P}$ be the class of probability measures defined on $\mathscr{A}$ for which T is invariant, i.e. if P is a probability measure defined on $\mathscr{A}$ then $P \in \mathscr{P}$ if and only if $PA = PTA$ for every $A \in \mathscr{A}$. Let $\mathscr{A}_1$ be the subclass of $\mathscr{A}$ which is invariant under T; a set $A \in \mathscr{A}$ belongs to $\mathscr{A}_1$ if and only if $A = TA$. It is trivial to verify that $\mathscr{A}_1$ is sub-σ-algebra of $\mathscr{A}$. Finally let $\mathscr{P}_1$ be the subclass of $\mathscr{P}$ for which T is ergodic, i.e. if $P \in \mathscr{P}$ then $P \in \mathscr{P}_1$ if and only if $PA = 0$ or $PA = 1$ for every $A \in \mathscr{A}_1$.

In § 2. several results are proved, concerning the structure of the class $\mathscr{P}$. These are not new, although several of them do not seem to have appeared in the literature. The main theorem of this paper is in § 3 where it is shown that each element of $\mathscr{P}$ can be represented as a convex combination of the extreme points of $\mathscr{P}$. Several consequences of this theorem are pointed out.

2. Some properties of the class $\mathscr{P}$.

THEOREM 1. *Let P and Q be elements of $\mathscr{P}$. Suppose $PA = QA$ for $A \in \mathscr{A}_1$ Then $P \equiv Q$.*

Proof. Let $\mu = P - Q$. Then μ is a completely additive set function defined on $\mathscr{A}$. If μ is not identically zero, there exists $A \in \mathscr{A}$ such $\mu(A) > 0$ and $\mu(A) \geqq \mu(B)$ for all $B \in \mathscr{A}$. This follows from the Hahn decomposition theorem. Write $\mu(A) = \alpha + \beta$, where $\alpha = \mu(A - A \cap TA)$ and $\beta = \mu(A \cap TA)$. Since $\mu(A - A \cap TA) = \mu(TA - A \cap TA)$ we have $\mu(A \cup TA) = 2\alpha + \beta$. Now if $\alpha < 0$, then $\mu(A \cap TA) > \mu(A)$ and A is not maximal, and if $\beta < 0$ then $\mu(A - A \cap TA) > \mu(A)$ and A is not maximal. Consequently $\alpha \geqq 0$ and $\beta \geqq 0$. But if A is maximal then $\alpha + \beta \geqq 2\alpha + \beta$. Hence $\alpha = 0$ and $\mu(A \cup TA) = \mu(A)$. By the same argument we show that $\mu(T^{-1}A \cup A \cup TA) = \mu(A)$ and it follows by in-

Received December 9, 1959.

[1] Supported in part by the Office of Ordnance Research, U. S. Army, under Contract No. DA-33-008-ORD-965.

[2] Work done while the author was a Standard Oil Co. Fellow at Indiana University.

duction that $\mu(B_n)=\mu(A)$ for every positive integer n, where $B_n = \bigcup_{i=-n}^{n} T^iA$. Now B_n is an increasing sequence of sets. Let $B = \lim_{n\to\infty} B_n$. Then $\mu(B)=\mu(A)>0$. But clearly $B = \bigcup_{i=-\infty}^{\infty} T^iA \in \mathscr{A}_1$ and μ is zero on $\mathscr{A}_1$. Consequently we have a contradiction and the theorem is proved.

Suppose now that $P \in \mathscr{P}_1$ and $Q \in \mathscr{P}$ and suppose also that Q is absolutely continuous with respect to P. Then if $A \in \mathscr{A}_1$ we have $PA = 0$ or $PA = 1$ and hence Q agrees with P on $\mathscr{A}_1$. Thus the theorem applies and we have

COROLLARY 1. *If $P \in \mathscr{P}_1, Q \in \mathscr{P}$, and Q is absolutely continuous with respect to P then $Q \equiv P$.*

Theorem 1 also furnishes an elegant proof of a result which was proved by Lamperti [3], and in a special situation by Harris [1]. Suppose P and Q are both ergodic, i.e. $P \in \mathscr{P}_1$ and $Q \in \mathscr{P}_1$. Then either P and Q are orthogonal or for each $A \in \mathscr{A}$ for which $PA = 1$ we have $Q(A) > 0$. Now suppose $A \in \mathscr{A}_1$ and $PA = 1$. Then if Q is not orthogonal to P and since $Q \in \mathscr{P}_1$ we must have $Q(A) = 1$ and it follows that $P = Q$ on $\mathscr{A}_1$. We have

COROLLARY 2. *If $P \in \mathscr{P}_1, Q \in \mathscr{P}_1$, then either $P \equiv Q$ or P is orthogonal to Q.*

In § 3, we shall show that this result can be considerably generalized.

THEOREM 2. *$\mathscr{P}$ is a convex set. $P \in \mathscr{P}_1$ if and only if P is an extreme point of $\mathscr{P}$.*

Proof. The first statement is obvious. Suppose $P \in \mathscr{P}_1$ and suppose we may represent P in the form $P \equiv \alpha P_1 + (1-\alpha)P_2$ where $0 < a < 1$ and $P_i \in \mathscr{P}, i = 1, 2$. Then clearly P_1 and P_2 are absolutely continuous with respect to P and it follows from Corollary 1 that $P_1 \equiv P_2 \equiv P$. Thus if $P \in \mathscr{P}_1$ it is an extreme point of $\mathscr{P}$. Conversely if $P \notin \mathscr{P}_1$ there exists a set $B \in \mathscr{A}_1$ with $0 < PB < 1$. Then we may write $P \equiv \alpha P_1 + (1-\alpha)P_2$ where $\alpha = PB$, and for $A \in \mathscr{A}$ we have $P_1(A) = P(A \cap B)/P(B)$ and $P_2(A) = P(A \cap B^c)/P(B^c)$. It is easily verified that P_1 and P_2 are invariant probability measures and it follows that P is not an extreme point of $\mathscr{P}$.

Theorem 2 strongly suggests that it may be possible to obtain the elements of $\mathscr{P}$ as convex combinations of the extreme points of $\mathscr{P}_1$. Under a rather mild assumption this is in fact true, as will be shown in the next section. Examples of the kind of theorem we have in mind were proved by Hewitt and Savage [2].

3. The representation theorem. Throughout part of this section we shall assume that if $A \in \mathscr{A}_1$ and if $PA = 0$ for every $P \in \mathscr{P}_1$ then

$PA = 0$ for every $P \in \mathscr{P}$. Clearly such a condition is necessary for a convex representation theorem and the condition can actually be verified in many examples of interest.

Suppose now that $P \in \mathscr{P}_1$. Theorem 1 tells us that P has a unique invariant extension from $\mathscr{A}_1$ to $\mathscr{A}$. This suggests that if $A \in \mathscr{A}$ we should be able to determine PA by knowing only the values of P on $\mathscr{A}_1$. A proof of this statement follows from the individual ergodic theorem.

THEOREM 3. *Let $A \in \mathscr{A}$. For every α with $0 \leqq \alpha \leqq 1$ there exists a set $A'_\alpha \in \mathscr{A}_1$ such that if $P \in \mathscr{P}_1$ then $PA = \alpha$ if and only if $PA'_\alpha = 1$.*

Proof. Let $f_s(x)$ be the set characteristic function of the set S. Let $A \in \mathscr{A}$, and α be given. For every positive integer n define $g_{n,A}(x) = 1/n \sum_{i=1}^{n-1} f_A(T^i x)$, and define $A'_\alpha = \{x \mid \lim_{n\to\infty} g_{n,A}(x) = \alpha\}$. Clearly $A'_\alpha \in \mathscr{A}_1$ and the individual ergodic theorem implies that $PA = \alpha$ if and only if $PA'_\alpha = 1$, whenever $P \in \mathscr{P}_1$.

Using the same technique we can prove

THEOREM 4. *Let $A \in \mathscr{A}$. For every α with $0 \leqq \alpha \leqq 1$ there exists a set $A_\alpha \in \mathscr{A}_1$ such that if $P \in \mathscr{P}_1$ then $PA \leqq \alpha$ if and only if $PA_\alpha = 1$.*

Let $A \in \mathscr{A}_1$. Define π_A by $\pi_A = \{P \in \mathscr{P}_1 \mid PA = 1\}$. Let Π be the collection of all such sets π_A i.e. $\Pi = \{\pi_A \mid A \in \mathscr{A}_1\}$. The following facts are easily verified:

(i) $\pi_\Omega = \mathscr{P}_1$

(ii) $[\pi_A]^c = \pi^c$

(iii) $\pi \bigcup_n A_n = \bigcup_n \pi A_n$

where A and each A_n is an element of $\mathscr{A}_1$. Since $\mathscr{A}_1$ is a σ-algebra it follows that Π is a σ-algebra. Now let $Q \in \mathscr{P}$. We define a set function μ_Q on Π by $\mu_Q(\pi_A) = Q(A)$.

We shall show that under the assumption at the beginning of this section μ_Q is in fact a probability measure defined on Π. Clearly $\mu_Q(\pi_A) \geqq 0$ for each π_A, and $\mu_Q(\mathscr{P}_1) = \mu_Q(\pi_\Omega) = Q(\Omega) = 1$. Now suppose $\{\pi_{A_n}\}$ is a sequence of disjoint elements of π. It is easily verified that this is the case if and only if $PA_n \cap A_m = 0$ for every pair of sets A_n, A_m in $\mathscr{A}_1$ with $n \neq m$ and for every $P \in \mathscr{P}_1$. It follows from the assumption that $Q(A_n \cap A_m) = 0$ for $n \neq m$. Hence $\mu_Q\{\bigcup_n \pi A_n\} = Q(\bigcup_n A_n) = \sum_n Q(A_n) = \sum_n \mu_Q\{\pi_{A_n}\}$ and we have shown that μ_Q is a probability measure defined on Π. We summarize in

THEOREM 5. *If Π and μ_Q are defined as above then Π is a σ-algebra of subsets of $\mathscr{P}_1$. Under the assumption at the beginning of this section μ_Q is a probability measure defined on Π.*

THEOREM 6. *Let $A \in \mathscr{A}$. Consider the function $f_A(P)$ defined on $\mathscr{P}_1$ and with values $f_A(P) = PA$. Then $f_A(P)$ is measurable with respect to Π.*

Proof. We must show that for every α with $0 \leqq \alpha \leqq 1$ we have $\{P \in \mathscr{P}_1 \mid f_A(P) \leqq \alpha\} = \{P \in \mathscr{P}_1 \mid PA \leqq \alpha\} \in \Pi$. But it follows from Theorem 4 that $\{P \in \mathscr{P}_1 \mid PA \leqq \alpha\} = \pi_{A-A_\alpha}$ where $A_\alpha \in \mathscr{A}_1$ is the set guaranteed by Theorem 4, and the theorem follows.

Since $f_A(P)$ is bounded and measurable it is clearly integrable with respect to any probability measure defined on Π. Now let $Q \in \mathscr{P}$ and μ_Q be the corresponding probability measure defined on Π. For each $A \in \mathscr{A}$ define $Q'(A)$ by

$$Q'(A) = \int_{\mathscr{P}_1} f_A(P) d\mu_Q = \int_{\mathscr{P}_1} PAd\mu_Q \,.$$

It follows immediately from this definition that Q' is an invariant probability measure defined on $\mathscr{A}$. But if $A \in \mathscr{A}_1$ we have $Q'(A) = \mu_Q\{\pi_A\} = Q(A)$. Hence $Q' = Q$ on $\mathscr{A}_1$ and it follows from Theorem 1 that $Q' \equiv Q$. Furthermore suppose we know that $Q(A) = \int_{\mathscr{P}_1} PAd\mu$, where μ is some probability measure defined on Π. Then if $A \in \mathscr{A}_1$ we have $Q(A) = \int_{\mathscr{P}_1} PAd\mu = \mu\{\pi_A\} = \mu_Q\{\pi_A\}$, i.e. $\mu \equiv \mu_Q$. We state these results in

THEOREM 7. *Suppose the assumption at the beginning of the section holds. Then for every $Q \in \mathscr{P}$ there exists a unique probability measure μ_Q defined on Π such that*

$$Q(A) = \int_{\mathscr{P}_i} P(A) d\mu_Q \text{ for every } A \in \mathscr{A} \,.$$

We shall refer to Theorem 7 as the representation theorem, and the rest of this section is devoted to exploring some consequences of this theorem. One immediate consequence is a generalization of Corollary 2 to Theorem 1.

THEOREM 8. *Let $Q_i \in \mathscr{P}$, $i = 1, 2$. Then Q_1 and Q_2 are orthogonal if and only if the corresponding measured μ_{Q_1} and μ_{Q_2} are orthogonal.*

Proof. Suppose Q_1 and Q_2 are orthogonal. Let B be a set such that $Q_1(B) = 1 = Q_2(B^c)$ and let $A = \bigcup_{i=-\infty}^{\infty} T^i B$. Then $A \in \mathscr{A}_1$ and $Q_1(A) = 1 = Q_2(A^c)$ and we obtain $1 = \mu_{Q_1}\{\pi_A\} = \mu_{Q_2}\{(\pi_A)^c\}$. Thus μ_{Q_1} and μ_{Q_2} are orthogonal. Conversely if μ_{Q_1} and μ_{Q_2} are orthogonal there is a set $A \in \mathscr{A}_1$ such that $1 = \mu_{Q_1}\{\pi_A\} = Q_1(A)$ and $0 = \mu_{Q_2}\{\pi_A\} = Q_2(A)$ and the theorem is proved.

Another interesting consequence of the theorem is the obvious fact that if $A \in \mathscr{A}$ and if $PA = 1$ for each $P \in \mathscr{P}_1$ then $Q(A) = 1$ for each $Q \in \mathscr{P}$. Thus the individual ergodic theorem for arbitrary invariant measures is an immediate consequence of that theorem for ergodic measures. Furthermore Theorem 7 throws some light on the evaluation of the limiting function in the individual ergodic theorem. Let $Q \in \mathscr{P}$ and let $f(x)$ be defined on Ω and measurable with respect to $\mathscr{A}$. Let $f_n(x) = 1/n \sum_{i=0}^{n-1} f(T^i x)$. Then if $f \in L_1(Q)$ the ergodic theorem states that $\lim_{n\to\infty} f_n(x) = f^*(x)$ say, exists on a set of Q-measure one. It is clear that f^* is invariant i.e. $f^*(Tx) = f^*(x)$ for all x for which f^* exists. If f is also integrable with respect to $P \in \mathscr{P}_1$ then f^* is constant on a set of P-measure one, and we have

$$Q\{x \mid f^*(x) \leqq u\} = \int_{\mathscr{P}_1} Px\{\mid f^*(x) \leqq u\} d\mu_Q = \mu_Q\{P \in \mathscr{P}_1 \mid f^* \leqq u\} ,$$

In particular we conclude f^* is a constant, say c, on a set of Q-measure one if and only if $\mu_P[P \in \mathscr{P}_1 \mid P\{x \mid f^*(x) = c\}] = 1$.

Finally, suppose f is again measurable with respect to $\mathscr{A}$. Let $Q \in \mathscr{P}$ and suppose $\mu_Q P\left\{ \in \mathscr{P}_1 \mid \int_\Omega |f| \, dP < \infty \right\} = 1$. Then we can easily prove

THEOREM 8. *If* $\int_\Omega |f| \, dP$ *is an integrable function of* P *(with respect to* μ_Q*) then* $f \in L_1(Q)$ *and*

$$\int_\Omega f dQ = \int_{\mathscr{P}_1} \left[\int_\Omega f dP \right] d\mu_Q .$$

REFERENCES

1. T. E. Harris, *On chains of infinite order*, Pacific J. Math., **5** (1955), 707–724.

2. E. Hewitt and L. J. Savage, *Symmetric measures on Cartesian products*, TAMS, **80** (1955), 470–501.

3. J. Lamperti, *Stationary measures for certain stochastic processes*, Pacific J. Math., **8** (1958), 127–132.

SANDIA CORPORATION AND
INTERNATIONAL BUSINESS MACHINES CORPORATION