J. Appl. Prob. **17**, 112–123 (1980)
Printed in Israel
0021-9002/80/010112-12 $01.45

UNIFORM CONDITIONAL STOCHASTIC ORDER

WARD WHITT,* *Bell Laboratories*

Abstract

One probability measure is less than or equal to another in the sense of UCSO (uniform conditional stochastic order) if a standard form of stochastic order holds for each pair of conditional probability measures obtained by conditioning on appropriate subsets. UCSO can be applied to the comparison of lifetime distributions or the comparison of decisions under uncertainty when there may be reductions in the set of possible outcomes. When densities or probability mass functions exist on the real line, then the main version of UCSO is shown to be equivalent to the MLR (monotone likelihood ratio) property. UCSO is shown to be preserved by some standard probability operations and not by others.

STOCHASTIC ORDER; CONDITIONAL PROBABILITY; MONOTONE LIKELIHOOD RATIO

1. Motivation and basic properties

The purpose of this paper is to introduce and investigate an interesting and potentially useful concept of stochastic order. Our framework is the set $\Pi(S)$ of all probability measures on a sample space S and a set $\mathcal{U}$ of real-valued 'evaluation' functions on $\Pi(S)$. The set $\mathcal{U}$ induces a partial order (reflexive and transitive binary relation) on $\Pi(S)$: $P_1 \leqq_{\mathcal{U}} P_2$ if $u(P_1) \leqq u(P_2)$ for all $u \in \mathcal{U}$. This general framework was used by Goroff and Whitt (1978) in a study of continuity properties of admissible sets in stochastic dominance. Usually, $u(P)$ represents the expectation with respect to P of a real-valued 'utility' function on S, and we use u to represent simultaneously the function on S and the function on $\Pi(S)$: $u(P) \equiv \int u dP$. For example, the set $\mathcal{U}$ of real-valued functions on S might contain all non-decreasing functions (assuming S is partially ordered), all non-decreasing concave functions (assuming S is also convex) or all concave functions, cf. Bawa (1975), Brumelle and Vickson (1975), Kamae, Krengel and O'Brien (1977), J. Meyer (1977), Chapter XI of P. A. Meyer (1966) and Strassen (1965). However, the set $\mathcal{U}$ might also contain functions on $\Pi(S)$ which cannot be represented as expectations of real-valued functions on S, e.g., the median and the variance.

Received 27 April 1978; revision received 26 January 1979.
* Postal address: Bell Laboratories, Holmdel, NJ 07733, U.S.A.

In this paper, we consider a stronger notion of stochastic order, which we call ***uniform conditional stochastic order*** (UCSO). In order to state the definition, let P_A represent the probability measure in $\Pi(S)$ obtained by conditioning on A, i.e., $P_A(B) = P(A \cap B)/P(A)$ for $P(A) > 0$.

Definition 1.1. Let $\mathscr{C}$ be a collection of measurable subsets of S. We say that P_1 is less than or equal to P_2 in $\Pi(S)$ in the sense of UCSO with respect to the pair $(\mathscr{U}, \mathscr{C})$, denoted by $P_1 \leqq_{\mathscr{U},\mathscr{C}} P_2$, if $P_{1A} \leqq_{\mathscr{U}} P_{2A}$ for all $A \in \mathscr{C}$ such that $P_1(A) > 0$ and $P_2(A) > 0$. We write $\leqq_{\mathscr{C}}$ when $\mathscr{U}$ is understood to be the set $\mathscr{U}_1$ of all non-decreasing bounded measurable real-valued functions on S, where S is endowed with a partial order $\leqq$; we write $\leqq_{\mathscr{U}}$ when $\mathscr{C}$ is understood to be $\{S\}$; and we write $\leqq_{\text{st}}$ when $\mathscr{U} = \mathscr{U}_1$ and $\mathscr{C} = \{S\}$.

In this section, we discuss the motivation for UCSO and some of its basic properties. In Section 2 we see to what extent UCSO is preserved under standard probability operations. Finally, we present all proofs in Section 3.

Obviously $P_1 \leqq_{\mathscr{U}_1,\mathscr{C}_1} P_2$ if $P_1 \leqq_{\mathscr{U}_2,\mathscr{C}_2} P_2$ with $\mathscr{U}_1 \subseteq \mathscr{U}_2$ and $\mathscr{C}_1 \subseteq \mathscr{C}_2$, which implies that $P_1 \leqq_{\text{st}} P_2$ whenever $P_1 \leqq_{\mathscr{C}} P_2$ and $S \in \mathscr{C}$. The following elementary example illustrates the difference between $\leqq_{\text{st}}$ and $\leqq_{\mathscr{C}}$ for $\mathscr{C} = \mathscr{S}$, the set of all measurable subsets of S.

Example 1.1. Let $S = \{1, 2, 3\}$ with the usual ordering. Let $P_1(\{1\}) = P_1(\{2\}) = P_1(\{3\}) = \frac{1}{3}$ and $P_2(\{1\}) = \frac{1}{4}$, $P_2(\{2\}) = \frac{1}{8}$ and $P_2(\{3\}) = \frac{5}{8}$. Obviously, $P_1 \leqq_{\text{st}} P_2$, but $P_{1A} \geqq_{\text{st}} P_{2A}$ for $A = \{1, 2\}$. The large mass at 3 in P_2 compensates for the smaller mass at 2 to get $P_1 \leqq_{\text{st}} P_2$.

To see how UCSO might be applied, consider the problem of choosing between two possible actions. Suppose the set of possible states of the world is a partially ordered set $(S, \leqq)$ and each action determines a probability measure on $(S, \leqq)$. The probability measure P_1 is preferred to the probability measure P_2 for all possible decision-makers if the expected utility with respect to P_1 is greater than or equal to the expected utility with respect to P_2 for all non-decreasing utility functions, i.e., if $P_1 \geqq_{\text{st}} P_2$. However, suppose additional information becomes available which limits the possible states of the world to a subset A in a class $\mathscr{C}$, so that the comparison is between P_{1A} and P_{2A}. If $P_1 \geqq_{\mathscr{C}} P_2$, then P_{1A} always remains preferred to P_{2A}, which is not true for ordinary stochastic order. Thus, just as ordinary stochastic order implies preference for all decision-makers with non-decreasing utility functions, so UCSO implies preference for all reductions of the space of possible outcomes in a specified class.

Another application is to lifetime distributions.

Example 1.2. Suppose that two items go into operation at the same time and we wish to compare their lifetime distributions. If $P_1 \leqq_{\mathscr{C}} P_2$ for $S = [0, \infty)$ and $\mathscr{C} = \{[t, \infty): t \geqq 0\}$, then all *conditional* lifetime distributions are ordered. Regardless of the age, the first item is more likely to fail first. Obviously, if $\bar{F}_i(t) = P_i\{[t, \infty)\}$, then $P_1 \leqq_{\mathscr{C}} P_2$ if and only if $\bar{F}_2(s)\bar{F}_1(t) \leqq \bar{F}_1(s)\bar{F}_2(t)$, $0 \leqq s < t$.

It is also worth mentioning another definition which might be useful in this setting. Let $P_{it}(A) = P_i(t + A \mid [t, \infty))$ for measurable $A \subseteq [0, \infty)$. Then $P_1 \leqq_{\mathscr{C}} P_2$ above is equivalent to $P_{1t} \leqq_{st} P_{2t}$ for all t. A stronger ordering would be $P_{1s} \leqq_{st} P_{2t}$ for all $s, t \geqq 0$, which says that the first item is more likely to fail first, regardless of the ages, even if the ages are different. For example this ordering holds for two exponential distributions. Note that the familiar new-better-than-used (NBU) property can be expressed as $P_{1t} \leqq_{st} P_{10}$ for all $t \geqq 0$ and the familiar decreasing failure rate (DFR) property can be expressed as $P_{1s} \leqq_{st} P_{1t}$ for $0 \leqq s < t$, cf. Barlow and Proschan (1975).

The following example illustrates UCSO applied to a different context.

Example 1.3. Let $S = (0, \infty)$ and let $\mathscr{C}$ be the set of all subintervals of S. Consider probability measures with continuous distribution functions F_i that are strictly increasing on their support. We say that F_1 is convex with respect to F_2 and write $F_1 \leqq_c F_2$ if $F_2^{-1}(F_1(x))$ is a convex function on the support of F_1; see p. 106 of Barlow and Proschan (1975). This example illustrates an ordering not defined in terms of a set $\mathscr{U}$ of evaluation functions, but we can still define UCSO. In this case, let $F_1 \leqq_{c,\mathscr{C}} F_2$ mean that $F_{2A_2}^{-1}(F_{1A_1}(x))$ is convex on the support of $F_{1A_1}(x)$ for all $A_1, A_2 \in \mathscr{C}$, where $F_{iA_i}(x)$ is the distribution function of the probability measure P_{iA_i}. It is not difficult to see that the ordering $\leqq_{c,\mathscr{C}}$ is strictly stronger than the ordering $\leqq_c$. To see this, suppose

$$F_1(x) = \begin{cases} \log(1+x), & 0 \leqq x \leqq e^{\frac{1}{2}} - 1, \\ x - e^{\frac{1}{2}} + \frac{3}{2}, & e^{\frac{1}{2}} - 1 \leqq x \leqq e^{\frac{1}{2}} - \frac{1}{2}, \end{cases}$$

and

$$F_2^{-1}(y) = \begin{cases} e^y - 1, & 0 \leqq y \leqq \frac{1}{2}, \\ y + e^{\frac{1}{2}} - \frac{3}{2}, & \frac{1}{2} \leqq y \leqq 1. \end{cases}$$

Then $F_2^{-1}(F_1(x)) = x$, $0 \leqq x \leqq e^{\frac{1}{2}} - \frac{1}{2}$ which is of course convex. However, $F_{2A_2}^{-1}(F_{1A_1}(x))$ is concave for $A_1 = [0, e^{\frac{1}{2}} - 1]$ and $A_2 = [e^{\frac{1}{2}} - 1, e^{\frac{1}{2}} - \frac{1}{2}]$. Notice that the exponential part of F_2^{-1} just compensates for the logarithm part of $F_1(x)$, which is lost with the conditioning. We now briefly indicate a few positive properties. For $A_i = (a_i, b_i]$,

$$F_{iA_i}(x)=\frac{F_i(x)-F_i(a_i)}{F_i(b_i)-F_i(a_i)}, \qquad a_i \leqq x \leqq b_i,$$

$$F_{iA_i}^{-1}(y)=F_i^{-1}([F_i(b_i)-F_i(a_i)]y+F_i(a_i)), \qquad 0 \leqq y \leqq 1,$$

and

$$F_{2A_2}^{-1}(F_{1A_1}(x))=F_2^{-1}(c+dF_1(x)), \qquad a_1 \leqq x \leqq b_1,$$

where

$$c=c(a_1,b_1,a_2,b_2)=\frac{F_2(a_2)F_1(b_1)-F_2(b_2)F_1(a_1)}{F_1(b_1)-F_1(a_1)}$$

and

$$d=d(a_1,b_1,a_2,b_2)=\frac{F_2(b_2)-F_2(a_2)}{F_1(b_1)-F_1(a_1)}.$$

In order to have $F_1 \leqq_{c,\mathscr{C}} F_2$, it is thus sufficient to have $F_2^{-1}(c+dF_1(x))$ convex for all c and d such that $0 \leqq c+dF_1(x) \leqq 1$. It is easy to see that it is sufficient to have F_1 convex on its support and F_2 concave on its support because the convexity is inherited by conditioning to subintervals and F_2^{-1} is convex whenever F_2 is concave.

We now return to the setting with a set $\mathscr{U}$ of evaluation functions. If the sample space S is a subset of the real line and the probability measures have densities or probability mass functions, then we shall show that UCSO using the set $\mathscr{U}_1$ of all non-decreasing functions is equivalent to the monotone likelihood ratio (MLR) property which arises in the study of uniformly most powerful statistical tests, cf. p. 208 of Ferguson (1967) and references there, which in turn is equivalent to total positivity of order 2 (TP_2), cf. Karlin (1968).

Definition 1.2. P_1 is less than or equal to P_2 in the sense of MLR, and we write $P_1 \leqq_r P_2$, if P_1 and P_2 are absolutely continuous with respect to a σ-finite measure μ on S with Radon–Nikodym derivatives p_1 and p_2 such that there exists $B \subseteq S$ with $P_1(S-B)=P_2(S-B)=0$ and $p_2(s)/p_1(s)$ is non-decreasing on B.

It is well known that MLR is stronger than ordinary stochastic order for probability measures on the real line, cf. Lehmann (1955), but we believe the relations between MLR and UCSO established in Theorems 1.1–1.3 below are new. For this discussion, assume S is a separable metric space with associated Borel σ-field $\mathscr{S}$.

Theorem 1.1. Suppose S is totally ordered and P_1 and P_2 are absolutely continuous with respect to a common σ-finite measure μ on S. If $P_1 \leqq_r P_2$, then $P_1 \leqq_{\mathscr{C}} P_2$ for $\mathscr{C}=\mathscr{S}$.

Since many parametric families of probability distributions are known to have the MLR property (p. 208 of Ferguson (1967)), we have many examples where UCSO holds. Unfortunately, this characterization does not extend to partially ordered subsets. With reasonable regularity conditions, UCSO with $\mathcal{U} = \mathcal{U}_1$ and $\mathcal{C} = \mathcal{S}$ implies MLR, but MLR fails to imply even ordinary stochastic order.

Example 1.4. Let $S = \{(0,1),(1,0),(1,1)\}$ with the usual ordering in the plane; let $p_1((0,1)) = 0.3$, $p_1((1,0)) = 0.4$, $p_1((1,1)) = 0.3$; and let $p_2((1,0)) = 0$, $p_2((0,1)) = p_2((1,1)) = 0.5$. Then $p_2(s)/p_1(s)$ is increasing, but $p_1 \leqq_{st} p_2$ fails to hold: $P_1(A) \geqq P_2(A)$ for $A = \{(1,0),(1,1)\}$.

Before turning to conditions under which UCSO implies MLR, it is important to note that UCSO is very difficult to achieve on non-totally ordered spaces if $\mathcal{C}$ is not restricted.

Example 1.5. Suppose S is an $n \times n$ integer lattice in the plane, i.e., $S = \{(i,j): 1 \leqq i \leqq n, 1 \leqq j \leqq n\}$ with the usual ordering in R^2. If p_1 is the uniform probability mass function on S, then to have $p_1 \leqq_{\mathcal{C}} p_2$ for $\mathcal{C} = \mathcal{S}$, where $p_1 \leqq_{\mathcal{C}} p_2$ means $P_1 \leqq_{\mathcal{C}} P_2$ for the measures associated with the mass functions, it is necessary and sufficient for a probability mass function p_2 to satisfy: (i) $p_2((i,j)) = p_2((2,1))$ for all (i,j) except $(1,1)$ and (n,n) and (ii) $p_2((1,1)) \leqq p_2((2,1)) \leqq p_2((n,n))$. This is easy to see by considering two point sets $\{(i,j),(k,l)\}$ with $i < k$ and $j > l$.

Example 1.5 suggests that if S is a lattice then $\mathcal{C}$ might be the set of all measurable sublattices or if $S = R^n$ then $\mathcal{C}$ might be the set of all possible measurable rectangles $A_1 \times \cdots \times A_n$, $A_i \subseteq R$. Even when $S \subseteq R^2$ and $\mathcal{C}$ is the set of all rectangles, strange behavior can occur: if p_1 is a probability mass function and an atom of mass is moved to a higher point, the new measure need not be bigger in the $\mathcal{C}$-order, as the following example shows.

Example 1.6. Let $S = \{(0,0),\ (0,1),\ (1,0),\ (1,1)\}$, $p_1((0,0)) = p_1((0,1)) = \frac{1}{2}$, $p_2((0,0)) = p_2((1,1)) = \frac{1}{2}$ and $\mathcal{C}$ be the set of rectangles. Then $p_1 \leqq_{\mathcal{C}} p_2$ does not hold: consider $A = \{(0,0),(0,1)\}$.

It is known that $(\Pi(S), \leqq_{st})$ is not necessarily a lattice when $\leqq$ is a partial order on S, even if $(S, \leqq)$ is a finite lattice, as the following example illustrates.

Example 1.7. Let S be as in Example 1.6. Let $p_1((0,0)) = p_1((0,1)) = p_2((0,0)) = p_2((1,0)) = p_3((1,0)) = p_3((0,1)) = p_4((0,0)) = p_4((1,1)) = \frac{1}{2}$. Then p_3 and p_4 are upper bounds for p_1 and p_2, but there is no least upper bound.

We have yet to determine whether $(\Pi(S), \leqq_{\mathcal{C}})$ is a lattice for appropriate S, $\leqq$ and $\mathcal{C}$. However, even when S is the set of positive integers, so that $(\Pi(S), \leqq_{st})$ is a lattice, $(\Pi(S), \leqq_r)$ need not be a lattice.

Example 1.8. Let S be the set of positive integers and let $P_1(\{2n\}) = P_2(\{2n-1\}) = 2^{-2n}$ and $P_1(\{2n-1\}) = P_2(\{2n\}) = 2^{-(2n-1)}$, $n \geqq 1$. Then P_1 and P_2 have no common upper bound in $(\Pi(S), \leqq_r)$ because any upper bound P_3 must have $P_3(\{n+1\})/P_3(\{n\}) \geqq \max P_1(\{n+1\})/P_1(\{n\}), P_2(\{n+1\})/P_2(\{n\}) \geqq 2$ for all $n \geqq 1$.

We now return to investigate the relationship between MLR and UCSO.

Theorem 1.2. If S is countable with a partial order, then the following are equivalent:
(i) $P_1 \leqq_r P_2$,
(ii) $P_1 \leqq_{\mathscr{C}} P_2$ for $\mathscr{C}$ consisting of all ordered two-point subsets, and
(iii) $P_1 \leqq_{\mathscr{C}} P_2$ for $\mathscr{C}$ consisting of all totally ordered subsets.

To state regularity conditions under which UCSO implies MLR on more general spaces, let $I(A)$ and $D(A)$ be the increasing and decreasing hull of A in the partially ordered space $(S, \leqq)$, defined by

$$I(A) = \{s \in S: s \geqq s' \quad \text{for some } s' \in A\}$$

and

$$D(A) = \{s \in S: s \leqq s' \quad \text{for some } s' \in A\}.$$

Also assume that $\leqq$ is a closed partial order, i.e., the graph $\Gamma = \{(s_1, s_2): s_1 \leqq s_2\}$ is a closed subset of S^2 in the product topology. Since the partial order is closed, $I(K)$ and $D(K)$ are closed subsets of S for each compact subset K, cf. p. 44 of Nachbin (1965). Let

$$I(s, \varepsilon) = \{s' \in I(\{s\}): d(s, s') \leqq \varepsilon\}$$

and

$$D(s, \varepsilon) = \{s' \in D(\{s\}): d(s, s') \leqq \varepsilon\},$$

where d is the metric on S. Since $I(\{s\})$ and $D(\{s\})$ are closed subsets of S, so are $I(s, \varepsilon)$ and $D(s, \varepsilon)$, which of course implies that they are measurable.

Theorem 1.3. Suppose P_1 and P_2 are absolutely continuous with respect to a σ-finite measure μ on S and there exists a subset B with $P_1(S-B) = P_2(S-B) = 0$ such that $\mu(I(s, \varepsilon)) > 0$ and $\mu(D(s, \varepsilon)) > 0$ for all $\varepsilon > 0$ and $s \in B$. Suppose $D(s_1, \varepsilon) \cup I(s, \varepsilon) \in \mathscr{C}$ for all $s_1 \leqq s_2$ and all $\varepsilon > 0$ sufficiently small. If either (i) the densities of P_1 and P_2 with respect to μ have continuous versions on B or (ii) S is locally compact, then $P_1 \leqq_r P_2$ whenever $P_1 \leqq_{\mathscr{C}} P_2$.

Remark. The regularity conditions in Theorem 1.3 (ii) are obviously satisfied when S is Euclidean space R^n, μ is Lebesgue measure and $\mathscr{C}$ contains unions of two ordered rectangles.

2. Preservation theorems

In this section we determine the extent to which UCSO is preserved under standard probability operations. Let each space S_i be a complete separable metric space with associated Borel σ-field $\mathcal{S}_i$ and closed partial order $\leqq$. Let all the functions in $\mathcal{U}_i$ be expectations of real-valued functions on the underlying space $\mathcal{S}_i$. First, to consider image measures, let Pv^{-1} be the image measure determined by a measurable function $v: S_1 \to S_2$ and a probability measure P on S_1, defined by

$$(Pv^{-1})(A) = P(v^{-1}(A)) = P(\{s_1 \in S_1 : v(s_1) \in A\}), \qquad A \in \mathcal{S}_2.$$

Theorem 2.1. Let $v: S_1 \to S_2$ be measurable. If $P_1 \leqq_{\mathcal{U}_1, \mathcal{C}_1} P_2$ on S_1, $v^{-1}(\mathcal{C}_2) \subseteq \mathcal{C}_1$ and $u \circ v \in \mathcal{U}_1$ for all $u \in \mathcal{U}_2$, then $P_1 v^{-1} \leqq_{\mathcal{U}_2, \mathcal{C}_2} P_2 v^{-1}$.

Let partial order be extended to product spaces in the usual way: $(s_{11}, s_{12}) \leqq (s_{21}, s_{22})$ in $S_1 \times S_2$ if $s_{11} \leqq s_{21}$ in S_1 and $s_{12} \leqq s_{22}$ in S_2. Since each space S_i is a complete separable metric space with a closed partial order, the same is true for $S_1 \times S_2$. Let $P_1 \times P_2$ represent the product probability measure on $S_1 \times S_2$ and let $\mathcal{C}_1 \times \mathcal{C}_2 = \{A_1 \times A_2 : A_1 \in \mathcal{C}_1, A_2 \in \mathcal{C}_2\}$. From the w.p.l.-representation of stochastic order, cf. Theorem 1 of Kamae, Krengel and O'Brien (1977), it is trivial that $P_1 \times P_2 \leqq_{st} Q_1 \times Q_2$ if $P_i \leqq_{st} Q_i$ for $i = 1, 2$. Here is an extension to UCSO.

Theorem 2.2. If $P_i \leqq_{\mathcal{C}_i} Q_i$ in $\Pi(S_i)$ for $i = 1, 2$, then $P_1 \times P_2 \leqq_{\mathcal{C}_1 \times \mathcal{C}_2} Q_1 \times Q_2$ in $\Pi(S_1 \times S_2)$.

It is also easy to see that the ordering $\leqq_r$ extends to products, i.e., $(P_1 \times P_2) \leqq_r (Q_1 \times Q_2)$ on $S_1 \times S_2$ if $P_1 \leqq_r Q_1$ and $P_2 \leqq_r Q_2$, but the following example illustrates that the ordering $\leqq_{\mathcal{C}}$ for $\mathcal{C} = \mathcal{S}$ does not extend to products. At the same time, this example shows that $\leqq_{\mathcal{C}}$ is not preserved by convolutions.

Example 2.1. Let $S_1 = S_2 = \{0, 1, 4\}$, $P_1(\{0\}) = 0.7$, $P_1(\{1\}) = 0.1$, $P_1(\{4\}) = 0.2$, $P_2(\{0\}) = 0.4$, $P_2(\{1\}) = 0.2$ and $P_2(\{4\}) = 0.4$. Obviously $P_1 \leqq_r P_2$, so that $P_1 \leqq_{\mathcal{C}} P_2$ for $\mathcal{C} = \mathcal{S}$. However, $(P_1 \times P_1)_A \leqq_{st} (P_2 \times P_2)_A$ fails for $A = \{(0, 4), (1, 4), (1, 0), (4, 0)\}$:

$$(P_1 \times P_1)_A(\{(4, 0)\}) = 14/37 > 16/48 = (P_2 \times P_2)_A(\{(4, 0)\}).$$

By Theorem 2.2, $(P_1 \times P_1)_{A_i} \leqq_{st} (P_2 \times P_2)_{A_i}$ for $A_1 = \{(0, 4), (1, 4)\}$ and $A_2 = \{(1, 0), (4, 0)\}$. However, as we saw above, $(P_1 \times P_1)_{A_1 \cup A_2} \leqq_{st} (P_2 \times P_2)_{A_1 \cup A_2}$ fails. Finally, $\leqq_r$ and $\leqq_{\mathcal{C}}$ are not preserved by convolutions, denoted by $*$. It is easy to see that both $P_1 * P_1 \leqq_r P_2 * P_2$ and $P_1 * P_2 \leqq_r P_2 * P_2$ fail here. Moreover, all these counterexamples persist if P_1 and P_2 are modified to have support $\{0, 1, 2, 3, 4\}$;

make the following modifications: $P_1(\{0\}) = 0.7 - 2\varepsilon$, $P_1(\{2\}) = P_1(\{3\}) = \varepsilon$, $P_2(\{0\}) = 0.4 - 4\varepsilon$, $P_2(\{2\}) = P_2(\{3\}) = 2\varepsilon$ for sufficiently small ε.

In order to state another positive result for product measures, let

$$\mathcal{U}_1 \times \mathcal{U}_2 = \{u: S_1 \times S_2 \to R \mid u(s_1, s_2) = u_1(s_1)u_2(s_2), u_1 \in \mathcal{U}_1, u_2 \in \mathcal{U}_2\}$$

$$\mathcal{U}_{12} = \left\{u: S_1 \times S_2 \to R \mid u = \sum_{j=1}^{k} a_j \bar{u}_j \text{ for some } k, a_j > 0, \bar{u}_j \in \mathcal{U}_1 \times \mathcal{U}_2\right\},$$

$$\hat{\mathcal{U}}_{12} = \{u: S_1 \times S_2 \to R \mid u(s_1, s_2) = \lim_{n\to\infty} u_n(s_1, s_2), u_{n+1} \geqq u_n \quad \text{in} \quad \mathcal{U}_{12}\}.$$

Theorem 2.3. (a) If $P_i \leqq_{\mathcal{U}_i, \mathcal{C}_i} Q_i$ on S_i for $i = 1, 2$, then

$$(P_1 \times P_2) \leqq_{\mathcal{U}_{12}, \mathcal{C}_1 \times \mathcal{C}_2} (Q_1 \times Q_2)$$

on $S_1 \times S_2$.

(b) If, in addition, the functions in $\mathcal{U}_1$ and $\mathcal{U}_2$ are non-negative, then $(P_1 \times P_2) \leqq_{\hat{\mathcal{U}}_{12}, \mathcal{C}_1 \times \mathcal{C}_2} (Q_1 \times Q_2)$.

Just as with convolutions, UCSO is not preserved under two-sided mixtures.

Example 2.2. Let $S = \{1, 2, 3, 4\}$, $P_1(\{1\}) = P_1(\{2\}) = P_2(\{3\}) = P_2(\{4\}) = \frac{1}{2}$, $Q_1(\{1\}) = Q_2(\{3\}) = \frac{1}{4}$ and $Q_1(\{2\}) = Q_2(\{4\}) = \frac{3}{4}$. Then $P_i \leqq_r Q_i$ for $i = 1, 2$, but $(\frac{1}{2})P_1 + (\frac{1}{2})P_2 \leqq_r (\frac{1}{2})Q_1 + (\frac{1}{2})Q_2$ fails to hold.

However, there is a positive result.

Theorem 2.4. If $P_1 \leqq_{\mathcal{U}, \mathcal{C}} Q$ and $P_2 \leqq_{\mathcal{U}, \mathcal{C}} Q$, then

$$\alpha P_1 + (1 - \alpha) P_2 \leqq_{\mathcal{U}, \mathcal{C}} Q$$

for all α, $0 \leqq \alpha \leqq 1$.

Corollary. If $P_i \leqq_{\mathcal{U}, \mathcal{C}} Q_j$ for $i = 1, 2$ and $j = 1, 2$, then

$$\alpha P_1 + (1 - \alpha) P_2 \leqq_{\mathcal{U}, \mathcal{C}} \beta Q_1 + (1 - \beta) Q_2$$

for all α and β, $0 \leqq \alpha \leqq 1$, $0 \leqq \beta \leqq 1$.

By Proposition 3 of Kamae, Krengel and O'Brien (1977), $\leqq_{st}$ is a closed partial order. We now give conditions under which the same is true for $\leqq_{\mathcal{U}, \mathcal{C}}$. Recall that a measurable subset A of S is a P-continuity set if $P(\partial A) = 0$, where ∂A is the boundary of A, i.e., the closure minus the interior. Here we do not require that the functions in $\mathcal{U}$ be expectations of real-valued functions on S.

Theorem 2.5. Assume:

(i) the functions in $\mathcal{U}$ are continuous on $\Pi(S)$ and

(ii) for each $A \in \mathcal{C}$, there is a sequence $\{A_k, k \geqq 1\}$ of subsets of S such that $A_k \in \mathcal{C}$ and A_k is a P_i-continuity set for $i = 1, 2$ for each k and $\bigcap_{k=1}^{\infty} A_k = A$.

If $P_{1n} \leqq_{\mathcal{U},\mathcal{C}} P_{2n}$ for all n, $P_{1n} \Rightarrow P_1$ and $P_{2n} \Rightarrow P_2$, then $P_1 \leqq_{\mathcal{U},\mathcal{C}} P_2$.

Condition (i) of Theorem 2.5 is satisfied if the functions in $\mathcal{U}$ are expectations of bounded continuous functions on S, cf. Billingsley (1968). In general, condition (i) does not hold in the setting of Kamae, Krengel and O'Brien (1977), i.e., for stochastic order on a complete separable metric space with a closed partial order, but under an extra condition stochastic order is determined by bounded continuous increasing functions.

Definition 2.1. An order topological space $(S, \leqq)$ is *normally ordered* if for each two disjoint closed subsets F_1 and F_2 such that F_1 is decreasing ($x \leqq y$ and $y \in F_1$ implies $x \in F_1$) and F_2 is increasing, there exists a continuous non-decreasing real-valued function f such that $f(s) = 0$ on F_1, $f(s) = 1$ on F_2 and $0 \leqq f(s) \leqq 1$ for all $s \in S$, cf. p. 28 of Nachbin (1965).

Theorem 2.6. Suppose $(S, \leqq)$ is normally ordered. If $\int f dP_1 \leqq \int f dP_2$ for all continuous bounded non-decreasing real-valued functions, then $P_1 \leqq_{st} P_2$.

Every compact ordered space is normally ordered, p. 48 of Nachbin (1965). Also, it is easy to see that Euclidean space R^n with the usual ordering is normally ordered.

Corollary. If $S = R^n$ and $\mathcal{C} = \{ \times_{i=1}^n [a_i, b_i]: a_i \leqq b_i, 1 \leqq i \leqq n\}$, then $\leqq_{\mathcal{C}}$ is a closed partial order.

3. Proofs

Proof of Theorem 1.1. To see that MLR implies UCSO for $\mathcal{U} = \mathcal{U}_1$ and $\mathcal{C} = \mathcal{S}$, suppose there is a subset B meeting the stipulated conditions with $p_2(s)/p_1(s)$ non-decreasing on B. For any $A \in \mathcal{C}$ with $P_1(A) > 0$ and $P_2(A) > 0$, $p_2(s)/p_1(s)$ is non-decreasing on $A \cap B$. Note that $p_{iA}(s) = p_i(s)/P_i(A)$ are Radon–Nikodym derivatives of P_{iA} with respect to μ on A. Since $p_2(s)/p_1(s)$ is non-decreasing on B, $p_{2A}(s)/p_{1A}(s)$ is non-decreasing on $A \cap B$. Choose s_1 in the closure of A such that $p_{2A}(s_0)/p_{1A}(s_0) \leqq 1 \leqq p_{2A}(s_2)/p_{1A}(s_2)$ for all $s_0 \leqq s_1 \leqq s_2$ with s_0, $s_2 \in A \cap B$. Such an s_1 must exist in order for P_{1A} and P_{2A} to be probability measures. For any $s_2 \geqq s_1$,

$$\int_{s_2}^{\infty} p_{1A}(s)\mu(ds) \leqq \int_{s_2}^{\infty} p_{1A}(s)[p_{2A}(s)/p_{1A}(s)]\mu(ds) = \int_{s_2}^{\infty} p_{2A}(s)\mu(ds),$$

and, for any $s_0 \leqq s_1$,

$$\int_{-\infty}^{s_0} p_{1A}(s)\mu(ds) \geqq \int_{-\infty}^{s_0} p_{1A}(s)[p_{2A}(s)/p_{1A}(s)]\mu(ds) = \int_{-\infty}^{s_0} p_{2A}(s)\mu(ds),$$

which implies that $P_{1A} \leqq_{st} P_{2A}$.

Proof of Theorem 1.2. The argument in the proof of Theorem 1.1 shows that (i) ⇒ (iii) and the implication (iii) ⇒ (ii) is trivial. Hence, it remains to show that (ii) ⇒ (i). Let A be any two-point subset $\{s_1, s_2\}$ of S such that $s_1 < s_2$, $P_1(A) > 0$ and $P_2(A) > 0$. Obviously, P_1 and P_2 are absolutely continuous with respect to the counting measure on S so that P_1 and P_2 have Radon–Nikodym derivatives, namely, the probability mass functions p_1 and p_2. Since $P_{1A} \leqq_{st} P_{2A}$, $p_{2A}(s)/p_{1A}(s)$ is non-decreasing on A. This in turn implies that $p_2(s)/p_1(s)$ is non-decreasing on A. Since this holds for all such two-point sets, $p_2(s)/p_1(s)$ is non-decreasing on the union of the supports of P_1 and P_2.

Proof of Theorem 1.3. (i) For any $s \in B$,

$$p_i(s) = \lim_{\varepsilon \to 0} \frac{P_i(I(s, \varepsilon))}{\mu(I(s, \varepsilon))} = \lim_{\varepsilon \to 0} \frac{P_i(D(s, \varepsilon))}{\mu(D(s, \varepsilon))}.$$

Now suppose $s_1 < s_2$ for s_1, $s_2 \in B$. Since $P_{1A_\varepsilon} \leqq_{st} P_{2A_\varepsilon}$ for each $\varepsilon > 0$, where $A_\varepsilon = D(s_1, \varepsilon) \cup (s_2, \varepsilon)$,

$$\frac{P_1(I(s_2, \varepsilon))}{P_1(D(s_1, \varepsilon))} \leqq \frac{P_2(I(s_2, \varepsilon))}{P_2(D(s_1, \varepsilon))}$$

for each $\varepsilon > 0$. Hence, we can let $\varepsilon \to 0$ to obtain $p_1(s_2)/p_1(s_1) \leqq p_2(s_2)/p_2(s_1)$, which establishes MLR.

(ii) Now suppose p_1 and p_2 are not continuous, but that S is locally compact. Then, by Urysohn's lemma, there exist sequences $\{p_{1n}\}$ and $\{p_{2n}\}$ of continuous densities such that

$$\int_S |p_{1n}(s) - p_1(s)| \mu(ds) \to 0 \quad \text{and} \quad \int_S |p_{2n}(s) - p_2(s)| \mu(ds) \to 0,$$

cf. p. 242 of Halmos (1950). Covergence in L^1 obviously implies convergence in measure, which in turn implies that there are subsequences $\{p_{1n_k}\}$ and $\{p_{2n_k}\}$ with common index set which converge almost everywhere $[\mu]$ to p_1 and p_2, cf. pp. 89, 93 of Halmos (1950). Since the preceding argument applies in the subsequence, $p_2(s)/p_1(s)$ is non-decreasing on a set whose complement has zero μ-measure.

Proof of Theorem 2.1. Apply a change of variables and the stated conditions.

Proof of Theorem 2.2. First note that $P_{1A_1} \times P_{2A_2} = (P_1 \times P_2)_{A_1 \times A_2}$. Then $(P_1 \times P_2)_{A_1 \times A_2} \leqq_{st} (Q_1 \times Q_2)_{A_1 \times A_2}$ if $P_{1A_1} \leqq_{st} Q_{1A_1}$ and $P_{2A_2} \leqq_{st} Q_{2A_2}$. Finally, $(P_1 \times P_2)(A_1 \times A_2) > 0$ if and only if $P_1(A_1) > 0$ and $P_2(A_2) > 0$, so that $(P_1 \times P_2) \leqq_{\mathscr{C}_1 \times \mathscr{C}_2} (Q_1 \times Q_2)$ as claimed.

Proof of Theorem 2.3. (a) For any $u \in \mathcal{U}_{12}$ and $A_1 \times A_2 \in \mathcal{C}_1 \times \mathcal{C}_2$,

$$\int u(s_1, s_2) d(P_1 \times P_2)_{A_1 \times A_2} = \sum_{j=1}^{k} a_j \int_{A_1} \bar{u}_{1j}(s_1) dP_{1A_1}(s_1) \int_{A_2} \bar{u}_{2j}(s_0) dP_{2A_2}(s_2)$$
$$\leqq \sum_{j=1}^{k} a_j \int_{A_1} \bar{u}_{1j}(s_1) dQ_{1A_1}(s_1) \int_{A_2} \bar{u}_{2j}(s_2) dP_{2A_2}(s_2)$$
$$= \int u(s_1, s_2) d(Q_1 \times Q_2)_{A_1 \times A_2}.$$

(b) Apply the monotone convergence theorem with Part (a).

Proof of Theorem 2.4. For any $A \in \mathcal{C}$ and $u \in \mathcal{U}$,

$$\int_S ud([\alpha P_1 + (1-\alpha)P_2]_A) = \frac{\alpha \int_A udP_1 + (1-\alpha)\int_A udP_2}{\alpha P_1(A) + (1-\alpha)P_2(A)}$$
$$\leqq Q(A)^{-1} \int_A udQ = \int_S udQ_A.$$

Proof of Theorem 2.5. By assumption, $u(P_{1n_{Ak}}) \leqq u(P_{2n_{Ak}})$ for all n, k and u. Use (i) and (ii) to take limits, first on n and then on k.

Proof of Theorem 2.6. Let G be an open increasing set. Thus, the complement G^c is closed and decreasing. Since S is a complete separable metric space, every probability measure on S is tight (Theorem 1.4 of Billingsley (1968)). Hence, for $\varepsilon > 0$ given, there exists a compact subset K of G such that $P_1(G) \leqq P_1(K) + \varepsilon$. Since $K \subseteq G$ and G is increasing, the increasing hull of K, denoted by $I(K)$, is a subset of G. Moreover, $I(K)$ is increasing and closed (p. 44 of Nachbin (1965)). Since G^c and $I(K)$ are disjoint closed sets, one decreasing and the other increasing, there exists a continuous non-decreasing real-valued function f which is 0 on G^c and 1 on $I(K)$ such that $0 \leqq f(s) \leqq 1$ for $s \in S$. Therefore,

$$P_1(G) \leqq P_1(K) + \varepsilon \leqq P_1(I(K)) + \varepsilon$$
$$\leqq \int_S fdP_1 + \varepsilon \leqq \int_S fdP_2 + \varepsilon \leqq P_2(G) + \varepsilon.$$

Since G and ε were arbitrary, $P_1(G) \leqq P_2(G)$ for all open increasing G. By Theorem 1 of Kamae, Krengel and O'Brien (1977), $P_1 \leqq_{st} P_2$.

Proof of corollary. By Theorem 2.6, $\leqq_{\mathcal{U}_1}$ order is determined by continuous bounded functions on $\mathcal{S}$, so that Condition (i) of Theorem 2.5 is satisfied. Condition (ii) of Theorem 2.5 is also satisfied because a product of intervals is either a P_i-continuity set for $i = 1, 2$ or the decreasing limit of products of intervals, all of which are P_i-continuity sets for $i = 1, 2$. To see this, note that a product of intervals $\times_{i=1}^{n} [a_i, b_i]$ necessarily is a P-continuity set if the ith one-dimensional marginal of P attaches no atoms of mass to the points a_i and b_i, $1 \leqq i \leqq n$.

Acknowledgements

This research was initiated in 1976 while the author was visiting the Department of Mathematics at the University of Nijmegen in the Netherlands on sabbatical leave of absence from the School of Organization and Management at Yale University. The research was completed at Bell Laboratories. I am grateful to Wim Vervaat of the University of Nijmegen for many helpful discussions about stochastic order. I am also grateful to the referee for identifying some mistakes.

References

BARLOW, R. E. AND PROSCHAN, F. (1975) *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart and Winston, New York.

BAWA, V. S. (1975) Optimal rules for ordering uncertain prospects. *J. Financial Economics* **2,** 95–121.

BILLINGSLEY, P. (1968) *Convergence of Probability Measures.* Wiley, New York.

BRUMELLE, S. L. AND VICKSON, R. G. (1975) A unified approach to stochastic dominance. *Stochastic Optimization Models in Finance*, ed. W. T. Ziemba and R. G. Vickson, Academic Press, New York, 101–113.

FERGUSON, T. S. (1967) *Mathematical Statistics, A Decision Theoretic Approach.* Academic Press, New York.

GOROFF, D. AND WHITT, W. (1978) Continuity of the admissible set in stochastic dominance. Economics Discussion Paper, Bell Laboratories, Holmdel, New Jersey 07733.

HALMOS, P. R. (1950) *Measure Theory.* Van Nostrand, Princeton, NJ.

KAMAE, T., KRENGEL, U. AND O'BRIEN, G. L. (1977) Stochastic inequalities on partially ordered spaces. *Ann. Prob.* **5,** 899–912.

KARLIN, S. (1968) *Total Positivity,* Vol. I. Stanford University Press, Stanford, CA.

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

MEYER, J. (1977) Choice among distributions. *J. Economic Theory* **14,** 326–336.

MEYER, P. A. (1966) *Probability and Potentials.* Blaisdell, Boston, MA.

NACHBIN, L. (1965) *Topology and Order.* Van Nostrand, Princeton, NJ.

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