TWO BASIC PARTIAL ORDERINGS FOR DISTRIBUTIONS DERIVED FROM SCHUR FUNCTIONS AND MAJORIZATION

Kumar Joag-Dev, University of Illinois and
Florida State University

and

Jayaram Sethuraman, Florida State University

Abstract

Researchers in applied fields have long recognized the usefulness of inequalities when exact results are not available. The use of inequalities allows us to say that one estimate is better than another, that one maintenance policy is better than another or that a certain selection procedure is better than another etc., even though, we may not know the best estimator, the best maintenance policy or the best selection procedure. Such results are generally obtained from inequalities between two probability measures or random variables. Inequalities between random variables are in turn obtained from deterministic inequalities or deterministic partial orderings.

Hardy, Littlewood and Pólya (1952) in their classical book entitled *Inequalities* have discussed various partial orderings in R^n, one of which is known as majorization. Majorization is intimately related to Schur functions. This partial ordering was used to derive the partial orderings of stochastic majorization and DT ordering among distributions in a series of papers by Proschan and Sethuraman (1977); Nevius, Proschan and Sethuraman (1977); Hollander, Proschan and Sethuraman (1977); and Hollander, Proschan and Sethuraman (1981). Even though many more partial orderings of this type have been studied in recent papers and books by Marshall and Olkin (1979), Tong (1980), Boland, Tong and Proschan (1987, 1988), Abouammoh, El-Neweihi and Proschan (1989), the above two partial orderings remain the centerpiece in this type of research endeavor. In this expository paper, we describe the essentials of stochastic majorization and DT ordering and demonstrate some applications. A new proof of a slight generalization of earlier result on DT functions in Hollander et al., 1981 is given.

Introduction

Researchers in applied fields have long recognized the usefulness of inequalities when exact results are not available. The use of inequalities allows

---

Research supported by the United States Army Research Office, Durham, under Grant No. DAAGLO3 86-K-0094. The United States Government is authorized to reproduce and distribute reprints for governmental purposes. FSU Technical Report Number M-814; USARO Technical Report Number D-109, September 1989.

us to say that one estimate is better than another, that one maintenance policy is better than another or that a certain selection procedure is better than another etc., even though, we may not know the best estimator, the best maintenance policy or the best selection procedure. Such results are generally obtained from inequalities between two probability measures or random variables. Inequalities between random variables are in turn obtained from deterministic inequalities or deterministic partial orderings.

Hardy, Littlewood and Pólya (1952) in their classical book entitled *Inequalities* have discussed various partial orderings in R^n, one of which is known as majorization. Majorization is intimately related to Schur functions. This partial ordering was used to derive the partial orderings of stochastic majorization and DT ordering among distributions in a series of papers by Proschan and Sethuraman (1977) [PS 77]; Nevius, Proschan and Sethuraman (1977) [NPS 77]; Hollander, Proschan and Sethuraman (1977) [HPS 77]; and Hollander, Proschan and Sethuraman (1981) [HPS 81]. Even though many more partial orderings of this type have been studied in recent papers and books by Marshall and Olkin (1979), Tong (1980), Boland, Tong and Proschan (1987, 1988), Abouammoh, El-Neweihi and Proschan (1989), the above two partial orderings remain the centerpiece in this type of research endeavor. In this expository paper, we describe the essentials of stochastic majorization and DT ordering and demonstrate some applications in the second and third sections. A new proof of a slight generalization of earlier result on DT functions is given in the third section.

Schur Functions

We begin by reviewing some basic concepts and results involving Schur functions. Given a vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, let $x_{[1]}, x_{[2]}, \ldots, x_{[n]}$ be a permutation of its co-ordinates satisfying $x_{[1]} \geq x_{[2]} \geq \ldots \geq x_{[n]}$. A vector $\mathbf{x}$ is said to *majorize* a vector $\boldsymbol{y}$, $\boldsymbol{x} \overset{m}{\geq} \boldsymbol{y}$ in symbols, if

$$\sum_{i=1}^{j} x_{[i]} \geq \sum_{i=1}^{j} y_{[i]}, \qquad j = 1, 2, \ldots, n-1,$$

and

$$\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}.$$

Majorization is not a true partial ordering on R^n since $\boldsymbol{x} \overset{m}{\geq} \boldsymbol{y}$ and $\boldsymbol{y} \overset{m}{\geq} \boldsymbol{x}$ implies only that the co-ordinate sequence of $\boldsymbol{x}$ is a permutation of the co-ordinate sequence of $\boldsymbol{y}$. However it is a partial ordering in the cone $\{\boldsymbol{x} : \boldsymbol{x} \in R^n, x_1 \geq x_2 \geq \ldots x_n\}$. In any case, $\boldsymbol{x} \overset{m}{\geq} \boldsymbol{y}$ means that the co-ordinates of $\boldsymbol{x}$ are more spread out than those of $\boldsymbol{y}$.

A measurable function f defined on R^n will be called a *Schur function* if it is either *Schur-convex*, that is, if $f(\boldsymbol{x}) \geq f(\boldsymbol{y})$ whenever $\boldsymbol{x} \overset{m}{\geq} \boldsymbol{y}$, or is *Schur-concave*, that is, if $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ whenever $\boldsymbol{x} \overset{m}{\geq} \boldsymbol{y}$. It is easy to construct Schur functions from the example below.

Example 1

Let $f(\boldsymbol{x}) = \sum_1^n g(x_i)$. Then $f(\boldsymbol{x})$ is Schur convex if and only if g is Schur convex.

A subset A of R^n is called *Schur increasing* if it satisfies:

$$\boldsymbol{x} \in \mathrm{A},\ \boldsymbol{y} \overset{m}{\geq} \boldsymbol{x} \Rightarrow \boldsymbol{y} \in \mathrm{A}.$$

Note that the indicator of a Schur increasing set is a Schur *convex* function and in fact such indicators are the building blocks of the class of Schur convex functions and act as their level sets.

A partial ordering for random vectors can be defined as follows using Schur increasing sets. Let X and X' be random n-vectors. Then X is said to *stochastically majorize* X' if for every Schur increasing set A in R^n, $P[X \in \mathrm{A}] \geq P[X' \in \mathrm{A}]$, or equivalently, $E[f(X)] \geq E[f(X')]$, for every bounded Schur convex function f on R^n. This is stated, in symbols, as $X \overset{st.m.}{\geq} X'$.

Stochastic majorization is a way of comparing distributions of random vectors in much the same way as the stochastic ordering is for comparing distributions functions of real random variables. In fact stochastic majorization can be equivalently defined as stochastic ordering between certain transformed random vectors. Recall that Z is said to be *stochastically larger* than Z' if for every bounded nondecreasing function h, $E[h(Z)] \geq E[h(Z')]$. Consider the transformation $\boldsymbol{y} = (y_1, y_2,\ldots,y_n) \overset{def}{=} T(\boldsymbol{x})$, where $y_i = \sum_{j=1}^{i} x_{[j]}$, $i = 1, 2,\ldots,n$. It is clear that $C \overset{def}{=} TR^n$ is a cone. Let X and X' be two random vectors and let $Y = TX$ and $Y' = TX'$. Then it is easy to see that $X \overset{st.m.}{\geq} X'$ if and only if $E[g(Y)] \geq E[g(Y')]$ for all bounded measurable functions g such that $g(\boldsymbol{y}) \geq g(\boldsymbol{y}')$ whenever $y_i \geq y_i'$, $i = 1, 2,\ldots,n-1$ and $y_n = y_n'$, that is, if and only if $Y \overset{st}{\geq} Y'$ and $Y_n \overset{st}{=} Y_n'$.

Oftentimes one shows that families of random variables are stochastically ordered by showing that they satisfy a stronger condition called TP_2 defined below. A function ϕ defined on R_2 is said to be *totally positive of order* 2 (TP_2) if it is nonnegative and satisfies

$$\phi(\lambda_1, x_1)\phi(\lambda_2, x_2) \geq \phi(\lambda_1, x_2)\phi(\lambda_2, x_1),$$

whenever $\lambda_1 < \lambda_2$, $x_1 < x_2$.

Let μ denote either the Lebesgue measure on $[0, \infty]$ or the counting measure on the set of non-negative integers. A function defined on $(0, \infty) \times [0, \infty)$ is said to possess a *semigroup property* in λ if

$$\phi(\lambda_1 + \lambda_2, x) = \int_0^\infty \phi(\lambda_1, x - y)\phi(\lambda_2, y)\,d\mu(y).$$

A class of theorems generally known as preservation theorems allows us to construct new Schur functions and understand their structure. The following is one of the first preservation theorems for Schur functions. We will see later that by using the TP_2 and Schur properties with a variety of preservation theorems, several commonly used parametric families of distributions possess interesting Schur properties.

Theorem 1

Let $f(x)$ be a Schur convex (Schur concave) function and let $\phi(\lambda, x)$ defined on $(0, \infty) \times [0, \infty)$ possess the TP_2 property and the semigroup property in λ. Let μ be the Lebesgue measure or the counting measure. Let the integral

$$h(\lambda) = \int \prod_{i=1}^{n} \phi(\lambda_i, x_i) f(x)\,d\mu(x)$$

be well defined. Then $h(\lambda)$ is Schur convex (Schur concave).

This theorem appears as the main theorem in [PS 77]. In the principal application of this theorem, one takes ϕ to be a probability density function and shows that the operation of taking the expected value of a Schur convex function transfers the Schur convexity to the parameter vector.

Theorem 2

Let X and X' be a pair of n-vectors and define $S = \sum_{i=1}^{n} X_i$ and $S' = \sum_{i=1}^{n} X'_i$. Then $X \overset{st.m.}{\geq} X'$ if and only if (a) $S \overset{st}{=} S'$ and (b) for each bounded Schur convex function f, $E[f(X)|S = s] \geq E[f(X')|S' = s]$, for all $s \in A_f$, where the distribution of S assigns probability one to A_f.

This theorem is one of the important tools to be found in [NPS 77]. The notion of a *Schur family* extends the concept of stochastic majorization to a family of random variables. Let X_λ be a family of random vectors with a distribution P_λ indexed by λ in R^n. The family X_λ and the family P_λ are said to be Schur families if $\lambda \overset{m}{\geq} \lambda'$ implies that $X_\lambda \overset{st.m.}{\geq} X_{\lambda'}$.

The following theorem shows that in Schur families, stochastic majorization is preserved among the posterior distributions when there is stochastic majorization among the prior distributions.

Theorem 3

Let $\{X_\lambda\}$ be a Schur family in λ. Let G_1 and G_2 be two prior distributions for λ, such that $G_1 \overset{st.m.}{\geq} G_2$. Then the posterior of X_λ under G_1 stochastically majorizes the posterior of X_λ under G_2.

Example 2. Shock Models.

Consider a system subject to a series of shocks and assume that the different types of shocks arrive in a Poissonian fashion. For example, suppose that $X_i(t)$ denote the number of shocks of the i^{th} type arriving in the interval $[0, t]$. Let $\bar{P}(\mathbf{k})$, where $\mathbf{k} = (k_1, k_2, \ldots, k_n)$, be the survival probability of the system surviving k_i shocks of the type i, $i = 1, 2, \ldots, n$. Suppose that for each i, the random variable $X_i(t)$ has a Poisson distribution with parameter $\lambda_i t$. Then it follows that the survival function of the system is given by

$$\bar{H}(t; \lambda) = E\left[\bar{P}\Big(X_1(t), X_2(t), \ldots, X_n(t)\Big)\right].$$

Assume further that $\bar{P}$ is Schur concave in $\mathbf{k}$. This assumption holds, for example, if the effects of shocks are independent and the $\bar{P}$ is the product of n survival functions, each of which is logconcave. The TP_2 property of Poisson density functions and Theorem 1 show that the survival function $\bar{H}(t; \lambda)$ is Schur concave in λ. For details see [PS 77].

Example 3. Schur Function of Partial Sums.

Let X_{ij}, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k_i$ be independent identically distributed random variables with common logconcave density function g. Let f be a Schur concave function and consider

$$h(\mathbf{k}) = E\left[f\left(\sum_{j_1=1}^{k_1} X_{1,j_1}, \sum_{j_2=1}^{k_2} X_{2,j_2}, \ldots, \sum_{j_n=1}^{k_n} X_{1,j_n}\right)\right].$$

According to a result of Karlin and Proschan (1960), the k-fold convolution $g^{(k)}(x)$ is TP_2 in k and x. Using this and Theorem 1, it follows that $h(\mathbf{k})$ is Schur concave in $\mathbf{k}$.

Example 4. Schur Concavity of Moments.

Let g be a Schur concave density with the support $[0, \infty]^n$. Let α_i, $i = 1, 2, \ldots, n$, be positive numbers and let

$$M(\alpha) = \int \cdots \int \frac{\prod_{i=1}^{n} x_i^{\alpha_i - 1}}{\prod_{i=1}^{n} \Gamma(\alpha_i)}\, g(\mathbf{x})\, d\mathbf{x}$$

be a *multivariate normalized moment.* One can rewrite the integrand as $\left[\prod_{i=1}^{n}\left\{x_i^{\alpha_i-1}e^{-x_i}/\Gamma(\alpha_i)\right\}g(\boldsymbol{x})exp\{\sum x_i\}\right]$. Note that $g(\boldsymbol{x})exp\{\sum x_i\}$ is Schur concave and that $\{x^{\alpha-1}e^{-x}/\Gamma(\alpha)\}$ is TP_2 in (α, x) and is a semigroup on $(0, \infty)$. From Theorem 1 it follows that $M(\alpha)$ is Schur concave. Note that there are examples where $M(\alpha)$ is Schur convex if the normalizing constant $\Gamma(\alpha)$ is omitted in the integrand.

Example 5. Schur Families.

A number of parametric families found in standard textbooks can be shown to be Schur families. To name a few: multinomial, multivariate negative binomial, multivariate hypergeometric, Dirichlet. Furthermore, families of independent random variables such as Poisson, Gamma etc. also form Schur families. A host of such examples are listed and demonstrated in [NPS 77].

Functions Decreasing in Transposition

The partial ordering of majorization can sometimes be better understood by a standard partial ordering on the space of permutation on the set of n integers $(1, 2, \ldots, n)$. This leads to the concept of functions which are *decreasing in transposition* (DT) which extends the concept of Schur functions.

Let $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$ denote a permutation of $(1, 2, \ldots, n)$. Let S denote the group of such permutations $\boldsymbol{\pi}$. Suppose that $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ differ only in two of their components, say the i^{th} and j^{th}, where $i < j$, $\pi_i < \pi_j$ and that $\pi_i' = \pi_j$, $\pi_j' = \pi_i$. We say that $\boldsymbol{\pi}'$ is a *simple transposition* of $\boldsymbol{\pi}$. If a member of S, say $\boldsymbol{\pi}''$ is obtained from $\boldsymbol{\pi}$ by successive simple transpositions, we say that $\boldsymbol{\pi}$ dominates $\boldsymbol{\pi}''$ in transposition and write $\boldsymbol{\pi} \underset{t}{\geq} \boldsymbol{\pi}''$. Clearly this relation establishes a partial ordering in S.

Suppose that the components of $\boldsymbol{x}$ are such that $x_1 \leq x_2 \leq \ldots \leq x_n$. A permutation obtained by composing it with $\boldsymbol{\pi}$ is denoted by $\boldsymbol{\pi} \circ \boldsymbol{x}$ and defined by

$$\boldsymbol{\pi} \circ \boldsymbol{x} = (x_{\pi_1}, x_{\pi_2}, \ldots, x_{\pi_n}).$$

The partial ordering defined above can be extended in an obvious way to the vectors obtained by permuting components of $\boldsymbol{x}$.

In many applications one considers two vectors, the first vector corresponding to a parameter and the second vector to an observed random variable. It is useful to describe in a mathematical fashion the fact that a random vector and its parameter vector increase and decrease together. Oftentimes one needs to study and compare the way in which two random vectors vary together. For instance, one use of rank correlation is to measure how similarly two random vectors vary together. We will see below that the partial ordering on permutations, defined above, provides a satisfactory way to compare how similarly two vectors, which may be random or deterministic, vary together.

Let Λ and Ξ be subsets of R. A function $g(\boldsymbol{\lambda}, \boldsymbol{x})$ is said to be *decreasing in transposition* (DT) on $\Lambda^n \times \Xi^n$ if $g(\boldsymbol{\lambda} \circ \boldsymbol{\pi}, \boldsymbol{x} \circ \boldsymbol{\pi}) = g(\boldsymbol{\lambda}, \boldsymbol{x})$, for every $\boldsymbol{\pi}$ (that is, g is invariant under the same permutation on the two vectors) and $g(\boldsymbol{\lambda}, \boldsymbol{x} \circ \boldsymbol{\pi}) \geq g(\boldsymbol{\lambda}, \boldsymbol{x} \circ \boldsymbol{\pi}')$, where $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$; $x_1 \leq x_2 \leq \ldots \leq x_n$ and $\boldsymbol{\pi} \overset{t}{\geq} \boldsymbol{\pi}'$.

When $g(x, y)$ is DT the function $g(x, y)$ gives larger values when the ranking in the pair (x, y) is more similarly ordered than when the ranking is less similarly ordered.

In certain applications there is only one vector and it is desirable to define functions of a single vector which exhibit a monotonicity under this partial ordering. Let h be defined on Ξ^n and suppose that the components of $\boldsymbol{x}$ are in increasing order. Then h is said to be DT if $h(\boldsymbol{x} \circ \boldsymbol{\pi}) \geq h(\boldsymbol{x} \circ \boldsymbol{\pi}')$ whenever $\boldsymbol{\pi} \overset{t}{\geq} \boldsymbol{\pi}'$.

DT functions occur quite frequently in statistics. The book of Marshall and Olkin (1979) has popularized the notion of DT functions under the more positive sounding name of *Arrangement Increasing* (AI) functions. The following theorem shows the relation between DT, Schur and TP_2 functions.

Theorem 4

(a) Suppose $g(\boldsymbol{\lambda}, \boldsymbol{x}) = h(\boldsymbol{\lambda} - \boldsymbol{x})$. Then g is DT on R^{2n} if and only if h is Schur concave.

(b) Suppose $g(\boldsymbol{\lambda}, \boldsymbol{x}) = h(\boldsymbol{\lambda} + \boldsymbol{x})$. Then g is DT on R^{2n} if and only if h is Schur convex.

(c) Suppose $g(\boldsymbol{\lambda}, \boldsymbol{x}) = \prod h(\lambda_i, x_i)$, then g is DT on R^{2n} if and only if h is TP_2.

The main result on DT functions is the following preservation theorem which states that the DT property is preserved under the operation of composition.

Theorem 5

Let g_i, $i = 1, 2$ be DT on R^{2n} and σ be a measure on R^n such that for every Borel set A in R^n, $\sigma(A) = \sigma(\boldsymbol{\pi} \circ A)$ for every $\boldsymbol{\pi}$. Suppose that

$$g(\boldsymbol{x}, z) = \int_A g_1(\boldsymbol{x}, \boldsymbol{y}) g_2(\boldsymbol{y}, \boldsymbol{x}) d\sigma(\boldsymbol{y}),$$

is well defined. Then g is DT on R^{2n}.

The proofs of the above two theorems can be found in [HPS 77].

Theorem 1 can be derived as a consequence of Theorem 5 and Theorem 4(b). Furthermore, the following result of Marshall and Olkin (1974) can also be obtained from Theorem 5 and Theorem 4(a).

Theorem 6

The convolution of two Schur concave functions is Schur concave.

Most of the families considered in the second section can also be shown to have DT property. In some sense this provides a better tool than Schur concavity because of the connections seen earlier. One of the interesting applications is the problem in ranking. Suppose the vector X has density $\phi(\boldsymbol{\lambda}, \boldsymbol{x})$ which is a DT function. Let $g(\boldsymbol{\lambda}, \boldsymbol{r})$ be the probability that the rank vector of X-observations is $\boldsymbol{r}$. By using Theorem 5 above, it can be shown that g is DT. This has important consequences in nonparametric statistics. For details of this please see [HPS 77].

It should be noted that the concept of Schur concavity is closely related to that of unimodality. From the above discussion it can be seen that a function defined on R^2 is Schur concave if and only if it is permutation invariant and its graph is such that it is unimodal on every section perpendicular to the line of equality. This definition can be extended to R_n by considering all bivariate sections obtained by fixing $(n-2)$ arguments and requiring Schur concavity for each section, in the sense just described.

The convolution of two symmetric univariate unimodal densities can be shown to be a symmetric unimodal density. This is known as Wintner's theorem. Using this result it follows that the convolution of two bivariate Schur concave densities is Schur concave. Again by considering sections, an alternative proof for Theorem 6 can be provided.

The condition that the set $\{\boldsymbol{x}: f(\boldsymbol{x}) \geq c\}$ be convex and permutation invariant, for every $c > 0$, is sufficient for all the required sections of an n-variate density $f(x)$ to be symmetric unimodal. Many results that follow from such basic unimodality have been explored in a book by Joag-Dev and Dharmadhikari (1988) which are useful in deriving various properties of Schur concave functions. For instance, consider a random vector whose density function is logconcave. The logconcavity implies that the set where the density exceeds a given constant is a convex set and hence it satisfies the condition state above. If the components of this random vector are also exchangeable, then the density function is Schur concave.

An important theorem for multivariate logconcave densities is due to Prékopa (1973) which is stated below.

Theorem 7

Let $Y = (Y_1, Y_2, \ldots, Y_m)$ have logconcave density. Then $Z = (Z_1, Z_2, \ldots, Z_k) \stackrel{def}{=} (\sum a_{1,i} Y_i, \sum a_{2,i} Y_i, \ldots, \sum a_{k,i} Y_i)$ also has a logconcave density. In particular all marginals have logconcave densities.

We will use this theorem to derive Schur concavity and DT properties of densities of some random vectors obtained as overlapping sums of random variables. We begin with a simple case before going to the general case because the notation can get quite complicated.

Theorem 8

Let X_{12}, X_{23}, X_{13} and X_{123} be random variables such that X_{12}, X_{23}, X_{13} are exchangeable. Define

$$X_1^{(2)} = X_{12} + X_{13},$$

$$X_2^{(2)} = X_{12} + X_{23},$$

$$X_3^{(2)} = X_{13} + X_{23},$$

$$T_1 = X_1^{(2)} + X_{123},$$

$$T_2 = X_2^{(2)} + X_{123},$$

$$T_3 = X_3^{(2)} + X_{123}.$$

Then the density of $T \overset{def}{=} (T_1, T_2, T_3)$ is Schur concave under either one of the following conditions:

(A) the joint density of X_{12}, X_{13}, X_{23}, X_{123} is log concave

(B) the random vector (X_{12}, X_{13}, X_{23}) has a logconcave density and is independent of the random variable X_{123}.

Proof. Note that T consists of overlapping sums of random variables. A more general case of overlapping sums will be considered later.

From the definition of T it is easy to see that it is exchangeable. The logconcavity of the density of T follows readily from Prékopa's theorem (Theorem 7) under condition (A). This establishes the Schur concavity of the density of T under (A). When condition (B) holds, Prékopa's theorem (Theorem 7) once again shows that the density $f(x_1, x_2, x_3)$ of $(X_1^{(2)}, X_2^{(2)}, X_3^{(2)})$ is Schur concave. The density function of T is given by

$$\int f(x_1 - y, x_2 - y, x_3 - y) g(y)\, dy$$

where $g(y)$ is the density function of $X_{1,2,3}$. Since a positive mixture of Schur concave functions is Schur concave, it follows that the density of T is Schur concave. □

We now generalize the above to random vectors in R^n. Let $J = \{1, 2, 3, \ldots, n\}$. For $k = 2, \ldots, n$, let

$$I_k = \{I : I \text{ is a subset of } J \text{ with cardinality } k\},$$

and

$$I^* = \bigcup_2^k \{I \in I_k\} \text{ and } I_{k,i} = \{I \in I_k : i \in I\}.$$

Let $\{X_i,\ i = 1, 2,\ldots,n\}$ and X_I, $I \in I^*$ be a collection of random variables. Let $W(k) = \{X_I : I \in I_k\}$, $X_i^{(k)} = \sum_{I \in I_{k,i}} X_I$ and $X^{(k)} = (X_1^{(k)}, X_2^{(k)},\ldots,X_n^{(k)})$ where $i = 1, 2,\ldots,n$ and $k = 2, 3,\ldots,n$. Thus $X_i^{(k)}$ is the sum of random variables, each having k subscripts, one of which is i.

Theorem 9

Let $X^{(1)} = (X_1, X_2,\ldots,X_n)$ be a random vector with probability density function which is DT. Suppose that the set $\{X_I,\ I \in I^*\}$ is independent of $X^{(1)}$ and one of the following conditions holds.

(A) The set of all variables $\{X_I,\ I \in I^*\}$ is exchangeable and has a logconcave joint density function.

(B) The collection of random variables in $W(k)$ has a logconcave density and is permutation invariant for $k = 2, 3,\ldots,n-1$, and the collections $W(2), W(3),\ldots, W(n)$ are independent.

Then the joint distribution of $Z = (Z_1, Z_2,\ldots,Z_n)$ is DT, where

$$Z_i = X_i + \sum_{k \geq 2} X_i^{(k)}.$$

Proof. The argument is similar to the proof of Theorem 8. Let $T_i = \sum_{k \geq 2} X_i^{(k)}$ and $T = \{T_1, T_2,\ldots,T_n\}$.

The density function of Z is the convolution of the density functions of T and $X^{(1)}$, the second of which is DT by assumption. If we can show that the density function of T is Schur concave, then it will follow that the density function of Z is DT from Theorems 4 and 5(a).

We will now show that condition (A) or (B) implies the Schur concavity of the density of T.

When condition (A) holds it is easy to see that Prékopa's theorem implies that the joint density of T is logconcave. The permutation invariance of this joint density follows from the exchangeability of $\{X_I,\ I \in I^*\}$. This establishes the joint density function satisfies the DT property.

When condition (B) holds, Prékopa's theorem once again shows that the density function of $X^{(k)}$ is log concave for $k = 2, 3,\ldots,n-1$ and is permutation invariant. From the independence of $W(k)$, $k = 2, 3,\ldots,n-1$ it follows that the density function of $T_1 - X_{\{1,\ldots,n\}},\ldots, T_n - X_{\{1,\ldots,n\}}$ is logconcave and permutation invariant and hence Schur concave. Notice that $W(n) = X_{\{1,\ldots,n\}}$ consists of a single random variable. From the same argument given in case (B) of Theorem 8, it follows that the density of T is Schur concave.

This completes the proof of Theorem 9. □

Theorem 9 generalizes Theorem 2.1 of [HPS 81] and contains a new proof. As an application of this theorem it can be shown that the density function of a generalized compound multivariate Poisson is DT. See [HPS 81] for details.

References

Abouammoh, A. M., El-Neweihi and Proschan, F. (1989): Schur structure functions, *Prob. Engg. Inform. Sc.* 3, 581-591.

Boland, P. J., Proschan, F. and Tong, Y. L. (1987): Fault diversity in software reliability, *Prob. Engg. Inform. Sc.* 1, 175-188.

Boland, P.J., Proschan, F. and Tong, Y. L. (1988): Moment and geometric probability inequalities arising from arrangement increasing functions, *Ann. Prob.* 16, 407-413.

Dharmadhikari, S. W. and Joag-Dev, K. (1988): *Unimodality, Convexity, and Applications*, Academic Press, New York.

Hardy, G. H., Littlewood, J.E. and Pólya, G. (1952): *Inequalities*, 2nd ed., Cambridge University Press, New York.

Hollander, M., Proschan, F. and Sethuraman, J. (1977): Functions decreasing in transposition and their applications in ranking problems, *Ann. Stat.* 5, 722-733.

Hollander, M., Proschan, F. and Sethuraman, J. (1981): Decreasing in transposition property of overlapping sums, and applications, *Jour. Mult. Anal.* 11, 50-57.

Karlin, S. and Proschan, F. (1960): Pólya type distributions of convolutions, *Ann. Math. Statist.* 31, 721-736.

Marshall, A. and Olkin, I. (1974): Majorization in multivariate distributions, *Ann. Statist.* 2, 1189-1200.

Marshall, A and Olkin, I. (1979): *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York.

Nevius, S.E., Proschan, F. and Sethuraman, J. (1977): Schur functions in Statistics II. Stochastic majorization, *Ann. Stat.* 5, 263-273.

Prékopa, A. (1973): On logarithmic concave measures and functions, *Acta Sci. Mat.* 34, 335-343.

Proschan, F. and Sethuraman, J. (1977): Schur functions in Statistics I. The preservation theorem, *Ann. Stat.* 5, 253-262.

Tong, Y. L. (1980): *Probability Inequalities in Multivariate Distributions*, Academic Press, New York.