July 1986

SOME MAJORIZATION INEQUALITIES IN MULTIVARIATE ANALYSIS AND THEIR APPLICATIONS*

Y. L. Tong

School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332
U.S.A.

This paper contains a review of certain majorization inequalities in multivariate analysis and a discussion of their applications. The results apply to a large class of distributions (including the multivariate normal distribution), and have implications in estimation, hypothesis testing, and other related problems.

1. Introduction and Summary

Let $\underset{\sim}{X} = (X_1,\ldots,X_n)$ be an n-dimensional random variable with density $f(\underset{\sim}{x})$ that is absolutely continuous w.r.t. the Lebesgue measure, and let $A \subset R^n$ be a measurable subset. In many problems in multivariate analysis the probability content of the form

$$P[\underset{\sim}{X} \in A] = \int_A f(\underset{\sim}{x})d\underset{\sim}{x} \tag{1.1}$$

is of great importance. For example,

*Research partially supported by National Science Foundation Grant DMS-8502346.
AMS 1980 Subject Classifications: 62H99, 60D15.
Key Words and Phrases: Majorization inequalities in multivariate analysis.

(i) When $A = \times_{i=1}^{n} (-\infty, a_i]$, the probability content in (1.1) is the cumulative distribution function of $\underset{\sim}{X}$.

(ii) When $A = \times_{i=1}^{n} [-a_i, a_i]$, it is the probability content of a n-dimensional rectangle, and is the cumulative distribution function of $|\underset{\sim}{X}| = (|X_1|, \ldots, |X_n|)$.

(iii) When $A = \{\underset{\sim}{x} = (x_1, \ldots, x_n) \,|\, \sum_{i=1}^{n} (x_i/a_i)^2 \le \lambda\}$, it is the probability content of an ellipsoid (or sphere).

Under the normal theory, the confidence region in estimation and the acceptance region in hypothesis testing for the mean vector are usually of the form given in (i)-(iii) so the probability in (1.1) is directly related to the confidence probability and the power of a test.

In many such applications, the numerical evaluation of the probability content is complicated (even for the normal distribution). Thus probability inequalities become useful. In particular, if a lower (or upper) bound on the true probability can be obtained and if the numerical value of the bound can be easily computed or is immediately available from an existing table, then a conservative (or liberal) solution for the estimation or hypothesis-testing problem can be obtained without knowing its true value. In addition, inequalities often provide certain monotonicty properties of the confidence probability function and the power function of a test, this in turn provides a better understanding of the nature of the underlying statistical procedure.

In this paper we provide a survey of probability inequalities in multivariate analysis via the notion of majorization, and discuss their applications. In order not to overload the paper, the proofs of the theorems are not given here. Instead the references from which the original proofs can be found are listed.

The notion of majorization involves the diversity (or dispersion) of the components of a vector, and plays an important role in the theory of inequalities. A convenient reference on this subject is, of course, the monograph by Marshall and Olkin (1979), and a brief treatment of the fundamental concepts in connection with Schur functions is given in Section 2.

Section 3 reviews certain existing majorization inequalities for exchangeable random variables which can be applied for the reduction of dimensionality in multivariate analysis. For instance, a special form of Theorem 3.3 says that if $X_1,\ldots,X_n$ are exchangeable, then the inequalities

$$P[X_i \le a,\ i = 1,\ldots,n] \ge \{P[X_i \le a,\ i = 1,\ldots,r]\}^{n/r},$$

$$P[|X_i| \le a,\ i = 1,\ldots,n] \ge \{P[|X_i| \le a,\ i = 1,\ldots,r]\}^{n/r}$$

hold for all a. Thus, the dimensionality involved is reduced from n to r. If in a given application the table value for the joint probability of r such variables is already available, then the numerical value of a lower bound on the joint probability of n variables can be obtained.

Section 4 deals with probability contents of certain classes of geometric regions when the density of $\underset{\sim}{X}$ is a Schur-concave function. The inequalities generally state that, when the underlying geometric region becomes less asymmetric in a certain fashion via majorization, then the probability content becomes larger. For example, Theorem 4.2 implies the following special result:

$$P[|X_i| \le a_i,\ i = 1,\dots,n] \le P[|X_i| \le \bar{a},\ i = 1,\dots,n], \quad (1.2)$$

where $\bar{a} = \frac{1}{n} \sum_{i=1}^{n} a_i$, for all Schur-concave random variables $\underset{\sim}{X} = (X_1,\dots,X_n)$. Note that the l.h.s. of (1.2) is the probability content of an n-dimensional rectangle and the r.h.s. is that of a cube while the perimeter is kept fixed. The probability content on the r.h.s. of (1.2) is easier to tabulate because it involves only the parameters n and $\bar{a}$ instead of n and $a_1, a_2, \dots, a_n$, and certain table values are already available for the cubes. Thus, once again, the numerical value of the (upper) bound on the true probability content of such an asymmetric geometric region can be rapidly obtained.

Section 5 contains other majorization inequalities which seem useful in multivariate analysis. The first result concerns peakedness in multivariate distributions, and yields a monotonicity property for the convergence of the sample mean vector to the population mean vector. The second result depends on the concept of arrangement increasing functions which is closely related to majorization.

Finally, in Section 6 we discuss briefly some of the applications of those inequalities in multivariate analysis, with special reference to the multivariate normal distribution.

2. Majorization and Schur Functions

For fixed $k > 1$, let

$$\underset{\sim}{a} = (a_1,\ldots,a_k), \quad \underset{\sim}{b} = (b_1,\ldots,b_k) \tag{2.1}$$

denote two real vectors, and let

$$a_{(1)} \geq a_{(2)} \geq \ldots \geq a_{(k)}, \quad b_{(1)} \geq b_{(2)} \geq \ldots \geq b_{(k)} \tag{2.2}$$

be their ordered components. $\underset{\sim}{a}$ is said to majorize $\underset{\sim}{b}$, in symbols, $\underset{\sim}{a} \succ \underset{\sim}{b}$, if

$$\sum_{i=1}^{m} a_{(i)} \geq \sum_{i=1}^{m} b_{(i)} \tag{2.3}$$

holds for $m = 1,2,\ldots,k-1$

and $\sum_{i=1}^{k} a_i = \sum_{i=1}^{k} b_i$. This notion provides a partial ordering of the diversity of the components of a vector, namely, $\underset{\sim}{a} \succ \underset{\sim}{b}$ implies that (for a fixed sum) the a_i's are more diverse than the b_i's. In particular, it is known that: (i) $\underset{\sim}{a} \succ (\bar{a},\ldots,\bar{a})$ for all $\underset{\sim}{a}$, where $\bar{a} = \frac{1}{k}\sum_{i=1}^{k} a_i$; (ii) $\underset{\sim}{a} \succ \underset{\sim}{b}$ implies that the variance of the a_i's is $\geq$ that of the b_i's, and straight inequality holds unless $\underset{\sim}{b}$ is a permutation of $\underset{\sim}{a}$.

There exist many results concerning (necessary and sufficient) conditions for the partial ordering of two vectors $\underset{\sim}{a},\underset{\sim}{b}$ via majorization. One such result is that

$\underset{\sim}{a} \succ \underset{\sim}{b}$ if and only if $\underset{\sim}{b} = \underset{\sim}{a}Q$ for some doubly stochastic matrix Q. Thus one may regard such a linear transformation as an averaging process. That is, by multiplying a vector $\underset{\sim}{a}$ and a doubly stochastic matrix Q the components of the new vector ($\underset{\sim}{b}$) are less diverse and in the meantime the sum (or arithmetic mean) of the components is unchanged.

A function $\phi: R^k \to R$ is said to be a Schur-concave (Schur-convex) function if $\underset{\sim}{a} \succ \underset{\sim}{b}$ implies $\phi(\underset{\sim}{a}) \le (\ge) \phi(\underset{\sim}{b})$. Intuitively speaking, the functional value of ϕ becomes larger (smaller) when the components of a vector in R^k are less diverse given their sum. Consequently, once a function is shown to be a Schur function, then a chain of inequalities may be obtained from the partial ordering of vectors via majorization. For a comprehensive reference on majorization and Schur functions, see Marshall and Olkin (1979).

3. Majorization Inequalities and Reduction of Dimensionality for Exchangeable Random Variables

Let $\underset{\sim}{X} = (X_1, \ldots, X_n)$ denote an n-dimensional random variable. In a number of applications one is interested in the following probabilities:

$$\beta_1(n) \equiv P[X_i \le a,\ i = 1,\ldots,n], \tag{3.1}$$

$$\beta_2(n) \equiv P[|X_i| \le a,\ i = 1,\ldots,n]. \tag{3.2}$$

If $\underset{\sim}{X}$ has a multivariate normal distribution with means 0, variances σ^2, and correlations $\rho \ge 0$, then it is well-known that $\beta_1(n), \beta_2(n)$ can be expressed as

$$\beta_1(n) = \int_{-\infty}^{\infty} \Phi^n((\sqrt{\rho}z + \delta)/\sqrt{1-\rho})\,d\Phi(z), \tag{3.3}$$

$$\beta_2(n) = \int_{-\infty}^{\infty}[\Phi((\sqrt{\rho}z+\delta)/\sqrt{1-\rho}) - \Phi((\sqrt{\rho}z-\delta)/\sqrt{1-\rho})]^n d\Phi(z), \tag{3.4}$$

where $\delta = a/\sigma$ and Φ is the N(0,1) c.d.f. Using those expressions the probability functions $\beta_1(n), \beta_2(n)$ can be calculated numerically and tables are now available. However, no matter how extensive the tables are, there is always a possibility that, in a given application, the true dimensionality n involved is outside the range of the table values. An earlier theorem of Tong (1970) says that, in that case, one can always obtain the numerical value of the lower bound on the basis of existing table values.

Theorem 3.1. Let $\beta_1(n), \beta_2(n)$ denote the multivariate normal probabilities defined in (3.3) and (3.4) where μ and σ^2 are arbitrary but fixed. If $\rho \geq 0$, then for all integers r,n satisfying $1 \leq r < n$

$$\beta_j(n) \geq [\beta_j(r)]^{n/r}, \quad j = 1,2.$$

Furthermore, the inequalities are straight unless $\rho = 0$.

The proof of this result depends on a simple argument: Defining the random variable $W = \Phi^r((\sqrt{\rho}z+\delta)/\sqrt{1-\rho})$, which is nonnegative, the r.h.s. of (3.3) is then the expectation of $W^{n/r}$. Since $w^{n/r}$ is a convex function of w for $w \geq 0$ and $n \geq r$, the proof for (3.3) follows from Jensen's inequality. The proof for (3.4) is similar.

Sidak (1973) adopted this argument to obtain a generalization of Theorem 3.1. His result is for any exchangeable random variables instead of just the positively-correlated normal variables. An infinite sequence $X_1, X_2, \ldots$ of random variables is said to be exchangeable if, for every finite n and every subset of positive integers $\{i_1, \ldots, i_n\}$, $(X_1, \ldots, X_n)$ and $(X_{i_1}, \ldots, X_{i_n})$ are identically distributed. A finite subset $(X_1, \ldots, X_n)$ of such an infinite sequence is called exchangeable. A well-known theorem of De Finetti says that $X_1, \ldots, X_n$ are exchangeable random variables if and only their joint distribution is a mixture of the form

$$F(x_1, \ldots, x_n) = \int \prod_{i=1}^{n} G_z(x_i) dH(z) \tag{3.5}$$

where $H(z)$ is a c.d.f. and $G_z(x)$ a conditional c.d.f. for every given z. Thus, using the same argument for proving Theorem 3.1, Sidak (1973) obtained

Theorem 3.2. Let $X_1, X_2, \ldots, X_n$ be exchangeable random variables, let $B \subset R$ be any given measurable subset, and define

$$\gamma(n) = P[X_i \in B,\ i = 1, \ldots, n]. \tag{3.6}$$

Then, for all $1 \leq r < n$,

$$\gamma(n) \geq \gamma(r)\gamma(n-r), \quad \gamma(n) \geq [\gamma(r)]^{n/r}. \tag{3.7}$$

Note that if $(X_1, \ldots, X_n)$ has a multivariate normal distribution with equal means, equal variances and equal correlations ρ, then the components are exchangeable iff $\rho \geq 0$.

Now if one defines $\gamma(0) \equiv 1$, then by the notion of majorization one can clearly see that, from (3.7),

(i) $(n,0) \succ (r,n-r)$ and $\gamma(n)\gamma(0) \geq \gamma(r)\gamma(n-r)$,

(ii) $(n,\ldots,n,0,\ldots,0) \succ (r,r,\ldots,r)$ and

$$(\gamma(n))^r(\gamma(0))^{n-r} \geq (\gamma(r))^n.$$

(The vectors in (ii) are $1 \times n$ and the first r components in the first vector take the value n.) A natural question one might ask is whether $(a_1,\ldots,a_k) \succ (b_1,\ldots,b_k)$ implies $\prod_{j=1}^{k} \gamma(a_j) \geq \prod_{j=1}^{k} \gamma(b_j)$. This question was answered in Tong (1977), and the proof of this theorem depends on an application of a moment inequality due to Muirhead (see e.g., Tong (1980), p. 119)).

Theorem 3.3. Let $X_1,\ldots,X_n$ be exchangeable random variables and let $B \subset R$ be any given measurable subset. Then, for $\gamma(n)$ defined in (3.6), $(a_1,\ldots,a_k) \succ (b_1,\ldots,b_k)$ implies $\prod_{j=1}^{k} \gamma(a_j) \geq \prod_{j=1}^{k} \gamma(b_j)$.

Theorem 3.3 is a generalization of Theorem 3.2. To consider an application of Theorem 3.3 for which Theorem 3.2 does not apply, simply consider the inequality

$$P[X_i \in B,\ i = 1,\ldots,5]P[X_6 \in B]$$
$$\geq P[X_i \in B,\ i = 1,\ldots,4]P[X_i \in B,\ i = 5,6].$$

This inequality follows because $(5,1) \succ (4,2)$.

A routine generalization of Theorem 3.3 to random vectors can be made following a simple notation change. That is, if $\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n$ are each m-dimensional random vectors and their joint density is a mixture of the form (3.5), where $G_z(\underset{\sim}{x})$ is the distribution of an m-dimensional random variable, and if $B \subset R^m$ is a measurable subset, then the statement in Theorem 3.3

remains true. Such an inequality can be applied for the reduction of dimensionality of m-dimensional random variables.

4. Majorization Inequalities for Asymmetric Geometric Regions

This section concerns majorization inequalities for the probability contents of a certain class of geometric regions. As an example, consider the probability contents of rectangles when the underlying distribution of $\underset{\sim}{X} = (X_1, X_2)$ is bivariate normal with equal means and equal variances. Let

$$A(a_1, a_2) = \{(x_1, x_2) \,\big|\, |x_1| \le a_1,\ |x_2| \le a_2\}$$

denote a rectangle with perimeter $4(a_1 + a_2)$. Now if $(a_1, a_2) \succ (b_1, b_2)$, then $a_1 + a_2 = b_1 + b_2$ and $|a_1 - a_2| \ge |b_1 - b_2|$ hold. Thus, $A(a_1, a_2)$ and $A(b_1, b_2)$ have common perimeter and $A(a_1, a_2)$ is more asymmetric (or $A(b_1, b_2)$ is closer to being the square $A(\bar{a}, \bar{a})$, $\bar{a} = \frac{1}{2}(a_1 + a_2)$). Consequently, since the joint density function of $\underset{\sim}{X}$ is permutation symmetric and unimodal, one might expect that the probability content of $A(b_1, b_2)$ is larger than that of $A(a_1, a_2)$. From Theorem 4.2 stated below we see that this indeed is true. In addition to rectangles, majorization can be used to provide a partial ordering of the asymmetry of other geometric regions. For example, a region

$$A(a_1, a_2) = \{(x_1, x_2) \,\big|\, (x_1/a_1)^2 + (x_2/a_2)^2 \le \lambda\}$$

defines an ellipse. If $(a_1^2, a_2^2) \succ (b_1^2, b_2^2)$ then, for fixed $c = a_1^2 + a_2^2 = b_1^2 + b_2^2$, $A(b_1, b_2)$ is closer to being a circle. Thus, one may expect that the probability content of $A(b_1,b_2)$ is larger for a permutation symmetric bivariate normal distribution. This again is true, as we shall see in Theorem 4.3 stated below.

The geometric regions considered in this section include (one-sided and two-sided) n-dimensional rectangles, ellipsoids, and a class of convex sets. The condition imposed on the density function of $\underset{\sim}{X}$ is Schur concavity. As stated in Section 2, a density function $f(\underset{\sim}{x})$ is a Schur concave function of $\underset{\sim}{x} \in R^n$ if $\underset{\sim}{a} \succ \underset{\sim}{b}$ implies $f(\underset{\sim}{a}) \leq f(\underset{\sim}{b})$. It is known that all Schur functions are permutation symmetric. Furthermore, it is known that the densities of most random variables which are permutation symmetric are Schur concave. In particular, the following statements are true:

(i) If $f(\underset{\sim}{x})$ is permutation symmetric and if $\log f(\underset{\sim}{x})$ is a concave function of $\underset{\sim}{x}$, then it is a Schur-concave function of $\underset{\sim}{x}$.

(ii) If $f(\underset{\sim}{x})$ is permutation symmetric and unimodal (i.e., the set $\{\underset{\sim}{x} | f(\underset{\sim}{x}) \geq c\}$ is a convex set for every $c > 0$), then $f(x)$ is a Schur concave function of $\underset{\sim}{x}$.

In particular, an n-dimensional multivariate normal density function with equal means, equal variances and equal correlations is a Schur-concave function.

In one of the earlier papers on majorization inequalities in multivariate analysis, Marshall and Olkin (1974) considered the probability contents of one-sided n-dimensional rectangles

$$A_0(\underset{\sim}{a}) \equiv \{\underset{\sim}{x} \mid x_i \le a_i,\ i = 1,\ldots,n\}. \tag{4.1}$$

where $\underset{\sim}{a} = (a_1,\ldots,a_n)$. Applying a fundamental convolution theorem in their paper they obtained

Theorem 4.1. If $f(\underset{\sim}{x})$ is a Schur-concave function of $\underset{\sim}{x}$, then

$$\begin{aligned}\phi_0(\underset{\sim}{a}) &= P[\underset{\sim}{X} \in A_0(\underset{\sim}{a})] \\ &= P[X_i \le a_i,\ i = 1,\ldots,n]\end{aligned} \tag{4.2}$$

is a Schur-concave function of $\underset{\sim}{a}$; that is, $\underset{\sim}{a} \succ \underset{\sim}{b}$ implies $\phi_0(\underset{\sim}{a}) \le \phi_0(\underset{\sim}{b})$.

This theorem yields a chain of inequalities for the distribution function of random variables with Schur concave density functions and an extreme case is that

$$\begin{gathered} P[X_i \le a_i,\ i = 1,\ldots,n] \le P[X_i \le \bar{a},\ i = 1,\ldots,n] \\ \bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i. \end{gathered} \tag{4.3}$$

Motivated by their result, Tong (1982) considered probability contents of two-sided rectangeles of the form

$$A_\infty(\underset{\sim}{a}) = \{\underset{\sim}{x} \,\big|\, |x_i| \le a_i,\ i = 1,\ldots,n\}, \tag{4.4}$$

and proved

Theorem 4.2. If $f(\underset{\sim}{x})$ is a Schur-concave density function of $\underset{\sim}{x}$, then

$$\phi_\infty(\underset{\sim}{a}) = P[|X_i| \le a_i,\ i = 1,\ldots,n] \tag{4.5}$$

is a Schur-concave function of $\underset{\sim}{a}$.

A special case of this theorem is that, among all n-dimensional rectangles with a fixed perimeter, the probability content is maximized when the region is a cube.

Tong (1982) also considered probability contents of ellipsoids of the form

$$A_2(\underset{\sim}{a}) = \{\underset{\sim}{x} \mid \sum_{i=1}^{n} (x_i/a_i)^2 \le \lambda\}, \quad \lambda > 0 \text{ fixed.} \tag{4.6}$$

Applying a similar argument he obtained

Theorem 4.3. If $f(\underset{\sim}{x})$ is a Schur-concave function of $\underset{\sim}{x}$, then

$$\phi_2(\underset{\sim}{a}) = P[\underset{\sim}{X} \in A_2(\underset{\sim}{a})] = P[\sum_{i=1}^{n} (X_i/a_i)^2 \le \lambda] \tag{4.7}$$

is a Schur-concave function of $(a_1^2, a_2^2, \ldots, a_n^2)$.

After proving Theorems 4.2 and 4.3, Tong (1982) then considered a larger class of geometric regions

$$A_m(\underset{\sim}{a}) = \{\underset{\sim}{x} \mid \sum_{i=1}^{n} (x_i/a_i)^m \le \lambda\}, \quad \lambda > 0 \text{ fixed} \tag{4.8}$$

for $m = 2,4,6,\ldots,\infty$, and conjectured that the probability content of $A_m(\underset{\sim}{a})$ is a Schur concave function of $(a_1^{m/(m-1)}, \ldots, a_n^{m/(m-1)})$. This conjecture was shown to be true by Karlin and Rinott (1983).

Theorem 4.4. If $f(\underset{\sim}{x})$ is a Schur concave function of $\underset{\sim}{x}$, then

$$\phi_m(\underset{\sim}{a}) = P[\underset{\sim}{X} \in A_m(\underset{\sim}{a})] = P[\sum_{i=1}^{n} (X_i/a_i)^m \le \lambda] \tag{4.9}$$

is a Schur-concave function of $(a_1^{m/(m-1)}, \ldots, a_n^{m/(m-1)})$ for every positive even integer m and for $m = \infty$.

The majorization inequality in Theorem 4.2 deals with n-dimensional rectangles centered at the origin. A result for arbitrary n-dimensional rectangles was obtained independently by Karlin and Rinott (1983) and Tong (1983), and the result is given via multivariate majorization.

Theorem 4.5. Assume that $f(\underset{\sim}{x})$ is a log-concave density function of $\underset{\sim}{X}$ (i.e., $\log f(\underset{\sim}{x})$ is a concave function); let

$$\underset{\sim}{a}_1 = (a_{11},\ldots,a_{1n}), \quad \underset{\sim}{a}_2 = (a_{21},\ldots,a_{2n}),$$
$$\underset{\sim}{b}_1 = (b_{11},\ldots,b_{1n}), \quad \underset{\sim}{b}_2 = (b_{21},\ldots,b_{2n})$$

denote n-dimensional real vectors and define

$$A(\underset{\sim}{a}_1,\underset{\sim}{a}_2) = \{\underset{\sim}{x} \mid a_{11} \le x_1 \le a_{21},\ldots,a_{1n} \le x_n \le a_{2n}\}.$$

If there exists a doubly stochastic matrix Q such that $\underset{\sim}{b}_1 = \underset{\sim}{a}_1 Q$ and $\underset{\sim}{b}_2 = \underset{\sim}{a}_2 Q$, then

$$P[\underset{\sim}{X} \in A(\underset{\sim}{a}_1,\underset{\sim}{a}_2)] \le P[\underset{\sim}{X} \in A(\underset{\sim}{b}_1,\underset{\sim}{b}_2)].$$

To illustrate an application of this theorem, consider the simple example given below:

Example. If (X_1,X_2) has a bivariate normal distribution with equal means and equal variances then, by taking Q to be the 2×2 matrix with elements 1/2, one has

$$P[1 \le X_1 \le 8,\ 5 \le X_2 \le 10] \le P[3 \le X_1 \le 9,\ 3 \le X_2 \le 9].$$

For the rectangular and elliptical regions defined in (4.4) and (4.6), their volumes (vol) are multiples of $\prod_{i=1}^{n} a_i$. Thus if $(a_1,\ldots,a_n) \succ (b_1,\ldots,b_n)$ and

$(a_1^2,\ldots,a_n^2) \succ (b_1^2,\ldots,b_n^2)$ hold, then

$$\mathrm{vol}(A_\infty(\underset{\sim}{a})) \le \mathrm{vol}(A_\infty(\underset{\sim}{b})), \quad \mathrm{vol}(A_2(\underset{\sim}{a})) \le \mathrm{vol}(\underset{\sim}{a}_2(\underset{\sim}{b})),$$

with strict inequality if $\underset{\sim}{a}$ is not a permutation of $\underset{\sim}{b}$. Consequently, in the inequalities stated in Theorems 4.2 and 4.3 the difference in probability contents could be partially due to the difference in the volumes of the subsets. In view of this fact Perlman (1982) suggested that a corresponding result would be of interest if the volumes of the subsets are kept fixed. This can be accomplished by inequalities via the majorization

$$(\log a_1,\ldots,\log a_n) \succ (\log b_1,\ldots,\log b_n).$$

Such a majorization inequality depends on the diversity of the elements of $\underset{\sim}{a}$ when the geometric mean (instead of the arithmetic mean) is kept fixed.

Shaked and Tong (1985) studied this problem for a class of geometric regions. They first showed in a counter-example that such a corresponding result is impossible under the sole assumption of Schur concavity of $f(\underset{\sim}{x})$. Then, using certain basic properties of the arrangement increasing functions obtained by Hollander, Proschan and Sethuraman (1977), they obtained the following theorem for the bivariate case:

<u>Theorem 4.5</u>. If (X_1,X_2) has a density $f(x_1,x_2)$ that is Schur concave and monotone unimodal, and if $f(x_1,-x_2)$ is Schur concave, then $P[(X_1/a_1,\ X_2/a_2) \in A]$ is a Schur-concave function of $(\log a_1,\ \log a_2)$ for all measurable subsets $A \subset R^2$ which are convex, permutation

symmetric, and symmetric about the origin.

After finding a proof for this theorem, they conjectured that an n-dimensional version of this statement is true. But a proof for the general case is not yet available. Note that Theorem 4.5 implies the result of Kunte and Kattihalli (1984) as a special case, and that the main theorem in Das Gupta and Rattihalli (1984) deals with a special case of this conjecture.

5. Other Related Inequalities

A few other majorization inequalities have been obtained recently. In this section we briefly review a partial ordering result on peakedness in multivariate distributions (Olkin and Tong (1984)) and certain geometric inequalities via the applications of arrangement increasing functions (Boland, Proschan and Tong (1985)).

Peakedness provides one of the principle descriptive indices of a distribution. In the univariate case, a random variable Y is said to be more peaked than X if

$$P[|Y| \le \lambda] \ge P[|X| \le \lambda]$$

holds for all λ. Now consider a sequence of i.i.d. random variables $Z_1, Z_2, \ldots$ with density $g(z)$ and mean μ. For $n = 1,2,\ldots$ let $\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$. The question of interest is whether or not $P[|\bar{Z}_n - \mu| \le \lambda]$ converges to 1 monotonically in n for all λ. This question can be rephrased as whether or not $\bar{Z}_n - \mu$ is more peaked than $\bar{Z}_{n-1} - \mu$. In an earlier paper Proschan (1965) proved the following theorem (for convenience we assume that

$\mu = 0$).

Theorem 5.1. Let $Z_1, Z_2, \ldots$ be a sequence of i.i.d. univariate random variables with density $g(z)$. If (i) $g(z) = g(-z)$, (ii) $\log g(z)$ is a concave function of z, and (iii) $(a_1, \ldots, a_n) \succ (b_1, \ldots, b_n)$, then $\sum_{i=1}^{n} b_i Z_i$ is more peaked than $\sum_{i=1}^{n} a_i Z_i$ for all n. Consequently (by taking $a_1 = \ldots = a_{n-1} = 1/(n-1)$, $a_n = 0$ and $b_1 = \ldots = b_n = 1/n$) $\bar{Z}_n$ is more peaked than $\bar{Z}_{n-1}$.

For N-dimensional random variables $\underset{\sim}{X}$ and $\underset{\sim}{Y}$, $\underset{\sim}{Y}$ is said to be more peaked than $\underset{\sim}{X}$ if

$$P[\underset{\sim}{Y} \in A] \geq P[\underset{\sim}{X} \in A]$$

holds for all measurable, compact, convex, and symmetric (about the origin) subsets $A \subset R^N$. Adopting this definition Olkin and Tong (1984) obtained a multivariate generalization of Theorem 5.1. Its proof depends on an application of Anderson's theorem and is different from Proschan's original proof.

Theorem 5.2. Let $\underset{\sim}{Z}_1, \underset{\sim}{Z}_2, \ldots$ be i.i.d. N-dimensional random variables with density $g(\underset{\sim}{z})$. If (i) $g(\underset{\sim}{z}) = g(-\underset{\sim}{z})$, (ii) $\log g(\underset{\sim}{z})$ is a concave function of $\underset{\sim}{z}$, and (iii) $(a_1, \ldots, a_n) \succ (b_1, \ldots, b_n)$, then $\sum_{i=1}^{n} b_i \underset{\sim}{Z}_i$ is more peaked than $\sum_{i=1}^{n} a_i \underset{\sim}{Z}_i$. Consequently, $\bar{\underset{\sim}{Z}}_n = \frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{Z}_i$ is more peaked than $\bar{\underset{\sim}{Z}}_{n-1}$.

This theorem deals with linear combinations of i.i.d. random vectors. The next theorem (Olkin and Tong (1984)) generalizes Theorem 5.1 from combining i.i.d. random variables to combining dependent variables.

Theorem 5.3. For fixed n let $Z_1,\ldots,Z_n$ have a density $g(z_1,\ldots,z_n)$ which is permutation symmetric, and assume that $(a_1,\ldots,a_n) \succ (b_1,\ldots,b_n)$. If the conditional density of

$$c(Z_1 + Z_2) + \sum_{i=3}^{n} a_i Z_i \,|\, Z_1 - Z_2 = v$$

is unimodal and symmetric about the origin for all fixed c, v, and $a_3,\ldots,a_n$, then $\sum_{i=1}^{n} b_i Z_i$ is more peaked than $\sum_{i=1}^{n} a_i Z_i$.

Certain closure properties of peakedness are also given in their paper. For example, it is shown there that the partial ordering of peakedness is preserved in the marginal distributions and the limiting distributions.

Some geometric and moment inequalities given by Boland, Proschan and Tong (1985) depend on a convolution result of arrangement increasing (AI) functions. (Such functions are closely related to majorization and Schur functions, and were treated extensively by Hollander, Proschan and Sethuraman (1977) and Marshall and Olkin ((1979), Chapter 6.) A main result in their paper states that

Theorem 5.4. Let $\underset{\sim}{X} = (X_1,\ldots,X_n)$ have density $f(\underset{\sim}{x})$ that is permutation symmetric. Let $h^{(1)}, h^{(2)}$ be AI functions on $R^n \times R^n$ and ϕ_1, ϕ_2: $R \to R$ be nondecreasing. Then, provided that the expectation exists,

$$\psi(\underset{\sim}{a},\underset{\sim}{b}) \equiv E[\phi_1(h^{(1)}(\underset{\sim}{a},\underset{\sim}{X}))\phi_2(h^{(2)}(\underset{\sim}{b},\underset{\sim}{X}))]$$

is an AI function in $(\underset{\sim}{a},\underset{\sim}{b}) \in R^n \times R^n$.

A special application of this theorem yields the following fact:

Fact. Let the density $f(\underset{\sim}{x})$ of $\underset{\sim}{X} = (X_1,\ldots,X_n)$ be permutation symmetric. Let

$$(a_1,a_2,\ldots,a_n), \quad (b_1,b_2,\ldots,b_n)$$

be given real vectors with ordered components

$$a_{(1)} \geq a_{(2)} \geq \ldots \geq a_{(n)}, \quad b_{(1)} \geq b_{(2)} \geq \ldots \geq b_{(n)}.$$

Then

$$\begin{aligned} &P[a_{(i)} \leq X_i \leq b_{(n-i+1)},\ i = 1,\ldots,n] \\ &\leq P[a_i \leq X_i \leq b_i,\ i = 1,\ldots,n] \\ &\leq P[a_{(i)} \leq X_i \leq b_{(i)},\ i = 1,\ldots,n]. \end{aligned}$$

This inequality again deals with probability contents of n-dimensional rectangles, and was obtained previously by Boland (1985).

6. Some Applications

In this section we outline a few applications of these inequalities in multivariate analysis. They are presented for the purpose of illustration; so obviously the list is not complete.

(a) In many applied problems the probability function involves exchangeable random variables. A simple example is the problem of selection and ranking. Another example is the confidence probability of a permutation symmetric confidence region. In this situation an application of the theorems in Section 3 results in a reduction of the dimensionality, and in certain cases the numerical value of a lower bound can be obtained from existing tables.

(b) In the estimation of the mean vector of a multivariate normal population, if the confidence region is rectangular or elliptical, then an upper bound on the confidence probability can be obtained from applying Theorems 4.1-4.3, and the numerical values of the upper bounds can be found from existing tables.

(c) Inequalities in Theorems 4.1-4.3 can be applied to obtain optimal solutions for allocation of sample sizes in estimation problems. One such application appeared in Tong (1982).

(d) When applying to the multivariate normal distribution, Theorem 4.3 yields the following inequality for convex combinations of dependent chi-squared variables: Let $(Z_1, Z_2, \ldots, Z_n)$ have a multivariate normal distribution with means μ, variances σ^2, and correlates $\rho \in (-\frac{1}{n-1}, 1)$. Then for $c_i > 0$ $(i = 1, \ldots, n)$ and arbitrary but fixed $\lambda > 0$ the probability $P[\sum_{i=1}^{n} c_i Z_i^2 \le \lambda]$ is a Schur-concave function of $(c_1^{-1}, \ldots, c_n^{-1})$. When taking $\rho = 0$, this yields a bound for convex combinations of independent chi-squared variables. Note that this result is similar to, but different from, the Okamoto-Marshall-Olkin inequality (Marshall and Olkin (1979), p. 303).

(e) An application of Theorem 4.5 to the normal distribution yields another inequality for convex combinations of chi-squared variables: If (Z_1, Z_2) has a bivariate normal distribution with means 0, variances $\sigma^2 > 0$, and any correlation ρ, then

$$P[c_1 Z_1^2 + c_2 Z_2^2 \le \lambda] \le P[\sqrt{c_1 c_2}(Z_1^2 + Z_2^2) \le \lambda].$$

When $\rho = 0$, then the same result also follows from the Okamoto-Marshall-Olkin equality.

(f) In most problems in multivariate analysis under the normal theory, the confidence region for the mean vector μ involves a subset that is compact, convex, and symmetric about the origin. Furthermore, in most hypothesis testing problems about μ the acceptance region involves a subset that is also compact, convex, and symmetric. Applying Theorem 5.2, one concludes that in this case the confidence probability is an increasing function of the sample size, and the type I error of such a test is a decreasing function of the sample size.

References

Boland, P. J. (1985) Integrating Schur concave densities on rectangles. Unpublished manuscript.

Boland, P. J., Proschan, F., and Tong, Y. L. (1985) Moment and geometric probability inequalities arising from arrangement increasing functions. Florida State University Statistics Report No. M708 (Submitted for publication).

Das Gupta, S. and Rattihalli, R. N. (1984) Probability content of a rectangle under normal distribution: Some inequalities. Sankhyā, Ser. A, 46, 454-457.

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. (1983) Comparison of measures, multivariate majorization, and applications to statistics. In Studies in Econometrics, Time Series and Multivariate Analysis, S. Karlin, et al. ed. New York: Academic Press, 465-489.

Kunte, S. and Rattihalli, R. N. (1984) Rectangular regions of maximum probability content. Ann. Statist., 1106-1108.

Marshall, A. W. and Olkin, O. (1974) Majorization in multivariate distributions. Ann. Statist., 2, 1189-1200.

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

Olkin, O. and Tong, Y. L. (1984) Peakedness in Multivariate distributions. Technical Report No. 200, Department of Statistics, Stanford University, Stanford, CA.

Perlman, M. D. (1982) Personal communication.

Proschan, F. (1965) Peakedness of distributions of convex combinations. Ann. Math. Statist., 36, 1703-1706.

Shaked, M. and Tong, Y. L. (1985) Inequalities for probability contents of convex sets via geometric average. Submitted for publication.

Sidak, Z. (1973) A chain of inequalities for some types of multivariate distributions with nine special cases. Apl. Math., 18, 110-118.

Tong, Y. L. (1970) Some probability inequalities of multivariate normal and multivariate t. J. Amer. Statist. Assoc., 65, 1243-1247.

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

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

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

Tong, Y. L. (1983) Probability inequalities for n-dimensional rectangles via multivariate majorization. Technical Report No. 189, Department of Statistics, Stanford University, Stanford, CA.

ATIONAL SCIENCE FOUNDATION
Washington, D.C. 20550

FINAL PROJECT REPORT
NSF FORM 98A

PLEASE READ INSTRUCTIONS ON REVERSE BEFORE COMPLETING

PART I—PROJECT IDENTIFICATION INFORMATION

Institution and Address	2. NSF Program	3. NSF Award Number
eorgia Institute of Technology tlanta, GA 30332	Statistics & Probab.	DMS-8502346
	4. Award Period From 7/85 To 11/88	5. Cumulative Award Amount $73,919

Project Title

nequalities in Multivariate Analysis and Reliability Theory

PART II—SUMMARY OF COMPLETED PROJECT *(FOR PUBLIC USE)*

. This project concerns research on theory and applications of inequalities nd stochastic orderings in multivariate statistical analysis and reliability heory. New results have been obtained via the applications of (a) theory of ajorization, (b) dependence in multivariate distributions, (c) convexity, earrangement inequalities and (d) concentration functions. The results in ultivariate statistical analysis include peakedness of multivariate distri- utions, inequalities for probability contents of n-dimensional rectangles nd of a large class of symmetric convex regions, dependence-related nequalities via mixture of distributions, crossing properties of distribu- ion functions, and a stochastic ordering of partial sums of positive i.i.d. andom variables with a log-concave density function. The results in reliabi- ity theory involve a partial ordering of distribution functions of the life- engths of certain parallel and series systems, the effects of positive ependence on the reliability function of a majority system, and a method for ptimal arrangement of components via pairwise rearrangements of components. he published results appear in the forms of research papers and a book on he multivariate normal distribution.

PART III—TECHNICAL INFORMATION *(FOR PROGRAM MANAGEMENT USES)*

ITEM *(Check appropriate blocks)*	NONE	ATTACHED	PREVIOUSLY FURNISHED	TO BE FURNISHED SEPARATELY TO PROGRAM: Check (✓)	Approx. Date
a. Abstracts of Theses	x				
b. Publication Citations		x			
c. Data on Scientific Collaborators	x				
d. Information on Inventions	x				
e. Technical Description of Project and Results		x			
f. Other *(specify)*					

. Principal Investigator/Project Director Name *(Typed)*	3. Principal Investigator/Project Director Signature	4. Date
ung L. Tong		2/10/89

Part III

b. Publication Citations (papers authored by Yung L. Tong)

1. Interval estimation of the critical value in a general linear model. *Annals of the Institute of Statistical Mathematics* **A39** (1987), 289-297.

2. Fault diversity in software reliability. *Probability in the Engineering and Informational Sciences* 1 (1987), 175-187 (with P. J. Boland and F. Proschan).

3. Moment and geometric probability inequalities arising from arrangement increasing functions. *The Annals of Probability* 16 (1988), 407-413 (with P. J. Boland and F. Proschan).

4. Inequalities for probability contents of convex sets via geometric average. *Journal of Multivariate Analysis* 24 (1988), 330-340 (with M. Shaked).

5. 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, (1988), 373-383 (with I. Olkin).

6. Some majorization inequalities in multivariate statistical analysis. *SIAM Review*, 30 (1988), 602-622.

7. Convexity of elliptically contoured distributions with applications. To appear in *Sankhya* (with S. Iyengar).

8. Dependence in majority systems. To appear in *Journal of Applied Probability* (with P. J. Boland and F. Proschan).

9. Crossing properties of mixture distributions. To appear in *Probability in the Engineering and Informational Sciences* (with P. J. Boland and F. Proschan).

10. Inequalities for a class of positively dependent random variables with a common marginal. To appear in *The Annals of Statistics.*

11. Optimal-partitioning inequalities in classification and multi-hypotheses testing. To appear in *The Annals of Statistics* (with T. P. Hill).

12. Optimal arrangement of components via pairwise rearrangements. To appear in *Naval Research Logistics* (with P. J. Boland and F. Proschan).

13. Probability inequalities for n-dimensional rectangles via multivariate majorization. To appear in the I. Olkin Volume.

14. Some majorization inequalities for functions of exchangeable random variables. Submitted to *Dependence in Statistics and Probability* edited by H. W. Block, A. R. Sampson and T. H. Savits (with P. J. Boland and F. Proschan).

15. Comparison of experiments for a class of positively dependent random variables. Submitted to *Canadian Journal of Statistics* (with M. Shaked).

16. *The Multivariate Normal Distribution.* To be published by Springer-Verlag in the Springer Series in Statistics (1989).

e. Technical Description of Project and Results

During the period of 1985-88 Professor Yung L. Tong had conducted research in the general area of probability inequalities in multivariate analysis and reliability theory. The resulted research publications are listed in the previous section of this report.

(1) The publications can be classified into four groups according to sources of support:

(i) Papers #1, 2, 3, 6, 10 and 14 were prepared under grant DMS-8502346.

(ii) Papers #4, 5, 7 and 13 were prepared under grants MCS 81-00775, A01 and DMS-8502346.

(iii) Papers #8, 9, 11, 12 and 15 were prepared under grants DMS-8502346 and DMS-8801327.

(iv) #16, a research book, was prepared under grants DMS-8502346 (chapters 1-4) and DMS-8801327 (chapters 5-9).

(2) Papers #3, 4, 5, 6, 7, 9, 10, 11, 13, 14, 15 and the book #16 mainly concern inequalities in multivariate analysis and multivariate distributions. The main results are derived via majorization (#5, 6, 10, 13, 14 and 16), dependence in multivariate distributions (#9, 10, 14, 15 and 16), convexity and rearrangements (#3, 4 and 7) and concentration functions (#11). Paper #5 contains two generalizations of a result of Proschan (*Ann. Math. Statist.*, 36 (1965), 1703-1706) from univariate distributions to multivariate distributions. Paper #6 provides a survey of majorization inequalities in multivariate analysis with special emphasis on results available after the publication of the Marshall-Olkin book. Paper #13 deals with inequalities for probability contents of n-dimensional rectangles using multivariate majorization as a tool. Papers #10 and 14 involve both majorization and dependence, and contain new results for exchangeable random variables and for partial sums of positive i.i.d. random variables with a log-concave density. Papers #9 and #15 deal with (i) positive dependence and crossing properties of distribution functions and (ii) information inequalities via positive dependence. Papers #3, 4 and 7 involve convexity properties of the elliptically contoured distributions (as functions of the correlation coefficients), probability contents of convex and symmetric geometric regions, and rearrangement inequalities in multivariate distributions. Paper #11 gives new results on probabilities of misclassifications in multi-hypothesis testing using concentration functions. The book #16 contains chapters (chapters 4, 5 and 7) on inequalities in multivariate distributions with special emphasis on the multivariate normal distribution.

(3) Papers #8, 9, 12 and 14 contain results on inequalities in reliability theory. Results in #9 and #14 involve partial orderings of distribution functions of lifelengths of certain parallel and series systems. Paper #8 studies how positive dependence affects the reliability of majority systems, and paper #12 introduces a notion of relative importance of components and provides a method for optimal arrangement of components via pairwise rearrangements.

PART IV - SUMMARY DATA ON PROJECT PERSONNEL

NSF Division Mathematical Sciences

The data requested below will be used to develop a statistical profile on the personnel supported through NSF grants. The information on this part is solicited under the authority of the National Science Foundation Act of 1950, as amended. All information provided will be treated as confidential and will be safeguarded in accordance with the provisions of the Privacy Act of 1974. NSF requires that a single copy of this part be submitted with each Final Project Report (NSF Form 98A); however, submission of the requested information is not mandatory and is not a precondition of future awards. If you do not wish to submit this information, please check this box ☐

Please enter the numbers of individuals supported under this NSF grant.
Do not enter information for individuals working less than 40 hours in any calendar year.

*U.S. Citizens/ Permanent Visa	PI's/PD's		Post-doctorals		Graduate Students		Under-graduates		Precollege Teachers		Others	
	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.	Male	Fem.
American Indian or Alaskan Native												
Asian or Pacific Islander	x											
Black, Not of Hispanic Origin												
Hispanic												
White, Not of Hispanic Origin												
otal U.S. Citizens	1											
on U.S. Citizens												
otal U.S. & Non-U.S. ..												
umber of individuals who have a handicap that limits a major life activity.												

*Use the category that best describes person's ethnic/racial status. (If more than one category applies, use the one category that most closely reflects the person's recognition in the community.)

AMERICAN INDIAN OR ALASKAN NATIVE: A person having origins in any of the original peoples of North America, and who maintains cultural identification through tribal affiliation or community recognition.

ASIAN OR PACIFIC ISLANDER: A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippine Islands and Samoa.

BLACK, NOT OF HISPANIC ORIGIN: A person having origins in any of the black racial groups of Africa.

HISPANIC: A person of Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race.

WHITE, NOT OF HISPANIC ORIGIN: A person having origins in any of the original peoples of Europe, North Africa or the Middle East.

THIS PART WILL BE PHYSICALLY SEPARATED FROM THE FINAL PROJECT REPORT AND USED AS A COMPUTER SOURCE DOCUMENT. DO NOT DUPLICATE IT ON THE REVERSE OF ANY OTHER PART OF THE FINAL REPORT.