CHAPTER - IV

RESULTS IN MULTIVARIATE DISTRIBUTIONS THROUGH MAJORIZATION

This chapter is based on a paper published by A.W.Marshall and Ingram Olkin [12]. The concepts could be viewed as an advancement on Anderson's theorem. It is shown that if the joint density f of $X = (X_1, \ldots, X_n)$ is Schur-concave, then $P(X \in A + \theta)$ is a Schur-concave function of θ, whenever the indicator function of A is Schur-concave.

4.A CONCEPTUAL BACKGROUND

Consider the joint distribution f of a set of random variables. Consider a set A. If one is interested to find conditions on f and A such that

$$\varphi(\theta) = \int_{A+\theta} f(x)\, dx = P\{X \in A + \theta\}$$

is Schur-concave function of θ; Marshall and Olkins theorem provides an answer to this problem. It is shown that to yield the aspaired result f should be Schur-concave and if $y \in A$ with $x < y \implies x \in A$.

4.B MARSHALL AND OLKIN'S THEOREM AND RELATED RESULTS

4.B.1 Theorem : Suppose that the random variables $X_1, \ldots, X_n$ have a joint density f that is Schur-concave. If $A \subset R^n$ is a Lebesgne-measurable set which satisfies

$$y \in A \quad \text{and} \quad x < y \implies x \in A$$

then

$$\int_{A+\theta} f(x)\,dx = P\{X \in A + \theta\} \quad \text{is a Schur-concave}$$

function of θ.

<u>Outline of the Proof</u> :

It is required to show that if $\theta \prec \xi$ then

$$\int_{A+\theta} f(x)\,dx \geq \int_{A+\xi} f(x)\,dx.$$

To achieve this we would show that the function f and the set A do satisfy the conditions of Anderson's theorem in some sense. The result follows immediately.

<u>Proof</u> : We must show that if $\theta \prec \xi$

then $$\int_{A+\theta} f(x)\,dx \geq \int_{A+\xi} f(x)\,dx.$$

The condition that f is Schur-concave implies f is permutation-symmetric or in otherwords the random variables $X_1, \ldots, X_n$ are exchangiable (1)

It is given that if $y \in A$

and $x \prec y \implies x \in A$ (2)

Hence we claim that A is permutation symmetric convex set.

If not let $x \in A$, $\Pi(x) \notin A$.

But $\Pi(x) \prec x$.

Hence by (2) $\Pi(x) \in A$.

Hence a contradiction. Hence A is permutation symmetric.

Also $$\int_{A+\theta} f(x)\,dx = \int_{A} f(x+\theta)\,dx$$

(This could be obtained by making a transformation as was explained in the proof of Anderson's theorem).

$$= \int_{A} f(\Pi x + \Pi \theta)\,dx$$

(Because f and A are permutation symmetric).

$$= \int_{A+\Pi\theta} f(\Pi x)\,dx$$

$$= \int_{A+\Pi\theta} f(x)\,dx.$$

Hence $$\int_{A+\theta} f(x)\,dx = \int_{A+\Pi\theta} f(x)\,dx.$$

Thus the result would be true for any permutation of θ. Hence we assume

$$\theta_1 \geq \theta_2 \geq \ldots\ldots \geq \theta_n$$

and

$$\xi_1 \geq \xi_2 \geq \ldots\ldots \geq \xi_n.$$

According to Hardy, Littlewood and Polya [8] θ can be derived from ξ by a finite number of pairwise averages of components.

Thus between θ and ξ we would find a finite number of vectors each differs from the preceeding one only by two components.

Thus if we can establish the result for any adjecent pairs, the same argument can be applied to successive pairs and the result can be arrived at.

Hence for all practical purpose it is logical to assume that θ and ξ differs only in two components. Call them j and k.

As $\qquad \theta < \xi$

$$\theta_j + \theta_k = \xi_j + \xi_k = 2\,\varepsilon .$$

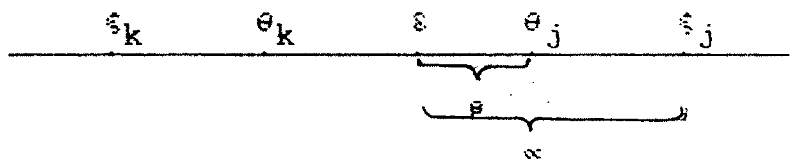


Fig. 1.

If $\qquad \xi_j = \varepsilon + \alpha \qquad$ and $\qquad \theta_j = \varepsilon + \beta$

then $\qquad \xi_k = \varepsilon - \alpha \qquad$ and $\qquad \theta_k = \varepsilon - \beta$

with $\qquad \alpha > \beta > 0$

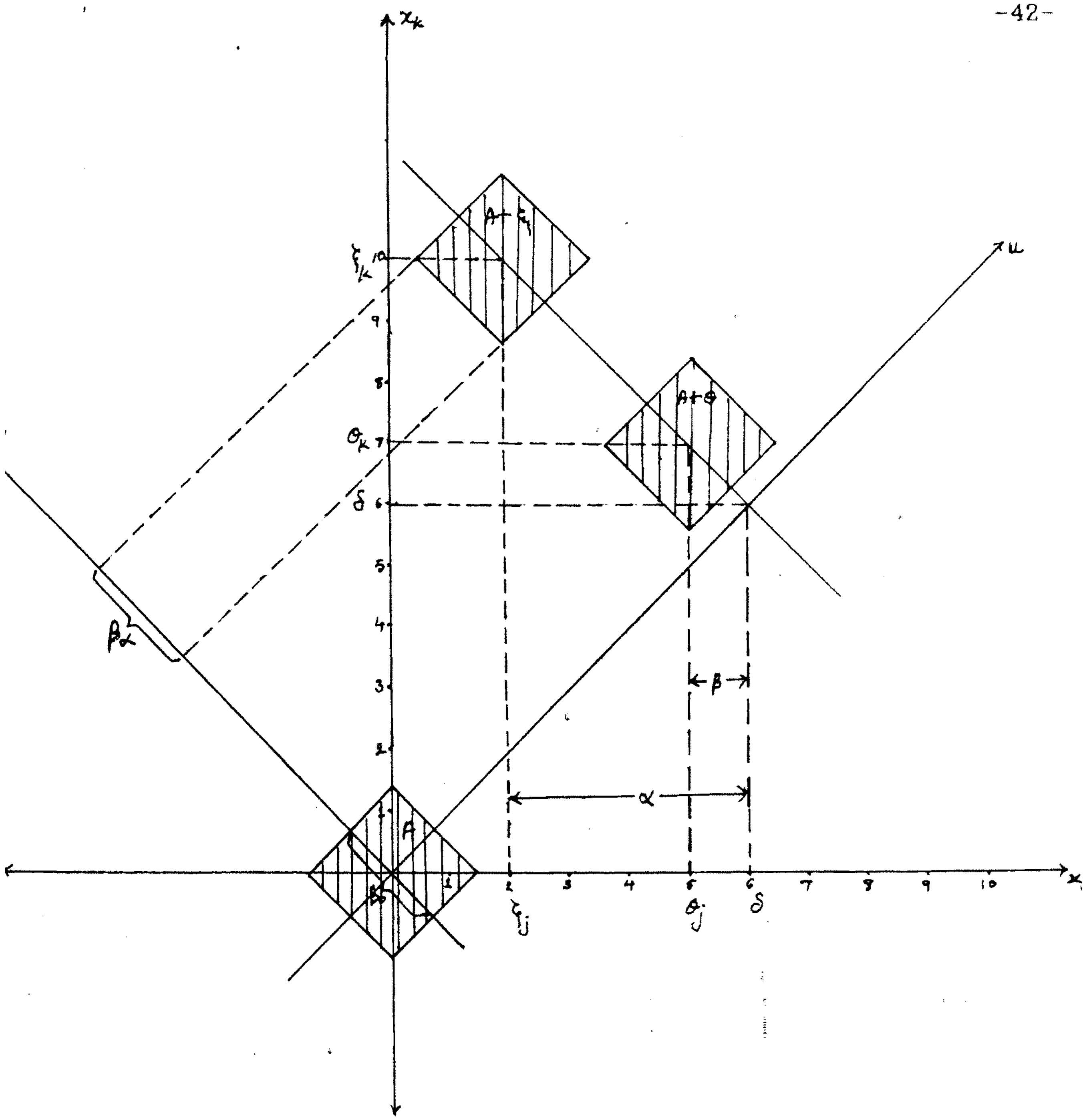
Let $\qquad u = x_j + x_k \qquad\qquad v = x_j - x_k$

To obtain the result integrate first on v conditionally with other variables held fixed.

Define the set

$$\beta_\alpha = \left\{ v : \left(x_1, \ldots, x_{j-1}, \frac{u+v}{2}, x_{j+1}, \ldots, x_{k-1}, \frac{u-v}{2}, x_{k+1}, \ldots, x_n\right) \in A + \xi \right\}$$

β_α is the set of v-values corresponding to points in $A + \xi$.



A typical example of the set A which is centered at the origin.

fig.2

Now ξ_j is at a distance α from the point ε on x_j- axis. For different values of α we get different $\mathcal{B}_\alpha$'s. When $\alpha = 0$, $A + \xi$ would locate itself on u-axis. And the corresponding $\mathcal{B}_\alpha$ would be denoted as $\mathcal{B}_0$. (This is shown in fig.2).

We claim that the set $\mathcal{B}_0$ is symmetric (1)

Let $V' \in \mathcal{B}_0$

Then

$$\left(x_1 - \xi_1, \ldots, x_{j-1} - \xi_{j-1}, \frac{(u-2\varepsilon) + v'}{2}, x_{j+1} - \xi_{j+1}, \ldots, x_{k-1} - \xi_{k-1}, \frac{(u-2\varepsilon) - v'}{2}, x_{k+1} - \xi_{k+1}, \ldots, x_n - \xi_n\right) \in A$$

As A is permutation symmetric

$$\left(x_1 - \xi_1, \ldots, x_{j-1} - \xi_{j-1}, \frac{(u-2\varepsilon) - v'}{2}, x_{j+1} - \xi_{j+1}, \ldots, x_{k-1} - \xi_{k-1}, \frac{(u-2\varepsilon) + v'}{2}, x_{k+1} - \xi_{k+1}, \ldots, x_n - \xi_n\right) \in A$$

$\Longrightarrow \quad -v' \in \mathcal{B}_0$.

Hence $\mathcal{B}_0$ is symmetric.

The set $\mathcal{B}_0$ is convex (2)

Let V' and $V'' \in \mathcal{B}_0$

Let $x' = \left(x_1 - \xi_1, \ldots, \frac{(u-2\varepsilon) + v'}{2}, \ldots, \frac{(u-2\varepsilon) - v'}{2}, \ldots x_n - \xi_n\right)$

$x'' = \left(x_1 - \xi_1, \ldots, \frac{(u-2\varepsilon) + v''}{2}, \ldots, \frac{(u-2\varepsilon) - v''}{2}, \ldots x_n - \xi_n\right)$

be two points from A corresponding to V' and V'' such that $x' < x''$ or $x'' < x'$ such a pair should always exist as per the definition of $\mathcal{B}_c$. Without loss of generality we assume that $x' < x''$.

To prove the convexity of $\mathcal{B}_c$ we have to show that

$$(\lambda V' + (1 - \lambda) V'') \in \mathcal{B}_c$$

i.e., to show that there exists at least one x call it x''' such that

$$x''' = \left[x_1 - \xi_1, \ldots\ldots, \frac{(u-2\xi) + (\lambda V' + (1-\lambda)V'')}{2}, \ldots\ldots, \frac{(u-2\xi) - (\lambda V' + (1-\lambda)V'')}{2}\right] \in A.$$

It is easy to varify that $x''' < x''$

Hence by the property of A

$$x''' \in A.$$

Hence $\mathcal{B}_c$ is convex.

As a function of v

$$f(x) = f(x_1, \ldots, x_{j-1}, \frac{u+v}{2}, x_{j+1}, \ldots, x_{k-1}, \frac{u-v}{2}, x_{k+1}, \ldots, x_n) \text{ is symmetric} \quad (3)$$

$$f(v) = f(x_1, \ldots, x_{j-1}, \frac{u+v}{2}, x_{j+1}, \ldots, x_{k-1}, \frac{u-v}{2}, x_{k+1}, \ldots, x_n)$$

$$= f(x_1, \ldots, x_{j-1}, \frac{u-v}{2}, x_{j+1}, \ldots, x_k + \frac{u+v}{2}, x_{k+1}, \ldots, x_n)$$

(because f is permutation invariant)

$$= f(-v)$$

Hence $f(x)$ is symmetric as a function of v.

$$f(x) = f(x_1, \dots, x_{j-1}, \frac{u+v}{2}, x_{j+1}, \dots, x_{k-1}, \frac{u-v}{2}, x_{k+1}, \dots, x_n)$$

is unimodel (4)

Because $f(x)$ is Schur-concave as we move on $v = x_j + x_k$ axis away from $u = x_j - x_k$ the value of the function decreases. Hence as a function of v, f is unimodel.

(1), (2), (3) and (4) satisfies the conditions of Anderson's theorem.

Thus we can write

$$\int_{\mathcal{B}_\beta} f(x)\, dv \geq \int_{\mathcal{B}_\alpha} f(x)\, dv.$$

This inequality would be preserved while integrating out the remaining variables to yield.

$$\int_{A+\theta} f(x)\, dx \geq \int_{A+\xi} f(x)\, dx.$$

<u>4.B.2 Remark</u> : Even though the theorem assumes the function to be a density function, the proof does not make use of it. Thus the result would be true for any function which satisfies the other conditions of the theorem.

<u>4.B.3 Example</u> :

Let $\varphi(a) = \sum_{i=1}^{2} (a_i)^{1/2}$ which is a Schur-concave function.

Let $A = \{(a_1, a_2) : 0 \leq a_1 \leq 2, \ 0 \leq a_2 \leq 2\}$

observe that φ is defined only on the first quadrint.

Also the conditions of the theorem are satisfied on A.

Let $\theta = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $\xi = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

Obviously $\theta < \xi$

A, A + θ and A + ξ are shown in fig. 3.

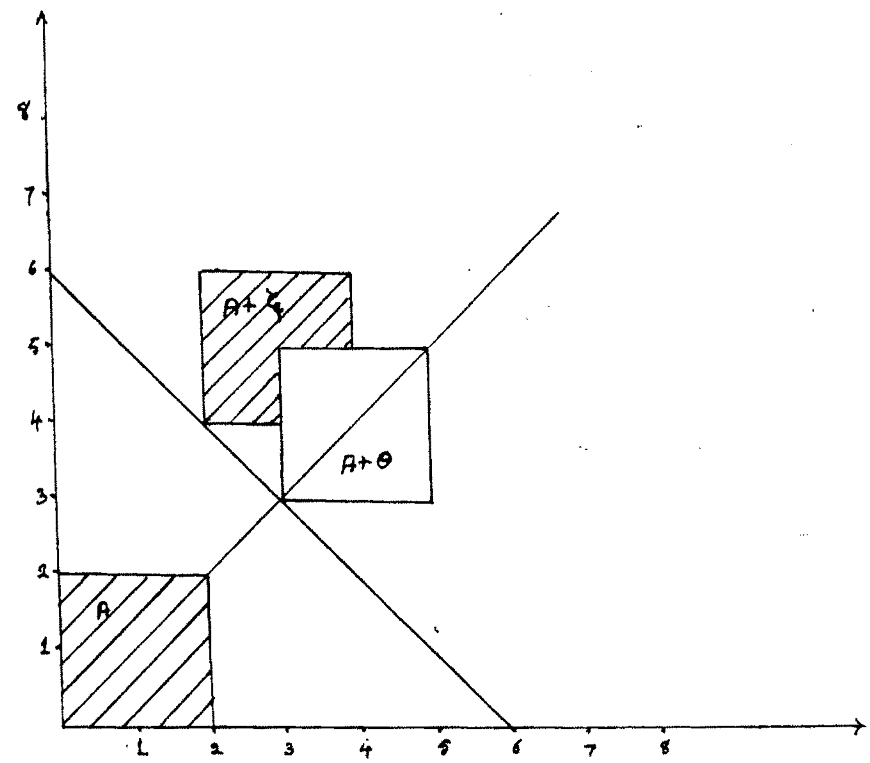


fig. 3

$$\int_{A+\theta} \varphi(a)\, da = \int_3^5 \int_3^5 (a_1^{1/2} + a_2^{1/2})\, da_1\, da_2$$

$$= 15.96$$

$$\int_{A+\xi} \varphi(a)\, da = \int_4^6 \int_2^4 (a_1^{1/2} + a_2^{1/2})\, da_1\, da_2$$

$$= 15.7$$

Hence $\int_{A+\theta} \varphi(a)\, da > \int_{A+\xi} \varphi(a)\, da.$

<u>Corollary</u> : If f_1 and f_2 are non-negative integrable Schur-concave functions defined on R^n then their convelution

$$f(a) = \int_{R^n} f_1(x)\ f_2(\theta - x)\ dx \quad .$$

is Schur-concave.

<u>Proof</u> :

Since f_2 is Schur-concave

$\tilde{f}_2(x) \equiv f_2(-x)$ is Schur-concave.

Hence by Marshall and Olkin's theorem

$$\int_{A+\theta} f_2(-x)\ dx = \int_{R^n} I_A(x)\ f_2(\theta - x)\ dx$$

is a Schur-concave function of θ. f_1 could be approximated by an increasing sequence of simple functions $\varphi_k = \Sigma\, \alpha_i\ I_{Ai}$, where the sets Ai satisfy the conditions of Marshall and Olkin's theorem and then use Lebesque's monotone convergence theorem.

<u>4.B.5 Remark</u> : The above corollary shows that the class of all non-negative integrable Schur-concave functions on R^n is closed under convolutions.

<u>4.B.6 Corollary</u> : If exchangable random variables $X_1,\ldots,X_n$ have a joint density which is Schur-concave and if φ is also permutation-symmetric non-negative and Schur-concave then

$$E\varphi(X-\theta) \quad \text{and} \quad P\{\varphi(X-\theta) \geq c\}$$

are Schur-concave functions of θ.

<u>Proof</u> :

Define $A = \{z : \varphi(z) \leq e\}$

As φ is Schur-concave $-\varphi$ is Schur-convex.

i.e., $x \prec y$ implies $\varphi(x) \leq \varphi(y)$

If $y \in A$ then $\varphi(y) \leq c$

i.e., $\varphi(x) \leq c$

Hence $x \in A$.

This implies the set A satisfies the condition of Marshall and Olkin theorem.

Let $f(x)$ be the joint density of $X_1,\ldots X_n$. Thus by the theorem

$$\int_{A+\theta} f(x)\,dx = P\{(X_1,\ldots,X_n) \in A+\theta\}$$

$$= P\{X-\theta \in A\}$$

$$= P\{-\varphi(X-\theta) \leq c\} \quad \text{(by definition of A)}$$

$$= P\{\varphi(X-\theta) \geq k\} \text{ is Schur-concave.}$$

Hence $P\{\varphi(X-\theta) \geq c\}$ is Schur-concave.

This implies

$$P\{\varphi(X - \theta) \geq c\} \leq P\{\varphi(X - \xi) \geq c\}.$$

for all C; whenever $\theta > \xi$

This implies $E\varphi(X - \theta) \leq E\varphi(X - \xi)$

Hence $E\varphi(X - \theta)$ is Schur-concave.

4.C APPLICATIONS :

4.C.1 : Let $X_1, \ldots, X_n$ be n standard normal random variables.

Consider the problem of locating the n-dimensional square $A = \{x : |x_i| \leq a\}$ of fixed size with optimal probability content.

It is easy to varify that Σx_i^2 is a Schur-convex function.

Hence $- \Sigma x_i^2$ is a Schur-concave function.

i.e., $\frac{1}{(2\pi)^n} e^{-1/2 \, \Sigma x_i^2}$ is a Schur-concave function.

Also the set A satisfies the condition of Marshall and Olkin theorem as it is permutation symmetric.

Hence by the theorem

$P\{X \in A + \theta\}$ is a Schur-concave function of θ.

Let $\Sigma \theta_i = b$

Hence $\frac{1}{n}(b, \ldots, b) < (\theta_1, \ldots, \theta_n)$

Thus the region of optimal probability content could be located at $\frac{1}{n}$ (b,...,b) for any choice of θ such that $\Sigma \theta_i = b$.

<u>4.C.2 Remark</u> : One main difficulty in applying the theorem is varification of the Schur-concavity of the function under consideration. Under such circumstances one can use the result due to Schur and Ostrowski, which states that a necessary and sufficient condition that a permutation - symmetric differentiable function be Schur-concave is that

$$\left[\frac{\partial \varphi(x)}{\partial x_i} - \frac{\partial \varphi(x)}{\partial x_j} \right] (x_i - x_j) \leq 0.$$

for all $i \neq j$.