ON EXTREME DOUBLY STOCHASTIC MEASURES

James R. Brown and Ray C. Shiflett

1. INTRODUCTION

This paper deals with the problem of extremality in the convex set of doubly stochastic measures defined on the unit square $X \times X$ with the Lebesgue structure. By m and m^2, we shall denote one- and two-dimensional Lebesgue measure. A doubly stochastic measure is a positive Borel measure μ defined on $X \times X$ such that

$$\mu(A \times X) = \mu(X \times A) = m(A)$$

for every measurable A.

The major results of this paper are called theorems, and the technical results are called propositions. We use the Douglas-Lindenstrauss Theorem:

Let μ *be a doubly stochastic measure. The measure* μ *is extreme if and only if the subspace consisting of all functions of the form* $h(x, y) = f(x) + g(y)$ $(f, g \in L_1(m))$ *is norm-dense in* $L_1(\mu)$.

R. G. Douglas [1] and Joram Lindenstrauss [4] discovered this theorem independently. It is the only known characterization of the extreme points of the set of doubly stochastic measures.

The authors would like to express their gratitude to the referee whose comments were very helpful.

2. RESULTS

Every example, known to the authors, of an extreme point in the set of doubly stochastic measures has the mass concentrated along line segments. In fact, there exist examples in which the mass seems to be concentrated on points.

Definition 1. The point (x, y) is a *point of density of the set* E, *relative to the measure* μ, if

$$\mu\,[([x - h, x + h] \times [y - h, y + h]) \cap E] > 0$$

for every $h > 0$.

The set of density points of E that are in E forms a closed subset of E in its relative topology. This collection is essentially the support of μ restricted to E. However, there may be points of density for E that are not in E. Proposition 1 and its corollaries are addressed to this larger collection of points.

Received April 26, 1969.

This work was supported by the National Science Foundation grants GP-6144 and GP-7474 in Applied Analysis. Part of this work comprised a portion of R. C. Shiflett's doctoral dissertation.

Michigan Math. J. 17 (1970).

Let $D(E, \mu)$ denote the set of all points of density of E relative to μ. Note that when $E \subset F$, we have the inclusion $D(E, \mu) \subset D(F, \mu)$. Let $X(E, \mu)$ denote the set of x such that (x, y) is in $E \cap D(E, \mu)$ for some y. Finally, let $Y(E, \mu)$ denote the set of y such that (x, y) is in $E \cap D(E, \mu)$ for some x.

PROPOSITION 1. *If* E *is measurable, then* $\mu(E) > 0$ *if and only if* $E \cap D(E, \mu) \neq \emptyset$.

Proof. If $\mu(E) = 0$, then $D(E, \mu) = \emptyset$. If $E \cap D(E, \mu) = \emptyset$, then for every (x, y) in E, there exists an $h(x, y) > 0$ with

$$\mu[([x - h, x + h] \times [y - h, y + h]) \cap E] = 0$$

for $h \geq h(x, y)$. Countably many such squares cover E. Using this countable collection, we have the inequalities

$$\mu(E) = \mu\Big[\bigcup([x - h, x + h] \times [y - h, y + h]) \cap E\Big]$$

$$\leq \sum \mu[([x - h, x + h] \times [y - h, y + h]) \cap E] = 0. \blacksquare$$

COROLLARY 1.1. *If* E *is measurable, then* $\mu[E - D(E, \mu)] = 0$; *that is, almost every point of* E *is a density point of* E.

Proof. Let $F = E - D(E, \mu)$. Then $F \subset E$ and $D(F, \mu) \subset D(E, \mu)$. Thus $F \cap D(F, \mu) = \emptyset$ and, by Proposition 1, $\mu(F) = 0$. ■

Definition 2. The *marginal measures* λ and ν of a measurable rectangle $A \times B$, determined by a doubly stochastic measure μ, are defined by the relations $\lambda(C) = \mu(C \times B)$, where C is a measurable subset of A, and $\nu(D) = \mu(A \times D)$, where D is a measurable subset of B. If $\mu(A \times B) > 0$, then $A \times B$ is called a *full rectangle*, relative to μ, if the marginal measures λ and ν are equivalent to m on A and B, respectively.

Note that $\lambda(C) \leq m(C)$ and $\nu(D) \leq m(D)$, for every C and D.

COROLLARY 1.2. *If* $A \times B$ *is full relative to* μ, *then*

$$m[A - X(A \times B, \mu)] = m[B - Y(A \times B, \mu)] = 0.$$

Proof. By Proposition 1, we have that

$$\lambda[A - X(A \times B, \mu)] = \mu([A - X(A \times B, \mu)] \times B) = 0.$$

Since λ is equivalent to m on A, we have that $m[A - X(A \times B, \mu)] = 0$. The other equality can be proved in the same way. ■

The idea of sequences of exceptional points has been used by Lindenstrauss [4] and Jaffa [3]. In this paper, a *path* is a sequence of points $\langle(x_i, y_i)\rangle$ for which $x_{2k} = x_{2k+1}$ and $y_{2k-1} = y_{2k}$. A path of length k exits in a set E if x_k is in E when k is even, or if y_k is in E when k is odd.

PROPOSITION 2. *Let* $\langle A_i \times B_i\rangle$ *be a sequence of full rectangles relative to the doubly stochastic measure* μ, *with* $A_{2k} = A_{2k+1}$ *and* $B_{2k-1} = B_{2k}$. *Let* $E_k = A_k$ *for even* k *or* $E_k = B_k$ *for odd* k. *If* E *is a set such that* $m(E_k - E) = 0$, *then for* $[m]$-*almost every* y_1 *in* B_1, *there exists an* x_1 *in* A_1 *such that* (x_1, y_1) *starts a* *path of length* k *with* (x_i, y_i) *in* $(A_i \times B_i) \cap D(A_i \times B_i, \mu)$, *and* [m]-*almost every point in* E_k *is an exit point for such a path.*

Proof. The proof is by induction. If $k = 1$, then the result follows from Corollary 1.2.

Suppose the result were true for $k = 2n - 1$. If $m(A_{2n} - E) = 0$, then $(A_{2n} \cap E) \times B_{2n}$ is full. Let E' be the collection of those y for which there exists an x such that (x, y) is in $(A_{2n} \cap E) \times B_{2n}$. By Corollary 1.2, we have that $m(B_{2n-1} - E') = 0$, since $B_{2n} = B_{2n-1}$. By the induction hypothesis, [m]-almost every element of B_1 connects to an element of E' by a path of length $2n - 1$, and therefore, by the definition of E', [m]-almost every element of B_1 starts a path of length 2n that exits in E.

Furthermore, the induction hypothesis states that [m]-almost every y_{2n} in $B_{2n} = B_{2n-1}$ is an exit point for a path of length $2n - 1$ that exits in E'. Therefore, [μ]-almost every point in the set of density points of $(A_{2n} \cap E) \times B_{2n}$ connects to a path of length $2n - 1$, forming a path of length 2n that exits in E. Hence [m]-almost all points of A_{2n} are exit points of paths of length 2n that exit in E.

A similar argument proves the case where $k = 2n$. ■

Definition 3. A sequence of rectangles $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$, each of which is full with respect to a doubly stochastic measure μ, is a *loop* if

$$B_1 = B_2,\ A_2 = A_3,\ B_3 = B_4,\ \cdots,\ B_{2n-1} = B_{2n}, \quad \text{and} \quad A_{2n} \subset A_1 .$$

We shall call a loop *invariant* if there exists a set $D_1 \subset B_1$ such that $m(D_1) > 0$ and such that for [m]-almost every y_1 in D_1, the path given by Proposition 2 satisfies the condition $x_1 = x_{2n}$.

THEOREM 1. *If* μ *is an extreme doubly stochastic measure, then* μ *has no invariant loops.*

Proof. Let $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$ be an invariant loop for the doubly stochastic measure μ. Let $D_1 \subset B_1$ be the set described above. Then, for [μ]-almost all (x_1, y_1) in the set of density points of $A_1 \times D_1$, there exists a path ending in the set of density points of $A_{2n} \times B_{2n}$.

If μ were extreme, the Douglas-Lindenstrauss Theorem states that we could approximate the characteristic function of $A_1 \times B_1$ in the $L_1(\mu)$-norm by functions $f(x) + g(y)$, where f and g are in $L_1(m)$. In this case, we could find a sequence $\langle f_n(x) + g_n(y) \rangle$ that converged in the mean to the characteristic function of $A_1 \times B_1$, and, consequently, we could find a subsequence that converged [μ]-almost everywhere. Therefore, [μ]-almost all points in the paths that start with a y_1 in D_1 are points of convergence of this subsequence. Choose one such path $(x_1, y_1), \cdots, (x_1, y_{2n})$. For this finite set of points, there exists, for each $\varepsilon > 0$, an N such that $f_m(x_i) + g_m(y_i)$ $(m > N)$ differs from the characteristic function of $A_1 \times B_1$ in absolute value by less than ε. For fixed M $(M > N)$, we have the inequalities

$$|1 - f_M(x_1) - g_M(y_1)| < \varepsilon, \quad |f_M(x_2) + g_M(y_1)| < \varepsilon,$$

$$|f_M(x_2) + g_M(y_3)| < \varepsilon, \quad \cdots, \quad |f_M(x_1) + g_M(y_{2n})| < \varepsilon .$$

Combining these inequalities, we obtain the relation $1 < 2n\varepsilon$ for arbitrarily small ε and fixed n. Thus the measure μ can not be extreme. ■

THEOREM 2. *If* μ *is an extreme doubly stochastic measure, then for every* $A \times B$ *with* $m(A)m(B) > 0$ *there exists a set* $(C \times D) \subset (A \times B)$ *with* $m(C)m(D) > 0$ *and* $\mu(C \times D) = 0$.

Proof. Suppose there are no such subsets in $A \times B$. Let $(C \times D) \subset (A \times B)$ be such that $m(A - C)m(B - D) > 0$. Form $C \times D$, $(A - C) \times D$, $(A - C) \times (B - D)$, and $C \times (B - D)$. Let λ and ν be the marginal measure of $C \times D$. Then $\lambda \leq m$ and $\nu \leq m$. If $\lambda(E) = 0$ $(E \subset C)$, then $\mu(E \times D) = 0$; thus $m(E)m(D) = 0$ by our assumption. Therefore $m \ll \lambda$. Similarly, $m \ll \nu$, and thus $C \times D$ is full. The same argument shows that $(A - C) \times D$, $(A - C) \times (B - D)$, and $C \times (B - D)$ are full.

By Proposition 2, for [m]-almost all y_1 in D, there exists a path (x_1, y_1), (x_2, y_1), (x_2, y_3), and (x_4, y_3). Since (x_1, y_1) is a density point of $C \times D$, we have that $m([x_1 - h, x_1 + h] \cap C) > 0$. Similarly, $m([y_3 - h, y_3 + h] \cap (B - D)) > 0$. Thus it follows from the assumption that

$$\mu[([x_1 - h, x_1 + h] \times [y_3 - h, y_3 + h]) \cap (C \times [B - D])] > 0.$$

Hence, for [m]-almost all y_1 in D, there is a path (x_1, y_1), (x_2, y_1), (x_2, y_3), and (x_1, y_3). By Theorem 1, we see that μ is not extreme. ■

The proof of the following corollary is now trivial.

COROLLARY 2.1. *If* μ *is an extreme doubly stochastic measure, then every nonempty open set* U *in* $X \times X$ *contains a rectangle* $E \times F$ *for which* $m(E)m(F) > 0$ *and* $\mu(E \times F) = 0$.

We need the next two technical results to obtain the final two major results. We use a concept similar to that of a loop.

Definition 4. A *near loop* is a sequence $A_1 \times B_1, \cdots, A_{2n} \times B_{2n}$ of full rectangles, with respect to some measure μ, such that $B_1 \subset B_2$, $A_2 \subset A_3$, $B_3 \subset B_4, \cdots, B_{2n-2} \subset B_{2n-1}$, and $m(A_{2n} \cap A_1) > 0$; moreover, for every other combination of i and j, we have the relation $m(A_i \cap A_j) = m(B_i \cap B_j) = 0$.

PROPOSITION 3. *If* $\mu(A \times B) > 0$, *then there exists a full rectangle* $C \times D$ *contained in* $A \times B$ *with* $\mu(C \times D) = \mu(A \times B)$.

Proof. Suppose λ and ν are the marginal measures of $A \times B$. Let C be the set of x $(x \in A)$ for which $d\lambda/dm$ is positive, and let D be the set of y $(y \in B)$ for which $d\nu/dm$ is positive. Then

$$\mu[(A - C) \times B] = \lambda(A - C) = \int_{A-C} (d\lambda/dm)dm = 0.$$

Similarly,

$$\mu[(A - C) \times (B - D)] = \mu[A \times (B - D)] = 0.$$

Thus $\mu(C \times D) = \mu(A \times B)$. The preceding argument also shows that λ and ν are equal to the marginal measures of $C \times D$ on C and D, respectively. Thus $C \times D$ is full. ■

PROPOSITION 4. *If* μ *has no near loops, then for each full rectangle* $A \times B$ *and each* $\varepsilon > 0$, *there exist two functions* f *and* g *in* $L_1(m)$ *such that*

$$\int_{X \times X} |\chi_{A \times B}(x, y) - f(x) - g(y)| \, d\mu < \varepsilon$$

and such that $f(x) = 0$ *for each* x *in* A.

Proof. Assume that μ has no near loops. Let $A_1 \times B_1$ be a full rectangle. Let λ_2 and ν_2 be the marginal measures of $(X - A_1) \times B_1$. If A_2 is the set of x in $X - A_1$ for which $d\lambda_2/dm$ is positive and B_2 is the set of y in B_1 for which $d\nu_2/dm$ is positive, then, as was shown in the proof of Proposition 3, $A_2 \times B_2$ is full, provided $\mu[(X - A_1) \times B_1] > 0$. If the measure of this set is zero, we take $f(x) = 0$ and g as the characteristic function of B_1. The conclusion follows.

In the same manner, we construct the full rectangle of $A_2 \times (X - B_2)$. We call this full rectangle $A_3 \times B_3$. In this manner, we construct a sequence of full rectangles such that

$$B_{2k} \subset B_{2k-1} \quad \text{and} \quad A_{2k+1} \subset A_{2k} \qquad (k = 1, 2, \cdots).$$

Since there are no near loops, we have for every other combination of i and j $(i \neq j)$ that $m(B_i \cap B_j) = m(A_i \cap A_j) = 0$.

Let $A_0 =$ A_i $(i = 2, 3, \cdots)$ and $B_0 = \bigcup B_i$ $(i = 1, 2, \cdots)$. Then

$$\mu[(X - A_0 - A_1) \times B_0] = \mu\left(\left[\left(X - \bigcup A_{2i}\right) \cap (X - A_1)\right] \times \bigcup B_{2i-1}\right)$$

$$\leq \sum \mu\left[\left([X - A_1] \cap \left[\bigcup X - A_{2i}\right]\right) \times B_{2i-1}\right]$$

$$\leq \mu[(X - A_1 - A_2) \times B_1] + \mu[(X - A_1 - A_2 - A_4) \times B_3] + \cdots = 0 \qquad (i = 1, 2, \cdots).$$

Similarly, $\mu[A_0 \times (X - B_0)] = 0$, and $\mu(A_1 \times B_j) = 0$ for $j > 2$. Now let $-f$ be the characteristic function of A_0, and let g be the characteristic function of B_0. Then $f(x) + g(y)$ equals the characteristic function of $A_1 \times B_1$ $[\mu]$-almost everywhere. ■

THEOREM 3. *If the doubly stochastic measure* μ *has no near loops, then* μ *is an extreme point.*

Proof. Let μ have no near loops, and suppose $\mu(A \times B) > 0$. By Proposition 3, there exists a full rectangle $A_1 \times B_1$ in $A \times B$ such that $\mu(A_1 \times B_1) = \mu(A \times B)$. Thus the characteristic functions of $A \times B$ and $A_1 \times B_1$ are equal $[\mu]$-almost everywhere. In the proof of Proposition 4, two sets, A_0 and B_0, were found, for which $-f$ was the characteristic function of A_0 and g was the characteristic function of B_0, and thus $f(x) + g(y)$ was equal to the characteristic function of $A_1 \times B_1$ $[\mu]$-almost everywhere. Hence, every simple function $\Phi(x, y)$ over the algebra of finite unions of measurable rectangles is of the form $\Psi(x) + \Gamma(y)$, where Ψ and Γ are simple functions. By the Douglas-Lindenstrauss Theorem, μ is extreme. ■

J. Feldman conjectured that if μ_1 and μ_2 are doubly stochastic measures with $\mu_1 \ll \mu_2$ and if μ_2 is extreme, then $\mu_1 = \mu_2$. Among those who have considered this problem is R. G. Douglas [2] who proved the following result:

If F is a vector-lattice that is weak*-dense in $L_\infty(\mu)$, if μ is extreme in the set of doubly stochastic measures, and if ν is a doubly stochastic measure with $\nu \ll \mu$ such that

$$\int_X f\,d\nu = \int_X f\,d\mu$$

for all bounded f in F, then $\nu = \mu$.

The next theorem proves that Feldman's conjecture holds for a class of extreme doubly stochastic measures that contains every example of an extreme doubly stochastic measure known to the authors.

THEOREM 4. *If* μ_1 *and* μ_2 *are doubly stochastic measures such that* μ_2 *has no near loops and* $\mu_1 \ll \mu_2$, *then* $\mu_1 = \mu_2$.

Proof. We form the doubly stochastic measure $t\mu_1 + (1 - t)\mu_2$, which is absolutely continuous with respect to μ_2 for each t in [0, 1]. It is simple to prove that if μ_2 has no near loops, then every measure absolutely continuous with respect to μ_2 has no near loops. Thus, $t\mu_1 + (1 - t)\mu_2$ has no near loops and is, by Theorem 3, extreme for every t in [0, 1]. Thus $\mu_1 = \mu_2$. ∎

REFERENCES

1. R. G. Douglas, *On extremal measures and subspace density*. Michigan Math. J. 11 (1964), 243-246.

2. ———, *On extremal measures and subspace density*. II. Proc. Amer. Math. Soc. 17 (1966), 1363-1365.

3. R. E. Jaffa, *Extremal doubly stochastic operators and related matrices*. Dissertation. University of California, Berkeley, 1964, p. 53.

4. J. Lindenstrauss, *A remark on extreme doubly stochastic measures*. Amer. Math. Monthly 72 (1965), 379-382.

Oregon State University
Corvallis, Oregon 97331
and
Wells College
Aurora - on - Cayuga, New York 13026