Multivariate Analysis and Its Applications
IMS Lecture Notes – Monograph Series (1994) Volume 24

SOME USEFUL NOTIONS FOR STUDYING STOCHASTIC INEQUALITIES IN MULTIVARIATE DISTRIBUTIONS

BY Y. L. TONG

Georgia Institute of Technology

In this expository paper we review some of the useful notions for studying stochastic inequalities in multivariate distributions. The notions have been classified into two categories: That which involve conditions on the joint density function of the random vector and that which involve certain positive dependence properties of the components of the random vector. Their possible implications and orderings are summarized, and examples of applications are given.

1. Introduction. Stochastic inequalities play an important role in many areas of statistics and probability. In the area of multivariate analysis, inequalities have become a useful tool for obtaining conservative confidence regions, establishing certain monotonicity properties of multivariate tests, finding probability bounds in multiple comparisons and related inference problems, etc. Such applications are well known and can be found in standard multivariate analysis books.

On the other hand, the theory of stochastic inequalities has intrinsic interest and importance, and need not rely only on applications. Although the general study of stochastic inequalities can be traced back to the days of C. F. Gauss, A. L. Cauchy, and P. L. Čebyšev, it is only recently that this area has experienced a rapid and more comprehensive growth. In this expository paper we discuss and review some of the mathematical notions that have been found useful in deriving stochastic inequalities in multivariate distributions. We will focus on a systematic treatment of the notions and "methods" instead of a description of existing results. Consequently, no attempts will be made

---

* Research supported in part by a grant from the U.S. AFOSR through NSF Grant DMS-9149151.

AMS 1991 Subject Classifications: 60E15, 62H99.

Key words and phrases: Arrangement increasing functions, association of random variables, exchangeability, log-concavity, majorization and Schur-concavity, MTP_2 density functions, multivariate probability inequalities, orthant dependence, stochastic inequalities, unimodality.

to provide a complete listing of the existing theorems in this area.

Throughout this paper, $\boldsymbol{X} = (X_1, \cdots, X_n)$ denotes an n-dimensional random vector with probability density function $f(\boldsymbol{x})$ and distribution $F(\boldsymbol{x})$, and B denotes a Borel-measurable set either in $\mathbb{R}$ or in $\mathbb{R}^n$ (to be specified). The probability content of $B \subset \mathbb{R}^n$ is then given by

$$P[\boldsymbol{X} \in B] = \int_B f(\boldsymbol{x}) d\boldsymbol{x}. \tag{1}$$

It appears that most of the existing results in the literature concerning inequalities and partial orderings of the probability content in (1) can be obtained through

(a) conditions on f (and B), such as the notions of unimodality, Schur–concavity, log–concavity and arrangement increasingness; or

(b) certain positive and negative dependence properties of the components of $\boldsymbol{X}$.

Related notions in (a) will be treated in Section 2 of this paper, and certain positive dependence properties will be discussed in Section 3. (In order not to overload this paper, notions of negative dependence will not be considered.)

2. Notions Concerning the Density Function.

2.1. Notion U: Unimodality. In the univariate case, the density function $f(x)$ of a continuous random variable X is said to be unimodal if the set $\{x : f(x) > \lambda\}$ is an interval for every given $\lambda > 0$. The following definition for n-dimensional random vectors, given by Anderson (1955), seems to be a natural generalization.

DEFINITION 1. *A density function $f(\boldsymbol{x}) : \mathbb{R}^n \to [0, \infty)$ is said to be A-unimodal if the set $\{\boldsymbol{x} : f(\boldsymbol{x}) \geq \lambda\}$ is convex for all $\lambda > 0$.*

In the univariate case, if $f(x)$ is unimodal and symmetric about the origin, and if $B \subset \mathbb{R}$ is an interval centered at the origin, then it is immediate that

$$P[X \in B + \alpha^*] \text{ is nonincreasing in } |\alpha^*|;$$

or equivalently, for every given $u \neq 0$ we have

$$P[X \in B + (\alpha u + (1-\alpha)(-u))] \geq P[X \in B + u]$$

for all $\alpha \in [0, 1]$. Anderson (1955) showed that a multivariate analog of this statement is true:

THEOREM 2. (Anderson, 1955) *Let $\boldsymbol{u}$ be any given non-zero vector in $\mathbb{R}^n$, and for $\alpha \in [0, 1]$ let $\boldsymbol{u}_\alpha = \alpha \boldsymbol{u} + (1-\alpha)(-\boldsymbol{u})$. If (i) f is A-unimodal* *and symmetric about the origin (i.e., $f(\boldsymbol{x}) = f(-\boldsymbol{x})$ holds for all $\boldsymbol{x}$), and (ii) $B \subset I\!R^n$ is convex and symmetric about the origin, then*

$$P[\boldsymbol{X} \in B + \boldsymbol{u}_\alpha] \geq P[\boldsymbol{X} \in B + \boldsymbol{u}] \quad (= P[\boldsymbol{X} \in B - \boldsymbol{u}]). \tag{2}$$

There have been several versions of generalizations of this theorem (see, e.g., Sherman (1955), Mudholkar (1966), and Eaton and Perlman (1977)), a comprehensive discussion on this topic can be found in Dharmadhikari and Joag-Dev (1988, Chapter 2).

An interesting question is whether the inequality in (2) also holds for other types of symmetric properties of f and B. In particular, it is of interest to know what can be said when f and B are both *permutation symmetric* instead. This consideration leads to the important notion of majorization and Schur functions.

2.2. Notion S: Schur-Concavity. The notion of majorization concerns the diversity of the components of a vector. Let

$$\boldsymbol{a} = (a_1, \cdots, a_n), \qquad \boldsymbol{b} = (b_1, \cdots, b_n),$$

denote two real vectors. Let

$$a_{[1]} \geq a_{[2]} \geq \cdots \geq a_{[n]}, \qquad b_{[1]} \geq b_{[2]} \geq \cdots \geq b_{[n]},$$

be their ordered components.

DEFINITION 3. *$\boldsymbol{a}$ is said to majorize $\boldsymbol{b}$, in symbols $\boldsymbol{a} \succ \boldsymbol{b}$, if*

$$\sum_{i=1}^{m} a_{[i]} \geq \sum_{i=1}^{m} b_{[i]} \quad \textit{holds for} \quad m = 1, 2, \cdots, n-1$$

and $\sum_{i=1}^{n} a_{[i]} = \sum_{i=1}^{n} b_i$.

This definition provides a partial ordering, namely, $\boldsymbol{a} \succ \boldsymbol{b}$ implies that (for a fixed sum) the a_i's are more diverse than the b_i's. From a geometric viewpoint, majorization is also closely related to convex combinations of the permutations of a vector. For a comprehensive treatment of this topic, see Marshall and Olkin (1979, Chapters 1–3) and Hardy, Littlewood and Pólya (1934, 1952; Chapters 2–3).

DEFINITION 4. *A function $f(\boldsymbol{x}) : I\!R^n \to I\!R$ is said to be a Schur–concave function if $\boldsymbol{x} \succ \boldsymbol{y}$ implies $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in I\!R^n$.*

(That is, $f(\boldsymbol{x})$ is a Schur-concave function if the functional value becomes larger when the components of $\boldsymbol{x}$ are less diverse in the sense of majorization.)

The following fact states how Schur-concavity, permutation symmetry, and unimodality are related:

FACT 5. *(a) All Schur-concave functions are permutation symmetric. (b) If $f(\boldsymbol{x})$ is permutation symmetric and A-unimodal, then it is a Schur-concave function of $\boldsymbol{x}$.*

In one of the earlier papers on majorization inequalities in multivariate distributions, Marshall and Olkin (1974) proved the following preservation theorem:

THEOREM 6. *Consider the integral of the form*

$$\psi(\boldsymbol{\theta}) = \int_{\mathbb{R}^n} g(\boldsymbol{\theta} - \boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}, \qquad \boldsymbol{\theta} \in \mathbb{R}^n. \tag{3}$$

If g and f are both Schur-concave functions defined on $\mathbb{R}^n$, then $\psi(\boldsymbol{\theta})$ is a Schur-concave function (provided that the integral exists).

This theorem has a number of interesting applications. As an illustration of the applications, we observe the following result (Marshall and Olkin, 1974): Since the distribution function $F(\boldsymbol{a})$ of an n-dimensional random vector $\boldsymbol{X}$ can be expressed in the form

$$\begin{aligned} F(\boldsymbol{a}) &= \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} f(\boldsymbol{x})\, d\boldsymbol{x} \\ &= \int_{\mathbb{R}^n} I_A(\boldsymbol{a} - \boldsymbol{x}) f(\boldsymbol{x})\, d\boldsymbol{x}, \end{aligned}$$

where I_A is the indicator function of the Schur-concave set

$$A = \{\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n, x_i \geq 0, i = 1, \cdots, n\},$$

it follows from Theorem 6 that:

APPLICATION 7. If $\boldsymbol{X} = (X_1, \cdots, X_n)$ has a Schur-concave density function, then its distribution function $F(\boldsymbol{a})$ is a Schur-concave function of $\boldsymbol{a} \in \mathbb{R}^n$.

2.3. Notion L: Log-Concavity. Log-concave density functions have many nice analytical properties, and play an important role in statistics.

DEFINITION 8. *A density function $f(\boldsymbol{x}) : \mathbb{R}^n \to [0, \infty)$ is said to be log–concave if*

$$f(\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) \geq [f(\boldsymbol{x})]^{\alpha} [f(\boldsymbol{y})]^{1-\alpha} \tag{4}$$

holds for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and all $\alpha \in [0, 1]$.

[If $f(\boldsymbol{x}) > 0$ *holds for all* $\boldsymbol{x} \in \mathbb{R}^n$*, then an equivalent condition is*

$$\log f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq \alpha \log f(\boldsymbol{x}) + (1-\alpha) \log f(\boldsymbol{y}).] \tag{4'}$$

The following result shows how the notion of log-concavity is related to notions of A-unimodality and Schur-concavity:

FACT 9. (a) *All log-concave density functions are* A*-unimodal.* (b) *If* $f(\boldsymbol{x})$ *is permutation symmetric and log–concave, then it is Schur-concave.*

In many applications, $X_1, \cdots, X_n$ are i.i.d. random variables with a common density function $h(x)$. The following fact explains how the log-concavity of h (defined on an interval in $\mathbb{R}$) and the log-concavity of $f(\boldsymbol{x})$ (defined on an interval in $\mathbb{R}^n$) are related.

FACT 10. *If* $h(x)$ *is a log-concave function of* $x \in I$ *(an interval in* $\mathbb{R}$*), then the joint density function* $f(\boldsymbol{x}) = \prod_{i=1}^n h(x_i)$ *is a log-concave function of* $\boldsymbol{x} \in I \times I \times \cdots \times I$.

A fundamental result for log-concave density functions, as given by Prékopa (1971), concerns probability contents of convex combinations of Borel sets in $\mathbb{R}^n$. Let A, B be subsets of $\mathbb{R}^n$. For arbitrary but fixed $\alpha \in [0,1]$ we define

$$\alpha A + (1-\alpha)B = \{\boldsymbol{z} : \boldsymbol{z} \in \mathbb{R}^n, \boldsymbol{z} = \alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y} \text{ for some } \boldsymbol{x} \in A \text{ and } \boldsymbol{y} \in B\}.$$

Prékopa's theorem states:

THEOREM 11. *If* $f(\boldsymbol{x})$ *is log–concave and* A*,* B*, and* $\alpha A + (1-\alpha)B$ *are all Borel measurable, then*

$$P[\boldsymbol{X} \in (\alpha A + (1-\alpha)B)] \geq \{P[\boldsymbol{X} \in A]\}^{\alpha} \{P[\boldsymbol{X} \in B]\}^{1-\alpha} \tag{5}$$

holds for all $\alpha \in [0,1]$*. If the probability contents are all positive, then*

$$\log P[\boldsymbol{X} \in (\alpha A + (1-\alpha)B)] \geq \alpha \log P[\boldsymbol{X} \in A] + (1-\alpha) \log P[\boldsymbol{X} \in B] \tag{5'}$$

holds for all $\alpha \in [0,1]$.

Related results and generalizations of Theorem 11 include Borell (1975), Rinott (1976), and Das Gupta (1980), and a comprehensive review can be found in Eaton (1987, Chapter 4). Some specific applications of this theorem will be illustrated in Section 2.6.

2.4. Notion AI: Arrangement Increasingness. Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)$ be two real vectors in $\mathbb{R}^n$. Let $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_n)$ be a permutation of the set of integers $\{1, 2, \cdots, n\}$ and denote

$$\boldsymbol{\pi}(\boldsymbol{x}) = (x_{\pi_1}, \cdots, x_{\pi_n}), \qquad \boldsymbol{\pi}(\boldsymbol{\mu}) = (\mu_{\pi_1}, \cdots, \mu_{\pi_n}).$$

Consider a function f of $2n$ variables of the form

$$f(\boldsymbol{x}, \boldsymbol{\mu}) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}.$$

We say that f is an arrangement permutation symmetric function of $(\boldsymbol{x}, \boldsymbol{\mu})$ if for every permutation $\boldsymbol{\pi}$ we have

$$f(\boldsymbol{\pi}(\boldsymbol{x}), \boldsymbol{\pi}(\boldsymbol{\mu})) = f(\boldsymbol{x}, \boldsymbol{\mu}) \quad \text{for all} \quad \boldsymbol{x}, \boldsymbol{\mu} \in \mathbb{R}^n. \tag{6}$$

To observe a more general result, a notion of the partial ordering of permutations is needed. Let $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ be two real vectors. We define "$\boldsymbol{\mu} \overset{p}{<} \boldsymbol{\nu}$" to mean that for some indices i, j, $1 \le i < j \le n$,

$$\mu_j = \nu_i < \nu_j = \mu_i \quad \text{and} \quad \mu_k = \nu_k \quad \text{for all} \quad k \neq i, k \neq j.$$

That is, $\boldsymbol{\nu}$ can be obtained from $\boldsymbol{\mu}$ by interchanging μ_i and μ_j such that they are now rearranged in an ascending order while all other components in $\boldsymbol{\mu}$ are held fixed. Furthermore, we define "$\boldsymbol{\mu} \overset{b}{<} \boldsymbol{\nu}$" to mean that there exists a finite number of vectors $\boldsymbol{\nu}_1, \cdots, \boldsymbol{\mu}_N$ such that

$$\boldsymbol{\mu} = \boldsymbol{\nu}_1 \overset{p}{<} \boldsymbol{\nu}_2 \overset{p}{<} \cdots \overset{p}{<} \boldsymbol{\nu}_{N-1} \overset{p}{<} \boldsymbol{\nu}_N = \boldsymbol{\nu}.$$

This is to say that $\boldsymbol{\nu}$ can be obtained from $\boldsymbol{\mu}$ by rearranging two components at a time in this fashion in a finite number of operations.

DEFINITION 12. *$f(\boldsymbol{x}, \boldsymbol{\mu}) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is said to be an arrangement increasing (AI) function of $(\boldsymbol{x}, \boldsymbol{\mu})$ if f is arrangement permutation symmetric and if*

$$f(\boldsymbol{x}\uparrow, \boldsymbol{\mu}) \le f(\boldsymbol{x}\uparrow, \boldsymbol{\nu}) \text{ holds for all } \boldsymbol{x} \text{ and all } \boldsymbol{\mu}, \boldsymbol{\nu} \text{ in } \mathbb{R}^n \text{ such that } \boldsymbol{\mu} \overset{b}{<} \boldsymbol{\nu},$$

where $\boldsymbol{x}\uparrow = (x_{\pi_1}, \cdots, x_{\pi_n})$ is such that $x_{\pi_1} \le \cdots \le x_{\pi_n}$.

Many useful functions are known to be arrangement increasing (see e.g., Hollander, Proschan and Sethuraman (1977) and Marshall and Olkin (1979, Section 6.F)). In particular, it is easy to see that

FACT 13. *Let $f(\boldsymbol{x};\boldsymbol{\mu},\Sigma)$ be the density function of an elliptically contoured distribution with location parameter vector $\boldsymbol{\mu}$. If Σ is of the form $\sigma^2\boldsymbol{R}$ where $\boldsymbol{R}$ is a correlation matrix such that $\rho_{ij}=\rho\in(-1/n-1,1)$ for all $i\neq j$ and $\sigma^2>0$ is arbitrary but fixed, then f is an arrangement increasing function of $(\boldsymbol{x},\boldsymbol{\mu})$.*

The following result, due to Hollander, Proschan, and Sethuraman (1977), illustrates how certain arrangement increasing functions and Schur–concave functions are related:

FACT 14. *Assume that $f(\boldsymbol{x},\boldsymbol{\theta}):\mathbb{R}^n\times\mathbb{R}^n\to[0,\infty)$ is of the form $f(\boldsymbol{x},\boldsymbol{\theta})=g(\boldsymbol{x}-\boldsymbol{\theta})$ for some $g:\mathbb{R}^n\to[0,\infty)$. Then f is arrangement increasing on $\mathbb{R}^n\times\mathbb{R}^n$ iff g is Schur-concave on $\mathbb{R}^n$.*

Hollander, Proschan and Sethuraman (1977) proved a fundamental preservation theorem for the integral of an arrangement increasing function. Applying their theorem, Boland, Proschan and Tong (1988) obtained the following result:

THEOREM 15. *Let $g_1,g_2:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ be arrangement increasing functions, and let $h_1,h_2:\mathbb{R}\to\mathbb{R}$ be nondecreasing. If $\boldsymbol{X}$ has a permutation symmetric density function $f:\mathbb{R}^n\to[0,\infty)$, then*

$$\psi(\boldsymbol{a},\boldsymbol{b})=E[h_1(g_1(\boldsymbol{a},\boldsymbol{X}))h_2(g_2(\boldsymbol{X},\boldsymbol{b}))] \tag{7}$$

is an arrangement increasing function of $(\boldsymbol{a},\boldsymbol{b})$.

A number of stochastic inequalities for permutation symmetric random variables have been obtained from this result. For details, see Boland, Proschan, and Tong (1988) or Tong (1990, pp. 172–173).

2.5. A Summary of the Implications. In Figure 1 below we summarize the directions of implications of the classes of density functions that are A-unimodal (U), Schur-concave (S), and log-concave (L), respectively. For convenience we let P denote the class of density functions that are permutation symmetric. The following figure is a summary of Facts 5 and 9; Fact 14 concerns a subclass of arrangement increasing density functions, and is not represented in the figure.

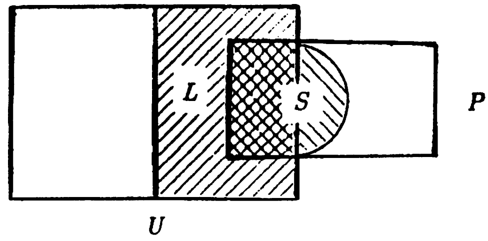

Figure 1. Summary of Implications

2.6. Examples of Applications and Some Open Problems. In the following we list some examples of applications of the mathematical notions treated above. The examples are for the purpose of illustration and, of course, they are not exhaustive.

EXAMPLE 16. (Concentration Inequalities). As discussed in Anderson (1955), Theorem 2 has some nice applications to concentration inequalities. One specific result concerns the multivariate normal distribution, and it states that: Let $\boldsymbol{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \Sigma)$, and let $B \subset I\!R^n$ be a convex set that is symmetric about the origin; let $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$ be two possible covariance matrices. If $\boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_1$ is positive semidefinite, then

$$P_{\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_1}[(\boldsymbol{X} - \boldsymbol{\mu}) \in B] \geq P_{\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_2}[(\boldsymbol{X} - \boldsymbol{\mu}) \in B]. \tag{8}$$

Inequality (8) has a number of applications in multivariate analysis, including results on the monotonicity property of certain tests under the assumption of normality.

EXAMPLE 17. (Probability Contents of Asymmetric Geometric Regions). For a large class of geometric regions in $I\!R^n$, a partial ordering of the degree of asymmetry may be obtained via the majorization ordering of the parameter vectors. For example, if we let

$$B_1(\boldsymbol{a}) = \{\boldsymbol{u} : \boldsymbol{u} \in I\!R^n, \boldsymbol{u} \leq \boldsymbol{a}\} \tag{9}$$

be a one sided n-dimensional rectangle, then $\boldsymbol{a} \succ \boldsymbol{b}$ implies that $B_1(\boldsymbol{b})$ is closer to being symmetric; and Application 7 states that if $f(\boldsymbol{x})$ (the density function of $\boldsymbol{X}$) is Schur–concave, then $\boldsymbol{a} \succ \boldsymbol{b}$ implies $P[\boldsymbol{X} \in B_1(\boldsymbol{a})] \leq P[\boldsymbol{X} \in B_1(\boldsymbol{b})]$. By the same token, if we consider two–sided rectangles of the form

$$B_2(\boldsymbol{a}) = \{\boldsymbol{u} : \boldsymbol{u} \in I\!R^n, |\boldsymbol{u}| \leq \boldsymbol{a}\}, \qquad \boldsymbol{a} > \boldsymbol{0} \tag{10}$$

or ellipsoids of the form

$$C(\boldsymbol{a}^2) = \left\{ \boldsymbol{u} : \boldsymbol{u} \in I\!R^n, \sum_{i=1}^{n} (u_i/a_i)^2 \leq \lambda \right\} \tag{11}$$

for $\boldsymbol{a}^2 = (a_1^2, \cdots, a_n^2) > \mathbf{0}$, where $\lambda > 0$ is arbitrary but fixed, then it seems reasonable to expect that similar results may hold. This was the motivation given by Tong (1982), and he showed that under the assumption that $f(\boldsymbol{x})$ is Schur-concave, (i) $\boldsymbol{a} \succ \boldsymbol{b}$ implies $P[\boldsymbol{X} \in B_2(\boldsymbol{a})] \leq P[\boldsymbol{X} \in B_2(\boldsymbol{b})]$, (ii) $\boldsymbol{a}^2 \succ \boldsymbol{b}^2$ implies $P[\boldsymbol{X} \in C(\boldsymbol{a}^2)] \leq P[\boldsymbol{X} \in C(\boldsymbol{b}^2)]$. A more general result was given by Karlin and Rinott (1983), and a comprehensive review on this topic may be found in Tong (1989) or Tong (1990, Section 7.4).

EXAMPLE 18. (Peakedness in Multivariate Distributions). Let $\{X_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. univariate random variables with density function $f^*(x)$, and mean μ, and for each fixed N consider the probability content

$$\beta_N = P[(\overline{X}_N - \mu) \in [-\lambda, \lambda]] = P[|\overline{X}_N - \mu| \leq \lambda], \qquad N = 1, 2, \ldots,$$

where $\overline{X}_N = N^{-1} \sum_{i=1}^{N} X_i$ and $\lambda > 0$ is arbitrary but fixed. The problem of interest, as considered by Proschan (1965), is when does $\{\beta_N\}$ converge to one *monotonically* in N. (Note that the weak law of large numbers does not yield this monotonicity property.) By realizing that $\overline{X}_{N-1} = \sum_{i=1}^{N} a_i X_i$, $\overline{X}_N = \sum_{i=1}^{N} b_i X_i$, where

$$\boldsymbol{a} = ((N-1)^{-1}, \ldots, (N-1)^{-1}, 0), \qquad \boldsymbol{b} = (N^{-1}, \cdots, N^{-1}, N^{-1}),$$

Proschan (1965) proved the following majorization inequality: If $f^*(x) : I\!R \to [0, \infty)$ is symmetric about μ and log-concave, and if $\boldsymbol{a} \succ \boldsymbol{b}$, then

$$P\left[\left|\sum_{i=1}^{N} a_i X_i - \mu\right| \leq \lambda\right] \leq P\left[\left|\sum_{i=1}^{N} b_i X_i - \mu\right| \leq \lambda\right] \qquad \text{for all } \lambda > 0. \tag{12}$$

A multivariate generalization of this result along two different directions was given by Olkin and Tong (1988), and some related results were obtained independently by Chan, Park, and Proschan (1989). The first Olkin–Tong result asserts that (12) holds when the joint density function of $(X_1, \cdots X_N)$ is permutation symmetric and satisfies some weaker conditions, but the X_i's are not necessarily i.i.d. The second result deals with a multivariate version of Proschan's (1965) result: If $\{\boldsymbol{X}_i\}_{i=1}^{\infty}$ is a sequence of i.i.d. n-dimensional random vectors with density function $f(\boldsymbol{x})$ that is symmetric about $\boldsymbol{\mu}$ and log-concave, and if $\boldsymbol{a} \succ \boldsymbol{b}$, then

$$P\left[\left(\sum_{i=1}^{N} a_i \boldsymbol{X}_i - \boldsymbol{\mu}\right) \in B\right] \leq P\left[\left(\sum_{i=1}^{N} b_i \boldsymbol{X}_i - \boldsymbol{\mu}\right) \in B\right] \tag{13}$$

holds for all convex sets $B \subset \mathbb{R}^n$ that are symmetric about the origin. The proof of this result depends on an application of Anderson's Theorem (Theorem 2), and is different in spirit from Proschan's original proof of the univariate case.

EXAMPLE 19. (Probability Contents for n-Dimensional Rectangles not Necessarily Centered at the Origin). Let

$$\boldsymbol{R} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \end{pmatrix} \equiv \begin{pmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \end{pmatrix}, \qquad \boldsymbol{S} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \end{pmatrix} \equiv \begin{pmatrix} \boldsymbol{s}_1 \\ \boldsymbol{s}_2 \end{pmatrix}$$

be two $2 \times n$ real matrices such that

$$r_{1j} < r_{2j}, \qquad s_{1j} < s_{2j} \qquad \text{for } j = 1, \cdots, n.$$

Let $B(\boldsymbol{R})$ be an n-dimensional rectangle given by

$$B(\boldsymbol{R}) = \{\boldsymbol{u} : \boldsymbol{u} \in \mathbb{R}^n, r_{1j} \leq u_j \leq r_{2j} \text{ for } j = 1, \cdots, n\},$$

and let $B(\boldsymbol{S})$ be defined similarly. Note that $B(\boldsymbol{R})$ is not centered at the origin unless $r_{1j} = -r_{2j}$ for all j. Further, note that if there exists a doubly stochastic matrix $\mathbf{Q}$ such that

$$\boldsymbol{S} = \boldsymbol{R}\boldsymbol{Q} \tag{14}$$

holds (which is stronger than saying that $\boldsymbol{r}_1 \succ \boldsymbol{s}_1$ and $\boldsymbol{r}_2 \succ \boldsymbol{s}_2$), then $B(\boldsymbol{S})$ is closer to being a cube. Karlin and Rinott (1983) and Tong (1983, 1989) independently proved the following result: If the density function $f : \mathbb{R}^n \to [0, \infty)$ of $\boldsymbol{X}$ is permutation symmetric and log-concave, and if (14) holds for some doubly stochastic matrix $\boldsymbol{Q}$, then

$$P[\boldsymbol{X} \in B(\boldsymbol{R})] \leq P[\boldsymbol{X} \in B(\boldsymbol{S})]. \tag{15}$$

The proof of Inequality (15) depends on a repeated application of Prékopa's theorem (Theorem 11).

OPEN PROBLEM 20. A conjecture given in both Karlin and Rinott (1983) and Tong (1983, 1989) is the following: If in Example 19 f is a Schur-concave function (which is weaker than saying that it is permutation symmetric and log-concave), and if (14) holds for some doubly stochastic matrix $\boldsymbol{Q}$, is Inequality (15) still valid? To our knowledge, this problem has not yet been solved.

OPEN PROBLEM 21. This open problem concerns a possible preservation property of Schur-concave density functions. Let $n = mk$ where $m \geq 2$,

$k \geq 2$ are arbitrary but fixed integers. Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ have density function f that is absolutely continuous w.r.t. Lebesgue measure, and define

$$Y_1 = \sum_{i=1}^{k} X_i, \qquad Y_2 = \sum_{i=k+1}^{2k} X_i, \; \cdots, \; Y_m = \sum_{i=(m-1)k+1}^{n} X_i. \tag{16}$$

Let $g(\boldsymbol{y}) : \mathbb{R}^m \to [0, \infty)$ be the density function of $\boldsymbol{Y} = (Y_1, \cdots, Y_m)$. The problem of interest is this: If f is a Schur-concave function in $\mathbb{R}^n$, is $g(\boldsymbol{y})$ also a Schur-concave function in $\mathbb{R}^m$? When f and g are absolutely continuous w.r.t. the counting measure, a counterexample has already been obtained (Boland, Proschan, and Tong (1991)), but an answer to the question stated above has not yet been found by the author.

3. Notions of Positive Dependence.

Notions of dependence have played a leading role in the study of stochastic inequalities in multivariate distributions, and they have been treated comprehensively in several books and monographs (see, e.g., Barlow and Proschan (1975, Section 2.2 and Chapter 5) and Tong (1980, Chapter 5; 1990, Chapter 5)). In this section we provide a brief survey of some of the notions and their implications.

3.1. Notion A: Association of Random Variables. This notion involves the positive dependence property resulted from monotone transformations of random variables. We first observe a classical result due to Čebyšev (1882, 1883):

THEOREM 22. *Let X be a univariate random variable. Then $Corr(g_1(X), g_2(X)) \geq 0$ (i. e., $Eg_1(X)g_2(X) \geq Eg_1(X)Eg_2(X)$) holds for all nondecreasing functions g_1 and g_2 such that the expectations exist.*

Intuitively speaking, if g_1, g_2 are both nondecreasing. Then $g_1(X), g_2(X)$ tend to take larger values together or smaller values together, thus are nonnegatively correlated. A multivariate generalization of this result involves the following definition of association of random variables:

DEFINITION 23. (Esary, Proschan, and Walkup, (1967)). *Random variables $X_1, \cdots, X_n$ are said to be (positively) associated (A) if*

$$Eg_1(X_1, \cdots, X_n)g_2(X_1, \cdots, X_n) \geq Eg_1(X_1, \cdots, X_n)Eg_2(X_1, \cdots, X_n) \tag{17}$$

holds, or equivalently,

$$Corr(g_1(X_1, \cdots, X_n), g_2(X_1, \cdots, X_n)) \geq 0$$

holds, for all nondecreasing functions g_1 and g_2.

The following theorem, due to Esary, Proschan, and Walkup (1967), is a multivariate generalization of Theorem 22.

THEOREM 24. (a) *A set consisting of a single random variable is a set of associated random variables.* (b) *Independent random variables are associated.*

3.2. Notion MTP_2: Multivariate Totally-Positive-of-Order-Two Density Functions. The following definition can be found in Karlin (1968, p. 11) and Karlin and Rinott (1980).

DEFINITION 25. *$f(\boldsymbol{x})$ is said to be multivariate-totally-positive-of-order-two (MTP_2) if the inequality*

$$f(\boldsymbol{y})f(\boldsymbol{y}^*) \leq f(\boldsymbol{x})f(\boldsymbol{x}^*) \tag{18}$$

holds for all $\boldsymbol{y} = (y_1, \cdots, y_n)$ and $\boldsymbol{y}^* = (y_1^*, \cdots, y_n^*)$ in the domain of f where

$$x_i^* = \min\{y_i, y_i^*\}, \qquad x_i = \max\{y_i, y_i^*\}$$

for $i = 1, \cdots, n$ and $\boldsymbol{x} = (x_1, \cdots, x_n)$, $\boldsymbol{x}^* = (x_1^*, \cdots, x_n^*)$.

Intuitively speaking, if the joint density function has the MTP_2 property, then the likelihood function takes a larger value when the components of the random vector $\boldsymbol{X}$ take smaller values together and larger values together. Note that this definition depends on a condition on the density function, while the notion of association in Definition 23 involves the expectations of functions of random variables. The following theorem, known as the FKG inequality, states how the two notions are related:

THEOREM 26. *If the density function of $(X_1, \cdots, X_n)$ has the MTP_2 property, then $X_1, \cdots, X_n$ are associated random variables.*

3.3. Notion OD: Orthant Dependence. In one of the earlier papers on stochastic inequalities in multivariate distributions, Lehmann (1966) introduced the following definition of orthant dependence (which depends only on the orthant probabilities):

DEFINITION 27. *Random variables $X_1, \cdots, X_n$ are said to be positively upper orthant dependent (PUOD) if $P[\cap_{i=1}^n \{X_i > a_i\}] \geq \prod_{i=1}^n P[X_i > a_i]$ holds for all $\boldsymbol{a} = (a_1, \cdots, a_n)$. There are said to be positively lower orthant dependent (PLOD) if $P[\cap_{i=1}^n \{X_i \leq a_i\}] \geq \prod_{i=1}^n P[X_i \leq a_i]$ holds for all $\boldsymbol{a}$.*

3.4. Notion NC: Nonnegatively Correlated Random Variables. The weakest condition on positive dependence may be given in terms of the correlation coefficients. Thus we observe the trivial definition:

DEFINITION 28. *$X_1, \ldots, X_n$ are said to be nonnegatively correlated random variables if Corr $(X_i, X_j) \geq 0$ for all $i \neq j$.*

3.5. Ordering of the Notions. The next theorem summarizes the orderings of the notions of positive dependence stated above.

THEOREM 29. *For the notions defined in Definitions 23, 25, 27, and 28, we have*

$$\mathrm{MTP}_2 \Rightarrow A \begin{matrix} \Rightarrow \mathrm{PUOD} \Rightarrow \\ \Rightarrow \mathrm{PLOD} \Rightarrow \end{matrix} \mathrm{NC}. \tag{19}$$

Note that the statement "$\mathrm{MTP}_2 \Rightarrow A$" is just Theorem 26 and the statements "PUOD $\Rightarrow$ NC," "PLOD $\Rightarrow$ NC" were proved by Lehmann (1966). The other two implications follow immediately by taking g_1, g_2 to be indicator functions of nondecreasing sets in $\mathbb{R}^n$.

3.6. Notion E: Exchangeability. Another form of positive dependence exists among exchangeable random variables. To see how exchangeability involves positive dependence, we first observe a definition of exchangeability for an infinite sequence of random variables:

DEFINITION 30. *Let $\{Y_i\}_{i=1}^{\infty}$ be an infinite sequence of univariate random variables. It is said to be a sequence of exchangeable random variables if, for every finite n and every permutation $\{\pi_1, \ldots, \pi_n\}$ of $\{1, \ldots, n\}$, $(Y_1, \cdots, Y_n)$ and $(Y_{\pi_1}, \cdots, Y_{\pi_n})$ are identically distributed.*

This definition leads to the next definition for an n-dimensional random vector $\boldsymbol{X} = (X_1, \cdots, X_n)$:

DEFINITION 31. *$X_1, \cdots, X_n$ are said to be exchangeable random variables if there exists a sequence of exchangeable random variables $\{X_i^*\}_{i=1}^{\infty}$ such that $(X_1, \cdots, X_n)$ and $(X_1^*, \cdots, X_n^*)$ are identically distributed.*

The key result for exchangeable random variables is de Finetti's theorem (see, e.g., Loéve (1963, p. 365)), which says that exchangeability is equivalent to a mixture of conditionally i.i.d. random variables. This mixing process then creates a positive dependence property among the random variables $X_1, \cdots, X_n$, and a number of results based on this dependence property have been developed. Further, the notion of exchangeability plays a key role in many applications. For example, the impacts of exchangeable random variables in Bayes theory and reliability theory are well known.

3.7. Examples of Applications. We now list some examples of applications of notions treated in Section 3.

APPLICATION 32. The important applications of the notion of association in reliability theory can be found in Barlow and Proschan (1975, Section 2.2).

APPLICATION 33. Exchangeability has an important application in reliability theory when the life length variables of the components of a system are exchangeable but not independent. This situation may arise, for example, when the components are subject to a common source of stress. Detailed results can be found in Barlow and Proschan (1975, Chapter 5) and Shaked (1977), and references for some recent applications can be found in Tong (1990, pp. 219–228).

APPLICATION 34. Multivariate probability and moment inequalities for exchangeable random variables via majorization ordering of the dimension vectors were given by vSidák (1973) and Tong (1970, 1977, 1989). Some applications of these results can be found in Tong (1990, pp. 197–198).

APPLICATION 35. If $\boldsymbol{X} = (X_1, \ldots, X_n)$ has a multivariate normal distribution with mean vector $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Sigma}$, and correlation matrix $\boldsymbol{R}$, then the conditions of MTP_2 and association can be made to depend on $\boldsymbol{\Sigma}$ or $\boldsymbol{R}$. Many authors have made important contributions to this area, and a convenient reference for the existing results is Tong (1990, Sections 5.1, 7.2, and 7.3).

REFERENCES

A. Books and Monographs

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

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

EATON, M. L. (1987). *Lectures on Topics in Probability Inequalities.* Centre for Mathematics and Computer Science, Amsterdam, Netherlands.

HARDY, G. H., Littlewood, J. E. and Pólya, G. (1934, 1952). *Inequalities*, 1st and 2nd eds., Cambridge University Press, Cambridge.

KARLIN, S. (1968). *Total Positivity*, Vol. 1. Stanford University Press, Stanford, CA.

LOÉVE, M. (1963). *Probability Theory,* 3rd ed. D. Van Nostrand Co., Princeton, NJ.

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

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

TONG, Y. L. (1990). *The Multivariate Normal Distribution.* Springer-Verlag, New York and Berlin.

B. Papers

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

BOLAND, P. J., Proschan, F. and Tong, Y. L. (1988). Moment and geometric probability inequalities arising from arrangement increasing functions. *Ann. Probab.* **16** 407–413.

BOLAND, P. J., Proschan, F., and Tong, Y. L. (1991). Some majorization inequalities for functions of exchangeable random variables. In *Topics in Statistical Dependence*, H. W. Block, A. R. Sampson, and T. H. Savits, eds. Institute of Mathematical Statistics, Hayward, CA, 85–91.

BORELL, C. (1975). Convex set functions in d-space. *Period. Math. Hungar.* **6** 111–136.

vCEBYVSEV, P. L. (1882). O priblivzennyh vyravzenijah odnih integralov vcerez drugie. *Soobvsćenija i protokoly zasedanivi Matemativ ceskogo obvsćestva pri Imperatorskom Har'kovskom Universitete* No. 2, 93–98; *Polnoe sobranie sovcinenii P. L. vCebyseva. Moskva-Leingrad* 1948a, 128–131.

vCEBYVSEV, P. L. (1883). Ob odnom rjade, dostavljajuvsćem predel'nye velivciny integralov pri razlovzenii podintegral'novi funkcii na mnovzeteli. *Prilovzeni k 57 tomu Zapisok Imp. Akad. Nauk, No. 4; Polnoe sobranie sovcinenii P. L. vCebyvseva.* Moskva-Leingrad, 1948b, 157–169.

CHAN, W., Park, D. H., and Proschan, F. (1989). Peakedness of weighted averages of jointly distributed random variables. In *Contributions to Probability and Statistics: Essays in Honor of Ingram Olkin.* L. J. Glaser, M. D. Perlman, S. J. Press, and A. R. Sampson, eds., Springer-Verlag, New York and Berlin, 58–62.

DAS Gupta, S. (1980). Brunn-Minkowski inequality and its aftermath. *J. Multivariate Anal.* **10** 296–318.

EATON, M. L. and Perlman, M. D. (1977). Reflection groups, generalized Schur functions, and geometry of majorization. *Ann. Probab.* **5** 829–860.

ESARY, J. D., Proschan, F. and Walkup, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38** 1466–1474.

HOLLANDER, M., Proschan, F. and Sethuraman, J. (1977). Functions decreasing in transposition and their applications in ranking problems. *Ann. Statist.* **5** 722–733.

KARLIN, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10** 467–498.

KARLIN, S. and Rinott, Y. (1983). Comparison of measures, multivariate majorization, and applications to statistics. In *Studies in Econometrics, Time Series and Multivariate Analysis,* S. Karlin, T. Amemiya and L. A. Goodman, eds., Academic Press, New York, 465–489.

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1153.

MARSHALL, A. W. and Olkin, I. (1974). Majorization in multivariate distributions. *Ann. Statist.* **2** 1189–1200.

MUDHOLKAR, G. S. (1966). The integral of an invariant unimodal function over an invariant convex set — an inequality and applications. *Proc. Amer. Math. Soc.* **17** 1327–1333.

OLKIN, I. and TONG, Y. L. (1988). Peakedness in multivariate distributions. In *Statistical Decision Theory and Related Topics, IV,* Vol. 2, S. S. Gupta and J. O. Berger, eds., Springer-Verlag, New York and Berlin, 373–383.

PRÉKOPA, A. (1971). Logarithmic concave measures with applications. *Acta Sci. Math.* **32** 302–316.

PROSCHAN, F. (1965). Peakedness of distributions of convex combinations. *Ann. Math. Statist.* **36** 1703–1706.

RINOTT, Y. (1976). On convexity of measures. *Ann. Probab.* **4** 1020–1026.

SHAKED, M. (1977). A concept of positive dependence for exchangeable random variables. *Ann. Statist.* **5** 505–515.

SHERMAN, S. (1955). A theorem on convex sets with applications. *Ann. Math. Statist.* **26** 763–766.

vSIDÁK, Z. (1973). A chain of inequalities for some types of multivariate distributions with nine special cases. *Appl. Math.* **18** 110–118.

TONG, Y. L. (1977). An ordering theorem for conditionally independent and identically distributed random variables. *Ann. Statist.* **5** 274–277.

TONG, Y. L. (1982). Rectangular and elliptical probability inequalities for Schur-concave random variables. *Ann. Statist.* **10** 637–642.

TONG, Y. L. (1983, 1989). Probability inequalities for n-dimensional rectangles via multivariate majorization. *Tech. Report No.* 189, (1983), Department of Statistics, Stanford University, Stanford, CA. Revised version in *Contributions to Probability and Statistics: Essays in Honor of Ingram Olkin.* L. J. Gleser, M. D. Perlman, S. J. Press and A. R. Sampson, eds. Springer-Verlag (1989), New York and Berlin, 146–159.

TONG, Y. L. (1988). Some majorization inequalities in multivariate statistical analysis. *SIAM Review* **30** 602–622.

TONG, Y. L. (1989). Inequalities for a class of positively dependent random variables with a common marginal. *Ann. Statist.* **17** 429–435.

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
225 NORTH AVE., N.W.
ATLANTA, GEORGIA 30332-0160
USA