



**HAL**  
open science

## Stochastic orderings for discrete random variables

A. Giovagnoli, H.P. Wynn

► **To cite this version:**

A. Giovagnoli, H.P. Wynn. Stochastic orderings for discrete random variables. *Statistics and Probability Letters*, 2010, 78 (16), pp.2827. 10.1016/j.spl.2008.04.002 . hal-00613919

**HAL Id: hal-00613919**

**<https://hal.science/hal-00613919>**

Submitted on 8 Aug 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Accepted Manuscript

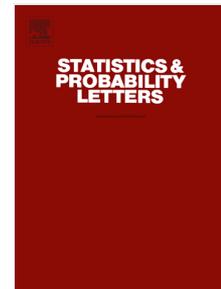
Stochastic orderings for discrete random variables

A. Giovagnoli, H.P. Wynn

PII: S0167-7152(08)00208-3  
DOI: 10.1016/j.spl.2008.04.002  
Reference: STAPRO 5047

To appear in: *Statistics and Probability Letters*

Received date: 3 December 2007  
Accepted date: 2 April 2008



Please cite this article as: Giovagnoli, A., Wynn, H.P., Stochastic orderings for discrete random variables. *Statistics and Probability Letters* (2008), doi:10.1016/j.spl.2008.04.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Stochastic orderings for discrete random variables

A Giovagnoli, University of Bologna, Italy  
 H.P. Wynn, London School of Economics, UK

December 3, 2007

## Abstract

A number of application areas of statistics make direct use of stochastic orderings. Here the special case of discrete distributions is covered. For a given partial ordering  $\preceq$  one can define the class of all  $\preceq$ -order preserving functions  $x \preceq y \Rightarrow g(x) \leq g(y)$ . Stochastic orderings may be defined in terms of  $\preceq$ :  $X \preceq^{st} Y \Leftrightarrow E_X g(X) \leq E_Y g(Y)$  for all order-preserving  $g$ . Alternatively they may be defined directly in terms of a class of functions  $\mathcal{F}$ :  $X \preceq^{st} Y \Leftrightarrow E_X g(X) \leq E_Y g(Y)$  for all  $f \in \mathcal{F}$ . For discrete distributions Möbius inversions plays a useful part in the theory and there are algebraic representations for the standard ordering  $\leq$  for integer grids. In the general case, based on  $\mathcal{F}$ , the notion of a dual cone is useful. Several examples are presented.

## 1 Introduction

Stochastic orderings and order preserving functions have a long history with perhaps the most famous being those associated with the Gini coefficient. Some general works are Shaked and Shanthikumar (1994) [13] and Müller and Stoyan (2002) [11]. Sub-areas include group invariant ordering: Marshall and Olkin (1979) [10], Eaton (1980)[3] and Giovagnoli and Wynn (1985) [5] and orderings related to association and dependence measures: Joe (1987) [8] and Haberman (1982) [7]. Applications are to general statistical theory, particularly testing: Bickel and Lehmann (1975) [1] and risk and insurance, eg Artzner et al (1999)[2].

One approach is to begin with a non-stochastic ordering between variables, consider the class of order preserving functions, and then make the variables stochastic and, via expectations, extend the ordering to random variables. These are sometimes called “integral” stochastic orderings and this is the starting point of the present paper.

Most of the existing literature focuses on real (continuous) univariate or multivariate random variables. There is less work on stochastic orderings for discrete random variables. We try to give a general theory that includes a large class of examples. Although one can consider some results for discrete random variables as specialisations of the continuous theory, there are nice features of the discrete case which can be used for some closer analysis, particularly the use of Möbius inversion.

## 2 Partial orderings and order-preserving functions

### 2.1 Lattices

We shall consider a partially ordered set  $\mathcal{X}$ , which is finite unless otherwise stated. We shall use  $\mathcal{X}$  both as the base set of the partial ordering and the support of a probability distribution.

**Definition 1** *By a partial ordering on the base set  $\mathcal{X}$  we mean a relationship  $\preceq$  which satisfies*

1. for any  $x \in \mathcal{X}$ ,  $x \preceq x$ ,
2. for  $x, y \in \mathcal{X}$ ,  $x \preceq y$  and  $y \preceq x$  implies  $x = y$ ,
3. for  $x, y, z \in \mathcal{X}$ ,  $x \preceq y$  and  $y \preceq z$  implies  $x \preceq z$ .

In this paper  $z \prec y$  will mean  $z \preceq y$ ,  $z \neq y$ . In some of the examples  $\mathcal{X}$  will be specialised to the product set

$$S = N_1 \times \cdots \times N_d$$

where  $N_i = \{0, \dots, n_i - 1\}$ . We shall call the entrywise order  $x \leq y \Leftrightarrow x_i \leq y_i$ , ( $i = 1, \dots, d$ ) the *standard order* on the discrete rectangular grid.

**Definition 2** *A total ordering on  $\mathcal{X}$  is a partial ordering  $\preceq$  with the the additional condition that for any  $x, y \in \mathcal{X}$  either  $x \preceq y$  or  $y \preceq x$*

**Definition 3** *For two partial orderings  $\preceq_\sigma, \preceq_\tau$  on  $\mathcal{X}$  we shall say that  $\preceq_\sigma$  is a refinement of  $\preceq_\tau$  if for all  $x, y \in \mathcal{X}$*

$$x \preceq_\tau y \Rightarrow x \preceq_\sigma y.$$

In the following schematic array we have  $d = 2$  and  $n_1 = n_2 = 4$ . We represent the total order  $S$  by the location of the numbers 1 to 16 from the lower left hand corner, representing the point  $(0, 0)$ , to the top right hand corner, representing  $(3, 3)$ .

$$\begin{array}{cccc} 7 & 11 & 15 & 16 \\ 4 & 8 & 12 & 14 \\ 2 & 6 & 10 & 13 \\ 1 & 3 & 5 & 9 \end{array}$$

We see that this total order is a refinement of the standard ordering. Taking any pair  $x, y$  with  $x \leq y$  we see that  $y$  is allocated a larger number than  $x$ .

We can define *generalized upper* and *lower quadrants* in  $\mathcal{X}$

**Definition 4** *A generalized upper quadrant is a subset of  $\mathcal{X}$  of the form*

$$Q_{x_0}^u = \{y \in \mathcal{X} : x_0 \preceq y\}$$

for a given  $x_0$  in  $\mathcal{X}$ . Similarly, a generalized lower quadrant is

$$Q_{x_0}^l = \{y \in \mathcal{X} : y \preceq x_0\}$$

We shall consider partial orderings such that for any two elements  $x, y$  there is always a “minimal element”  $x \wedge y$  such that  $x \wedge y \preceq x$ ,  $x \wedge y \preceq y$  and if  $z \preceq x$ ,  $z \preceq y$ , then  $z \preceq x \wedge y$ . Similarly for a “maximal element”  $x \vee y$ : thus the set  $\mathcal{X}$  equipped with partial order  $\preceq$  becomes a lattice. The following holds

**Remark 1.** The intersection of generalized upper (lower) quadrants is an upper (lower) quadrant, since

$$\begin{aligned} Q_x^u \cap Q_y^u &= Q_{x \vee y}^u \\ Q_x^l \cap Q_y^l &= Q_{x \wedge y}^l \end{aligned}$$

The term *ideal* is used in lattice theory, but we prefer quadrant because of its geometric interpretation in the case of  $S$  defined above.

A more general notion is *upper* and *lower sets*.

**Definition 5** A set  $U \subseteq \mathcal{X}$  is an upper set if for all  $x \in U$  and all  $y \in \mathcal{X}$ , if  $x \preceq y$ , then  $y \in U$ . Similarly for a lower set  $L$ .

It is straightforward to see that all upper (lower) quadrants are upper (lower) sets, but not conversely.

**Remark 2** A set  $U \subseteq \mathcal{X}$  is an upper set if and only if it is a union of upper quadrants; that is to say there is a set  $I^*$  such that

$$U = \cup_{x \in I^*} Q_x^u$$

Similarly a lower set is a union of lower quadrants.

## 2.2 Order preserving functions

**Definition 6** Given a partial ordering  $\preceq$ , a  $\preceq$ -preserving function  $g$  on  $\mathcal{X}$  is one which satisfies

$$x \preceq y \Rightarrow g(x) \leq g(y)$$

for all  $x, y \in \mathcal{X}$ . A  $\preceq$ -decreasing function is one which satisfies

$$x \preceq y \Rightarrow g(x) \geq g(y)$$

for all  $x, y \in \mathcal{X}$ .

The characterisation of order preserving functions is a key result. We first need a definition.

**Definition 7** For any set  $U \subseteq \mathcal{X}$  we define the indicator functions

$$\delta_U(x) = \begin{cases} 1 & \text{if } x \in U \\ 0 & \text{otherwise} \end{cases}$$

**Theorem 1** Let  $\mathcal{U}$  be the collection of all upper sets. A real function  $\phi(x)$  is  $\prec$ -order preserving if and only if there are constants  $\alpha_0$  and  $\alpha_U \geq 0$  such that

$$\phi(x) = \alpha_0 + \sum_{U \in \mathcal{U}} \alpha_U \delta_U(x) \quad (1)$$

*Proof.* If  $\phi(x)$  is expressible in the given form then it is clearly order preserving since each  $\delta_U(x)$  is order preserving from Remark 2 above.

In the other direction we proceed by construction in the case when all  $\phi(x)$  values are distinct. The proof is easily adapted for the case of “ties”. Let  $\phi(x)$  be a  $\preceq$ -increasing function. Label the  $x$  to be increasing by increasing  $\phi(x)$ :

$$\phi(x_0) < \phi(x_1) < \dots < \phi(x_{n-1})$$

For any  $x_i \in \mathcal{X}$  define

$$U_{x_i} = \{y \in \mathcal{X} : \phi(x_i) \leq \phi(y)\}.$$

Since the total ordering induced by  $\phi(x)$  is a refinement of  $\preceq$ , every  $U_{x_i}$  is an upper set:

$$x_i \preceq y \Rightarrow \phi(x_i) \leq \phi(y) \Rightarrow y \in U_{x_i}$$

Set  $\alpha_0 = \phi(x_0)$  and

$$\alpha_i = \phi(x_i) - \phi(x_{i-1}), \quad i = 1, \dots, n-1,$$

and  $\alpha_U = 0$  for all other  $U \in \mathcal{U}$ . Then the construction is

$$\alpha_0 + \sum_{i=1}^{n-1} \alpha_i U_{x_i}(x_j) = \begin{cases} \phi(x_0) & \text{for } j = 0 \\ \phi(x_0) + \sum_{i=1}^j \{\phi(x_i) - \phi(x_{i-1})\} = \phi(x_j) & \text{for } j = 1, \dots, n-1 \end{cases}$$

This completes the proof.

### 3 Stochastic orderings

Order preserving functions with respect to a partial ordering lead naturally to a certain type of stochastic ordering. Let  $X$  and  $Y$  be two random variables defined with respect to the same base measure on  $\mathcal{X}$ . We shall use the notation  $p_X(x) = \Pr\{X = x\}$  and for a subset  $A \subset \mathcal{X}$  we shall sometimes write  $p_X(A) = \Pr\{X \in A\}$ . It is convenient to take the base measure as uniform. We define a partial ordering  $\preceq$  as above.

**Definition 8** For two random variables  $X, Y$  on  $\mathcal{X}$ , and a partial ordering,  $\preceq$ , on  $\mathcal{X}$  we define an integral stochastic orderings of Type I as the stochastic ordering  $\preceq^{st}$  by

$$X \preceq^{st} Y \Leftrightarrow \mathbb{E}_X\{g(X)\} \leq \mathbb{E}_Y\{g(Y)\}$$

for all  $\preceq$ -preserving functions  $g$  on  $\mathcal{X}$ .

The characterisation of order preserving functions in Theorem 2 gives a characterisation of all  $\preceq^{st}$ .

**Theorem 2** For a partial ordering  $\preceq$  on  $\mathcal{X}$  and random variables  $X, Y$  we have  $X \preceq^{st} Y$  if and only if for all upper sets  $U \subseteq \mathcal{X}$

$$\Pr\{X \in U\} \leq \Pr\{Y \in U\}$$

*Proof.* Using the characterisation (1) we have

$$\mathbb{E}_X\{\phi(X)\} = \alpha_0 + \sum_{U \in \mathcal{U}} \alpha_U p_X(U)$$

If  $p_X(U) \leq p_Y(U)$  then it follows, since all  $\alpha_U \geq 0$ , that  $\mathbb{E}_X\{\phi(X)\} \leq \mathbb{E}_Y\{\phi(Y)\}$ . The converse follows by recalling that the indicator functions  $\delta_U(x)$  are order preserving.

The dual ordering  $\tilde{\preceq}$  is given by

$$x \tilde{\preceq} y \Leftrightarrow y \preceq x$$

(the authors find it confusing to write  $\succeq$ ).

**Definition 9** For two random variables  $X, Y$  on  $\mathcal{X}$  and  $\preceq$  on  $\mathcal{X}$  define the dual Type I integral stochastic  $\preceq$ -ordering by

$$X \tilde{\preceq}^{st} Y \Leftrightarrow \mathbb{E}_X\{g(X)\} \leq \mathbb{E}_Y\{g(Y)\}$$

for all  $\preceq$ -decreasing functions  $g$  on  $\mathcal{X}$ .

We obtain the characterisation of  $\tilde{\preceq}^{st}$  as in Theorem 1, by replacing upper sets by lower sets.

Type I stochastic ordering is a special case of a more general type of integral stochastic ordering based on some class of functions  $\mathcal{F}$ .

**Definition 10** For two random variables  $X, Y$  on  $\mathcal{X}$  and  $\preceq$  on  $\mathcal{X}$  we define integral stochastic orderings of Type II as the stochastic ordering  $\preceq^{st}$  by

$$X \preceq^{st} Y \Leftrightarrow \mathbb{E}_X\{g(X)\} \leq \mathbb{E}_Y\{g(Y)\}$$

for all function  $g$  in some class  $\mathcal{F}$ .

For this more general class the characterisation of the ordering in terms of the probability function  $p(x)$  or the cdf  $F(x)$  may be harder. There is an important technique for doing this, in some cases, which is that of dual cones. We shall return to this later.

## 4 Möbius inversion

Let  $\preceq$  be a partial ordering on a lattice  $\mathcal{X}$ . Define the zeta function to be

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x \preceq y \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

(see for example [6]). Define the Möbius function  $\mu$  by  $\mu(x, x) = 1$  and

$$\mu(x, y) = - \sum_{x \preceq z \prec y} \mu(x, z). \quad (3)$$

Möbius inversion gives the inversion of functions of a particular kind on  $\mathcal{X}$ .

**Theorem 3** (Möbius inversion) For functions  $f, g$  on  $S$  such that

$$g(x) = \sum_{z \preceq x} f(z) = \sum_{z \in S} \zeta(z, x) f(z)$$

we have

$$f(x) = \sum_{z \preceq x} \mu(z, x) g(z). \quad (4)$$

Thus,  $\mu$  is the inverse of  $\zeta$ : let

$$\delta(x, y) = \begin{cases} 1, & \text{if } x = y \\ \text{otherwise} & \end{cases},$$

then

$$\sum_{x \preceq z \preceq y} \mu(x, z) = \sum_{z \in S} \mu(x, z) \zeta(z, y) = \delta(x, y)$$

**Definition 11** The cumulative distribution function (cdf) relative to  $\preceq$  is

$$F_X(x) = \Pr\{X \preceq x\} = \Pr\{X \in Q_x^l\}$$

and the dual cdf is

$$\tilde{F}_X(x) = \Pr\{x \preceq X\} = \Pr\{X \in Q_x^u\}$$

Since

$$F(x) = \sum_{z \preceq x} p_X(z) = \sum_{z \preceq x} \zeta(z, x) p_X(z),$$

Equation (??) gives

$$p_X(x) = \sum_{z \preceq x} F(z) \mu(z, x).$$

The above formulae have duals

$$p_X(x) = \sum_{z \preceq x} \tilde{\mu}(z, x) \tilde{F}_X(x). \quad (5)$$

It is a little easier to use a matrix notation at this stage. Thus, with a suitable ordering of elements of  $S$ , we define the elements of  $[p]$  and  $[F]$  for the vector of values of functions  $f$  and  $F$  consistent with the ordering. Similarly, we can think of  $\zeta$  and  $\mu$  as matrices  $[\zeta]$  and  $[\mu]$ . Then

$$[\mu] = [\zeta]^{-1}$$

and the equation

$$[F] = [\zeta][p] \quad (6)$$

gives, from Theorem 3

$$[p] = [\mu][F] \quad (7)$$

We have analogous formulae for the dual

$$\begin{aligned} [\tilde{F}] &= [\tilde{\zeta}][p], \\ [p] &= [\tilde{\mu}][\tilde{F}] \end{aligned}$$

where  $\tilde{\zeta}$  and  $\tilde{\mu}$  is for the dual ordering  $\tilde{\preceq}$ . Moreover, if we keep the order of the elements in the matrix rows and columns as in the  $\preceq$  ordering, the duals are obtained by taking the transpose:

$$[\tilde{\mu}] = [\tilde{\zeta}]^{-1} = [\zeta]^{T-1} = [\mu]^T.$$

It is useful to be able to switch between  $F$  and  $\tilde{F}$ :

$$[\tilde{F}] = [\tilde{\zeta}][\mu][F] = [\zeta]^T[\mu][F] = ([\mu]^T)^{-1}[\mu][F]$$

#### 4.1 Cone ordering

Using the above representations stochastic ordering,  $\preceq^{st}$ , can be represented as a matrix cone ordering. We recall that a cone  $\mathcal{C}$  in  $\mathbb{R}^d$  with apex at the origin is a set with the property

$$u \in \mathcal{C} \Rightarrow \rho u \in \mathcal{C}$$

for any scalar  $\rho \geq 0$ . Define the cone ordering

$$u \preceq_{\mathcal{C}} v \Leftrightarrow v - u \in \mathcal{C}$$

This is a generalisation of the standard ordering for which the cone is the non-negative quadrant in  $\mathbb{R}^d$ :  $\{u : u \geq 0\}$ . The columns of a  $d \times m$  matrix generate a cone by taking all non-negative linear combinations.

Starting with the representation in Theorem 2 we collect together all the statements that  $\Pr_X(U) \leq \Pr_Y(U)$  for all  $U \in \mathcal{U}$  into a single matrix inequality:

$$A[p_X] \leq A[p_Y].$$

This can be written

$$A([p_Y] - [p_X]) \geq 0.$$

This, in turn, says that

$$[p_Y] - [p_X] \in \mathcal{C}^*(A)$$

the dual cone to the cone formed by the columns of  $A$ . From Möbius inversion we also have a cone for  $F$

$$[F_Y] - [F_X] \in \mathcal{C}^*(A[\mu]).$$

The dual cone idea is useful for characterising some cases of Type II integral stochastic orderings. This is when the class of function  $\mathcal{F}$  is itself defined by a collection of linear inequalities. These may be derived from the original lattice representation, as in the case of Type I ordering above or directly in  $\mathbb{R}^d$ .

As a small example consider the restriction of the so-called convex ordering on the non-negative integers:  $\mathcal{N} = \{0, 1, \dots, N-1\}$ . Discrete convexity can be defined in terms of second differences:

$$\mathcal{F} = \{f : f(i) - 2f(i+1) + f(i+2) \geq 0, \quad i = 0, 1, \dots, N-3\}$$

This can be written in matrix form:  $A[f] \geq 0$ .

Consider any ordering which has such a representation. From Definition 10, and using  $[p_X]$  and  $[p_Y]$  for the vector of probabilities for  $X$  and  $Y$ , the required condition for  $X \prec_{st} Y$  is

$$[p_X]^T [f] \geq A[p_Y]^T [f] \geq 0 \Leftrightarrow ([p_Y] - [p_X])^T [f] \geq 0,$$

for all  $[f] : A[f] \geq 0$ . These condition says that  $[f] \in \mathcal{C}^*(A)$ , the dual cone of  $\mathcal{C}(A)$  and that  $[p_Y] - [p_X] \in \mathcal{C}^{**}(A) = \mathcal{C}(A)$ . Thus without computing  $\mathcal{C}^*(A)$  we can characterise the ordering in terms of constraints on the differences:

$$[d] = [p_Y] - [p_X] = A^T[\alpha]$$

for some  $[\alpha] \geq 0$ .

Much of the effort of characterising the ordering reduces to a minimal representation, which is equivalent to finding the extreme generators of the cone  $\mathcal{C}(A)$ ; some of the rows of  $A$  may be redundant. In the above example of convexity the representation, in terms of the components of the differences  $[d]$ , is

$$d_0 = \alpha_0, \quad d_1 = -2\alpha_0 + \alpha_1, \dots, \quad d_i = \alpha_{i-2} - 2\alpha_{i-1} + \alpha_i, \dots, \quad d_{N-2} = d_{N-2} - 2\alpha_{N-1}, \quad d_{N-1} = \alpha_{N-1}$$

The condition is that differences of probabilities are the backward second differences of the nonnegative  $\alpha_i$  sequence.

Where there is a natural partial ordering the cone conditions on the  $[d]$  induce cone conditions on  $D = [F_Y] - [F_X]$  which we can write

$$[D] = [\zeta][d] = [\zeta]A^T[\alpha].$$

In the convex case we have the representation in terms of first differences:

$$D_0 = \alpha_0, \quad D_1 = \alpha_1 - \alpha_0, \dots, \quad D_i = \alpha_i - \alpha_{i-1}, \dots, \quad D_{N-2} = -\alpha_{N-1}, \quad D_{N-1} = 0.$$

## 5 Algebra for $(\mathbb{Z}^+)^d$

Because of its importance in many fields we study in more detail the infinite integer lattice  $(\mathbb{Z}^+)^d$ , namely the set of all non-negative integer  $d$ -vectors and the standard ordering  $\leq$ . We show how to “algebraize” the problem.

We use a dummy variable  $x = (x_1, \dots, x_d)$  and code  $\alpha \in (\mathbb{Z}^+)^d$  by the monomial

$$x^\alpha = x_1^{\alpha_1} \dots x_d^{\alpha_d}.$$

A set  $V \subset (\mathbb{Z}^+)^d$ , or equivalently its indicator function  $\delta_V(x)$  can be represented by a (multivariate) polynomial generating function

$$V(x) = \sum_{\alpha \in V} x^\alpha$$

We first note that

$$(\mathbb{Z}^+)^d(x) = \frac{1}{\prod_{i=1}^d (1 - x_i)}$$

For upper and lower quadrants algebraically we have:

$$Q_\alpha^l(x) = \sum_{0 \leq \beta \leq \alpha} x^\beta = \prod_{i=1}^d \frac{1 - x_i^{\alpha_i+1}}{1 - x_i}$$

and

$$Q_\alpha^u(x) = \sum_{\beta \geq \alpha} x^\beta = \frac{x^\alpha}{\prod_{i=1}^d (1 - x_i)}$$

Motivation is provided by the identity in one-dimension:

$$1 + x + x^2 + \dots + x^{n-1} = \frac{1 - x^n}{1 - x} = \frac{1}{1 - x} - \frac{x^{n+1}}{1 - x}$$

The left hand side of this identity is  $V(x)$  where  $V = \{0, 1, \dots, n-1\}$ . The right hand side has a term  $x^n$  in the numerator, with an exponent one more than the largest exponent of the left hand side. This is expressed as the difference between the generating function of the upper quadrant, in this case a integer “line” starting at  $n+1$ , and that of  $\mathbb{Z}^+$ .

Möbius inversion is the relation between a single  $x^\alpha$  and the generating function for the lower quadrant  $Q_\alpha^l$  given by the expansion of the numerator in

$$x^\alpha = \frac{x^\alpha \prod_{i=1}^d (1 - x_i)}{\prod_{i=1}^d (1 - x_i)}$$

We are now in a position to describe the upper sets algebraically. This development is partly based on Giglio and Wynn (2005) [4], in the context of system reliability. For two vectors  $\alpha, \alpha' \in \mathbb{Z}^+$

$$\alpha \vee \alpha' = (\max(\alpha_1, \alpha'_1), \dots, \max(\alpha_d, \alpha'_d)).$$

The relation  $Q_\alpha^u \cap Q_{\alpha'}^u = Q_{\alpha \vee \alpha'}^u$  carries over to the generating functions:

$$Q_\alpha^u \cap Q_{\alpha'}^u(x) = Q_{\alpha \vee \alpha'}^u(x)$$

Now consider an upper quadrant  $U_\alpha = \cup_{\alpha \in U^*} Q_\alpha^u$ . We can use an inclusion-exclusion to obtain the generating function  $U(x)$ . First index the elements of  $U^*$  in some order:  $\alpha^{(1)}, \dots, \alpha^{(|U^*|)}$ . Then the inclusion-exclusion lemma is:

$$U = \sum_i U_{\alpha_i}^u - \sum_{i \neq j} U_{\alpha_i}^u \cap U_{\alpha_j}^u + \dots$$

which can be written

$$U(x) = \sum_i U_{\alpha_i}^u(x) - \sum_{i \neq j} U_{\alpha_i \vee \alpha_j}^u(x) + \dots$$

The key construction in this section is to form a special polynomial  $W_{U^*}(x)$  which holds the term such as  $\alpha_i \vee \alpha_j$  in the above expansion, together with the sign given by the inclusion-exclusion. Thus

$$W_{U^*}(x) = \sum_i x^{\alpha_i} - \sum_{i \neq j} x^{\alpha_i \vee \alpha_j} + \dots$$

In the algebraic theory this is called the *Taylor polynomial*.

**Theorem 4** For an upper set  $U = \cup_{\alpha \in U^*} Q_\alpha^u$  and associated Taylor polynomial  $W_{U^*}(x)$  the following identity holds

$$U^c(x) = \frac{1 - W_{U^*}(x)}{\prod_{i=1}^d (1 - x_i)},$$

where  $U^c = Z^+ \setminus U$ .

We do not give a formal proof of this but a simple example will make the identity clear. Let  $d = 2$  and  $U^* = \{(3, 0), (2, 1), (0, 2)\}$ . Then

$$W_{U^*}(x) = x_1^3 + x_1^2 x_2 + x_2^2 - x_1^3 x_2 - x_1^2 x_2^2 - x_1^3 x_2^2 + x_1^3 x_2^2 = x_1^3 + x_1^2 x_2 + x_2^2 - x_1^3 x_2 - x_1^2 x_2^2$$

Notice the redundancy in the inclusion-exclusion lemma. The formula in Theorem 4 gives

$$1 + x_1 + x_2 + x_1^2 + x_1 x_2 = \frac{1 - (x_1^3 + x_1^2 x_2 + x_2^2 - x_1^3 x_2 - x_1^2 x_2^2)}{(1 - x_1)(1 - x_2)}$$

In general and using a compact notation we have

$$U^c(x) = (\mathbb{Z}^+)^d(x) - U(x)$$

The formula in Theorem 4 becomes

$$U^c(x) = (\mathbb{Z}^+)^d(x)(1 - W(x))$$

Eliminating  $U^c(x)$  we have

$$U(x) = W(x) (\mathbb{Z}^+)^d(x)$$

Again this can be verified for the example. In essence the proof of Theorem 4 is an elaboration of this formula. These formulae lead to an algebraic characterisation of the standard  $\leq$  and  $\leq^{st}$ -orderings.

**Definition 12** The generating function (gf) of a function  $g(x)$  on  $Z^+$  is defined to be

$$g^*(s) = \sum_{\alpha \geq 0} g(\alpha) s^\alpha$$

We also obtain a formula for the generating function of an order preserving function. Thus, for a  $\leq$ -preserving function  $\phi(x)$  with characterisation  $\phi(x) = \alpha_0 + \sum_{U \in \mathcal{U}} \alpha_U \delta_U(x)$  of Theorem 1,

$$\phi^*(s) = \alpha_0 + \sum_{U \in \mathcal{U}} \alpha_U U(s) = \alpha_0 + \left\{ \sum_{U \in \mathcal{U}} \alpha_U W_U(s) \right\} (\mathbb{Z}^+)^d(s)$$

This follows simply by recalling that the generating function of an upper quadrant  $Q_\alpha^u$  is  $Q_\alpha^u(x)$ .

## 5.1 Moments

We now take  $x \in (\mathbb{Z}^+)^d$  and consider the matrix indexed by  $x$  and  $\alpha$

$$\xi(x, \alpha) = x^\alpha$$

Consider  $x$  as a support point of a random variable  $X$  and let the general  $\alpha$ -moment be  $m_\alpha = E(X^\alpha)$ , then in the matrix notation:

$$[m] = [\xi]^T f = [\xi]^T [\mu][F]$$

This gives another equivalent condition to stochastic ordering, in terms of moments:

$$[m_Y] - [m_X] = [\xi]^T [\mu][\alpha]$$

for  $[\alpha] \geq 0$ . This says that  $[m_Y] - [m_X]$  lies in the cone generated by the columns of  $[\xi]^T [\mu]$ . Thus, because of the linear relationship between  $[M_\alpha]$  and  $[p]$  for both Type I and II we can express stochastic orderings in terms of moments. In the infinite support case it is known that moments do not necessarily determine distributions, so we must restrict to the finite support case or take care with convergence.

## 6 Examples

**Example 1.** The standard ordering on  $S = \{0, \dots, n-1\}$ . The stochastic ordering is the standard univariate stochastic ordering

$$X \leq^{st} Y \Leftrightarrow F_Y(z) \leq F_X(z) \Leftrightarrow \tilde{F}_X(z) \leq \tilde{F}_Y(z),$$

for all  $z \in S$ . Here the stochastic ordering and its dual are equivalent since  $\tilde{F}(z) = 1 - F(z-1)$ . The extension to the multivariate case  $S = N_1 \times \dots \times N_d$  has been dealt with in Section 5.

**Example 2.** Lexicographic total ordering  $\prec_\tau$ . In this case we fix an order of the variables (dimensions). For ease of explanation take the standard order. Then favouring the earlier entries we have in  $S = N_1 \times \dots \times N_d$

$$\begin{aligned} (0, 0, \dots, 0) &\prec_\tau (1, 0, \dots, 0) \prec_\tau \dots \prec_\tau (n_1, 0, \dots) \prec_\tau \\ (n_1, 1, \dots, 0) &\prec_\tau \dots \prec_\tau (n_1, n_2, 0, \dots, 0) \prec_\tau \dots \prec_\tau \\ (n_1, n_2, \dots, n_d - 1) &\prec_\tau (n_1, \dots, n_d) \end{aligned}$$

The table below shows the order of the entries for when  $d = 2$  and  $n_1 = n_2 = 4$

13	14	15	16
9	10	11	12
5	6	7	8
1	2	3	4

The general statement is that  $x \preceq y$  if and only if the first nonnegative entry in  $y - x$  is positive. The stochastic version considers (for  $X$  and  $Y$ ), the probability that the first non-zero entry of  $X - z$  (or  $Y - z$ ) is positive. We could name this the lexicographic stochastic ordering.

**Example 3.** We give an example of Type II integral stochastic ordering which cannot naturally be represented as Type I ordering, namely the supermodular stochastic ordering. Define, for a finite lattice the class of real supermodular functions namely functions satisfying

$$\mathcal{F} = \{f : f(x \wedge y) + f(x \vee y) - f(x) - f(y) \geq 0\},$$

for all  $x, y$  in the lattice  $\mathcal{X}$ , where  $\wedge$  and  $\vee$  are the join and meet on the lattice. For a recent paper on this case see Promislow and Young (2005) [13]. This leads to a special matrix  $A$  as in the discussion in Subsection 4.1. The dual cone analysis there leads to a representation, similar in spirit to the convex function case. Each row of  $A$  has two ones and two minus ones corresponding to the condition, and there is one such row for every pair  $(x, y)$  in the lattice. Let  $\alpha(x, y) \geq 0$  be the weighting for such a row. The representation is: for each  $z$  in the lattice:

$$d(z) = p_Y(z) - p_X(z) = \sum_{(x,y)} \alpha(x,y) \{\delta(x \wedge y, z) + \delta(x \vee y, z) - \delta(x, z) - \delta(y, z)\},$$

where the sum is over distinct pairs  $(x, y)$ , ie every pair is visited once. As before we can induce an ordering on the cdf's: which is neatly express by simply replacing the  $\delta$  functions by  $\zeta$  functions:

$$D(z) = F_Y(z) - F_X(z) = \sum_{(x,y)} \alpha(x,y) \{\zeta(x \wedge y) + \zeta(x \vee y) - \zeta(x) - \zeta(y)\}$$

**Example 4.** In generalising univariate to multivariate orderings one is naturally led to the ‘‘quadrant orderings’’. We can express these in terms of the cdf's,  $F$ , for the lower quadrant and  $\tilde{F}$  for the upper quadrant ordering. Thus the lower quadrant ordering is

$$X \prec_Q Y \Leftrightarrow F_X(z) \leq F_Y(z)$$

for all  $z$  in the lattice. The task is to express this as a Type II ordering by finding the class  $\mathcal{F}$  in Definition 10. We use the dual cone argument in reverse. Since the condition on the distributions is

$$[F_X] - [F_Y] = [\zeta][d] \geq 0,$$

we see that the function class is

$$[f] = [\zeta]^T[\alpha] = [\tilde{\zeta}][\alpha]$$

for  $[\alpha] \geq 0$ . This shows that  $f$  is the upper accumulation of a non-negative function and we can find the conditions on  $f$  by Möbius inversion:

$$[\tilde{\mu}][f] = [\mu]^T[f] = [\alpha] \geq 0$$

In summary, the class of function  $\mathcal{F}$  for lower quadrant ordering is the cone of all positive multiples all upper quadrant cdf's. In some sense this is a canonical case of a cone ordering in which Möbius inversion performs operation of computing the relevant dual cone.

### References

- [1] P. J. Bickel and E.L. Lehmann. (1975). Descriptive statistics for non-parametric models. i introduction. *Ann. Statist*, 3:1038–1044.
- [2] P. Artzner, F. Delbaen, J.M. Eber and D. Heath. (1999) Coherent measures of risk. *Math. Finance*, 9:203–228.
- [3] M.L. Eaton. Lectures on topics in probability inequalities. CWI Tract 35, Amsterdam, 1980.
- [4] B. Giglio and H.P. Wynn (2005). Monomial ideals and the Scarf complex for coherent systems in reliability theory. *Ann. Statist.* 32:1289–1311.
- [5] A. Giovagnoli and H.P. Wynn. (1985). G-majorization with applications to matrix orderings. *Linear Algebra Appl.*, 67:111–135.
- [6] J.E. Graver and M.E. Watkins. (1977). Combinatorics with emphasis on the theory of graphs. Springer-Verlag, New York.
- [7] S.J. Haberman. (1982) Measures of association. *Encyclopedia of Statistics* 1:130–137.
- [8] H. Joe. (1997). Multivariate models and dependence concepts. Chapman and Hall Ltd, London.
- [9] E.L. Lehmann. (1966). Some concepts of dependence. *Ann Math Statist*, 37:1137–1153.
- [10] A.W. Marshall and I. Olkin. (1979). Inequalities: Theory of majorization and its applications. Academic Press, New York.
- [11] A. Müller and D. Stoyan. (2002). Comparison Methods for Stochastic Models and Risks. Wiley, New York.
- [12] S. Promislow and V. Young.(2005). Supermodular functions on finite lattices. *Order* 22. 389–413.
- [13] M. Shaked and J.G. Shanthikumar. (1994). Stochastic orders and their applications. Academic Press, New York.