Ordering of Distributions and Rearrangement of Functions
Author(s): Ludger Ruschendorf
Source: *The Annals of Probability*, Apr., 1981, Vol. 9, No. 2 (Apr., 1981), pp. 276-283
Published by: Institute of Mathematical Statistics
Stable URL: https://www.jstor.org/stable/2243459

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve and extend access to *The Annals of Probability*

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

The Annals of Probability
1981, Vol. 9, No. 2, 276–283

ORDERING OF DISTRIBUTIONS AND REARRANGEMENT OF FUNCTIONS

BY LUDGER RÜSCHENDORF

Aachen Technical University

Some characterizations of semiorders defined on the set of all probability measures on R^n by the set of Schur-convex functions and by some subsets of all convex functions are proved. A connection of these results to the theorem of Hardy, Littlewood and Polya on the rearrangement of functions is discussed. Furthermore, by means of the results on the ordering of probability measures a generalization of a theorem on doubly stochastic linear operators due to Ryff is proved.

1. Ordering by Schur convex functions. The Schur-ordering in R^n is defined for $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$ by

(1) $\quad a \prec b \quad$ if and only if $\sum_{i=1}^k a_{(i)} \leq \sum_{i=1}^k b_{(i)}$, $1 \leq k \leq n-1$, $\quad$ and $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$

where $a_{(1)} \geq \ldots \geq a_{(n)}$ and $b_{(1)} \geq \ldots \geq b_{(n)}$ are the components of a, b rearranged in decreasing order. Hardy, Littlewood, Polya [5], [6], Rado [11] and Mirsky [9] proved the following equivalences

(2)
- (i) $a \prec b$.
- (ii) There exists a doubly stochastic matrix Q with $Qb = a$;
- (iii) a lies in the convex hull of $\{b_\pi; \pi \in \gamma_n\}$ (where $b_\pi = (b_{\pi(1)}, \ldots, b_{\pi(n)})$ and where γ_n is the symmetric group of order n);
- (iv) $\varphi(a) \leq \varphi(b)$ for all symmetric, convex functions φ on R^n (φ symmetric means $\varphi(b_\pi) = \varphi(b)$ for all $\pi \in \gamma_n$);
- (v) $\sum_{i=1}^n \varphi(a_i) \leq \sum_{i=1}^n \varphi(b_i)$ for all convex functions φ on R^1.

$\prec$ is a semiorder on R^n (it is not antisymmetric). The monotonically nondecreasing functions on R^n w.r.t. $\prec$ are called Schur-convex the monotonically nonincreasing functions are called Schur-concave.

Nevius, Proschan, Sethuraman [10] introduced the notion of stochastic majorization and discussed some applications. They defined for $P_1, P_2 \in \mathscr{M}(R^n)$—the set of all probability measures on $(R^n, \mathfrak{B}^n)$—(cf. [10], Theorem 2.2)

$$P_1 \leq_1 P_2 \quad \text{if and only if} \quad \int f\, dP_1 \leq \int f\, dP_2 \tag{3}$$
$$\text{for all} \quad f \in M_1 = \{f: R^n \to R^1;\ f \text{ bounded, measurable, Schur-convex}\}.$$

$\leq_1$ is a semiorder on $\mathscr{M}(R^n)$. Some equivalent conditions for $\leq_1$ are given in Theorem 2.2. of Nevius, Proschan, Sethuraman [10].

An element $A \in \mathfrak{B}^n$ is called Schur-convex if $1_A \in M_1$. Denote by $d(x, y)$ the Euclidean distance on R^n and define

$$d(x, A) = \inf\{d(x, y), y \in A\} \qquad \text{for} \quad A \subseteq R^n.$$

LEMMA 1. *If $A \subseteq R^n$ is a Schur-convex set then $d(\cdot, A)$ is Schur-concave.*

Received October 4, 1979.
AMS 1970 *subject classification*. Primary 60B99, 62H99.
Key words and phrases. Stochastic semiorder, rearrangement, convex functions, diffusion.

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

PROOF. For $x, y \in R^n$ with $x \prec y$ and for $z \in A$ we have to show that there exists a $z' \in A$ such that

$$d(y, z') \leq d(x, z).$$

Since $d(x, z) = d(x_\pi, z_\pi)$ for all $\pi \in \gamma_n$ and $z_\pi \in A$ for all $\pi \in \gamma_n$ $d(\cdot, A)$ is symmetric. Furthermore, $x \prec y$ is equivalent to $x_\pi \prec y_\nu$ for all $\pi, \nu \in \gamma_n$ so that we can assume that $x_1 \geq \ldots \geq x_n$ and $y_1 \geq \ldots \geq y_n$.

Defining $z_{(\,)} = (z_{(1)}, \ldots, z_{(n)})$ we obtain $d(x, z) \geq d(x, z_{(\,)})$ since $\sum_{i=1}^n x_i z_i \leq \sum_{i=1}^n x_i z_{(i)}$ and, therefore, we can assume that $z_1 \geq \ldots \geq z_n$.

Defining $z' = z + d$, where $d = y - x$ we obtain

$$\textstyle\sum_{i=1}^k d_i \geq 0,\ 1 \leq k \leq n-1 \qquad \text{and} \qquad \sum_{i=1}^n d_i = 0.$$

This implies

$$\textstyle\sum_{i=1}^k z_i' = \sum_{i=1}^k (z_i + d_i) \geq \sum_{i=1}^k z_i, \qquad 1 \leq k \leq n-1$$

and

$$\textstyle\sum_{i=1}^n z_i' = \sum_{i=1}^n (z_i + d_i) = \sum_{i=1}^n z_i$$

and, therefore,

$$\textstyle\sum_{i=1}^k z_{(i)}' \geq \sum_{i=1}^k z_i, \qquad 1 \leq k \leq n-1 \qquad \text{so that} \quad z \prec z'.$$

Since A is Schur-convex $z' \in A$ and $d(x, z) = d(y, z')$. □

We now obtain

THEOREM 2. *Let $P_1, P_2 \in \mathscr{M}(R^n)$. Then*

$$P_1 \leq_1 P_2 \qquad \textit{if and only if} \quad \int f\, dP_1 \leq \int f\, dP_2$$

$$\textit{for all} \quad f \in M_1' = \{f \in M_1; f \textit{ continuous and bounded}, f \geq 0\}. \tag{4}$$

PROOF. As in the proof of the equivalence of (1) and (6) of Theorem 1 of Kamae, Krengel and O'Brien [7], $P_1 \leq_1 P_2$ is equivalent to $P_1(A) \leq P_2(A)$ for all closed Schur-convex sets (cf. also Theorem 2.2 of [8]).

Assume the right-hand condition of (4), let A be closed and Schur-convex and define

$$f_k(t) = \begin{cases} 0 & \text{if} \quad d(t, A) \geq \dfrac{1}{k} \\ 1 - kd(t, A) & \text{if} \quad d(t, A) \leq \dfrac{1}{k} \end{cases}, \qquad k \in N.$$

Then f_k is continuous, $f_k \geq 0$ and by Lemma 1 f_k is Schur-convex. Furthermore,

$$\lim_{k\to\infty} f_k(t) = 1_A(t)$$

since $A = \bar{A}$. Therefore, by the monotone convergence theorem

$$P_1(A) = \lim_{k\to\infty} \int f_k(t)\ dP_1(t) \leq \lim_{k\to\infty} \int f_k(t)\ dP_2(t) = P_2(A)$$

which implies $P_1 \leq_1 P_2$. □

Theorem 2 implies that $\leq_1$ is a closed semiorder on $\mathscr{M}(R^n)$ when we consider the topology of weak convergence. This remark strengthens Theorem 3.6 of Nevius, Proschan, Sethuraman [10].

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

The following theorem gives a pointwise characterization of $\leq_1$. For a Markov kernel T on $(R^n, \mathfrak{B}^n)$ and $P \in \mathcal{M}(R^n)$ define $TP \in \mathcal{M}(R^n)$ by $TP(A) = \int T(x, A)\, dP(x)$ and define ϵ_x to be the one point measure in x.

THEOREM 3. *Let $P_1, P_2 \in \mathcal{M}(R^n)$. Then the following conditions* (a), (b), (c) *are equivalent*:

(a) $P_1 \leq_1 P_2$;

(b) *there exists a Markov kernel T on $(R^n, \mathfrak{B}^n)$ with $P_2 = TP_1$ and $\epsilon_x \leq_1 T(x, \cdot)[P_1]$*;

(c) *there exists a probability space $(\Omega, \mathcal{A}, P)$ and random vectors X, Y on $(\Omega, \mathcal{A})$ with $P^X = P_1$, $P^Y = P_2$ and $X \prec Y[P]$.*

PROOF. Define $\omega = \{(x, y) \in R^n \times R^n; x \prec y\}$ and define $\pi_1, \pi_2: R^n \times R^n \to R^n$ by $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$. If $U \subset R^n$ is open then $\pi_1(\omega \cap (R^n \times U)) = U_+$ where $U_+ = \{y \in R^n; \exists\, x \in U \text{ with } x \prec y\}$. This implies assuming (a) that

$$P_1(U) \leq P_1(U_+) \leq P_2(U_+) = P_2(\pi_1(\omega \cap (R^n \times U))).$$

By Theorem 11 of Strassen [15] there exists a probability measure P on $(R^{2n}, \mathfrak{B}^{2n})$ with $P^{\pi_i} = P_i$, $i = 1, 2$ and

$$P(\omega) = P(\pi_1 \prec \pi_2) = 1.$$

So (a) implies (c). (c) implies (b) defining T to be the regular conditional distribution of Y given X. If (b) holds true, then for each Schur-convex bounded function φ we obtain

$$\int \varphi\, dP_2 = \int \varphi\, dTP_1 = \int \left(\int \varphi(y) T(x, dy) \right) dP_1(x) \geq \int \varphi(x)\, dP_1(x).$$

So (b) implies (a). □

REMARK 1.

(a) Theorem 3 implies Theorem 2.4 of Nevius, Proschan and Sethuraman [10].

(b) The results proved by Kamae, Krengel and O'Brien [7] for partial ordered polish spaces hold also true in the case of the Schur-order (which is only a semiorder). This allows us to order some types of stochastic processes w.r.t. $\leq_1$.

The following corollary answers a question put by Nevius, Proschan and Sethuraman [10] in connection with their Theorem 2.9.

COROLLARY 4. *Let X, Y be two random vectors with values in R^n and define*

$$S = \textstyle\sum_{i=1}^n X_i, \qquad S' = \sum_{i=1}^n Y_i.$$

Then $P^X \leq_1 P^Y$ if and only if

(a) $P^S = P^{S'}$

and

(b) $P^{X|S} \leq_1 P^{Y|S}\, [P^S]$.

2. Some subsets of convex functions. Define the following sets of functions

$$\begin{aligned} &M_2 = \{f: R^n \to R^1; f \text{ convex}\};\\ (5)\qquad &M_3 = \{f \in M_2; f \text{ is monotonically nondecreasing w.r.t. the componentwise partial order on } R^n\};\\ &M_4 = \{f \in M_2; f(x_\pi) = f(x), \quad \forall \pi \in \gamma_n, \quad \forall x \in R^n\}. \end{aligned}$$

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

Furthermore, define for $P_1, P_2 \in \mathscr{M}(R^n)$ and $i = 2, 3, 4$

$$(6) \quad P_1 \leq_i P_2 \quad \text{if} \quad \int f\,dP_1 \leq \int f\,dP_2 \quad \text{for all} \quad f \in M_i \quad \text{such that the integrals exist.}$$

A main difference to the ordering by means of Schur-convex functions is that $\leq_i$ for $i = 2$, 3, 4, is not induced by a closed semiorder as $\leq_1$ is and, therefore, Theorem 11 of Strassen [15] does not apply to these cases.

A Markov kernel T on $(R^n, \mathfrak{B}^n)$ is called a M_i-diffusion $i = 2, 3, 4$ if

$$(7) \qquad \epsilon_x \leq_i T(x, \cdot), \qquad \forall x \in R^n.$$

For $P \in \mathscr{M}(R^n)$ with existing first moments we define EP to be the vector of the first moments of P. In the following lemma we determine the set of all M_i-diffusions.

Lemma 5. *Let T be a Markov kernel on $(R^n, \mathfrak{B}^n)$ and assume that $ET(x, \cdot)$ exists for all $x \in R^n$ (and that $\int y_{(\,)} T(x, dy)$ exists if $i = 4$). Then T is a M_i-diffusion if and only if*

$$(8) \qquad \begin{aligned} x &= ET(x, \cdot) && \text{for all} \quad x \in R^n \quad \text{if} \quad i = 2 \\ x &\leq ET(x, \cdot) && \text{for all} \quad x \in R^n \quad \text{if} \quad i = 3 \\ x &\prec \textstyle\int y_{(\,)} T(x, dy) && \text{for all} \quad x \in R^n \quad \text{if} \quad i = 4. \end{aligned}$$

Proof. We consider at first the case $i = 2$. If $x = ET(x, \cdot)$ for all $x \in R^n$ and if f is convex and such that $\int f(y)\,T(x, dy)$ exists then by Jensen's inequality

$$f(x) = f\left(\int yT(x, dy)\right) \leq \int f(y)\,T(x, dy)$$

which implies

$$\epsilon_x \leq_2 T(x, \cdot).$$

The other direction follows from the fact that the linear functions are convex and also concave.

The case $i = 3$ is analogous to the case $i = 2$. For the case $i = 4$ assume that $x \prec \int y_{(\,)} T(x, dy)$ and that f is convex and symmetric and that $\int f(y)\,T(x, dy)$ exists. Then f is also Schur-convex and by Jensen's inequality we obtain that

$$f(x) \leq f\left(\int y_{(\,)} T(x, dy)\right) \leq \int f(y_{(\,)})\,T(x, dy) = \int f(y)\,T(x, dy)$$

by the symmetry of f. This implies that $\epsilon_x \leq_4 T(x, \cdot)$.

For the other direction define

$$f_k(x) = \textstyle\sum_{i=1}^k x_{(i)}, \qquad 1 \leq k \leq n.$$

If $x, y \in R^n$, $\alpha \in (0, 1)$ then for all $\pi \in \gamma_n$

$$\textstyle\sum_{i=1}^k (\alpha x_{\pi(i)} + (1-\alpha) y_{\pi(i)}) \leq \sum_{i=1}^k (\alpha x_{(i)} + (1-\alpha) y_{(i)}) = \alpha f_k(x) + (1-\alpha) f_k(y)$$

which implies

$$f_k(\alpha x + (1-\alpha) y) \leq \alpha f_k(x) + (1-\alpha) f_k(y)$$

so that f_k are convex and symmetric. Therefore, assuming $\epsilon_x \leq_4 T(x, \cdot)$ we obtain

$$\begin{aligned} f_k(x) = \textstyle\sum_{i=1}^k x_{(i)} &\leq \int \textstyle\sum_{i=1}^k y_{(i)}\,T(x, dy) \\ &= \textstyle\sum_{i=1}^k \int y_{(i)}\,T(x, dy) = f_k\left(\int y_{(\,)}\,T(x, dy)\right), \qquad 1 \leq k \leq n-1 \end{aligned}$$

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

and

$$f_n(x) = \int f_n(y_{(\,)})\,T(x, dy) = f_n\left(\int y_{(\,)}\,T(x, dy)\right)$$

since f_n is also Schur-concave. This implies that

$$x \prec \int y_{(\,)}\,T(x, dy). \qquad \square$$

Lemma 5 shows that $\leq_4$ is different from $\leq_1$ in the general case. This is somewhat surprising since by (2) $\leq_4$ and $\leq_1$ are identical for one-point measures. The following theorem gives the pointwise characterization of $\leq_i$, $i = 2, 3, 4$.

THEOREM 6. *Let $P_1, P_2 \in \mathscr{M}(R^n)$ have first moments. Then the following conditions are equivalent for $i = 2, 3, 4$.*

(a) $P_1 \leq_i P_2$;

(b) *there exists a M_i-diffusion T on $(R^n, \mathfrak{B}^n)$ with $TP_1 = P_2$;*

(c) *there exists a probability space $(\Omega, \mathscr{A}, P)$ and random vectors X, Y on $(\Omega, \mathscr{A})$ with $P^X = P_1$, $P^Y = P_2$ and such that*

$$E(Y|X) = X \quad [P] \qquad \textit{if} \quad i = 2$$

$$E(Y|X) \geq X \quad [P] \qquad \textit{if} \quad i = 3$$

$$E(Y_{(\,)}|X) \succ X\,[P] \qquad \textit{if} \quad i = 4.$$

PROOF. We at first consider the case $i = 2$. For $x \in R^n$ define K_x to be the set of all $P \in \mathscr{M}(R^n)$ such that $\epsilon_x \leq_2 P$. Then K_x is convex and closed w.r.t. the topology of weak convergence. For $f \in C_b(R^n)$—the set of all bounded continuous functions on R^n—define

$$h_f(x) = \sup\left\{\int f\,dP;\, P \in K_x\right\}.$$

Since $\epsilon_x \in K_x$ we have

$$f(x) \leq h_f(x) \leq \sup_{y \in R^n} |f(y)| \leq \infty.$$

For $x, y, \in R^n$ and $\alpha \in (0, 1)$ we have

$$\epsilon_{\alpha x + (1-\alpha)y} \leq_2 \alpha\epsilon_x + (1 - \alpha)\epsilon_y.$$

Therefore, $P, P' \in \mathscr{M}(R^n)$ and $\epsilon_x \leq_2 P$, $\epsilon_y \leq_2 P'$ implies $\epsilon_{\alpha x+(1-\alpha)y} \leq_2 \alpha P + (1 - \alpha)P'$. This implies that

$$h_f(\alpha x + (1 - \alpha)y) \geq \alpha h_f(x) + (1 - \alpha)h_f(y)$$

and, therefore, h_f is a concave, bounded function on R^n.

If $P_1 \leq_2 P_2$ we obtain

$$\int f\,dP_2 \leq \int h_f\,dP_2 \leq \int h_f\,dP_1.$$

But this implies by Theorem 3 of Strassen [15] the existence of a Markov kernel T on $(R^n, \mathscr{L}^n)$ with $P_2 = TP_1$ and $T(x, \cdot) \in K_x[P_1]$. We modify $T(x, \cdot)$ on the exceptional nullset by $T(x, \cdot) = \epsilon_x$. From our assumption on P_1, P_2 it follows that T has a.s. (w.r.t. P_1) existing first moments and so T is an M_2-diffusion which yields (b). The equivalence of (b) and (c) follows from the characterization of the M_2-diffusions given in Lemma 5 similarly to the proof of Theorem 2. (a) is immediate from (c) by Jensen's inequality.

The proof for $\leq_3$, $\leq_4$ is analogous to observing that h_f (in the obvious modified form) is monotonically nonincreasing if $i = 3$ and symmetric if $i = 4$. $\square$

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

REMARK 2.

(a) In the case $i = 2, 3$ Theorem 6 has already been proved (in a different way) by Strassen [15], Theorems 2, 8, 9. The result for $i = 2$ is a generalization of a famous theorem due to Hardy, Littlewood, Polya, Blackwell, Stein, Sherman, Cartier, Fell and Meyer (cf. Strassen [15], Theorem 2). Our proof is on the lines of the theory of balayage cf., f.i. Theorem 53 of Meyer [8]. The function h_f defined in the proof of Theorem 6 is the least concave majorant of f. In the compact case consider Proposition 26.13 of Choquet [2].

(b) The method of proving Theorem 6 can be applied in many further cases. Let, for example, M_5 be the set of all $f: R^n \to R^1$ increasing in absolute value that means if $|x| \le |y|$ then $f(x) \le f(y)$ where $|\ |$ is any norm on R^n. Then a Markov kernel T is a M_5-diffusion if and only if for all $x \in R^n$ $T(x, \{y; |x| \le |y|\}) = 1$. Therefore, for $P_1, P_2 \in \mathscr{M}(R^n)$ $P_1 \le_5 P_2$ is equivalent to the existence of random vectors X, Y on $(\Omega, \mathscr{A}, P)$ with $P^X = P_1$, $P^Y = P_2$ and $|X| \le |Y|\,[P]$.

3. Connections to the rearrangement of functions. Let $L^1 = L^1([0, 1], \lambda^1)$ be the space of all integrable functions on $([0, 1], \mathscr{L}^1[0, 1], \lambda^1)$. Hardy, Littlewood and Polya [6] and Chong [1] generalized the Schur-order to elements of L^1 defining for $f, g \in L^1$

(9)
$$\begin{aligned} &f << g \quad \text{if} \quad \int_0^s f^*(u)\,du \le \int_0^s g^*(u)\,du,\ s \in (0, 1) \quad \text{and} \\ &f < g \quad \text{if} \quad f << g \quad \text{and} \quad \int_0^1 f^*(u)\,du = \int_0^1 g^*(u)\,du \end{aligned}$$

where f^*, g^* are the monotonically nonincreasing (equimeasurable) rearrangements of f, g. Similarly to the finite-dimensional case

(10)
$$f << g \quad \text{is equivalent to} \quad \int \varphi \circ f\, d\lambda^1 \le \int \varphi \circ g\, d\lambda^1$$

for all convex, monotonically nondecreasing functions φ such that the integrals exist and

(11)
$$f < g \quad \text{is equivalent to} \quad \int \varphi \circ f\, d\lambda^1 \le \int \varphi \circ g\, d\lambda^1$$

for all convex functions such that the integrals exist. (cf., Theorems 2.3, 2.5 of Chong [1]).

If $P_1, P_2 \in \mathscr{M}(R^1)$ have distribution functions F_1, F_2 (10) leads to the equivalence

(12)
$$P_1 \le_3 P_2 \quad \text{if and only if} \quad \int_{[x,\infty)} (t - x)\, dP_1(t) \le \int_{[x,\infty)} (t - x)\, dP_2(t), \qquad \forall\, x \in R^1$$

and (11) leads to the equivalence

(13)
$$P_1 \le_2 P_2 \quad \text{if and only if} \quad P_1 \le_3 P_2 \quad \text{and} \quad \int t\, dP_1(t) = \int t\, dP_2(t).$$

This follows from the observation that for $f \in L^1(P_i)$

$$\int f\, dP_i = \int_{[0,1]} f \circ F_i^{-1}\, d\lambda^1, \qquad i = 1, 2$$

and Theorem 1.6 of Chong [1]. (12), (13) are identical to criterions given by Stoyan [14] for the order relations $\le_2$, $\le_3$ in the case $n = 1$. So ordering distributions w.r.t. $\le_2$, $\le_3$ is equivalent (for $n = 1$) to the rearrangement of the pseudoinverses of their distribution functions. (For $n \ge 1$ sufficient conditions for $\le_3$ have been given by Franken and Stoyan [4], Satz 5, and Franken and Kirstein [3], Satz 4.3).

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

Now assume that $a, b \in R^n$. Then by (2) $a \prec b$ is equivalent to $(1/n)\sum_{i=1}^n \epsilon_{\{a_i\}} \leq_2 (1/n)\sum_{i=1}^n \epsilon_{\{b_i\}}$. By Theorem 6 there exists a Markov kernel T with

$$T\left(\frac{1}{n}\sum_{i=1}^n \epsilon_{\{a_i\}}\right) = \frac{1}{n}\sum_{i=1}^n \epsilon_{\{b_i\}} \quad \text{and} \quad \int yT(x, dy) = x\left[\frac{1}{n}\sum_{i=1}^n \epsilon_{\{a_i\}}\right]. \tag{14}$$

Defining $Q = (T(a_i, \{b_j\}))_{1\leq i\leq n; 1\leq j\leq n}$ (14) implies that Q is doubly stochastic and that $Qb = a$. So Theorem 6 implies the first equivalence in (2) which is due to Hardy, Littlewood and Polya [6], Theorem 46.

Ryff [12], [13] generalized the notion of doubly stochastic matrices. A linear operator $T: L^1 \to L^1$ is called doubly stochastic if $Tf \prec f$ for all $f \in L^1$. Ryff [12] also gave some equivalent definitions of doubly stochastic operators and proved the following generalization of the theorem of Hardy, Littlewood and Polya [6] (cf., [13], Theorem 3).

If $f, g \in L^1$ then

(15) $f \prec g$ if and only if there exists a doubly stochastic operator T such that $Tg = f$.

We want to prove a generalization of (15) by means of Theorem 6. Let $(\Omega, \mathscr{A}, P)$ be a probability space and define

$$L_n^1(P) = \{f: \Omega \to R^n: f \text{ is integrable w.r.t. } P\}.$$

For $f, g \in L_n^1(P)$ we define

$$f \prec g \quad \text{if} \quad \int \varphi\circ f\, dP \leq \int \varphi\circ g\, dP$$

for all convex functions φ such that the integrals exist. A linear operator $T: L_n^1(P) \to L_n^1(P)$ is called doubly stochastic if $Tf \prec f$ for all $f \in L_n^1(P)$.

THEOREM 7. *Let $f, g \in L_n^1(P)$. Then $f \prec g$ if and only if there exists a doubly stochastic operator $T: L_n^1(P) \to L_n^1(P)$ with $Tg = f[P]$.*

PROOF. $f \prec g$ is by definition equivalent to $P^f \leq_2 P^g$; so by Theorem 6 there exists a Markov kernel K on $(R^n, \mathscr{L}^n)$ such that $KP^f = P^g$ and $\int yK(x, dy) = x[P^f]$. Define a linear operator

$$T': L_n^1(\mathscr{A}(g), P) \to L_n^1(\mathscr{A}(f), P)$$

by

$$T'(h\circ g)(y) = \int K(f(y), dx)h(x).$$

If $h\circ g \in L_n^1(\mathscr{A}(g), P)$ then

$$\int h\circ g\, dP = \int \left(\int K(f(y), dx)h(x)\right) dP$$

which implies that $\int K(f(y), dx)h(x)$ exists P a.s. If, furthermore, $h_1\circ g = h_2\circ g[P]$ then

$$0 = \int |h_1\circ g - h_2\circ g|\, dP = \int \left(\int K(f(y), dx)|h_1(x) - h_2(x)|\right) dP$$

which implies $T'(h_1\circ g) = T'(h_2\circ g)$ and so T' is well defined.

$$T'g(y) = \int K(f(y), dx)x = f(y)[P].$$

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms

For each convex function $\varphi : R^n \to R^1$ such that the following integrals exist we obtain

$$\int \varphi(T'(h\circ g))\,dP = \int \varphi\Big(\int T(x,dy)h(y)\Big)\,dP^f \le \int\Big(\int \varphi\circ h(y)T(x,dy)\Big)\,dP^f$$
$$= \int \varphi\circ h(y)\,dTP^f = \int \varphi\circ h(y)\,dP^g = \int \varphi(h\circ g)\,dP$$

and, therefore, $T'(h\circ g) \prec h\circ g$ for all $h\circ g \in L_n^1(\mathscr{A}(g), P)$. We now extend T' from $L_n^1(\mathscr{A}(g), P)$ to $L_n^1(P)$. Define

$$T'' : L_n^1(P) \to L_n^1(\mathscr{A}(g), P)$$

to be the projection

$$T''(h) = E(h \mid g).$$

Then $T''(h) \prec h$ by Jensen's inequality and $T''(h\circ g) = h\circ g$ for all $h\circ g \in L_n^1(\mathscr{A}(g), P)$. Now defining

$$T : L_n^1(P) \to L_n^1(P)$$

by

$$T(h) = T'(T''(h))$$

we obtain the conclusion of Theorem 7. □

REFERENCES

[1] CHONG, KONG-MING. (1974). Some extensions of a theorem of Hardy, Littlewood and Polya and their applications. *Canad. J. Math* **26** 1321–1340.

[2] CHOQUET, G. (1969). *Lectures on Analysis, Vol.* II. *Representation Theory.* W. A. Benjamin, Inc.

[3] FRANKEN, P. and KIRSTEIN, B. M. (1977). Zur Vergleichbarkeit zufälliger Prozesse. *Math. Nachr.* **78** 197–205.

[4] FRANKEN, P. and STOYAN, D. (1975). Einige Bemerkungen über monotone und vergleichbare markowsche Prozesse. *Math. Nachr.* **66** 201–209.

[5] HARDY, G. H., LITTLEWOOD, J. E. and POLYA, G. (1929). Some simple inequalities satisfied by convex functions. *Mess. of Math.* **58** 145–152.

[6] HARDY, G. H., LITTLEWOOD, J. E. and POLYA, G. (1934). *Inequalities.* Cambridge Univ. Press.

[7] KAMAE, T., KRENGEL, U. and O'BRIEN, G. L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probability* **5** 899–912.

[8] MEYER, P. A. (1966). *Probability and Potentials.* Blaisdell Publishing Company.

[9] MIRSKY, L. (1959). Inequalities for certain classes of convex functions. *Proc. Edinburgh Math. Soc.* **11** 231–235.

[10] NEVIUS, S. E., PROSCHAN, F. and SETHURAMAN, J. (1977). Schur functions in statistics. II. Stochastic majorization. *Ann. Statist.* **5** 263–273.

[11] RADO, R. (1952). An inequality. *J. London Math. Soc.* **27** 1–6.

[12] RYFF, J. V. (1963). On the representation of doubly stochastic operators. *Pacific J. Math.* **13** 1379–1386.

[13] RYFF, J. A. (1965). Orbits of L^1-functions under doubly stochastic transformations. *Trans. Amer. Math. Soc.* **117** 92–100.

[14] STOYAN, D. (1972). Halbordnungsrelationen für Verteilungsgesetze. *Math. Nachr.* **52** 315–331.

[15] STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.

INSTITUT FÜR STATISTIK UND WIRTSCHAFTSMATHEMATIK
WÜLLNERSTR. 3
5100 AACHEN
W-GERMANY

This content downloaded from
73.74.251.164 on Thu, 20 May 2021 22:08:29 UTC
All use subject to https://about.jstor.org/terms