PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 5, Fasc. 2 (1985), p. 225-233

UNIQUENESS AND EXTREMALITY FOR A CLASS OF MULTIPLY-STOCHASTIC MEASURES

BY

RAE MICHAEL SHORTT, (HOUGHTON, MICHIGAN)

0. Introduction. Since the fundamental papers of Douglas [3] and Lindenstrauss [8], much has been written concerning the convex set of doubly-stochastic measures (probability measures on the unit square with uniform marginals), most of it an attempt to obtain a clearer understanding of this set's extreme points; the problem is hard, and success has been limited. We present a three-dimensional analogue for the class of doubly-stochastic measures related to some earlier work on property (V) in Shortt [12]. There the concern was with the existence of probabilities on a product $X \times Y \times Z$ with prescribed marginals on $X \times Y$ and $Y \times Z$, and a rather detailed measure-theoretic description of those separable spaces X, Y, Z for which such a construction is possible was obtained. The question of uniqueness was not adressed, nor that of determining the extreme points of the solution set; this is the substance of theorems 1 and 2 infra.

Note that if Y reduces to a single point, then the problem becomes doubly-stochastic; the point of theorem 2 is to analyse three-dimensional extreme points in terms of two-dimensional extreme "slices".

Before treating these results, we offer a preliminary section touching on some basic facts about separable spaces and reviewing the notion of a Markov transition kernel. This is a natural and important tool, but its use requires care; in Shiflett [11], for example, the measurability of $P(x, B(x))$ in the proof of Theorem 3 is by no means assured. The argument may be completed by an appeal to the Blackwell and Ryll-Nardzewski [1] selection theorem, much as in the proof of theorem 2, claim 3 below. In what follows, we have been to some degree fastidious concerning measurability questions.

1. Preliminaries. We deal exclusively with *separable* spaces, i.e. measurable spaces $(X, \mathscr{B})$ for which:

(a) there is a metric d on X under which (X, d) becomes a separable metric space, and

(b) $\mathscr{B}$ is the σ-algebra of Borel sets generated by the topology of (X, d). Then d is said to be a *metric* for $(X, \mathscr{B}(X))$.

By a result of Marczewski [9], a measurable space $(X, \mathscr{B})$ is separable if and only if $\mathscr{B}$ is countably generated and contains all singletons drawn from X. We shall often suppress the notation of a σ-algebra, calling the space X and, where necessary, indicating its measurable structure with $\mathscr{B} = \mathscr{B}(X)$. If A is a subset of a separable space X, then A itself becomes a separable space with the relativised σ-algebra $\mathscr{B}(A) = \{A \cap B : B \in \mathscr{B}(X)\}$.

A separable space X is *standard* if there is a metric d for X under which (X, d) becomes a *Polish space* (i.e. a topologically complete and separable metric space). It is well-known that if X is an uncountable standard space, then X is Borel-isomorphic with the unit interval $I = [0, 1]$ under its usual Borel structure. (This equivalence allows one, in questions of pure measure theory, to replace any uncountable standard space with the interval I.) A subset A of a standard space X is itself standard if and only if A is a member of $\mathscr{B}(X)$.

For proofs of these facts, as well as additional information on standard spaces, we refer the reader to Cohn [2], Chapter 8; for the classical treatment of these topics, see Kuratowski [6], Volume I, Chapter 3.

If X is a separable space, let $\mathscr{M}(X)$ denote the linear space of all finite signed measures on $(X, \mathscr{B}(X))$. If d is a metric for X, then $\mathscr{M}(X)$ may be given the topology of weak convergence of measures; this topology is separable and metrisable, so that under the corresponding Borel structure, $\mathscr{M}(X)$ becomes a separable space. As is shown in Varadarajan [13], p. 193, this separable Borel structure on $\mathscr{M}(X)$ is generated by all mappings of the form $\mu \to \mu(B)$ as B ranges over $\mathscr{B}(X)$; thus the Borel structure on $\mathscr{M}(X)$ does not depend on the choice of d. The probability measures in $\mathscr{M}(X)$ will be called *laws*, and $\mathscr{P}(X)$ will denote the collection of all laws on X.

Let X and Y be separable spaces; following Neveu [10], Chapter II.2, we call a real function $P(x, B)$ on $X \times \mathscr{B}(Y)$ a *Markov (transition) kernel* over the space X if

(1) $P(x, \cdot)$ is a law on Y for each $x \in X$, and

(2) $P(x, B)$ is measurable in x for each $B \in \mathscr{B}(Y)$.

Condition 2 is equivalent to:

(2') the mapping $x \to P(x, \cdot)$ is measurable from X to $\mathscr{P}(Y)$ under the separable structure it inherits from $\mathscr{M}(Y)$.

The following two lemmas are essentially a restatement of Neveu [10] Proposition III.2.1:

LEMMA 1. *Suppose that $P(x, B)$ is a Markov kernel over X. If P_1 is any law on X, then the set function*

$$P(A \times B) = \int_A P(x, B)\, dP_1(x),$$

defined for all rectangles $A \times B$ with $A \in \mathscr{B}(X)$, $B \in \mathscr{B}(Y)$, extends uniquely to a law P on $X \times Y$ with marginal P_1 on X.

Remark. In this situation, we call $P(x, \cdot)$ a Markov kernel for P over (X, P_1), or simply over X.

LEMMA 2. *If $P(x, \cdot)$ is a Markov kernel for a law P over X, then*

(a) *the mapping $x \to P(x, B_x)$ is measurable on X for each $B \in \mathscr{B}(X \times Y)$; here B_x denotes the section of B over $x \in X$, and*

(b) *$P\big((A \times Y) \cap B\big) = \int_A P(x, B_x)\, dP_1(x)$ for all $A \in \mathscr{B}(X)$, $B \in \mathscr{B}(X \times Y)$, where P_1 is the marginal of P on X.*

If $P(x, B)$ is a Markov kernel for P over (X, P_1), then the function $Q\big((x, y), B\big) = P(x, B_x)$ is a proper conditional distribution for P given projection onto X, as defined in Blackwell and Ryll-Nardzewski [1]. Conversely, if $Q\big((x, y), B\big)$ is such a proper conditional distribution, its values do not depend on Y and $P(x, B) = Q\big((x, y), X \times B\big)$ defines a Markov kernel for P. Essentially, the two notions coincide here, and theorems regarding the existence of one may be used to prove the existence of the other.

LEMMA 3. *Let P be a law on $X \times Y$ such that there is a standard subset $S \subset Y$ with $P(X \times S) = 1$. Then there is a Markov kernel for P over X.*

Proof. Easily adapted from Blackwell and Ryll-Nardzewski [1] Theorem 1 or Shortt [12] Theorem 4.

LEMMA 4. *Let P be a law on $X \times Y$ with marginal P_1 on X and suppose that $P(x, B)$ and $Q(x, B)$ are Markov kernels for P over X; then $P(x, \cdot) = Q(x, \cdot)$ as laws on Y a.s. (P_1).*

Remark. From condition 2′ in the definition of Markov kernel, it follows that $\{x\colon P(x, \cdot) = Q(x, \cdot)\}$ belongs to $\mathscr{B}(X)$.

Proof. Let $\mathscr{A}(Y)$ be a countable algebra generating $\mathscr{B}(Y)$. For each $B \in \mathscr{A}(Y)$, $P(x, B) = Q(x, B)$ for P_1-almost all x: if, say, $P(x, B) > Q(x, B)$ for all $x \in A$ with $P_1(A) > 0$, then

$$P(A \times B) = \int_A P(x, B)\, dP_1(x) > \int_A Q(x, B)\, dP_1(x) = P(A \times B),$$

a contradiction. From the countablility of $\mathscr{A}(Y)$ it follows that $P(x, B) = Q(x, B)$ for all $B \in \mathscr{A}(Y)$ a.s. (P_1). Since $\mathscr{A}(Y)$ generates $\mathscr{B}(Y)$, the lemma is obtained, q.e.d.

LEMMA 5. *Let $P(x, B)$ be a real function on $X \times \mathscr{B}(Y)$ such that*

(a) *$P(x, \cdot)$ is a law on Y for each $x \in X$, and*

(b) *$P(x, A)$ is measurable in x for all $A \in \mathscr{A}(Y)$, where $\mathscr{A}(Y)$ is some algebra generating $\mathscr{B}(Y)$.*

Then $P(x, B)$ is a Markov kernel over X.

Proof. Inspecting the collection of those $B \in \mathscr{B}(Y)$ for which $P(\cdot, B)$ is measurable, we find it to be closed under increasing unions and decreasing intersections. The usual monotone class argument finishes the proof, q.e.d.

2. The Principal Results. Let X, Y, Z be separable spaces with X and Z uncountable and standard. As noted in Section 1, metrics may be chosen for X and Z under which both become homeomorphic with the unit interval $I = [0, 1]$; without hesitation, X and Z are written interchangeably with I.

Suppose now that P and Q are laws on $X \times Y$ and $Y \times Z$, respectively, sharing a common marginal P_1 on Y. Our primary object of study will be $V(P, Q)$, the set of all laws on $X \times Y \times Z$ having marginals P on $X \times Y$ and Q on $Y \times Z$. $V(P, Q)$ is clearly a convex set, and Shortt [12], Theorem 5 implies that it is non-empty. The two questions before us are:

1. What are the extreme points of $V(P, Q)$?
2. When is $V(P, Q)$ a singleton set?

The answers to these are the content of theorems 1 and 2 below.

Before proceeding, however, it might be worthwhile to consider the case where X, Y, Z are finite sets; the marginal problem then becomes one of finding a 3-dimensional array $a(i, j, k)$ of non-negative numbers subject to the constraint that

$$\sum_i a(i, j, k) \quad \text{and} \quad \sum_k a(i, j, k)$$

are given matrices. It is not too hard to prove that a solution is extreme if and only if for each fixed j, $a(i, j, k)$ is an extreme matrix with specified row and column sums

$$\sum_i a(i, j, k) \quad \text{and} \quad \sum_k a(i, j, k).$$

Also, the solution is unique precisely when for each fixed j, either $\sum_i a(i, j, k) > 0$ for at most one k, or $\sum_i a(i, j, k) > 0$ for at most one i.

The conditions we obtain will in fact be "continuous" versions of these.

If M_1 and M_2 are laws on X and Z, respectively, then by $S(M_1, M_2)$ is denoted the set of all laws on $X \times Z$ with marginals M_1 on X and M_2 on Z. $S(M_1, M_2)$ is convex and, for the topologies we have chosen for X and Z, compact under weak convergence (it is a closed subset of $\mathscr{P}(X \times Z) = \mathscr{P}(I \times I)$, which is compact).

By lemma 3, there are Markov kernels $P(y, \cdot)$ and $Q(y, \cdot)$ for P and Q over (Y, P_1). We assume that these kernels are fixed for the extent of this section. Again by lemma 3, any law R in $V(P, Q)$ also has a Markov kernel $R(y, \cdot)$ over (Y, P_1). If $R(y, \cdot) \in S(P(y, \cdot), Q(y, \cdot))$ for each $y \in Y$, then call $R(y, \cdot)$ a *canonical* kernel for R. It follows easily from lemma 4 that any Markov kernel for R may be altered on a set of $y \in N$ with $P_1(N) = 0$ to obtain a canonical kernel.

Let $R(y, \cdot)$ be a canonical kernel for a law R in $V(P, Q)$; define $E \subset Y$ by

$$E = \{y \in Y\colon\ R(y, \cdot) \text{ is an extreme point of } S(P(y, \cdot), Q(y, \cdot))\}.$$

LEMMA 6. *E is a Borel subset of Y.*

Proof. For each μ in $K = \mathscr{P}(X \times Z)$, denote by μ_x and μ_z the marginals of μ on X and on Z. Define

$$K_0 = \{\mu \in K\colon \mu \text{ is not an extreme point of } S(\mu_x, \mu_z)\},$$

$$M = \{(\mu_1, \mu_2) \in K \times K\colon \mu_1 \text{ and } \mu_2 \text{ have the same marginals on } X \text{ and } Z\},$$

$$M_0 = \{(\mu_1, \mu_2) \in M\colon \mu_1 \neq \mu_2\}.$$

Then M is compact, and M_0 is open in M, and therefore σ-compact. Define $G\colon M \to K$ by $G(\mu_1, \mu_2) = \frac{1}{2}(\mu_1 + \mu_2)$. Then K_0 is the image of M_0 under G; since G is continuous, K_0 is σ-compact.

Finally, E is the inverse image of $K \backslash K_0$ under the mapping $y \to R(y, \cdot)$ and so is measurable in Y, q.e.d.

LEMMA 7. *Let K_0 be as in the preceding proof; then there is a measurable function $f\colon K_0 \to \mathscr{M}(X \times Z)$ such that for all $\mu \in K_0$*

(1) $\mu + f(\mu) \in K$,

(2) $\mu - f(\mu) \in K$,

(3) $f(\mu) \neq 0$,

(4) $f(\mu)_x = f(\mu)_z = 0$.

Remark. Compare Karlin and Studden, [5] p. 270-274.

Proof. Define the sets

$$L = \left\{\frac{1}{2}(\mu - \tilde{\mu})\colon \mu, \tilde{\mu} \in K\right\} \quad \text{and} \quad L_0 = L \backslash \{0\}.$$

Then L is a continuous image of K and is therefore compact, whilst L_0 is open in L and so is σ-compact. Also define

$$A = \{(\mu, \tilde{\mu}) \in K_0 \times L_0\colon (\mu + \tilde{\mu}, \mu - \tilde{\mu}) \in K \times K \text{ and } \tilde{\mu}_x = \tilde{\mu}_z = 0\}.$$

Then A is relatively closed in $K_0 \times L_0$ and hence is σ-compact. If $\mu \in K_0$, then $\mu = \frac{1}{2}(\mu_1 + \mu_2)$ with $\tilde{\mu} = \frac{1}{2}(\mu_1 - \mu_2) \in L_0$ and $(\mu, \tilde{\mu}) \in A$; thus each section of A over points $\mu \in K_0$ is non-empty.

Hoffmann-Jørgensen [4], p. 137 Theorem 3, implies the existence of a measurable function $f\colon K_0 \to L_0$ whose graph is contained in A, q.e.d.

LEMMA 8. *Let K and K_0 be as above; then there are measurable functions $F_1\colon K_0 \to K$ and $F_2\colon K_0 \to K$ such that for each $\mu \in K_0$*

(1) $F_1(\mu)_x = F_2(\mu)_x = \mu_x$ *and* $F_1(\mu)_z = F_2(\mu)_z = \mu_z$,

(2) $(F_1(\mu) + F_2(\mu))/2 = \mu$,

(3) $F_1(\mu) \neq F_2(\mu)$.

Proof. Let $f\colon K_0 \to \mathscr{M}(X \times Z)$ be as in lemma 7 and set $F_1(\mu) = \mu + f(\mu)$ and $F_2(\mu) = \mu - f(\mu)$.

THEOREM 1. *A law R in $V(P, Q)$ is an extreme point of $V(P, Q)$ if and only if for each canonical kernel $R(y, \cdot)$ for R over (Y, P_1), $R(y, \cdot)$ is an extreme point of $S(P(y, \cdot), Q(y, \cdot))$* a.s. (P_1), i.e. $P_1(E) = 1$, *where E is the subset of Y in lemma 6.*

Proof. Suppose that R is not extreme; then $R = \alpha R_1 + (1-\alpha) R_2$ for distinct R_1, R_2 in $V(P, Q)$, $0 < \alpha < 1$. Let $R_1(y, \cdot)$ and $R_2(y, \cdot)$ be canonical kernels for R_1 and R_2; then $R(y, \cdot) = \alpha R_1(y, \cdot) + (1-\alpha) R_2(y, \cdot)$ defines a canonical kernel for R. Since $R_1 \neq R_2$, the laws $R_1(y, \cdot)$ and $R_2(y, \cdot)$ differ for all y in some set of positive P_1-measure; $R(y, \cdot)$ is not extreme in $S(P(y, \cdot), Q(y, \cdot))$ for all such y.

Suppose, conversely, that there is a canonical kernel $R(y, \cdot)$ for R for which $P_1(E) < 1$. Define K_0 and K as before, and let F_1: $K_0 \to K$ and F_2: $K_0 \to K$ be as in lemma 8. Define R_1 and R_2 on $Y \times \mathscr{B}(X \times Z)$ by

$$R_1(y, B) = \begin{cases} F_1(R(y, \cdot))(B) & y \in Y \backslash E, \\ R(y, B) & y \in E; \end{cases}$$

$$R_2(y, B) = \begin{cases} F_2(R(y, \cdot))(B) & y \in Y \backslash E, \\ R(y, B) & y \in E. \end{cases}$$

Then $R_1(y, \cdot)$ and $R_2(y, \cdot)$ are canonical kernels for laws R_1 and R_2 over (Y, P_1); R_1 and R_2 are in $V(P, Q)$, and $\frac{1}{2}(R_1 + R_2) = R$. Now for each $y \notin E$, $R(y, \cdot)$ is in K_0, so that $R_1(y, \cdot) \neq R_2(y, \cdot)$ for such y. If $R_1 = R_2$, then by lemma 4, $R_1(y, \cdot) = R_2(y, \cdot)$ a.s. (P_1), a contradiction. Thus $R_1 \neq R_2$, and R is not extreme in $V(P, Q)$, q.e.d.

We maintain the notations used in theorem 1, recalling that the Markov kernels $P(y, \cdot)$ and $Q(y, \cdot)$ were fixed. Define the sets

$$D(P) = \{y\colon\ P(y, \cdot) \text{ is not a point mass}\}$$

and

$$D(Q) = \{y\colon\ Q(y, \cdot) \text{ is not a point mass}\}.$$

These are members of $\mathscr{B}(Y)$: for example, $D(P)$ is the inverse image of $\{\mu \in \mathscr{P}(X)\colon\ \mu \text{ is not a point mass}\}$ under the mapping $y \mapsto P(y, \cdot)$; this last set is open for the topology of weak convergence in $\mathscr{P}(X)$, so that $D(P)$ *is measurable.*

THEOREM 2. *The following statements are equivalent*:

1. $V(P, Q)$ *is a singleton set*;
2. $P_1(D(P) \cap D(Q)) = 0$.

Proof. *Statement 1 implies statement 2.* We prove the contrapositive, supposing $P_1(D(P) \cap D(Q)) > 0$. Recall that X and Z are metrised as the interval I and written interchangeably with it. We define the Markov kernels $R_1(y, \cdot)$ and $R_2(y, \cdot)$ on $Y \times \mathscr{B}(X \times Z)$ as follows:

$R_1(y, \cdot)$ and $R_2(y, \cdot)$ are laws on $X \times Z$ with cumulative distribution functions

$$R_1(y, [0, x] \times [0, z]) = P(y, [0, x]) \cdot Q(y, [0, z]),$$

$$R_2(y, [0, x] \times [0, z]) = P(y, [0, x]) \wedge Q(y, [0, z]),$$

where $a \wedge b = \text{minimum } \{a, b\}$. Thus the mappings $x \to R_1(y, A)$ and $y \to R_2(y, A)$ are measurable for each $A \in \mathscr{A}(X \times Z)$, the algebra generated by all sets of the form $[0, x] \times [0, y]$; since this algebra generates $\mathscr{B}(X \times Z)$, lemma 5 implies that $R_1(y, \cdot)$ and $R_2(y, \cdot)$ are in fact Markov kernels for laws R_1 and R_2 on $X \times Y \times Z$ over (Y, P_1).

Claim 1. Both R_1 and R_2 have marginals P on $X \times Y$ and Q on $Y \times Z$, i.e. are members of $V(P, Q)$. (This is easily checked.)

We shall now endeavour to show that $R_1 \neq R_2$, so that statement 1 does not hold.

Define the functions g_1: $X \times Y \to I$ and g_2: $Y \times Z \to I$ by $g_1(t, y) = P(y, [0, t])$ and $g_2(y, t) = Q(y, [0, t])$.

Claim 2. The functions g_1 and g_2 are jointly Borel-measurable: this follows e.g. from Lange [7], Lemma 3.1 and the right-continuity of $g_1(t, y)$ and $g_2(y, t)$ in t for each fixed y.

Define the Borel sets

$$S = \{(t, y) \in X \times Y\colon\ 0 < g_1(t, y) < 1\}$$

and

$$T = \{(y, t) \in Y \times Z\colon\ 0 < g_2(y, t) < 1\},$$

noting that $D(P)$ and $D(Q)$ are the projections to Y of S and T, respectively.

Claim 3. There are measurable functions h: $D(P) \to X$ and k: $D(Q) \to Z$ whose graphs lie entirely within S and T.

This follows from Blackwell and Ryll-Nardzewski [1], Theorem 2 in this way: the sections of S and T over a point $y \in Y$ are given by

$$S_y = \{t \in X\colon\ 0 < g_1(t, y) < 1\}, \qquad T_y = \{t \in Z\colon\ 0 < g_2(y, t) < 1\},$$

so that

$$P(y, S_y) > 0 \text{ whenever } y \in D(P),$$

and

$$Q(y, T_y) > 0 \text{ whenever } y \in D(Q).$$

The claim follows.

We extend the functions h and k (measurably) to all of Y by setting

$$h(y) = 1 \text{ for } y \in Y \setminus D(P) \quad \text{and} \quad k(y) = 1 \text{ for } y \in Y \setminus D(Q).$$

Define the Borel subsets $H \subset X \times Y$ and $K \subset Y \times Z$ by

$$H = \{(t, y)\colon\ t \leqslant h(y)\} \quad \text{and} \quad K = \{(y, t)\colon\ t \leqslant k(y)\}.$$

Then the sections of H and K over a point $y \in Y$ are simply $H_y = [0, h(y)]$ and $K_y = [0, k(y)]$, so that

$$(*) \qquad \begin{cases} 0 < P(y, H_y) < 1 & \text{whenever } y \in D(P), \\ 0 < Q(y, K_y) < 1 & \text{whenever } y \in D(Q). \end{cases}$$

Define

$$B=(H\times Z)\cap(X\times K) \quad \text{and} \quad C=\big(X\times(D(P)\cap D(Q))\times Z\big)\cap B$$

in $\mathscr{B}(X\times Y\times Z)$; then according to lemma 2,

$$\begin{aligned} R_1(C) &= \int_{D(P)\cap D(Q)} R_1(y, B_y)\,dP_1(y) \\ &= \int_{D(P)\cap D(Q)} R_1(y, [0, h(y)]\times[0, k(y)])\,dP_1(y) \\ &= \int_{D(P)\cap D(Q)} P(y, H_y)\,Q(y, K_y)\,dP_1(y), \end{aligned}$$

and likewise

$$R_2(C) = \int_{D(P)\cap D(Q)} P(y, H_y)\wedge Q(y, K_y)\,dP_1(y).$$

But from (*), $P(y, H_y)Q(y, K_y) < P(y, H_y)\wedge Q(y, K_y)$ for all y in $D(P)\cap D(Q)$, a set assumed to have positive P_1 - measure. Thus $R_1(C) < R_2(C)$, and the result is proved.

Statement 2 implies statement 1. Suppose that $P_1(D(P)\cap D(Q))=0$ and that R is any law in $V(P, Q)$. Put $A_1 = Y\backslash D(P)$ and $A_2 = Y\backslash(D(Q)\cup A_1)$; these are disjoint sets such that $P_1(A_1)+P_1(A_2)=1$.

Define functions $f_1\colon A_1\to X$ and $f_2\colon A_2\to Z$ as follows: for each $y\in A_1$, $P(y, \cdot)$ is a point mass at $f_1(y)$; for each $y\in A_2$, $Q(y, \cdot)$ is a point mass at $f_2(y)$. These functions are measurable:

$$f_1^{-1}(B)=\{y\in A_1\colon P(y, B)=1\} \quad \text{for any } B\in\mathscr{B}(X),$$

and

$$f_2^{-1}(C)=\{y\in A_2\colon Q(y, C)=1\} \quad \text{for any } C\in\mathscr{B}(Z).$$

Let $G_1\subset X\times Y$ and $G_2\subset Y\times Z$ be the (measurable) graphs of the functions f_1 and f_2; then $G_1\times Z$ and $X\times G_2$ are disjoint sets with $R(G_1\times Z)+ +R(X\times G_2)=P_1(A_1)+P_1(A_2)=1$.

Define

$$F_1\colon A_1\times Z\to G_1\times Z \quad \text{and} \quad F_2\colon X\times A_2\to X\times G_2$$

by $F_1(y, z)=(f_1(y), y, z)$ and $F_2(x, y)=(x, y, f_2(y))$; F_1 and F_2 are easily seen to be Borel isomorphisms onto $G_1\times Z$ and $X\times G_2$.

If now B is a Borel subset of $X\times A_1\times Z$, then

$$B\cap(G_1\times Z)=\big[X\times F_1^{-1}(B\cap(G_1\times Z))\big]\cap(G_1\times Z),$$

so that $R(B\cap(G_1\times Z)) = R\big(X\times F_1^{-1}(B\cap(G_1\times Z))\big) = Q\big(F_1^{-1}(B\cap(G_1\times Z))\big)$;

thus the value of R on subsets of $G_1\times Z$ is wholly determined by Q.

Likewise, if B is a Borel subset of $X\times A_2\times Z$, then $R(B\cap(X\times G_2))$

$= P(F_2^{-1}(B \cap (X \times G_2)))$, and the value of R on subsets of $X \times G_2$ is determined by P. Since $R(G_1 \times Z) + R(X \times G_2) = 1$, this establishes the uniqueness of R, q.e.d.

REFERENCES

[1] D. Blackwell and C. Ryll-Nardzewski, *Non-existence of everywhere proper conditional distributions*, Annals Math. Stat. 34 (1963), p. 223 - 225.

[2] D. L. Cohn, Measure theory, Birkhäuser, Boston 1980.

[3] R. G. Douglas, *On extremal measures and subspace density*, Michigan Journal Math. 11 (1964), p. 243 - 246.

[4] J. Hoffman-Jørgensen, *The theory of analytic spaces*, Various Publications Series No. 10 (1970), Aarhus Universitat, Aarhus.

[5] S. Karlin and W. J. Studden, *Tchebycheff systems with applications to analysis and statistics*, Interscience 1966.

[6] K. Kuratowski, *Topology*, Vol. I, Academic Press, New York 1966.

[7] K. Lange, *Borel sets of probability measures*, Pacific Journal Math. 48 (1) (1973), p. 141 - 161.

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

[9] E. Marczewski (Szpilrajn), *The characteristic function of a sequence of sets and some of its applications*, Fund. Math. 31 (1938), p. 207 - 223.

[10] J. Neveu, *Mathematical foundations of the calculus of probability*, Holden - Day, Sand Francisco 1965.

[11] R. C. Shiflett, *Extreme stochastic measure and Feldman's conjecture*, Journal Math. Analysis and Appl. 68 (1979), p. 111 - 117.

[12] R. M. Shortt, *Universally measurable spaces: an invariance theorem and diverse characterisations*, Fund. Math. (to appear) (1982).

[13] V. S. Varadarajan, *Groups of automorphisms of Borel spaces*, Transactions AMS 109 (1963), p. 191 - 220.

Department of Mathematical and Computer Sciences
Michigan Technological University
Houghton, Michigan 49931, USA

Received on 18. 2. 1983