Chapter 28. Muirhead's Theorem-Related Moment and Probability Inequalities

by

Frank Proschan¹ and Yung L. Tong²
Florida State University and Georgia Institute of Technology

FSU Technical Report Number M-804
AFOSR Technical Report Number 89-237

April, 1989

Department of Statistics
Florida State University
Tallahassee, Florida 32306-3033

¹ Support provided under Air Force Office of Scientific Research Grant AFOSR 88-0040.

² Support provided by the National Science Foundation under Grant NSF-DMS-8801327.

Key Words: Muirhead's Theorem, moment inequalities, probability inequalities, Schur functions, positive dependence, exchangeable random variables, majorization, semigroup property, multivariate normal.

A.M.S. Classification Number 62H99
Abstract

This report represents a chapter in a forthcoming book by J. Pecaric, F. Proschan, and Y. L. Tong on convexity and some partial orderings useful in statistics.

The report first treats a moment inequality that can be obtained by applying Muirhead's theorem. The moment inequality is applied in Section 2 proving dimension-related inequalities for exchangeable random variables. In Section 3 we use the same moment inequality for a partial ordering of positive dependence of random variables with a common marginal distribution. Such a partial ordering applies to exchangeable random variables, distributions with the semigroup property, and the multivariate normal distribution among others.
28. Muirhead's Theorem-Related Moment and Probability Inequalities

1. MOMENT INEQUALITIES

Applying Muirhead's Theorem (Theorem 4.2, Chap. 11) to permutation-symmetric non-negative random variables, one obtains the following moment inequality (Proshan and Sethuraman, 1977, Tong, 1977):

1.1. Theorem. Let \((Z_1, \ldots, Z_N)\) have a permutation symmetric density function such that 
\[ P\left[ \bigcap_{j=1}^N \{Z_j \geq 0\} \right] = 1, \] and let \(a = (a_1, \ldots, a_N)\) and \(b = (b_1, \ldots, b_N)\) be two real vectors. If \(a \succ b\), then
\[
E \prod_{j=1}^N Z_j^{a_j} \geq E \prod_{j=1}^N Z_j^{b_j}. \tag{1.1}
\]

Proof. By Muirhead's theorem, for every \(\omega\) in the sample space the inequality
\[
\sum_{s} \prod_{j=1}^N Z_{s_j}^{a_j}(\omega) \geq \sum_{s} \prod_{j=1}^N Z_{s_j}^{b_j}(\omega) \tag{1.2}
\]
holds, where the summations are taken over all \(s = (s_1, \ldots, s_N)\), the permutations of \((1, \ldots, N)\). The conclusion follows by taking expectations on both sides of (1.2) and the symmetry property of the distribution function.

A special case of Theorem 1.1 is

1.2. Corollary. For \(\beta \in [0, \infty)\) let \(\mu_\beta\) denote the \(\beta^{th}\) moment of the random variable \(Z\). If \(Z \geq 0\) a.s. and \(a \succ b\), then
\[
\prod_{j=1}^N \mu_{a_j} \geq \prod_{j=1}^N \mu_{b_j}. \tag{1.3}
\]

First Proof. Immediate from Theorem 1.1 by letting \(Z, Z_1, Z_2, \ldots, Z_N\) be i.i.d. random variables.

For the special case in which \(a_j \geq 0\) and \(b_j \geq 0\) for all \(j\), a different proof is possible:

Second Proof. Without loss of generality it may be assumed that \(a_1 \leq b_1 \leq a_2 \leq b_2\), \(a_1 + a_2 = b_1 + b_2\) and \(a_i = b_i\) \((i = 3, \ldots, n)\). The conclusion then follows immediately from Theorem 2.3 of Chap. 27.
Corollary 1.2 asserts that the function $\psi(\mathbf{a}) = \prod_{j=1}^{N} \mu_{a_j}$ is a Schur-convex function of $\mathbf{a}$. This property will be used in the rest of the chapter to derive probability inequalities for a class of positively dependent random variables.

2. ADDITIONAL INEQUALITIES FOR EXCHANGEABLE RANDOM VARIABLES

Applying Corollary 1.2 to exchangeable random variables (defined in Definition 3.2 of Chap. 27), one obtains the following result (Tong, 1977):

2.1. Theorem. Let $X_1, \ldots, X_n$ be exchangeable random variables and let $\mathbf{a} = (a_1, \ldots, a_N)$ and $\mathbf{b} = (b_1, \ldots, b_N)$ be vectors of nonnegative integers such that $\sum_{i=1}^{N} a_i = \sum_{i=1}^{N} b_i = n$. Let $B \subset \mathbb{R}$ be an arbitrary but fixed Borel-measurable set, and denote

$$\gamma(k) = \mathbb{P}\left[ \bigcap_{i=1}^{k} \{X_i \in B\} \right], \quad k = 1, \ldots, n,$$

as defined in (3.3) of Chap. 27, where $\gamma(0) \equiv 1$. If $\mathbf{a} \succ \mathbf{b}$, then

$$\prod_{i=1}^{N} \gamma(a_i) \geq \prod_{i=1}^{N} \gamma(b_i). \quad (2.1)$$

Proof. Following the line of argument given in the proof of Theorem 3.7 of Chap. 27, we have from Corollary 1.2,

$$\prod_{j=1}^{N} \gamma(a_j) = \prod_{j=1}^{N} \mathbb{E} \left[ a_j \prod_{i=1}^{N} I_B(g(U_i,v)) | V = v \right]$$

$$= \prod_{j=1}^{N} \mathbb{E} \tau^{a_j}(V)$$

$$\geq \prod_{j=1}^{N} \mathbb{E} \tau^{b_j}(V) = \prod_{j=1}^{N} \mathbb{E} \left[ b_j \prod_{i=1}^{N} I_B(g(U_i,v)) | V = v \right]$$

$$= \prod_{j=1}^{N} \gamma(b_j)$$

where $\tau(v) = \mathbb{E}\{I_B(g(U_1,v)) | V = v\}$ is the conditional expectation. □

Note that Theorem 2.1 implies Theorem 3.7 of Chap. 27, but the converse is false. To see this, consider the inequality
\[ \gamma(n-1) \gamma(1) \geq \gamma(n-2) \gamma(2), \quad n \geq 3 \]

which follows from \((n-1,1) \succ (n-2,2)\) and Theorem 2.1. But Theorem 3.7 of Chap. 27 fails to apply.

A special result from Theorem 2.1 and \((n,0,...,0) \succ (1,1,...,1)\) is that

\[ P\left[ \bigcap_{i=1}^{n} \{ X_i \in B \} \right] \geq \prod_{i=1}^{n} P[X_i \in B]. \tag{2.2} \]

Since the r.h.s. of (2.2) corresponds to the case of independent \(X_1,...,X_n\), this result can be restated as:

**2.2. Fact.** Let \(X_1,...,X_n\) be exchangeable random variables and let \(Y_1,...,Y_n\) be i.i.d. random variables such that \(X_i,Y_i\) have a common marginal distribution. Then

\[ P\left[ \bigcap_{i=1}^{n} \{ X_i \in B \} \right] \geq P\left[ \bigcap_{i=1}^{n} \{ Y_i \in B \} \right]. \tag{2.3} \]

holds for all Borel-measurable sets \(B \subset \mathbb{R}\).

A question of interest is whether the inequality in (2.3) also holds when the \(Y_i\)'s are less positively dependent than the \(X_i\)'s in a certain fashion. This leads to the problem of the partial ordering of types of positive dependence of exchangeable random variables, a natural extension from the comparison of positively dependent random variables and i.i.d. random variables to the comparison of two sets of exchangeable random variables.

Rinott and Pollak (1980) and Shaked and Tong (1985), among others, studied this problem recently, and obtained results under reasonable assumptions. A special result they obtained is for exchangeable normal variables (Rinott and Pollak, 1980, for \(n=2\), Shaked and Tong, 1985, for general \(n\):

**2.3. Fact.** Let \(X_1,...,X_n\) be exchangeable normal variables with means \(\mu\), variances \(\sigma^2\), and correlation coefficients \(\rho_2\). Let \(Y_1,...,Y_n\) be exchangeable normal variables with means \(\mu\), variances \(\sigma^2\), and correlation coefficients \(\rho_1\). If \(0 \leq \rho_1 < \rho_2\), then \(E \prod_{i=1}^{n} \phi(X_i) \geq E \prod_{i=1}^{n} \phi(Y_i)\) holds for all Borel-measurable functions \(\phi: \mathbb{R} \rightarrow [0,\infty)\) such that the expectations exist. Consequently (2.3) holds for all Borel-measurable sets \(B\).

Note that for \(n=2\), \(E \prod_{i=1}^{2} \phi(X_i) \geq E \prod_{i=1}^{2} \phi(Y_i)\) holds iff
Corr(\(\phi(X_1),\phi(X_2)\)) \geq Corr(\(\phi(Y_1),\phi(Y_2)\)) \tag{2.4}

holds, which is the motivation given by Rinott and Pollak (1980) for studying this problem. Furthermore, if \(\phi_1,\phi_2\) are not necessarily identical but are both monotonically nondecreasing, then a result corresponding to (2.4) can be found in Remark 1.5 of Chap. A.

3. INEQUALITIES FOR A CLASS OF POSITIVELY DEPENDENT RANDOM VARIABLES

We now show how to use the Muirhead-related moment inequality given in Theorem 1.1 to obtain more general results. We consider n-dimensional random vectors \(X=(X_1,...,X_n)\) and \(Y=(Y_1,...,Y_n)\) such that the \(X_i\)'s are not necessarily exchangeable and the \(Y_i\)'s are not necessarily exchangeable, but they all have a common marginal distribution. We show that if the \(X_i\)'s are more positively dependent than the \(Y_i\)'s in a certain fashion, then the inequality in (2.3) holds.

To obtain sufficient conditions for such a partial ordering we first consider a sequence of i.i.d. random variables \(\{U_i\}_{i=1}^n\), another independent sequence of i.i.d. random variables \(\{V_i\}_{i=1}^n\), and an independent random variable \(W\) as "building blocks." Then for a given Borel measurable function \(g: \mathbb{R}^3 \to \mathbb{R}\) and a fixed n-dimensional vector of nonnegative integers

\[
k = (k_1,...,k_r,0,...,0), \quad 1 \leq r \leq n, \quad k_j \geq 1 \text{ for } j \leq r \quad \text{and} \quad \sum_{j=1}^r k_j = n, \tag{3.1}
\]

we define an n-dimensional random vector \(\xi = (\xi_1,...,\xi_n)\) given by

\[
\xi_1 = g(U_1,V_1,W), \quad \xi_{k_1} = g(U_{k_1},V_1,W),
\]

\[
\xi_{k_1+1} = g(U_{k_1+1},V_2,W), \quad \xi_{k_1+k_2} = g(U_{k_1+k_2},V_2,W),
\]

and

\[
\xi_{k_1+...+k_{r-1}+1} = g(U_{k_1+k_{r-1}+1},V_r,W), \quad \xi_n = g(U_n,V_r,W). \tag{3.2}
\]

That is, each of the \(\xi_i\)'s depends on the common variable \(W\) and on a different variable \(U_i\). Furthermore, the first \(k_1\) of them depend on the common variable \(V_1\), the next \(k_2\) of them depend on the common variable \(V_2\), and so on. The vector \((\xi_1,...,\xi_n)\) will be denoted by \(\xi(k)\).

It is easy to verify that \(\xi_1,...,\xi_n\) have a common marginal distribution. Furthermore, the vector \(k\) plays an important role in the positive dependence of \(\xi(k)\). Consider the following two special cases.

Case 1: (i) \(W\) is a singular random variable and (ii) \(k=(1,1,...,1)\). Clearly \(\xi_1,...,\xi_n\) are i.i.d.
random variables. Case 2: (i) \( P[U_i = u_i] = 1 \) \((i = 1, \ldots, n)\) and (ii) \( k=(n,0,\ldots,0) \). Then \( P[\xi_1 = \cdots = \xi_n] = 1 \), so that \( \text{Corr}(\xi_i, \xi_{i'}) = 1 \) for \( i \neq i' \). This illustrates the fact that for given random variables \( \{U_i\}, \{V_i\} \) and \( W \), the “strength” of the positive dependence of the components of \( \xi(k) \) can be partially determined via the diversity of the components of \( k \). Next we make this idea more precise using Corollary 1.2 (Tong, 1989).

3.1. Theorem. For fixed \( n \geq 2 \) assume that (i) \( \{U_i\}_{i=1}^n \), \( \{V_i\}_{i=1}^n \), and \( W \) are stochastically independent, \( U_1, \ldots, U_n \) are i.i.d., \( V_1, \ldots, V_n \) are i.i.d., (ii) \( g: \mathbb{R}^3 \rightarrow \mathbb{R} \) is any Borel-measurable function, and (iii) \( k \) and \( k' \) are two real vectors of the form given in (3.1). Let \( X = \xi(k) \) and \( Y = \xi(k') \) be two random vectors as defined in (3.2). If \( k > k' \), then

\[
E \prod_{i=1}^n \phi(X_i) \geq E \prod_{i=1}^n \phi(Y_i)
\]

(3.3)

holds for all Borel-measurable functions \( \phi: \mathbb{R} \rightarrow [0, \infty) \) such that the expectations exist. Consequently (2.3) holds.

Proof. For every given \( \phi \geq 0 \) such that the expectations exist we can write

\[
E \prod_{i=1}^n \phi(X_i) = E \prod_{j=1}^n \mathbb{E} \left[ \prod_{i=1}^n \phi(g(U_i, V_j, W))(V_j, W) \right] | W = \mathbb{E} \prod_{j=1}^n \mathbb{E} \left[ \tau^{k_j}(V_j, W) | W \right],
\]

(3.4)

where

\[
\tau(v, w) = \mathbb{E} \{ \phi(g(U_1, V_j, W)) | (V_j, W) = (v, w) \}
\]

(3.5)

denotes the conditional expectation. Now for every given \( W = w \) the random variables \( \tau(V_1, w), \ldots, \tau(V_n, w) \) are i.i.d. and are \( \geq 0 \) a.s. By defining \( \mu_{k_j} \) to be the \( k_j \)-th moment of \( \tau(V_j, w) \) and applying Corollary 1.2, we see that

\[
\prod_{j=1}^n \mathbb{E} \tau^{k_j}(V_j, w) \geq \prod_{j=1}^n \mathbb{E} \tau^{k_j'}(V_j, w)
\]

(3.6)

holds for every fixed \( w \). Thus

\[
E \prod_{i=1}^n \phi(X_i) = E \prod_{j=1}^n \mathbb{E} \tau^{k_j}(V_j, W) | W \geq E \prod_{j=1}^n \mathbb{E} \tau^{k_j'}(V_j, W) | W = E \prod_{i=1}^n \phi(Y_i).
\]

In the following corollary we show that if the elements in \( k \) and \( k' \) are even integers (including 0), then the condition that \( \phi \geq 0 \) can be dropped.
3.2. Corollary. Let \( \{ U_i \}^n_1, \{ V_i \}^n_1, W, \) and \( g \) satisfy the conditions stated in Theorem 3.1. Let \( \underline{k}, \underline{k}' \) be two real vectors such that their components are nonnegative even integers. If \( \underline{k} \succ \underline{k}' \), then (3.3) holds for all Borel-measurable functions \( \phi: \mathbb{R} \to \mathbb{R} \) such that the expectations exist.

Proof. The proof is as for Theorem 3.1 since \( \frac{1}{2} \underline{k} \succ \frac{1}{2} \underline{k}' \).

In certain applications the special case \( \underline{k}' = (1, \ldots, 1) \) is of great interest. In the following corollary we show that if the vector \( \underline{k} \) contains only even integers, then again the condition that \( \phi \geq 0 \) can be removed.

3.3. Corollary. Let \( \{ U_i \}^n_1, \{ V_i \}^n_1, W, \) and \( g \) satisfy the conditions stated in Theorem 3.1. Let \( \underline{k}, \underline{k}' \) be two \( n \)-dimensional real vectors such that \( \underline{k}' = (1, \ldots, 1) \) and the components of \( \underline{k} \) are nonnegative even integers such that \( \sum_{i=1}^n k_i = n \). Then (3.3) holds for all Borel-measurable functions \( \phi: \mathbb{R} \to \mathbb{R} \).

Proof. For every fixed \( W = w \) the function \( \tau \) defined in (3.5) satisfies (again by Corollary 1.2)

\[
\prod_{j=1}^n \mathbb{E}^{\underline{k}'_j}(V_j, w) \geq [\mathbb{E}^{\tau^2}(V_1, w)]^{n/2} \geq [\mathbb{E}^{\tau}(V_1, w)]^n.
\]

The conclusion then follows after unconditioning.

As a special consequence, we observe

3.4. Corollary. Let \( \{ U_i \}^n_1, \{ V_i \}^n_1, W, \) and \( g \) satisfy the conditions stated in Theorem 3.1 and let \( \xi(\underline{k}) \) be the random vector defined in (3.1). Let \( \underline{k} = (n, 0, \ldots, 0), \underline{k}' = (1, \ldots, 1), \) and \( n \) be an even positive integer. Then (3.3) holds for all Borel-measurable functions \( \phi: \mathbb{R} \to \mathbb{R} \).

In certain applications to be discussed in Section 4 we restrict our attention to a family of random variables such that \( \xi(\underline{k}) \) is obtained by choosing \( \underline{k} = (s, 1, \ldots, 1, 0, \ldots, 0) \) in (3.1). For notational convenience we shall denote this random vector \( \xi(s) \). The following corollary shows how the components of \( \xi(s) \) depend on \( s \). Its proof follows immediately from Theorem 3.1 and is omitted.

3.5. Corollary. Let \( \{ U_i \}^n_1, \{ V_i \}^n_1, W, \) and \( g \) satisfy the conditions stated in Theorem 3.1. For given \( s \geq 1 \), let \( \xi(s) = (\xi_1, \ldots, \xi_n) \) be the random vector obtained according to (3.1) by choosing
\[ k_1 = s, \quad k_2 = \cdots = k_{n-s+1} = 1, \quad k_{n-s+2} = \cdots = k_n = 0. \] (3.6)

Let \( X = \xi(s+1) \) and \( Y = \xi(s) \). Then (3.3) holds for all such \( \phi \geq 0 \), for all \( n \) and all \( s < n \).

4. APPLICATIONS TO SPECIAL FAMILIES OF RANDOM VARIABLES AND DISTRIBUTIONS

In this section we apply the main results in Section 3 to obtain inequalities via this partial ordering of positive dependence for several families of random variables and distributions.

4(A). Exchangeable Random Variables

Consider the random variables defined for \( i = 1, \ldots, n < \infty \) by

\[ X_i = g(U_i, V_i, W), \quad Y_i = g(U_i, V_i, W). \]

Then \( X_1, \ldots, X_n \) are exchangeable and \( Y_1, \ldots, Y_n \) are exchangeable. But for \( k = (n, 0, \ldots, 0) \) and \( k' = (1, \ldots, 1) \) we have \( X \overset{d}{=} \xi(k) \) and \( Y \overset{d}{=} \xi(k') \). Thus a partial ordering of positive dependence can be obtained by applying Theorem 3.1 and the related results given in Section 3.

By applying this result to the exchangeable normal, t, chi-square, gamma, F, and exponential variables, we obtain many useful inequalities as special cases. The multivariate normal variables will be treated separately in this section. Exchangeable exponential variables can be obtained by taking \( f(u, v, w) = \min(u, v, w) \) as considered previously by Marshall and Olkin (1967), and have an important application in reliability theory.

4(B). Distributions with the Semigroup Property

Let \( \{f_{\phi}(x); \phi \in \Omega\} \) denote a family of density functions, and assume that \( \Omega \) is an interval of real numbers or an interval of integers. It is said to possess the semigroup property (see, e.g., Proschan and Sethuraman, 1977) if \( \phi', \phi'' \in \Omega \) implies \( \phi' + \phi'' \in \Omega \) and the convolution \( f_{\phi'}(x) * f_{\phi''}(x) = f_{\phi' + \phi''}(x) \).

4.1. Application. Let \( X_{\theta_1}, \ldots, X_{\theta_n} \) denote i.i.d. random variables with density \( f_{\theta}(x) \), and for fixed \( \theta_0 \) and \( \theta_0 - \theta \in \Omega \) let \( X_{\theta_0 - \theta}(x) \) denote another independent random variable with density \( f_{\theta_0 - \theta}(x) \). Next define an \( n \)-dimensional random vector \( \chi(\theta) = (X_1, \ldots, X_n) \) such that \( X_i = X_{\theta_1}, i = 1, \ldots, n \). If \( \{f_{\theta}(x); \theta \in \Omega\} \) possesses the semigroup property and if \( \theta_1, \theta_2 \in \Omega \), \( \theta_1 \neq \theta_2 \) implies \( |\theta_1 - \theta_2| \in \Omega \), then (a)
$E_{\theta_1,\theta_2} \prod_{i=1}^n \phi(X_i)$ is a nonincreasing function of $\theta$ for $\theta < \theta_0$ for all Borel-measurable functions $\phi \geq 0$ (provided that the expectation exists); (b) $E_{\theta_1,\theta_2} \prod_{i=1}^n \phi(X_i)$ is a nonincreasing function of $\theta$ for $\theta < \theta_0$ for all even positive integers $n$ and all Borel-measurable functions $\phi$; (c) $P_{\theta}[X_1 \in B, \ldots, X_n \in B]$ is a nonincreasing function of $\theta$ for $\theta < \theta_0$ for all Borel-measurable sets $B \subset \mathbb{R}$.

Proof. For every fixed $\theta_1, \theta_2 \in \Omega$ such that $\theta_1 < \theta_2 < \theta_0$ define $U_i = X_{\theta_1,i}$, $V_i = X_{\theta_2 - \theta_1,i}$ ($i = 1, \ldots, n$) and $W = X_{\theta_0 - \theta_2}$. The conclusions follow from Theorem 3.1.

Note that Application 4.1 applies to the binomial, gamma, Poisson distributions, the Poisson process, and several other distributions.

4(C). The Multivariate Normal Distribution

Applying Theorem 3.1 we now show how the positive dependence of a multivariate normal variable with a common marginal distribution can be partially ordered via the correlation coefficients.

4.2. Application. Let $0 \leq \rho_1 < \rho_2 \leq 1$, $\underline{k}$ and $\underline{k}'$ be two vectors of nonnegative integers as given in (3.1), and $R = R(\underline{k}) = (\rho_{ij})$ be a correlation matrix such that for $i \neq j$,

$$
\rho_{ij} = \begin{cases} 
\rho_2 & \text{if } 1 \leq i, j \leq k_1, k_1 + 1 \leq i, j \leq k_1 + k_2, \ldots, \sum_{m=1}^{r-1} k_m + 1 \leq i, j \leq n; \\
\rho_1 & \text{otherwise.}
\end{cases}
$$

[That is, the random variables $X_1, \ldots, X_n$ are partitioned into $r$ groups of sizes $k_1, \ldots, k_r$, respectively; the correlation coefficients of the variables within the same group are $\rho_2$, and the correlation coefficients between groups are $\rho_1$.] Let $X \sim \mathcal{N}(\mu, \sigma^2 R(\underline{k}))$ and $Y \sim \mathcal{N}(\mu, \sigma^2 R(\underline{k}'))$, where $\mu = (\mu, \ldots, \mu)$. (a) If $\underline{k} > \underline{k}'$, then (3.3) holds for all such $\phi \geq 0$ and thus the probability inequality (2.3) holds. (b) If $\underline{k} > \underline{k}'$ and the components of $\underline{k}, \underline{k}'$ are even integers, then (3.3) holds for all such Borel-measurable functions $\phi: \mathbb{R} \to \mathbb{R}$.

Proof. The result follows immediately by choosing

$$
g(u,v,w) = \mu + \sigma \left( \sqrt{1 - \rho_2} u + \sqrt{\rho_2 - \rho_1} v + \sqrt{\rho_1} w \right)
$$

in Theorem 3.1 and Corollary 3.2.
4.3. Example. Let \( X = (X_1, X_2, X_3, X_4) \) have a multivariate normal distribution with equal means, equal variances, and a correlation matrix \( R \). Let

\[
R_3 = \begin{pmatrix}
1 & \rho_2 & \rho_2 & \rho_1 \\
\rho_2 & 1 & \rho_2 & \rho_1 \\
\rho_2 & \rho_2 & 1 & \rho_1 \\
\rho_1 & \rho_1 & \rho_1 & 1
\end{pmatrix}
\]

\[
R_2 = \begin{pmatrix}
1 & \rho_2 & \rho_1 & \rho_1 \\
\rho_2 & 1 & \rho_1 & \rho_1 \\
\rho_1 & \rho_1 & 1 & \rho_2 \\
\rho_1 & \rho_1 & \rho_2 & 1
\end{pmatrix}
\]

Then

\[
E_{\rho_{ij} = \rho_2} \prod_{i=1}^{4} \phi(X_i) \geq E_{R=R_3} \prod_{i=1}^{4} \phi(X_i) \geq E_{R=R_2} \prod_{i=1}^{4} \phi(X_i) \geq E_{\rho_{ij} = \rho_1} \prod_{i=1}^{4} \phi(X_i)
\]

holds for all \( \phi \geq 0 \) and all \( 0 \leq \rho_1 < \rho_2 \leq 1 \).

**Proof.** The conclusion follows from \((4,0,0,0) \succ (3,1,0,0) \succ (2,2,0,0) \succ (1,1,1,1)\). \(\square\)

A special consequence of Application 4.2 is the result given in Fact 2.3 for exchangeable normal variables.

**REFERENCES**


Chapter 29. Log-Concavity Property of Probability Measures

by

Frank Proschan\textsuperscript{1} and Yung L. Tong\textsuperscript{2}
Florida State University and Georgia Institute of Technology

FSU Technical Report Number M-805
AFOSR Technical Report Number 89-238

April, 1989

Department of Statistics
Florida State University
Tallahassee, Florida 32306-3033

\textsuperscript{1} Support provided by the Air Force Office of Scientific Research Grant AFOSR 88-0040.

\textsuperscript{2} Support provided by the National Science Foundation under Grant NSF-DMS-8801327.

Key Words: Log-concavity, Brunn-Minkowski inequality, convolution, unimodality, Schur-concavity, increasing failure rate, series system, doubly-stochastic matrix.

A.M.S. Classification Number 62H99
Abstract

This report represents a chapter in a forthcoming book by J. Pecaric, F. Proschan, and Y. L. Tong on convexity and some partial orderings useful in statistics.

The report concerns inequalities via the log-concavity property of certain probability measures. In particular, we treat the Brunn-Minkowski Inequality and the theorems of Prékopa and Borell and their generalizations. Some statistical applications in multivariate distributions and reliability theory are then discussed.
29. Log-Concavity Property of Probability Measures

In this chapter we study inequalities via the log-concavity property of certain probability measures. In particular, we treat the Brunn-Minkowski Inequality and the theorems of Prékopa (1971), Borell (1975) and their generalizations (Riott, 1976, Das Gupta, 1980, and Karlin and Rinott, 1983). We then describe some applications of these results in statistics.

Let $\mathcal{B}$ denote the class of all Borel-measurable sets in $\mathbb{R}^n$, and let $P$ be a measure defined on $\mathcal{B}$. For arbitrary but fixed $\alpha \in [0,1]$ and $B_1, B_2 \in \mathcal{B}$ let us define

$$\alpha B_1 + (1-\alpha) B_2 = \{ \bar{x} : \bar{x} \in \mathbb{R}^n, \bar{x} = \alpha x_1 + (1-\alpha)x_2 \text{ for some } x_1 \in B_1 \text{ and } x_2 \in B_2 \}.$$ 

Throughout this chapter we assume that $B_1, B_2$ are in $\mathcal{B}$, are nonempty, and $(\alpha B_1 + (1-\alpha) B_2 \in \mathcal{B}^+$. Let $X = (X_1, ..., X_n)$ have a probability density function $f(x)$ that is absolutely continuous w.r.t. Lebesgue measure. The inequalities we study here concern lower bounds on $P(\alpha B_1 + (1-\alpha) B_2)$ in terms of $P(B_1)$ and $P(B_2)$ where, if $P$ is a probability measure, $P(B)$ stands for $P(X \in B)$.

1. BRUNN-MINKOWSKI INEQUALITY

The classical Brunn-Minkowski Inequality states:

1.1. Theorem. If $P$ is the Lebesgue measure, then

$$P(\alpha B_1 + (1-\alpha) B_2) \geq [\alpha(P(B_1))^{1/n} + (1-\alpha)(P(B_2))^{1/n}]^n$$

holds for all $B_1, B_2 \in \mathbb{R}^n$, all $\alpha \in [0,1]$, and all $n=1,2,\ldots$.

When $B_1, B_2$ are convex sets (which implies that $\alpha B_1 + (1-\alpha) B_2$ is also convex), this inequality was first proved by Brunn (1887). Minkowski (1919) derived conditions for equality to hold when $B_1, B_2$ are convex. Later Lusternik (1935) generalized this result to any Borel-measurable sets, and his conditions for equality to hold were corrected by Henstock and Macbeath (1953). A treatment of the historical developments related to this inequality can be found in Das Gupta (1980). Das Gupta (1980) also provided a generalization which includes a previous generalization of Henstock and Macbeath (1953) as a special case.

*Karlin (1983) provided a counterexample showing that $B_1, B_2 \in \mathcal{B}$ does not necessarily imply $(\alpha B_1 + (1-\alpha) B_2 \in \mathcal{B}$ for all $\alpha \in [0,1]$. Thus this assumption is needed.*
2. A CLASS OF LOG-CONCAVE PROBABILITY MEASURES

If $P$ is a probability measure such that

$$P(B) = \int_B f(x) dx \quad \text{for all } B \in \mathcal{B},$$

then $(\mathbb{R}^n, \mathcal{B}, P)$ is a probability space. Prékopa (1971) considered probability measures with the following property:

$$P(\alpha B_1 + (1-\alpha)B_2) \geq (P(B_1))^\alpha (P(B_2))^{1-\alpha}, \quad \alpha \in [0,1].$$

If all the probabilities are positive, then (2.2) is equivalent to

$$\log P(\alpha B_1 + (1-\alpha)B_2) \geq \alpha \log P(B_1) + (1-\alpha) \log P(B_2),$$

which is just the log-concavity of $P$. He studied conditions on $f(x)$ for (2.2) to hold, and proved that a sufficient condition is the log-concavity of $f$.

2.1. Definition. Let $f(x): \mathbb{R}^n \rightarrow [0,\infty)$ be a probability density function that is absolutely continuous w.r.t Lebesgue measure. $f$ is said to be log-concave if

$$f(\alpha x_1 + (1-\alpha)x_2) \geq (f(x_1))^\alpha (f(x_2))^{1-\alpha}$$

holds for all $\alpha \in [0,1]$ and all $x_1, x_2 \in \mathbb{R}^n$. When $f(x) > 0$ for all $x \in \mathbb{R}^n$, then (2.3) is equivalent to

$$\log f(\alpha x_1 + (1-\alpha)x_2) \geq \alpha \log f(x_1) + (1-\alpha) \log f(x_2).$$

The main theorem of Prékopa (1971) states:

2.2. Theorem. Let $P$ be defined as in (2.1). If $f$ is log-concave, then (2.2) holds for all $B_1, B_2 \subset \mathbb{R}^n$ and all $\alpha \in [0,1]$.

An immediate generalization of Theorem 2.2 is:

2.3. Corollary. Let $X$ be an n-dimensional random vector with probability density function $f(x)$, and
let \( B_1, \ldots, B_k \) (\( k \geq 2 \)) be subsets of \( \mathbb{R}^n \). If \( f(x) \) is a log-concave function of \( x \), then

\[
P\left[ x \in \sum_{i=1}^{k} \alpha_i B_i \right] \geq \prod_{i=1}^{k} \{P[X \in B_i]\}^{\alpha_i}
\]

holds for all \( \alpha_i \) such that \( \alpha_i > 0 \) and \( \sum_{i=1}^{k} \alpha_i = 1 \).

Proof. \[P\left[ x \in \sum_{i=1}^{k} \alpha_i B_i \right] = P\left[ x \in \alpha_1 B_1 + (1-\alpha_1) \sum_{i=2}^{k} (\alpha_i/(1-\alpha_1)) B_i \right]
\geq \{P[X \in B_1]\}^{\alpha_1} \{P[X \in B_2]\}^{1-\alpha_1} \geq \{P[X \in B_1]\}^{\alpha_1} \{P[X \in B_2]\}^{\alpha_2} \ldots \geq \prod_{i=1}^{k} \{P[X \in B_i]\}^{\alpha_i}.
\]

Prékopa's original result is given when \( B_1, B_2 \) are convex sets only; his original proof depends on an application of the Brunn-Minkowski Inequality and is quite lengthy. His result was generalized to a larger class of probability measures for convex sets by Borell (1975). Borell considered measures with the property that

\[
P(B) \geq [\alpha(P(B_1))^s + (1-\alpha)(P(B_2))^s]^{1/s}
\]

holds for all convex sets \( B_1, B_2 \subset \mathbb{R}^n \), all \( \alpha \in [0,1] \) and for some \( s \in [-\infty, \frac{1}{n}] \). When \( s = 0 \), then by continuity (2.6) reduces to (2.2). When \( s = -\infty \), then the right-hand side of (2.6) is just \( \min\{P(B_1), P(B_2)\} \). Borell (1975) first showed that

2.4. Theorem. If the probability measure \( P \) satisfies (2.6) for all \( B_1, B_2 \subset \mathbb{R}^n \), all \( \alpha \in [0,1] \) and for some \( s \in [-\infty, \frac{1}{n}] \), then it is absolutely continuous w.r.t Lebesgue measure.

As a consequence of Theorem 2.4, the class of measures satisfying (2.6) for some \( s \in [-\infty, \frac{1}{n}] \)
must be of the form (2.1) for some \( f \) that is absolutely continuous w.r.t Lebesgue measure. Borell's (1975) main theorem concerns a characterization of the class of probability measures satisfying (2.6). His original result concerns convex sets only. In the following theorem the convexity condition on \( B_1, B_2 \) is removed:

2.5. Theorem. Let \( F: \mathbb{R}^n \to [0, \infty) \) be a probability density function that is absolute continuous w.r.t Lebesgue measure, and let \( P \) be the probability measure defined in (2.1). Then the following statements are equivalent:

(a) \( P \) satisfies (2.6) for all sets \( B_1, B_2 \subset \mathbb{R}^n \), all \( \alpha \in [0, 1] \) and for some \( s \in [-\infty, \frac{1}{n+1}] \).

(b) There exists a Borel-measurable function \( g: \mathbb{R}^n \to \mathbb{R} \) such that \( f(x) = g(x) \) almost everywhere and

(i) if \( s \in (-\infty, 0) \), then \( g(x) \) is convex,

(ii) if \( s = 0 \), then \( \log g(x) \) is concave,

(iii) if \( s \in (0, \frac{1}{n+1}) \), then \( g(x) \) is concave.

Sketch of Proof. The original proof of Borell (1975) is lengthy and is for convex sets only. In the following we adopt the proof given by Rinott (1976). Rinott’s proof involves replacing integrals on \( f \) over sets in \( \mathbb{R}^n \) by a certain measure \( \mu \) of epigraphs in \( \mathbb{R}^{n+1} \). Thus his proof is valid only for \( s \in [-\infty, \frac{1}{n+1}] \). Let \( B^* \) be any set in \( \mathbb{R}^{n+1} \). Consider a measure \( \mu \) given by

\[
d\mu(B^*) = \begin{cases} \frac{(1-ns)/s}{(x_{n+1})^{(s-1-(n+1))}} \prod_{i=1}^{n+1} dx_i, & x_{n+1} > 0, \quad \text{for } s \neq 0, s \in [-\infty, \frac{1}{n+1}], \\ e^{-x_{n+1} \prod_{i=1}^{n+1} dx_i}, & s = 0. \end{cases} \tag{2.7}
\]

Then it can be shown that, for \( B_1^*, B_2^* \subset \mathbb{R}^{n+1} \),

\[
\mu(\alpha B_1^* + (1-\alpha)B_2^*) \geq \begin{cases} \left( \alpha (\mu(B_1^*))^{s} + (1-\alpha) (\mu(B_2^*))^{s} \right)^{1/s} & \text{for } s \neq 0, s \in [-\infty, \frac{1}{n+1}], \\ (\mu(B_1^*))^{\alpha} (\mu(B_2^*))^{1-\alpha} & \text{for } s = 0. \end{cases} \tag{2.8}
\]

The inequality in (2.8) can be obtained by first showing it for rectangular sets, then extending it to any Borel measurable sets by using the argument in Borell (1975).

(b) \( \Rightarrow \) (a): For arbitrary but fixed \( B \subset \mathbb{R}^n \) and \( g: \mathbb{R}^n \to \mathbb{R} \) let us define the epigraph in \( \mathbb{R}^{n+1} \):

\[
A(B, g) = \{ (x, \lambda); x \in B, \lambda \in \mathbb{R}^1, g(x) \leq \lambda \}.
\]

Let us first consider the case \( s = 0 \) and assume that \( f \) satisfies (2.3). Denoting \( B^* = A(B, -\log f) \subset \mathbb{R}^{n+1} \),
we can easily verify that for measure $\mu$ defined in (2.7) we have

$$P(B) = \mu(B^*)$$  \hspace{1cm} (2.9)$$

for all $B \subset \mathbb{R}^n$ and

$$[\alpha B_1 + (1 - \alpha) B_2]^* \supseteq \alpha B_1^* + (1 - \alpha) B_2^*$$  \hspace{1cm} (2.10)$$

for all $B_1, B_2 \subset \mathbb{R}^n$. Thus (2.8), (2.9), and (2.10) together imply

$$P[\alpha B_1 + (1 - \alpha) B_2] = \mu[\alpha B_1 + (1 - \alpha) B_2]^* \supseteq \mu[\alpha B_1^* + (1 - \alpha) B_2^*]$$

$$\geq (\mu(B_1))^\alpha (\mu(B_2))^{1 - \alpha} = (P(B_1))^\alpha (P(B_2))^{1 - \alpha},$$

which completes the proof for $s = 0$. For $s < 0$ we can repeat the above arguments with

$B^* = A(B, f^{\omega/(1 - ns)})$, and for $s \in (0, \frac{1}{n + 1})$ we can define $B^* = H(B, f^{\omega/(1 - ns)})$ where the hypograph $H(B, g)$ is given by

$$H(B, g) = \{(x, \lambda): x \in B, \lambda \in \mathbb{R}^1, \text{ and } 0 \leq \lambda \leq g(x)\}.$$

(a) $\Rightarrow$ (b): First consider the case $s = 0$ and assume that (2.2) holds. Let $S_k(\bar{x}) = \{y: |y - \bar{x}| \leq \frac{1}{k}\}$ and define

$$f_k(\bar{x}) = \left[\int_{S_k(\bar{x})} f(y)dy\right] / \int_{S_k(\bar{x})} dy.$$

Then (2.2) implies

$$f_k(\alpha \bar{x}_1 + (1 - \alpha) \bar{x}_2) \geq (f_k(\bar{x}_1))^{\alpha} (f_k(\bar{x}_2))^{1 - \alpha} \text{ for } \alpha \in [0, 1].$$

Consequently the function $g(x) = \lim_{k \to \infty} \inf f_k(\bar{x})$ satisfies

$$g(\alpha \bar{x}_1 + (1 - \alpha) \bar{x}_2) \geq (g(\bar{x}_1))^{\alpha} (g(\bar{x}_2))^{1 - \alpha} \text{ for } \alpha \in [0, 1].$$

By differentiation of the integral argument we have $f = g$ almost everywhere. Thus the proof is
complete for \( s=0 \). For \( s= -\infty \), a similar argument yields \( g(\alpha x_1 + (1-\alpha) x_2) \geq \min \{ g(x_1), g(x_2) \} \). For \( s \in (-\infty, 0) \), we can consider the sets \( C_1, C_2 \subset \mathbb{R}^{n+1} \) defined by \( C_i = B_i \times (c_i, \infty) \) where the \( B_i \)’s are spheres in \( \mathbb{R}^n \) and \( c_i > 0 \) (i=1,2). If \( B_1, B_2 \) are chosen to satisfy

\[
\left( \int_{B_1} \frac{1}{1} \, dx_i / \int_{B_2} \frac{1}{1} \, dx_i \right) = \left( \frac{c_1}{c_2} \right)^n,
\]

then \( C_1, C_2 \) satisfy (2.8) for some \( s \in (-\infty, 0) \). Suppose \( f^{s/(1-ns)} \) is not convex. We choose \( B_1, B_2 \) in this fashion to satisfy \( B_i \subset A(B_i, f^{s/(1-ns)}) \) and

\[
\alpha B_1 + (1-\alpha) B_2 \supset A(\alpha B_1 + (1-\alpha) B_2, f^{s/(1-ns)}).
\]

Thus we have

\[
P[\alpha B_1 + (1-\alpha) B_2] < \mu(\alpha C_1 + (1-\alpha) C_2) = \left[ \alpha (\mu(C_1))^s + (1-\alpha) (\mu(C_2))^s \right]^{1/s}
\]

\[
< \left[ \alpha (P(B_1))^s + (1-\alpha) (P(B_2))^s \right]^{1/s},
\]

which is a contradiction. For \( s \in (0, \frac{1}{n+1}) \) the proof proceeds similarly by choosing \( C_i = B_i \times (0, c_i) \) (i=1,2).

\( \square \)

**3. SOME PROPERTIES OF LOG-CONCAVE DENSITY FUNCTIONS**

Log-concave density functions which satisfy (2.3) play an important role in statistics and probability. In the following we observe some known facts concerning this class of densities.

**3.1. Fact.** Let \( X_1, \ldots, X_n \) be i.i.d. univariate random variables with a common density function \( h(x) \). If \( h(x) \) is a log-concave function of \( x \) for \( x \in \mathbb{R}^1 \), then the joint density function of \( (X_1, \ldots, X_n) \) is a log-concave function of \( \bar{x} \) for \( \bar{x} \in \mathbb{R}^n \).

**3.2. Fact.** If \( f(\bar{x}) = g(T(\bar{x})) \) where \( g: \mathbb{R}^1 \to [0, \infty) \) is decreasing and \( T(\bar{x}) \) is a convex function of \( \bar{x} \) for \( \bar{x} \in \mathbb{R}^n \), then \( f \) is a log-concave function of \( \bar{x} \) for \( \bar{x} \in \mathbb{R}^n \).

The following theorem, due to Brascamp and Lieb (1975), shows that the integral of a log-concave function is log-concave:
3.3. Theorem. Let \( f(x,y) : \mathbb{R}^{n+m} \to [0,\infty) \) be a log-concave function of \((x,y)\) for \( x \in \mathbb{R}^n \) and \( y \in \mathbb{R}^m \). Then the function \( g : \mathbb{R}^n \to [0,\infty) \) given by

\[
g(x) = \int_{\mathbb{R}^m} f(x,y) \, dy
\]

is log-concave.

Proof. We adopt the proof in Brascamp and Lieb (1975). First note that it suffices to prove the theorem for \( m=n=1 \) because the general case follows by Fubini's theorem and induction. Let \( x_1, x_2 \) be two points in \( \mathbb{R}^1 \) such that \( g(x_1)g(x_2) \neq 0 \). For convenience we may assume that

\[
\sup_y f(x,y) = \sup_y f(x',y);
\]

for otherwise, we can replace \( f(x,y) \) by \( e^{b(x,y)} \) for suitably chosen \( b \) and the problem remains unchanged. For each fixed \( \lambda > 0 \) denote

\[
C_1(\lambda) = \{(x,y) : f(x,y) \geq \lambda \} \subset \mathbb{R}^2,
\]

\[
C_2(x,\lambda) = \{y : f(x,y) \geq \lambda \} \subset \mathbb{R}^1.
\]

Then, by log-concavity of \( f \), \( C_1(\lambda) \) is convex and \( C_2(x,\lambda) \) is an interval. (For the convexity of \( C_1(\lambda) \) see Fact 3.5 below.) Letting \( \nu(x,\lambda) = \int_{C_2(x,\lambda)} dy \) be the Lebesgue measure of the set \( C_2(x,\lambda) \) we have, by Theorem 1.1

\[
\nu(\alpha x_1 + (1-\alpha)x_2,\lambda) \geq \alpha \nu(x_1,\lambda) + (1-\alpha) \nu(x_2,\lambda) \quad \text{for all } \alpha \in [0,1].
\]

Since \( g(x) \) can be expressed as \( g(x) = \int_0^\infty \nu(x,\lambda) d\lambda \), we have

\[
g(\alpha x_1 + (1-\alpha)x_2) \geq \alpha g(x_1) + (1-\alpha)g(x_2) \geq (g(x_1))^\alpha (g(x_2))^{1-\alpha}
\]

for all \( \alpha \in [0,1] \), where the second inequality follows from the arithmetic mean-geometric mean inequality.

A simple application of Theorem 3.3 is (Brascamp and Lieb (1975)):
3.4. Corollary. The convolution of two log-concave density functions in \( \mathbb{R}^n \) is log-concave.

**Proof.** Let \( f_1, f_2 \) be log-concave density functions. Then \( f(x) = f_1(x-y)f_2(y) \) is jointly log-concave in \( (x,y) \in \mathbb{R}^{2n} \). Thus by Theorem 3.3

\[
g(x) = \int_{\mathbb{R}^n} f_1(x-y)f_2(y)dy
\]

is log-concave. \( \Box \)

A density function \( f \) is said to be unimodal if the set

\[
D_\lambda = \{ x : x \in \mathbb{R}^n, f(x) \geq \lambda \}
\]

is a convex set in \( \mathbb{R}^n \) for all \( \lambda > 0 \). The following facts show how log-concavity and unimodality are related.

3.5. Fact. If \( f : \mathbb{R}^n \rightarrow [0, \infty) \) is a probability density function that is absolutely continuous w.r.t. Lebesgue measure, then log-concavity of \( f \) implies unimodality of \( f \).

**Proof.** Let \( x_1, x_2 \in \mathbb{R}^n \) be in \( D_\lambda \). Then for every \( \alpha \in [0,1] \) we have by (2.4),

\[
\log f(\alpha x_1 + (1-\alpha)x_2) \geq \alpha \log f(x_1) + (1-\alpha) \log f(x_2) \geq \alpha \log \lambda + (1-\alpha) \log \lambda \geq \log \lambda.
\]

Thus \( \alpha x_1 + (1-\alpha)x_2 \) is also in \( D_\lambda \). \( \Box \)

A function \( f \) is said to be Schur-concave if \( x \succ y \) implies \( f(x) \leq f(y) \) for all \( x, y \in \mathbb{R}^n \) (see Chap. 11). It is known that all Schur-concave functions are permutation symmetric. Furthermore, the following is known:

3.6. Fact. If \( f : \mathbb{R}^n \rightarrow [0, \infty) \) is a permutation-symmetric and log-concave function of \( x \in \mathbb{R}^n \), then it is a Schur-concave function of \( x \in \mathbb{R}^n \).

**Proof.** Without loss of generality it may be assumed that \( x, y \) are of the form
\[ \chi = (a_1, a_2, a_3, \ldots, a_n), \quad \gamma = (b_1, b_2, a_3, \ldots, a_n) \]

where \( a_2 < b_2 \leq b_1 < a_1 \) and \( a_1 + a_2 = b_1 + b_2 \). Let \( \chi' = (a_2, a_1, a_3, \ldots, a_n) \). Then there exists an \( \alpha \in (0,1) \) such that \( \gamma = \alpha \chi + (1-\alpha)\chi' \). Thus by the permutation-symmetry and log-concavity properties of \( f \) we have

\[
\log f(\gamma) = \log f(\alpha \chi + (1-\alpha)\chi') \geq \alpha \log f(\chi) + (1-\alpha) \log f(\chi') = \log f(\chi).
\]

4. SOME STATISTICAL APPLICATIONS

In this section we describe some of the applications of Prekopa's theorem (Theorem 2.2) and Borell's theorem (Theorem 2.5) in statistics and reliability theory. Application 4.1 is easy. Applications 4.2–4.4 were given in Rinott (1976), and Application 4.7 was given independently by Karlin and Rinott (1983) and Tong (1983, 1989).

4.1. Application. Let \( \chi = (X_1, \ldots, X_n) \) have a probability density function \( f(\chi) \) and distribution function \( F(\chi) \). If \( f(\chi) \) satisfies the condition in (b) of Theorem 2.5 for some \( s \in [-1,1] \), then

\[
F(\alpha x_1 + (1-\alpha)x_2) \geq \begin{cases} 
(\alpha (F(x_1))^s + (1-\alpha)(F(x_2))^s)^{1/s} & \text{for } s \neq 0, \\
(F(x_1))^\alpha (F(x_2))^{1-\alpha} & \text{for } s = 0.
\end{cases}
\]

Proof. Choosing \( B_i = \{ \chi : x_i \in \mathbb{R}^n, \chi \leq \chi_i \} \) for \( i = 1, 2 \) we have

\[
\alpha B_1 + (1-\alpha)B_2 = \{ \chi : x \in \mathbb{R}^n, x \leq \alpha \chi_1 + (1-\alpha)\chi_2 \}.
\]

Thus the inequality follows from Theorem 2.5.

For \( s = 0 \) this result states that the distribution function of an \( n \)-dimensional random vector is log-concave if its density function is log-concave.

4.2. Application. Let \( (X_1, \ldots, X_n) \) have density function \( f(\chi) \) such that \( f \) satisfies the condition in Theorem 2.5. Let \( \eta_1(t_1, \ldots, t_{n-1}), \eta_2(t_1, \ldots, t_{n-1}) \) be the implicit functions given by the equations

\[
P[X_1 \leq t_1, \ldots, X_{n-1}, X_n \leq \eta_1(t_1, \ldots, t_{n-1})] = \alpha_1,
\]

\[
P[X_1 \geq t_1, \ldots, X_{n-1}, X_n \geq \eta_2(t_1, \ldots, t_{n-1})] = \alpha_2,
\]
for arbitrary but fixed $\alpha_j \in (0,1)$ ($j=1,2$). Then for $j=1,2$, $\eta_j(t_1,\ldots,t_{n-1})$ is concave on the convex regions where it is defined in $\mathbb{R}^{n-1}$.

When $n=2$ and $f$ is the bivariate normal density function, this result was given earlier by Tihansky (1972).

4.3. Application. Consider the problem of hypothesis testing concerning a location parameter $\theta \in \mathbb{R}^n$ of a density function $f(x-\theta)$. If (i) the acceptance region is a convex set in $\mathbb{R}^n$, (ii) the test is $\alpha_0$-similar on the boundary of $\mathcal{H}_0$ (under the null hypothesis), and (iii) $f(x)$ satisfies the condition in Theorem 2.5 for some $s \in [-\infty, \frac{1}{3}]$, then the test is unbiased with level $\alpha_0$.

4.4. Application. Let $f(x)$, the density function of $X=(X_1,\ldots,X_n)$, satisfy (2.3). Let $\phi(x): \mathbb{R}^n \rightarrow \mathbb{R}$ be a concave function of $x$ and define $\bar{G}(t)=P[\phi(X) \geq t]$. Then Theorem 2.2 implies

$$\bar{G}(\alpha t_1+(1-\alpha)t_2) \geq (\bar{G}(t_1))^\alpha (\bar{G}(t_2))^{1-\alpha}$$

for all $t_1,t_2 \geq 0$ and all $\alpha \in [0,1]$. Since $\phi(x)=\min\{x_1,\ldots,x_n\}$ is a concave function, this result implies that if the joint density function of the life lengths of $n$ components in a series system is log-concave, then the distribution of the life length of the system has the increasing failure rate property.

The final application (Application 4.7) given below is for the special case in which $f(x)$ is both log-concave and permutation symmetric. We first observe

4.5. Fact. Let the probability density function of $X=(X_1,\ldots,X_n)$ be permutation symmetric and log-concave. Then

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

where

$$\xi(B) = \{\chi: \chi=\varpi(\chi), \chi \in B\}$$

and $\varpi(x)$ is any given permutation of $x$.

By applying this result to $n$-dimensional rectangles, we obtain the following result: Let $B$ be an $n$-dimensional rectangle given by
$B = B(\xi_1, \xi_2) = \{ \bar{x}: \bar{x} \in \mathbb{R}^n, c_{1i} \leq x_i \leq c_{2i} \text{ for } i = 1, \ldots, n \}$,

where

$\xi_1 = (c_{11}, \ldots, c_{1n}), \quad \xi_2 = (c_{21}, \ldots, c_{2n})$,

and $c_{1i} < c_{2i} \ (i = 1, \ldots, n)$. Let $\pi = (\pi_1, \ldots, \pi_n)$ be a permutation of $(1, \ldots, n)$ and let $\delta$ be the inverse permutation of $\pi$. Then it is easy to see that

$\pi(B) = \{ \bar{x}: \bar{x} \in \mathbb{R}^n, c_{1\pi_i} \leq x_{\pi_i} \leq c_{2\pi_i} \text{ for } i = 1, \ldots, n \}$

$= \{ \bar{x}: \bar{x} \in \mathbb{R}^n, c_{1\delta_i} \leq x_{\delta_i} \leq c_{2\delta_i} \text{ for } i = 1, \ldots, n \}$.

Thus

$\alpha B + (1 - \alpha)\pi(B) = \{ \bar{x}: \bar{x} \in \mathbb{R}^n, \alpha c_{1i} + (1 - \alpha)c_{1\pi_i} \leq x_i \leq \alpha c_{2i} + (1 - \alpha)c_{2\pi_i} \text{ for } i = 1, \ldots, n \}$,

which is also an $n$-dimensional rectangle. Now let $\{\pi_1, \ldots, \pi_{n!}\}$ be the group of $n!$ permutations and for $\alpha_i \geq 0, \sum_{i=1}^{n!} \alpha_i = 1$ consider the rectangle given by

$S = \sum_{i=1}^{n!} \alpha_i \pi_i(B)$. \hspace{1cm} (4.1)

Then we have (Karlin and Rinott, 1983, and Tong, 1983, 1989):

4.6. Fact. \textit{S is of the form (4.1) iff there exists an $n \times n$ doubly stochastic matrix $Q$ such that}

$d_1 = (d_{11}, \ldots, d_{1n}) = \xi_1 Q, \quad d_2 = (d_{21}, \ldots, d_{2n}) = \xi_2 Q$

and

$S = \{ \bar{x}: \bar{x} \in \mathbb{R}^n, d_{1i} \leq x_i \leq d_{2i} \text{ for } i = 1, \ldots, n \}$.

For notational convenience we may write $S = Q(B)$. Note that when $-\infty < c_{1i} < c_{2i} < \infty \ (i = 1, \ldots, n)$, the perimeter of $S$ is equal to the perimeter of $B$, but $S$ is closer to being an $n$-dimensional
cube.

Using Theorem 2.2 and Fact 4.5, Karlin and Rinott (1983) and Tong (1983, 1989) independently obtained:

4.7. Application. Let $f(x)$, the density function of $X = (X_1, ..., X_n)$, be permutation symmetric and log-concave. Then for every given $n$-dimensional rectangle $B$ we have

$$P[X \in B] \leq P[X \in Q(B)]$$

for every $n \times n$ doubly stochastic matrix $Q$. In particular,

$$P \left[ \bigcap_{i=1}^n \{ c_{1i} \leq X_i \leq c_{2i} \} \right] \leq P \left[ \bigcap_{i=1}^n \{ \bar{c}_1 \leq X_i \leq \bar{c}_2 \} \right]$$

(4.2)

holds, where $\bar{c}_j = \frac{1}{n} \sum_{i=1}^n c_{ji}$ is the arithmetic mean ($j = 1, 2$).

4.8. Remarks. (a) When $-\infty < c_{1i} < c_{2i} < \infty$ for each $i = 1, ..., n$, the region on the r.h.s. of (4.2) is a cube; and the inequality holds when $Q$ is chosen to be the matrix with all elements being $\frac{1}{n}$.

(b) When $c_{1i} = -\infty$ ($i = 1, ..., n$), Application 4.7 states that the distribution function of an $n$-dimensional random vector with a permutation symmetric log-concave density is Schur-concave. In view of Fact 3.6, this also follows from a result in Marshall and Olkin (1974).

(c) When $c_{1i} = -c_{2i} < 0$, the statement in Application 4.7 is a special case of a result in Tong (1982).

REFERENCES


