J. Appl. Prob. **19**, 695–701 (1982)
Printed in Israel
0021-9002/82/030695-07 $00.95

MULTIVARIATE MONOTONE LIKELIHOOD RATIO AND UNIFORM CONDITIONAL STOCHASTIC ORDER

WARD WHITT,* *Bell Laboratories*

Abstract

Karlin and Rinott (1980) introduced and investigated concepts of multivariate total positivity (TP_2) and multivariate monotone likelihood ratio (MLR) for probability measures on R^n. These TP and MLR concepts are intimately related to supermodularity as discussed in Topkis (1968), (1978) and the FKG inequality of Fortuin, Kasteleyn and Ginibre (1971). This note points out connections between these concepts and uniform conditional stochastic order (UCSO) as defined in Whitt (1980). UCSO holds for two probability distributions if there is ordinary stochastic order for the corresponding conditional probability distributions obtained by conditioning on subsets from a specified class. The appropriate subsets to condition on for UCSO appear to be the sublattices of R^n. Then MLR implies UCSO, with the two orderings being equivalent when at least one of the probability measures is TP_2.

STOCHASTIC ORDER; CONDITIONAL PROBABILITY; UNIFORM CONDITIONAL STOCHASTIC ORDER; TOTAL POSITIVITY; MONOTONE LIKELIHOOD RATIO; SUPERMODULARITY; FKG INEQUALITY

In Whitt (1980) we introduced and investigated the concept of uniform conditional stochastic order (UCSO). One probability measure P is less than or equal to another Q in UCSO if P_A is stochastically less than or equal to Q_A for all subsets A in a designated class, where P_A and Q_A are conditional measures, i.e., $P_A(B) = P(A \cap B)/P(A)$, $P(A) > 0$ (Definition 5 below). A satisfactory theory was shown to be possible for probability measures on totally ordered spaces largely because in that setting UCSO, conditioning on all subsets, is equivalent to monotone likelihood ratio (MLR) order. The purpose of this note is to show that similar connections exist for probability measures on partially ordered spaces if appropriate definitions are used.

We do not discuss applications here, but it is clear that the two notions of stochastic order and conditioning can fruitfully be applied together. In some cases it will be useful to condition on all subsets; in others it will be useful to condition only on special subsets. For example, to compare stochastic processes,

Received 25 March 1981; revision received 25 September 1981.
* Postal address: Bell Laboratories, WB-1A350, Bell Laboratories, Holmdel, NJ 07733, U.S.A.

it is natural to condition on the histories; see Arjas (1981) and Whitt (1981). For measures on the real line, it is useful to condition on all semi-infinite subintervals of the form $(-\infty, x]$ or $[x, \infty)$ as well as on all intervals (which is equivalent to all subsets); see Keilson and Sumita (1983), Whitt (1981a) and references there. Of course, when the MLR property holds, there are many known applications; see Karlin and Rinott (1980) and references there. The MLR property was applied in Whitt (1979) to study the effect of a sample on the posterior distribution; also see Fahmy et al. (1982). For related work in economics, see Milgrom (1981) and Milgrom and Weber (1982).

Here our goal is to make a connection between UCSO and MLR for probability measures on partially ordered spaces. For simplicity, attention is restricted to R^n with the usual partial order: $x \leqq y$ for $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$ if $x_i \leqq y_i$ for each i. For UCSO, it seems appropriate to condition on all sublattices of R^n; for MLR, it seems appropriate to use the definition of multivariate MLR order recently introduced by Karlin and Rinott (1980), Definition 2 below, which is intimately related to the FKG inequality; see Fortuin, Kasteleyn and Ginibre (1971) and Kemperman (1977). In general, MLR is stronger than UCSO with these definitions (Theorem 2), but these orderings are equivalent and the theory simplifies when at least one probability measure is totally positive (Theorem 3), also as defined by Karlin and Rinott (1980); see Definition 3 below. The results here follow quite easily from previous ones, but they seem important supplements to Whitt (1980). Additional properties of UCSO (for probability measures on R) with many applications are contained in Keilson and Sumita (1983).

Let P and Q be probability measures on R^n. We begin with the monotone likelihood ratio (MLR) orderings, which involve comparisons between densities p and q of P and Q. We assume these densities are either with respect to Lebesgue measure or a counting measure on R^n, the latter making p and q probability mass functions. We start with the weaker form considered in Whitt (1980).

For simplicity, we assume the probability measures P and Q have common support. When P and Q do not have common support, there is a natural extra condition to impose in the following MLR and UCSO definitions in order for P to be less than or equal to Q. Let $s(P)$ be the support of P and let $A \geqq B$ for sets A and B hold if $a \geqq b$ for all $a \in A$, $b \in B$. The natural extra condition is: $s(Q) - s(P) \geqq s(P)$ and $s(Q) \geqq s(P) - s(Q)$.

Definition 1 (weak MLR order). $P \leqq_r Q$ if there exist densities p and q such that

$$p(y)q(x) \leqq p(x)q(y) \tag{1}$$

whenever $x \leqq y$.

Next we give the stronger definition in Karlin and Rinott (1980); see (1.18) and references there to the FKG inequality. Let $\vee$ and $\wedge$ be the usual lattice operations on R^n, i.e., for $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$,

$$x \wedge y = (\min\{x_1, y_1\}, \cdots, \min\{x_n, y_n\})$$

and

$$x \vee y = (\max\{x_1, y_1\}, \cdots, \max\{x_n, y_n\}).$$

Definition 2 (strong MLR order). $P \leqq_{\text{tp}} Q$ if there exist densities p and q such that

$$p(y)q(x) \leqq p(x \wedge y)q(x \vee y) \tag{2}$$

for all x, y.

Obviously $\leqq_{\text{tp}}$ is stronger than $\leqq_{\text{r}}$ but equivalent for probability measures on R. The tp subscript is explained by the connection to multivariate total positivity; see (1.4) in Karlin and Rinott (1980).

Definition 3 (multivariate total positivity). P is TP_2 if $P \leqq_{\text{tp}} P$.

Definition 3 brings out the fact that, unlike $\leqq_{\text{r}}$, $\leqq_{\text{tp}}$ is not reflexive in general. A probability density p on R^n is TP_2 if and only if $\log p$ is supermodular in the sense of Topkis (1968), (1978). As a consequence, Theorems 3.1 and 3.2 of Topkis (1978) imply that p is TP_2 if and only if $p(x + \varepsilon u^i)/p(x)$ is non-decreasing in x_j for each $x \in R^n$ and $i \neq j$, where $x = (x_1, \cdots, x_n)$, u^i is the ith unit vector in R^n and $\varepsilon > 0$. Also, if p is twice differentiable, then p is TP_2 if and only if $\partial^2 \log p(x)/\partial x_i \partial x_j \geqq 0$ for each $x \in R^n$ and $i \neq j$. Preservation theorems for supermodularity and, equivalently, multivariate total positivity are given by Topkis (1978) and Karlin and Rinott (1980). The arguments of Theorems 3.1 and 3.2 in Topkis (1978) also easily apply to characterize $\leqq_{\text{tp}}$ order for two densities p and q in terms of comparisons involving only two variables at a time as follows.

Theorem 1. The strong MLR ordering (2) for densities p and q holds if and only if

$$\frac{q(x + \varepsilon_1 u^i)}{p(x)} \leqq \frac{q(x + \varepsilon_1 u^i + \varepsilon_2 u^j)}{p(x + \varepsilon_2 u^j)} \tag{3}$$

for all $x \in R^n$, positive ε_1 and ε_2, and $i \neq j$, with $p(x + \varepsilon_2 u^j) > 0$, where u^i is the ith unit vector in R^n.

Now recall the following standard definition of stochastic order.

Definition 4. $P \leqq_{\text{st}} Q$ if

$$\int_{R^n} f dP \leqq \int_{R^n} f dQ$$

for all non-decreasing measurable real-valued functions f on R^n.

We now introduce a version of uniform conditional stochastic order (UCSO) considered in Whitt (1980). For this purpose, recall that a subset A of R^n is a lattice if $x \wedge y$, $x \vee y \in A$ whenever $x, y \in A$. Let $\mathscr{L}$ be the set of all lattice subsets of R^n. For any subset A with $P(A) > 0$, let the conditional probability measure be P_A, defined by

$$P_A(B) = P(A \cap B)/P(A), \qquad B \in R^n.$$

We refer to the following ordering as uniform conditional stochastic order with respect to sublattices (UCSOL).

Definition 5 (UCSOL). $P \leqq_{\mathscr{L}} Q$ if $P_A \leqq_{\mathrm{st}} Q_A$ for all $A \in \mathscr{L}$ with $P(A) > 0$ and $Q(A) > 0$.

Let the relations $\leftrightarrow$, $\rightarrow$, and $\not\leftrightarrow$ between orderings have the obvious interpretation, i.e., of equivalence, implication, and non-implication, respectively. From existing results, it is easy to establish the following relationships.

Theorem 2. $\leqq_{\mathrm{tp}} \not\rightleftarrows \leqq_{\mathscr{L}} \not\rightleftarrows \leqq_{\mathrm{r}}$.

Proof. The implication $\leqq_{\mathscr{L}} \rightarrow \leqq_{\mathrm{r}}$ is covered by Theorem 1.3 (ii) of Whitt (1980). To have $D(s_1, \varepsilon) \cup I(s_2, \varepsilon) \in \mathscr{L}$ as required there, let the metric d on R^n be the metric associated with the L^∞ or supremum norm, defined by

$$d(x, y) = \max\{|x_1 - y_1|, \cdots, |x_n - y_n|\}$$

for $x = (x_1, \cdots, x_n)$ and $y = (y_1, \cdots, y_n)$.

A minor modification of Example 1.4 in Whitt (1980) shows that $\leqq_{\mathrm{r}} \not\leftrightarrow \leqq_{\mathrm{st}}$. On $\{(0,0), (0,1), (1,0), (1,1)\}$ let $p_1((0,0)) = 0.1$, $p_1((0,1)) = p_1((1,0)) = p_1((1,1)) = 0.3$ and $p_2((0,0)) = 0.01$, $p_2((1,0)) = 0.09$, $p_2((0,1)) = p_2((1,1)) = 0.45$. Then $P_1 \leqq_{\mathrm{r}} P_2$, but $P_1(A) \geqq P_2(A)$ for $A = \{(0,1), (1,1)\}$.

Next we show that $\leqq_{\mathrm{tp}} \rightarrow \leqq_{\mathscr{L}}$. First, it is known that $P \leqq_{\mathrm{st}} Q$ if $P \leqq_{\mathrm{tp}} Q$; see (1.19) of Karlin and Rinott (1980). Next, it is easy to see that $P \leqq_{\mathrm{tp}} Q$ if and only if $P_A \leqq_{\mathrm{tp}} Q_A$ for all $A \in \mathscr{L}$ with $P(A) > 0$ and $Q(A) > 0$ because such subsets A are sublattices and the relation (2) is preserved when p and q are multiplied by positive constants. Hence, $P_A \leqq_{\mathrm{st}} Q_A$ when $P \leqq_{\mathrm{tp}} Q$, as claimed. Finally, to see that the two orderings $\leqq_{\mathrm{tp}}$ and $\leqq_{\mathscr{L}}$ are not equivalent, let P be any probability measure on R^n that is not TP_2. Then $P \leqq_{\mathscr{L}} P$, but not $P \leqq_{\mathrm{tp}} P$.

The theory greatly simplifies if all the probability measures are TP_2 or even just one is.

Theorem 3. If either P or Q is TP_2, then for P and Q $\leqq_{\mathrm{tp}} \leftrightarrow \leqq_{\mathscr{L}} \leftrightarrow \leqq_{\mathrm{r}}$.

Proof. It suffices to show that $\leqq_{\mathrm{r}} \rightarrow \leqq_{\mathrm{tp}}$. For this purpose, suppose P is TP_2 and $P \leqq_{\mathrm{r}} Q$. (The argument is similar when Q is TP_2.) Then there exist densities p_1, p_2, p_3, and q such that

$$p_1(x)p_2(y) \leqq p_1(x \vee y)p_2(x \wedge y) \tag{4}$$

and

$$p_3(x \vee y)q(x) \leqq q(x \vee y)p_3(x), \tag{5}$$

for all x, y. Now let $p(x) = p_1(x)$ for those x for which $p_1(x) = p_2(x) = p_3(x)$, and let $p(x) = 0$ elsewhere (a null set). Then (4) and (5) hold with p instead of p_i, so that (4) and (5) can be combined to yield the desired relation (2).

Suppose S is a countable sublattice in R^n, $\mathcal{T}$ is the set of all ordered two-point subsets, and $\mathcal{O}$ is the set of all totally ordered subsets. Let $\leqq_{\mathcal{T}}$ and $\leqq_{\mathcal{O}}$ be UCSO with respect to $\mathcal{T}$ and $\mathcal{O}$, respectively. The following corollary is an immediate consequence of Theorem 1.2 in Whitt (1980) and Theorem 3 above.

Corollary. Let P and Q be probability measures on the countable sublattice S. If either P or Q is TP$_2$, then

$$\leqq_{\text{tp}} \leftrightarrow \leqq_{\mathcal{L}} \leftrightarrow \leqq_{\text{r}} \leftrightarrow \leqq_{\mathcal{O}} \leftrightarrow \leqq_{\mathcal{T}}.$$

We conclude by exhibiting some lattice structure associated with the orderings $\leqq_{\text{r}}$ and $\leqq_{\text{tp}}$, and some algorithms for calculating the upper and lower bounds, for the case of finite lattices. Let S be a finite sublattice of R^n with a least element s_0 and a greatest element s_1. Let N_s be the subset of nearest lower neighbors of s in S, i.e.,

$$N_s = \{x \in S : x < s, x < y < s \text{ for no } y \in S\}.$$

Let N^s be the corresponding subset of nearest higher neighbors of s.

For probability mass functions p and q with support on S, i.e., with $p(s)$, $q(s) > 0$ for all $s \in S$, define $p \vee_{\text{r}} q$ and $p \wedge_{\text{r}} q$ recursively by

$$\frac{(p \vee_{\text{r}} q)(s)}{(p \vee_{\text{r}} q)(s_0)} = \max_{x \in N_s} \left\{ \frac{(p \vee_{\text{r}} q)(x)}{(p \vee_{\text{r}} q)(s_0)} \max \left\{ \frac{p(s)}{p(x)}, \frac{q(s)}{q(x)} \right\} \right\},$$

and

$$\frac{(p \wedge_{\text{r}} q)(s)}{(p \wedge_{\text{r}} q)(s_0)} = \min_{x \in N_s} \left\{ \frac{(p \wedge_{\text{r}} q)(x)}{(p \wedge_{\text{r}} q)(s_0)} \min \left\{ \frac{p(s)}{p(x)}, \frac{q(s)}{q(x)} \right\} \right\},$$

with $(p \vee_{\text{r}} q)(s_0)$ and $(p \wedge_{\text{r}} q)(s_0)$ being normalization constants determined by the condition that $p \vee_{\text{r}} q$ and $p \wedge_{\text{r}} q$ be probability mass functions. It is easy to check the following.

***Theorem* 4.** (i) The space $\Pi(S)$ of all probability mass functions with support on the finite lattice S is itself a lattice with the partial order $\leqq_{\text{r}}$, greatest lower bound $\wedge_{\text{r}}$ and least upper bound $\vee_{\text{r}}$.

(ii) The subset of totally positive probability mass functions in $\Pi(S)$ is a sublattice in which $\leqq_{\text{r}} \leftrightarrow \leqq_{\text{tp}}$.

For any probability mass function $p \in \Pi(S)$, the set of all probability mass functions with $p(s) > 0$ for all $s \in S$, let p_u and p_l be defined recursively by

$$\frac{p_u(s)}{p_u(s_0)} = \max\left\{\frac{p_u(x)}{p_u(s_0)}\frac{p(y)}{p(x \wedge y)}: x \in N_s,\ y \in N_s \cup \{s\},\ x \vee y = s\right\}$$

and

$$\frac{p_l(s_1)}{p_l(s)} = \min\left\{\frac{p_l(s_1)}{p_l(x)}\frac{p(x \vee y)}{p(y)}: x \in N^s,\ y \in N^s \cup \{s\},\ x \wedge y = s\right\},$$

with $p_u(s_0)$ and $p_l(s_1)$ the normalizations so that the total probability is 1 in each case. Let

$$p \vee_{\text{tp}} q = p_u \vee_{\text{r}} q_u \quad \text{and} \quad p \wedge_{\text{tp}} q = p_l \wedge_{\text{r}} q_l.$$

Finally, it is not difficult to establish the following result.

Theorem 5. (i) p_l and p_u are TP_2;
(ii) $p_l \leqq_{\text{tp}} p \leqq_{\text{tp}} p_u$;
(iii) $p \leqq_{\text{tp}} q$ if and only if $p_u \leqq_{\text{tp}} q_l$;
(iv) $\Pi(S)$ with $\leqq_{\text{tp}}$, $\vee_{\text{tp}}$ and $\wedge_{\text{tp}}$ is a lattice;
(v) $p_u \vee_{\text{r}} q_u = (p \vee_{\text{r}} q)_u$ and $p_l \wedge_{\text{r}} q_l = (p \wedge_{\text{r}} q)_l$.

Acknowledgment

I am grateful to Adrian Kester of Bell Laboratories for several helpful suggestions.

References

[1] Arjas, E. (1981) A stochastic process approach to multivariate reliability systems: notions based on conditional stochastic order. *Math. Operat. Res.* **6**, 263–276.

[2] Fahmy, S., de B. Pereira, C. A., Proschan, F. and Shaked, M. (1982) The influence of the sample on the posterior distribution. *Commun. Statist.* To appear.

[3] Fortuin, C. M., Kasteleyn, P. W. and Ginibre, J. (1971) Correlation inequalities on some partially ordered sets. *Commun. Math. Phys.* **22**, 89–103.

[4] Karlin, S. and Rinott, Y. (1980) Classes of orderings of measures and related correlation inequalities: I. Multivariate totally positive distributions. *J. Multivariate Analysis* **10**, 467–498.

[5] Keilson, J. and Sumita, U. (1983) Uniform stochastic ordering and related inequalities. *Canad. J. Statist.* To appear.

[6] Kemperman, J. H. B. (1977) On the FKG inequality for measures on a partially ordered space. *Indag. Math.* **39**, 313–331.

[7] Milgrom, P. R. (1981) Good news and bad news: representation theorems and applications. *Bell J. Econom.* **12**, 380–391.

[8] Milgrom, P. R. and Weber, R. J. (1982) A theory of auctions and competitive bidding. *Econometrica.* To appear.

[9] Topkis, D. M. (1968) Ordered Optimal Solutions. Ph.D. dissertation, Department of Operations Research, Stanford University.

[10] Topkis, D. M. (1978) Minimizing a submodular function on a lattice. *Operat. Res.* **26**, 305–321.

[11] Whitt, W. (1979) A note on the influence of the sample on the posterior distribution. *J. Amer. Statist. Assoc.* **74**, 424–426.

[12] Whitt, W. (1980) Uniform conditional stochastic order. *J. Appl. Prob.* **17**, 112–123.

[13] Whitt, W. (1981) Comparing counting processes and queues. *Adv. Appl. Prob.* **13**, 207–220.

[14] Whitt, W. (1981a) The renewal-process stationary-excess operator. Bell Laboratories.