Journal of Statistical Planning and Inference 62 (1997) 193–205

journal of statistical planning and inference

Orthant orderings of discrete random vectors

Rainer Dyckerhoff*, Karl Mosler

Lehrstuhl für Statistik und Ökonometrie, Universität zu Köln, Meister-Ekkehart-Str. 9, 50923 Köln, Germany

Received 6 December 1995; revised 15 August 1996

Abstract

We investigate four orthant stochastic orderings between random vectors X and Y that have finitely discrete probability distributions in $\mathbb{R}^k$. They have applications to multiattribute decision under risk, dependency of random vectors, and the statistical comparison of two k-variate samples. For each of the orderings we present conditions that are necessary and sufficient for dominance of Y over X. Given their distributions numerically, these conditions can be checked in an efficient way. © 1997 Elsevier Science B.V.

AMS classification: primary 60E15; secondary 90A43; 62H20

Keywords: Multivariate stochastic orders; Dependence ordering; Decision under risk; Bivariate risk aversion; Comparison of empirical distribution functions

1. Introduction

Multivariate stochastic orders are frequently used in applied probability and statistics. An early example is Lehmann's order, which is called the *usual multivariate stochastic order*. Lehmann (1955) defines that a random vector Y in $\mathbb{R}^k$ is *stochastically larger* than another random vector X, $X \preceq_{\text{st}} Y$, if

$$E(\phi(X)) \leqslant E(\phi(Y)) \tag{1}$$

holds for every function $\phi : \mathbb{R}^k \to \mathbb{R}$ that is increasing[1] in the componentwise ordering $\leqslant$ of $\mathbb{R}^k$ and for which both expectations exist. Equivalently, $P[X \in U] \leqslant P[Y \in U]$ for every upper set U in $\mathbb{R}^k$. (U is upper in $(\mathbb{R}^k, \leqslant)$ if $x \leqslant y$, $x \in U$, implies $y \in U$.) This holds also in more general spaces (Kamae et al., 1977). Properties and applications of

* Corresponding author.

[1] "Increasing" is meant in the weak sense.

0378-3758/97/$17.00
PII S0378-3758(96)00184-X

the usual and many other multivariate stochastic orderings are collected in two recent books; see Shaked and Shanthikumar (1994) and Mosler and Scarsini (1991a).

Stochastic orders occur in different contexts. Either they arise as structural properties of a given stochastic model, and their presence may be derived from the model assumptions. Or a certain stochastic order is hypothesized between two distributions that are given numerically. Then the hypothesis may be confirmed by checking the data and, possibly, doing some statistical inference.

It is the second situation that we have in mind in this paper. Given a finite subset $\{s^1, s^2, \ldots, s^n\}$ of $\mathbb{R}^k$ and two vectors of probabilities $p = (p_1, p_2, \ldots, p_n)$ and $q = (q_1, q_2, \ldots, q_n)$ where $p_i = P[X = s^i]$ and $q_i = P[Y = s^i]$, $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1$, we ask whether Y is larger than X in some stochastic ordering. For $\preceq_{\mathrm{st}}$ this is equivalent to finding a feasible solution of the linear program

$$\sum_{j=1}^n \pi_{ij} p_i = q_j \quad \forall j, \qquad \sum_{\{j: s^i \leqslant s^j\}} \pi_{ij} = 1 \quad \forall i, \quad \sum_{j=1}^n \pi_{ij} = 1 \quad \forall i, \quad \pi_{ij} \geqslant 0 \quad \forall i, j,$$

which is easily done by standard methods.

In this paper we investigate four orderings that are weaker than $\preceq_{\mathrm{st}}$. Let $]-\infty, a]$ denote the k-dimensional interval $\{x \in \mathbb{R}^k : x \leqslant a\}$, which we call the *lower orthant at a.* Similarly, $[a, \infty[$ is called the *upper orthant at a.* The orderings are based on the following conditions.

$$P[X \leqslant a] \geqslant P[Y \leqslant a] \quad \text{for all } a \in \mathbb{R}^k, \tag{LO}$$

$$\int_{]-\infty, a]} P[X \leqslant z]\, \mathrm{d}z \geqslant \int_{]-\infty, a]} P[Y \leqslant z]\, \mathrm{d}z \quad \text{for all } a \in \mathbb{R}^k, \tag{LOCC}$$

$$P[X \geqslant a] \leqslant P[Y \geqslant a] \quad \text{for all } a \in \mathbb{R}^k, \tag{UO}$$

$$\int_{[a, \infty[} P[X \geqslant z]\, \mathrm{d}z \leqslant \int_{[a, \infty[} P[Y \geqslant z]\, \mathrm{d}z \quad \text{for all } a \in \mathbb{R}^k. \tag{UOCX}$$

Each of the four conditions defines a preorder (transitive and reflexive) between random variables in $\mathbb{R}^k$ and a partial order (antisymmetric preorder) between the underlying k-variate probability distribution functions.

Definition. If one of the four conditions holds we say that Y is larger than X in the sense of *upper orthant ordering* (UO), *upper orthant convex ordering* (UOCX), *lower orthant ordering* (LO), *lower orthant concave ordering* (LOCC). In symbols, $X \preceq_{\mathrm{uo}} Y$, $X \preceq_{\mathrm{uocx}} Y$, $X \preceq_{\mathrm{lo}} Y$, and $X \preceq_{\mathrm{locc}} Y$.

These orderings have interpretations and applications in many fields. See Bergmann (1978, 1991), Stoyan (1983), Mosler (1984), Scarsini (1988), and others. They can be characterized by Eq. (1) with certain classes of functions ϕ. For example, for (UO) this is the class of functions that are Δ-monotone (Rüschendorf, 1980). Then ϕ may be interpreted as a von Neumann–Morgenstern utility function, and the ordering as the unanimous expected utility preference of Y over X. Other classes of ϕ will be surveyed in Section 2.

Further applications of the orthant orderings include the modelling of positive dependence of a random vector, multivariate slippage of ordinal data, a notion of increasing multivariate peakedness, and percolation. See Section 2 for references and some details.

A principal problem in checking for an orthant ordering, say (UO), is that the definition involves an infinite number of k-variate orthants $[a, \infty[$. A similar problem arises when an expected utility characterization (1) is considered.

In this paper we will show that, for two given distributions having support in a finite set $\{s^1, s^2, \ldots, s^n\} \subset \mathbb{R}^k$, it is sufficient to check the defining condition at a relatively small set of points $a \in \mathbb{R}^k$. For (UO) and (LO), this set is a subset of the k-dimensional grid generated by $\{s^1, s^2, \ldots, s^n\}$. When checking for (UOCX) and (LOCC), in addition, similar conditions have to be verified with the lower-dimensional marginal distributions.

Our results allow for the construction of efficient procedures to test whether two numerically given distributions are ordered in one of the four orthant orderings. In particular, as the algorithms are very fast, statistical inference by resampling procedures becomes possible.

In Section 2 we survey some basic properties of the orderings and their applications. Section 3 presents the main theorems, which are proved in Section 4. Section 5 concludes the paper.

2. Basic properties and applications

Since $[a, \infty[$ and the complement of $]-\infty, a]$ are upper sets, $X \preceq_{st} Y$ implies $X \preceq_{lo} Y$ and $X \preceq_{uo} Y$. Moreover, the next proposition is easily derived from the definitions.

Proposition 1.
(i) $X \preceq_{st} Y \;\Rightarrow\; X \preceq_{uo} Y \;\Rightarrow\; X \preceq_{uocx} Y$,
(ii) $X \preceq_{st} Y \;\Rightarrow\; X \preceq_{lo} Y \;\Rightarrow\; X \preceq_{locc} Y$,
(iii) $X \preceq_{lo} Y \;\Leftrightarrow\; -Y \preceq_{uo} -X$,
(iv) $X \preceq_{locc} Y \;\Leftrightarrow\; -Y \preceq_{uocx} -X$.

When $k = 1$, both $\preceq_{uo}$ and $\preceq_{lo}$ coincide with $\preceq_{st}$. However, when $k \geqslant 2$, no other reverse implications are generally true in Proposition 1. Also, $X \preceq_{uo} Y$ and $X \preceq_{lo} Y$ together do not imply $X \preceq_{st} Y$. This is shown by an example.

Example 1. Let $k = 2$ and

$$P[X = (0, 0)] = 0.5, \quad P[X = (0, 2)] = 0.4, \quad P[X = (1, 1)] = 0.1,$$

$$P[Y=(0,1)]=0.3, \quad P[Y=(1,0)]=0.3, \quad P[Y=(2,2)]=0.4.$$

It can easily be checked that $X \preceq_{\mathrm{lo}} Y$ as well as $X \preceq_{\mathrm{uo}} Y$ hold, but $X \preceq_{\mathrm{st}} Y$ is not true.

In the univariate case, $X \preceq_{\mathrm{locc}} Y$ ($X \preceq_{\mathrm{uocx}} Y$) means that (1) holds for all ϕ that are increasing and concave (convex).

Note that if each of the random vectors X and Y consists of stochastically independent components, the problem of checking for the orderings becomes a trivial one. Then each of the four orderings is equivalent to the ordering of all univariate marginals in the usual stochastic order respectively the usual concave (convex) order.

$X \preceq_{\mathrm{uo}} Y$ ($X \preceq_{\mathrm{uocx}} Y$) is equivalent to (1) for all ϕ that are of the form

$$\phi(x_1, \ldots, x_k) = \prod_{i=1}^{k} \phi_i(x_i),$$

where the ϕ_i are real functions that are nonnegative and increasing (resp. nonnegative, increasing and convex). On the other hand, $X \preceq_{\mathrm{lo}} Y (X \preceq_{\mathrm{locc}} Y)$ holds if and only if (1) for all ϕ of the form increasingness and

$$\phi(x_1, \ldots, x_k) = -\prod_{i=1}^{k} (-\phi_i(x_i)),$$

where the ϕ_i are nonpositive and increasing (resp. nonpositive, increasing and concave). Moreover, $X \preceq_{\mathrm{uo}} Y$ is equivalent to (1) for all ϕ that are Δ-monotone. See Bergmann (1978), Rüschendorf (1980), Mosler (1984), Mosler and Scarsini (1991b) for these results. Note that, in case $k=2$, a Δ-monotone function is called 2-monotone and characterized by increasingness and

$$\phi(x_1, x_2) + \phi(y_1, y_2) \geqslant \phi(x_1, y_2) + \phi(y_1, x_2)$$

whenever $x_1 < y_1$ and $x_2 < y_2$. For these and related classes of utility functions, see Mosler (1984) and Scarsini (1985, 1988).

Each of the four orderings is preserved under projections. More precisely, for a given random vector X on $\mathbb{R}^k$ and a nonempty subset $I \subset \{1, \ldots, k\}$ let X_I denote the random vector $(X_i)_{i \in I}$. Then the following proposition holds.

Proposition 2. $X \preceq Y \Rightarrow X_I \preceq Y_I$, *where* $\preceq$ *is any of the four orthant orderings* $\preceq_{\mathrm{lo}}$, $\preceq_{\mathrm{uo}}$, $\preceq_{\mathrm{locc}}$, *and* $\preceq_{\mathrm{uocx}}$.

Proof. Assume that $X \preceq_{\mathrm{uocx}} Y$ holds. This is equivalent to $E(\prod_{j=1}^{k} \phi_j(X_j)) \leqslant E(\prod_{j=1}^{k} \phi_j(Y_j))$ for all nonnegative, increasing and convex functions $\phi_j : \mathbb{R} \to \mathbb{R}$. With $\phi_j(x_j) = 1$ for all $j \notin I$, we conclude $E(\prod_{j \in I} \phi_j(X_j)) \leqslant E(\prod_{j \in I} \phi_j(Y_j))$ for all nonnegative, increasing and convex functions ϕ_j. Hence $X_I \preceq_{\mathrm{uocx}} Y_I$. The proofs for the other three orderings are similar. □

Further, each of the four orderings implies the ordering of the corresponding expectation vectors in the componentwise partial order on $\mathbb{R}^k$.

Proposition 3. *$X \preceq Y \Rightarrow E(X) \leqslant E(Y)$, where $\preceq$ is any of the four orthant orderings $\preceq_{\text{lo}}$, $\preceq_{\text{uo}}$, $\preceq_{\text{locc}}$, $\preceq_{\text{uocx}}$, and $\leqslant$ is the usual partial order on $\mathbb{R}^k$.*

Proof. For $i = 1, \ldots, k$, consider the ith projection π_i. Clearly, $\pi_i = \prod_{j=1}^k \phi_j$ where $\phi_i(x) = x$ for all x and $\phi_j = 1$ if $j \neq i$. Each ϕ_j is nonnegative, increasing, and convex. Therefore, $X \preceq_{\text{uocx}} Y$ implies $E(X_i) = E(\pi_i(X)) \leqslant E(\pi_i(Y)) = E(Y_i)$ for all i, hence $E(X) \leqslant E(Y)$. For the other three orderings the proofs are similar. □

Proposition 3 gives a necessary condition for two random variables to be ordered in any of the four orthant orderings, that can easily be checked. We proceed with three applications.

Positive dependence of random variables. Let $k = 2$. X is *less positive dependent* than Y, $X \preceq_{\text{PD}} Y$, if $X \preceq_{\text{uo}} Y$ and $X_i =_{\text{st}} Y_i$ holds for $i = 1$ and 2, i.e. the univariate marginal distributions of X and Y are the same. An equivalent definition is $X \succeq_{\text{lo}} Y$ and $X_i =_{\text{st}} Y_i$, $i = 1, 2$. In particular, the random vector Y is called *positive dependent* if $F_{Y_1} F_{Y_2} \preceq_{\text{uo}} F_Y$. Here F_Y denotes the distribution function of Y, and F_{Y_1}, F_{Y_2} the marginals. Note that every sensible notion of an increasing dependence ordering for bivariate distributions (in the sense of Kimeldorf and Sampson, 1987) implies the above ordering $\preceq_{\text{PD}}$ and almost every known statistic that measures the degree of dependence in a bivariate random vector is increasing with respect to them. See also Nguyen and Sampson (1987). Bergmann (1978, 1991) investigates all four orthant orderings and their relations to increasing dependence.

Comparison of two samples with respect to location and dispersion: When the random vectors X and Y represent ordinal data, $X \preceq_{\text{uo}} Y$ can be seen as an ordinal shift in location from X to Y. The same interpretation applies to $X \preceq_{\text{lo}} Y$. Checking for these orderings then means checking for a location hypothesis. The same orderings may be used to describe increasing multivariate dispersion about a center. Let a and b two vectors in $\mathbb{R}^k$ that describe the location of X and Y, respectively. For example, a and b may be the expectations or proper medians. For $z \in \mathbb{R}^k$, $|z|$ denotes the vector $(|z_1|, \ldots, |z_k|)$. We say that X is *rectangular less peaked about a than Y about b* if

$$|Y - b| \preceq_{\text{lo}} |X - a|.$$

This notion is weaker than, i.e., implied by, Anderson's (1955) notion of being less peaked about a given point (which refers to all symmetric convex sets instead of rectangles only; see also Birnbaum (1948) for $k = 1$), but much easier to handle.

Percolation. The orderings (UO) and (LO) have also applications to the comparison of percolation probabilities. See Rüschendorf (1982).

So far we have considered arbitrary distributions. The subsequent results of the paper provide characterizations of the four orthant orderings when the distributions of X and Y are finitely discrete. However, when general distributions of X and Y are given, it is easily seen that the lower and the upper orthant ordering imply the same orderings for proper discretizations of X and Y. More precisely, let X follow a general distribution in $\mathbb{R}^k$, and let $a_j(i) \in \mathbb{R} \cup \{-\infty, \infty\}$, $i = 1, \ldots, m_j, j = 1, \ldots, k$, with $a_j(i) < a_j(i+1)$, $D_{i_1, \ldots, i_k} = \{x \in \mathbb{R}^k : a_j(i_j) \leqslant x_j < a_j(i_j + 1) \forall j\}$,

$$X_{\mathscr{D}} = (i_1, \ldots, i_k) \quad \text{if} \quad X \in D_{i_1, \ldots, i_k}, \quad \mathscr{D} = (D_{i_1, \ldots, i_k})_{i_j = 1, \ldots, m_j, j = 1, \ldots, k}.$$

The next proposition, the proof of which is obvious, tells that lower orthant as well as upper orthant ordering between X and Y carry over to their discretizations $X_{\mathscr{D}}$ and $Y_{\mathscr{D}}$.

Proposition 4. (i) $X \preceq_{\text{lo}} Y \;\Rightarrow\; X_{\mathscr{D}} \preceq_{\text{lo}} Y_{\mathscr{D}}$,
(ii) $X \preceq_{\text{uo}} Y \;\Rightarrow\; X_{\mathscr{D}} \preceq_{\text{uo}} Y_{\mathscr{D}}$.

3. Main results

In this section we present our characterization theorems. They give necessary and sufficient conditions for the four dominance relations when X and Y have finitely discrete distributions. Proofs are postponed to Section 4.

We start with an example which shows that for $k \geqslant 2$, contrary to the univariate case, checking (LO) only on the combined support[2] of X and Y, is not sufficient.

Example 2. Let $k = 2$ and

$$P[X = (0, 0)] = 0.5, \quad P[X = (2, 2)] = 0.5,$$

$$P[Y = (0, 2)] = 0.3, \quad P[Y = (1, 1)] = 0.4, \quad P[Y = (2, 0)] = 0.3.$$

If (LO) is checked on the combined support we get the following.

a	(0, 0)	(0, 2)	(1, 1)	(2, 0)	(2, 2)
$P[X \leqslant a] - P[Y \leqslant a]$	0.5	0.2	0.1	0.2	0.0

Thus, (LO) is satisfied on the combined support. However, for $a = (1, 2)$ we get

$$P[X \leqslant (1, 2)] - P[Y \leqslant (1, 2)] = 0.5 - 0.7 = -0.2.$$

This shows that $X \preceq_{\text{lo}} Y$ does not hold. We have to check additional points when to decide on dominance between X and Y.

We introduce some notation. Let $S \subset \mathbb{R}^k$ be the combined support of X and Y. For every $s \in S$ and $a \in \mathbb{R}^k$ we denote $\delta_s = P[X = s] - P[Y = s]$, $L(a) = \{s \in S : s \leqslant a\}$, $U(a) = \{s \in S : s \geqslant a\}$. We introduce Δ^{lo} and Δ^{uo} by

$$\Delta^{\text{lo}}(a) = \sum_{s \in L(a)} \delta_s = P[X \leqslant a] - P[Y \leqslant a],$$

[2] We use the term combined support to denote the union of the supports of X and Y.

$$\Delta^{\mathrm{uo}}(a) = \sum_{s \in U(a)} \delta_s = P[X \geqslant a] - P[Y \geqslant a].$$

Then, $X \preceq_{\mathrm{lo}} Y$ holds if and only if $\Delta^{\mathrm{lo}}(a) \geqslant 0$ for all a.

Let $K = \{1, \ldots, k\}$ and consider a nonempty subset $I \subset K$. For $a \in \mathbb{R}^k$, let $a_I = (a_i)_{i \in I}$. Similarly, for a given random vector X, we denote $X_I = (X_i)_{i \in I}$. The combined support of X_I and Y_I is denoted by S_I. Further, let $L_I(a) = \{s \in S \mid s_I \leqslant a_I\}$, $U_I(a) = \{s \in S \mid s_I \geqslant a_I\}$ and

$$\Delta_I^{\mathrm{locc}}(a) = \sum_{s \in L_I(a)} \delta_s \cdot \prod_{i \in I} (a_i - s_i),$$

$$\Delta_I^{\mathrm{uocx}}(a) = \sum_{s \in U_I(a)} \delta_s \cdot \prod_{i \in I} (s_i - a_i).$$

Given $x, y \in \mathbb{R}^k$, $x \vee y$ and $x \wedge y$ are defined by

$$x \vee y = (\max\{x_1, y_1\}, \ldots \max\{x_k, y_k\}), \tag{2}$$

$$x \wedge y = (\min\{x_1, y_1\}, \ldots \min\{x_k, y_k\}). \tag{3}$$

With the two operations $\wedge$ and $\vee$, $\mathbb{R}^k$ is a lattice. Let T be a finite set in $\mathbb{R}^k$, $T = \{t^1, t^2, \ldots t^m\}$. The vector

$$\bigvee_{t \in T} t = (\max\{t_1^1, t_1^2, \ldots, t_1^m\}, \max\{t_2^1, t_2^2, \ldots, t_2^m\}, \ldots, \max\{t_k^1, t_k^2, \ldots, t_k^m\}) \tag{4}$$

is called the *join* of T, while

$$\bigwedge_{t \in T} t = (\min\{t_1^1, t_1^2, \ldots, t_1^m\}, \min\{t_2^1, t_2^2, \ldots, t_2^m\}, \ldots, \min\{t_k^1, t_k^2, \ldots, t_k^m\}) \tag{5}$$

is called the *meet* of T.

The *join-semilattice* $J(S)$ of S is the smallest set containing S that is closed under the join operation $\vee$; see Birkhoff (1940). It consists of the joins of all finite sets in S. The *meet-semilattice* $M(S)$ of S is the smallest set that contains S and is closed under the meet operation $\wedge$. $M(S)$ consists of the meets of all finite sets in S.

Example 3. Let $S = \{(1, 3), (2, 2), (3, 4), (4, 1)\}$. Then the join-semilattice $J(S)$ consists of S and the additional points in $J(S) \setminus S = \{(2, 3), (4, 2), (4, 3), (4, 4)\}$. This is shown in the following figure.

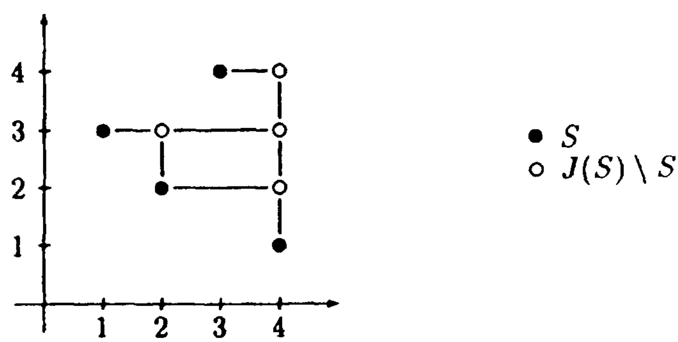

Fig. 1. The join-semilattice $J(S)$ generated by S.

The following two theorems present our main results. Theorem 1 says that, in checking (LO), $a \in J(S)$ is enough.

Theorem 1. *Let X and Y have finitely discrete distributions on* $\mathbb{R}^k$ *with combined support S. Then the following two conditions are equivalent:*

1. $X \preceq_{\mathrm{lo}} Y$.
2. $\Delta^{\mathrm{lo}}(a) \geqslant 0$ *for all* $a \in J(S)$.

In checking (LOCC), $a \in J(S)$ is not enough. We will demonstrate this by

Example 4. Let $k = 2$ and

$$P[X = (0,0)] = 0.5, \quad P[X = (0,1)] = 0.5, \quad P[Y = (1,0)] = 1.$$

Then, $J(S) = \{(0,0), (0,1), (1,0), (1,1)\}$. (LOCC) holds for all $a \in J(S)$. But for $a = (3,1)$ we get

$$\int_{]-\infty,(3,1)]} P[X \leqslant z] - P[Y \leqslant z] \, \mathrm{d}z = -0.5.$$

This shows that $X \preceq_{\mathrm{locc}} Y$ does not hold.

As we know from Proposition 2, $X \preceq_{\mathrm{locc}} Y$ implies $X_I \preceq_{\mathrm{locc}} Y_I$ for all marginals X_I and Y_I. Therefore, a necessary condition for $X \preceq_{\mathrm{locc}} Y$ is that for every I (LOCC) holds for X_I and Y_I at all $a \in J(S_I)$. Theorem 2 says that this condition is also sufficient:

Theorem 2. *Let X and Y have finitely discrete distributions on* $\mathbb{R}^k$ *with combined support S. Then the following two conditions are equivalent:*

1. $X \preceq_{\mathrm{locc}} Y$.
2. $\Delta_I^{\mathrm{locc}}(a) \geqslant 0$ *for all* $a \in J(S_I)$ *and all nonempty subsets I of K.*

Concerning the upper orderings $\preceq_{\mathrm{uo}}$ and $\preceq_{\mathrm{uocx}}$ we get two similar theorems.

Theorem 3. *Let X and Y have finitely discrete distributions on* $\mathbb{R}^k$ *with combined support S. Then the following two conditions are equivalent:*

1. $X \preceq_{\mathrm{uo}} Y$.
2. $\Delta^{\mathrm{uo}}(a) \leqslant 0$ *for all* $a \in M(S)$.

Theorem 4. *Let X and Y have finitely discrete distributions on* $\mathbb{R}^k$ *with combined support S. Then the following two conditions are equivalent:*

1. $X \preceq_{\mathrm{uocx}} Y$.
2. $\Delta_I^{\mathrm{uocx}}(a) \leqslant 0$ *for all* $a \in M(S_I)$ *and all nonempty subsets I of K.*

4. Proofs

Proof of Theorem 1. $X \preceq_{\text{lo}} Y$ is equivalent to $\Delta^{\text{lo}}(a) \geqslant 0$ for all $a \in \mathbb{R}^k$. This follows from (LO) and the equation

$$\Delta^{\text{lo}}(a) = \sum_{s \in L(a)} \delta_s = \sum_{s \in L(a)} P[X = s] - \sum_{s \in L(a)} P[Y = s] = P[X \leqslant a] - P[Y \leqslant a].$$

We have to show that $\Delta^{\text{lo}}(a) \geqslant 0$ for all $a \in J(S)$ is sufficient for $X \preceq_{\text{lo}} Y$ to hold.

For $a \in \mathbb{R}^k$ define $\underline{a}$ by

$$\underline{a} = \begin{cases} \bigvee_{s \in L(a)} s & \text{if } L(a) \neq \emptyset, \\ -\infty & \text{if } L(a) = \emptyset. \end{cases}$$

If $L(a) = \emptyset$ then clearly $\Delta^{\text{lo}}(a) = 0$. In the case $L(a) \neq \emptyset$ we have $L(a) = L(\underline{a})$ and therefore $\Delta^{\text{lo}}(a) = \Delta^{\text{lo}}(\underline{a})$. This implies that $X \preceq_{\text{lo}} Y$ if and only if

$$\Delta^{\text{lo}}(\underline{a}) \geqslant 0 \quad \text{for all } a \in \mathbb{R}^k \text{ such that } L(a) \neq \emptyset.$$

It remains to show that the set of all these $\underline{a}$ is given by $J(S)$. If $L(a) \neq \emptyset$, then $\underline{a}$ is the join of finitely many points in S and thus is contained in $J(S)$. Conversely, if $a \in J(S)$ then $a = \bigvee_{t \in T} t$ for some nonempty subset T of S. But in this case $a = \underline{a}$ and thus a is contained in the set of all $\underline{a}$. □

Proof of Theorem 2. We proceed in two steps.

Step 1. $X_I \preceq_{\text{locc}} Y_I$ is equivalent to $\Delta_I^{\text{locc}}(a) \geqslant 0$ for all $a \in \mathbb{R}^k$. This can be seen from (LOCC) and the following equation.

$$\int_{]-\infty_I, a_I]} P[X_I \leqslant z] - P[Y_I \leqslant z] \, \mathrm{d}z = \Delta_I^{\text{locc}}(a).$$

In the rest of the proof we use the following notation. If $I = K \setminus \{m\}$, we write shortly a_{-m}, X_{-m}, S_{-m} in place of $a_{K \setminus \{m\}}$, $X_{K \setminus \{m\}}$, $S_{K \setminus \{m\}}$, respectively.

Step 2. $X \preceq_{\text{locc}} Y$ if and only if the following two conditions hold:

1. $\Delta_K^{\text{locc}}(a) \geqslant 0$ for all $a \in J(S)$,
2. $\Delta_{-m}^{\text{locc}}(a) \geqslant 0$ for all $a \in \mathbb{R}^k$, $m = 1, \ldots, k$.

Since the second condition is equivalent to $X_{-m} \preceq_{\text{locc}} Y_{-m}$, $m = 1, \ldots, k$, Theorem 2 then follows by induction.

We only have to show the 'if'-part. The 'only if'-part follows from Proposition 2. We introduce some notation. For $i \in \{0. \ldots n\}$ let

$$C_i = \{x \in \mathbb{R}^k \mid \exists c \in J(S): c \leqslant x, c_j = x_j, j = 1, \ldots, i\}.$$

Clearly,

$$J(S) = C_k \subset C_{k-1} \subset \cdots C_0 = \{x \in \mathbb{R}^k \mid \exists c \in J(S): c \leqslant x\}.$$

Since $\Delta_K^{\text{locc}}(a) = 0$ for all $a \in \mathbb{R}^k \backslash C_0$, it suffices to demonstrate that $\Delta_K^{\text{locc}}(a) \geqslant 0$ holds for $a \in C_0$. For $a \in C_k$ it is assumed that $\Delta_K^{\text{locc}}(a) \geqslant 0$. Now assume that $\Delta_K^{\text{locc}}(a) \geqslant 0$ is true for all $a \in C_m$, where $1 \leqslant m \leqslant k$. We have to show that this implies that $\Delta_K^{\text{locc}}(a) \geqslant 0$ for all $a \in C_{m-1}$. Then the assertion follows by induction.

Let $a \in C_{m-1}$. We consider the sets

$$U = \{x \in J(S) \mid x \leqslant a, x_j = a_j, j = 1, \ldots, m-1\},$$

$$V = \{x \in J(S) \mid x_j = a_j, j = 1, \ldots, m-1, x_m > a_m, x_j \leqslant a_j, j = m+1, \ldots, k\}.$$

Since $a \in C_{m-1}$ the set U is nonempty. We choose $u \in U$ such that u_m is maximum. V may be empty or not. Thus we have to distinguish two cases.

Case 1. $V \neq \emptyset$. We choose $v \in V$ such that v_m is minimum. For $\lambda \in [0, 1]$ define a^λ to be

$$a^\lambda = (a_{-m}, (1-\lambda)u_m + \lambda v_m).$$

Note that $a^\circ \leqslant a \leqslant a^1$ and $a^0, a^1 \in C_m$.

Let $\lambda \in [0, 1]$ be such that $a = a^\lambda$. Then

$$\Delta_K^{\text{locc}}(a) = (1-\lambda) \sum_{s \in L(a)} \delta_s \prod_{j=1}^{k} (a_j^0 - s_j) + \lambda \sum_{s \in L(a)} \delta_s \prod_{j=1}^{k} (a_j^1 - s_j).$$

To complete the proof in Case 1 it remains to show that

$$\sum_{s \in L(a)} \delta_s \prod_{j=1}^{k} (a_j^0 - s_j) = \Delta_K^{\text{locc}}(a^0), \tag{6}$$

$$\sum_{s \in L(a)} \delta_s \prod_{j=1}^{k} (a_j^1 - s_j) = \Delta_K^{\text{locc}}(a^1). \tag{7}$$

Since a^0, $a^1 \in C_m$ it then follows from the induction hypothesis that $\Delta_K^{\text{locc}}(a^0)$, $\Delta_K^{\text{locc}}(a^1) \geqslant 0$ and thus that $\Delta_K^{\text{locc}}(a) \geqslant 0$

To prove (6) it suffices to show that $L(a) = L(a^0)$. a and a^0 coincide in all coordinates except the mth coordinate $a_m^0 = u_m$. If there was an $s \in L(a) \backslash L(a^0)$ this would contradict the maximality of u_m.

In the case of (7) it would again suffice to show that $L(a^1) = L(a)$. Unfortunately, this need not hold. $L(a^1)$ may be strictly larger than $L(a)$. However, for $s \in L(a^1) \backslash L(a)$ it holds that $s_m = v_m$ and thus $\prod_{j=1}^k (a_j^1 - s_j) = 0$. Thus (7) holds true.

Case 2. $V = \emptyset$. Let u_m be as in Case 1. For $\lambda \in [0, \infty)$ we define a^λ to be

$$a^\lambda = (a_{-m}, u_m + \lambda).$$

Again, $a^0 \in C_m$. Now, choose $\lambda \geqslant 0$ such that $a = a^\lambda$. Then

$$\Delta_K^{\text{locc}}(a) = \sum_{s \in L(a)} \delta_s \prod_{j=1}^{k} (a_j^0 - s_j) + \lambda \sum_{s \in L(a)} \delta_s \prod_{j \neq m} (a_j - s_j).$$

As in Case 1 we prove that

$$\sum_{s\in L(a)} \delta_s \prod_{j=1}^{k} (a_j^0 - s_j) = \Delta_K^{\mathrm{locc}}(a^0) \geqslant 0.$$

It remains to show that

$$\sum_{s\in L(a)} \delta_s \prod_{j\neq m} (a_j - s_j) = \Delta_{-m}^{\mathrm{locc}}(a). \tag{8}$$

Since $\Delta_{-m}^{\mathrm{locc}}(a) \geqslant 0$ for all $m = 1, \ldots, k$ and $a \in \mathbb{R}^k$, it then follows that $\Delta_K^{\mathrm{locc}}(a) \geqslant 0$. To prove (8) it is sufficient to show that $L(a) = L_{-m}(a)$. But since V is empty, there is no $s \in S$ such that $s_{-m} \leqslant a_{-m}$ and $s_m > a_m$. This completes the proof of Step 2. □

In the proof of Theorem 3 and Theorem 4 we use the following notation. For a set $A \subset \mathbb{R}^k$, we denote $-A = \{x \in \mathbb{R}^k \mid -x \in A\}$.

Proof of Theorem 3. According to Proposition 1, $X \preceq_{\mathrm{uo}} Y \Leftrightarrow -Y \preceq_{\mathrm{lo}} -X$. The combined support of $-X$ and $-Y$ is given by $-S$. Thus it follows from Theorem 1 that

$$X \preceq_{\mathrm{uo}} Y \quad\Leftrightarrow\quad \sum_{s\in L^-(a)} -\delta_{-s} \geqslant 0 \quad \text{for all } a \in J(-S),$$

where $L^-(a) = \{s \in -S \mid s \leqslant a\}$. Since $-x \vee -y = -(x \wedge y)$, it is obvious that $J(-S) = -M(S)$. Therefore

$$X \preceq_{\mathrm{uo}} Y \quad\Leftrightarrow\quad \sum_{s\in L^-(-a)} \delta_{-s} \leqslant 0 \quad \text{for all } a \in M(S).$$

But $L^-(-a) = \{s \in -S \mid s \leqslant -a\} = -\{s \in S \mid -s \leqslant -a\} = -U(a)$. Thus

$$X \preceq_{\mathrm{uo}} Y \Leftrightarrow \sum_{s\in -U(a)} \delta_{-s} \leqslant 0 \quad \text{for all } a \in M(S)$$

$$\Leftrightarrow \sum_{s\in U(a)} \delta_s \leqslant 0 \quad \text{for all } a \in M(S),$$

as was to be shown. □

Proof of Theorem 4. Again from Proposition 1 it follows that

$$X \preceq_{\mathrm{uocx}} Y \quad\Leftrightarrow\quad \sum_{s\in L_I^-(a)} -\delta_{-s} \cdot \prod_{i\in I} (a_i - s_i) \geqslant 0 \quad \text{for all } a \in J(-S_I), \quad \emptyset \neq I \subset K.$$

As in the proof of Theorem 3 we get

$$X \preceq_{\text{uocx}} Y \quad \Leftrightarrow \quad \sum_{s \in L_I^-(-a)} \delta_{-s} \cdot \prod_{i \in I} (-a_i - s_i) \leqslant 0 \quad \text{for all } a \in M(S_I), \emptyset \neq I \subset K$$

and finally

$$X \preceq_{\text{uocx}} Y \quad \Leftrightarrow \quad \sum_{s \in U_I(a)} \delta_s \cdot \prod_{i \in I} (s_i - a_i) \leqslant 0 \quad \text{for all } a \in M(S_I), \quad \emptyset \neq I \subset K,$$

which completes the proof of Theorem 4. □

5. Conclusions

We start this section with two remarks on the size and structure of the set $J(S)$. Analogous remarks apply to the set $M(S)$. As has been said before, $J(S)$ consists of all joins of finitely many points in S. To construct $J(S)$ we thus have to consider all subsets of S and construct the join of each subset. However, this task is simplified by making use of the special structure of $J(S)$. Let $\{t^1, \ldots, t^m\}$ be a subset of S. The join of the m points is given by (4):

$$\bigvee_{i=1}^{m} t^i = (\max\{t_1^1, t_1^2, \ldots, t_1^m\}, \ldots, \max\{t_k^1, t_k^2, \ldots, t_1^m\})$$

For each coordinate there is one of the t^i's that maximizes this coordinate. Thus we can construct a set of at most k points that has the same join as $\{t^1, \ldots, t^m\}$. This gives rise to the following remark.

Remark. Let S be a finite subset of $\mathbb{R}^k$. Then $J(S)$ consists of all joins of at most k elements in S.

Thus to every subset of S with at most k elements there corresponds an element of $J(S)$. This shows that the size of $J(S)$ cannot exceed $\sum_{d=1}^{k} \binom{n}{d}$ where n is the number of points in the combined support S. If k is small compared to n, checking (LO) only on $J(S)$ (and not on the whole grid) results in a considerable reduction of computations. In checking (LOCC) much more work has to be done since all marginal distributions have to be checked. There are $\binom{k}{i}$ marginal distributions of dimension i and for every such marginal distribution at most $\sum_{d=1}^{i} \binom{n}{d}$ checks have to be made. Thus the maximum number of checks does not exceed $\sum_{i=1}^{k} \sum_{d=1}^{i} \binom{k}{i} \binom{n}{d}$. Based on the results of Section 3 a fast algorithm has been developed to check whether, given two finitely discrete k-variate random vectors X and Y, X dominates Y in one of the four orthant orderings; see Dyckerhoff, Holz and Mosler (1995). In a setting of multiattribute decision under risk the algorithm has been used to determine an efficient (with respect to one of the orderings) set of distributions from a given finite set of distributions. The algorithm has also been used to perform statistical tests on $\preceq_{lo}$ and $\preceq_{uo}$ that employ resampling techniques.

Acknowledgements

We would like to thank an anonymous referee and an associate editor for their careful reading and valuable remarks.

References

Anderson, T.W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6**, 170–176.

Bergmann, R. (1978). Some classes of semi-ordering relations for random vectors and their use for comparing covariances. *Math. Nachr.* **82**, 103–114.

Bergmann, R. (1991). Stochastic orders and their applications to a unified approach to various concepts of dependence and association. In: *Stochastic Orders and Decision under Risk*. Eds., K. Mosler and M. Scarsini, Institute of Mathematical Statistics, Hayward, CA, 48–73.

Birkhoff, G. (1940). *Lattice Theory*. American Mathematical Society, Providence, RI.

Birnbaum, Z.W. (1948). On random variables with comparable peakedness. *Ann. Math. Statist.* **19**, 76–81.

Dyckerhoff, R., H. Holz and K. Mosler (1995). Checking for orthant orderings between discrete multivariate distributions. Algorithm. Mimeo.

Kamae, T., U. Krengel and G.L. O'Brian (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5**, 899–912.

Kimeldorf, G. and A.R. Sampson (1987). Positive dependence orderings. *Ann. Inst. Statist. Math.* **39**, 113–128.

Lehmann, E.L. (1955). Ordered families of distributions. *Ann. Math. Statist.* **26**, 399–419.

Mosler, K. (1984). Stochastic dominance decision rules when the attributes are utility independent. *Management Science* **30**, 1311–1322.

Mosler, K. and M. Scarsini, Eds. (1991a). *Stochastic Orders and Decision under Risk*. Institute of Mathematical Statistics, Hayward, CA.

Mosler, K. and M. Scarsini (1991b). Some theory of stochastic dominance. In: K. Mosler and M. Scarsini, Eds., *Stochastic Orders and Decision under Risk*, Institute of Mathematical Statistics, Hayward, CA, 261–284.

Nguyen, T.T. and A.R. Sampson (1987). Testing for positive quadrant dependence in ordinal contingency tables. *Naval Res. Logist.* **34**, 859–877.

Rüschendorf, L. (1980). Inequalities for the expectations of Δ-monotone functions. *Z. Wahrscheinlichkeitstheorie verwandte Gebiete* **54**, 341–349.

Rüschendorf, L. (1982). Comparison of percolation probabilities. *J. Appl. Probab.* **19**, 864–868.

Scarsini, M. (1985). Stochastic dominance with pair-wise risk aversion. *J. Math. Econom.* **14**, 187–201.

Scarsini, M. (1988). Dominance conditions for multivariate utility functions. *Management Sci.* **34**, 454–460.

Shaked, M. and J.G. Shanthikumar (1994). *Stochastic Orders and Their Applications*. Academic Press, Boston.

Stoyan, D. (1983). *Comparison Methods for Queues and Other Stochastic Models*. Wiley, Chichester.