The Annals of Statistics
1977, Vol. 5. No. 6, 1228–1234

AN INEQUALITY FOR MULTIVARIATE NORMAL PROBABILITIES WITH APPLICATION TO A DESIGN PROBLEM

BY YOSEF RINOTT[1] AND THOMAS J. SANTNER[2]

University of Chicago, The Hebrew University; and Cornell University

Some results from the theory of total positivity and Schur convexity are applied in deriving inequalities for multivariate normal probabilities having a certain covariance matrix. The result is applied to determine an optimal experimental design in an analysis of covariance model when selection of the best treatment is desired.

1. Introduction. Probabilistic inequalities for the multivariate normal distribution as a function of the covariance matrix have frequently been studied in the literature. Anderson (1955) considered probabilities of symmetric convex sets and proved that they are a monotone decreasing function of the covariance matrix with partial ordering of covariance matrices defined by $\Sigma \succeq \Psi$ if $\Sigma - \Psi$ is positive semidefinite. More precisely let $\mathbf{X} = (X_1, \cdots, X_p)$ and $\mathbf{Y} = (Y_1, \cdots, Y_p)$ be p-variate normally distributed random vectors with common mean vector zero and covariance matrices Σ and Ψ respectively and let E be a convex set symmetric about the origin. Then $\Sigma \succeq \Psi$ implies

$$P[\mathbf{X} \in E] \leqq P[\mathbf{Y} \in E]\,. \tag{1.1}$$

Slepian (1962) studied the probabilities of events defined by one-sided inequalities (quadrants) and showed they are monotone increasing in the covariances of the random variables. Specifically let $\mathbf{X}$ and $\mathbf{Y}$ be p-variate normally distributed with common mean vector and covariance matrices $\sum = (\sigma_{ij})$ and $\Psi = (\psi_{ij})$ respectively. If $\sigma_{ii} = \psi_{ii}$, $1 \leqq i \leqq p$ and $\sigma_{ij} \leqq \psi_{ij}$, $1 \leqq i \neq j \leqq p$ then

$$P[X_1 \leqq c_1, \cdots, X_p \leqq c_p] \leqq P[Y_1 \leqq c_1, \cdots, Y_p \leqq c_p] \tag{1.2}$$

for any $c_1, \cdots, c_p$. For related inequalities and references see Das Gupta et al. (1972).

In this paper we combine aspects of (1.1) and (1.2) by considering quadrant probabilities with special covariance matrices where the ordering involved is similar to that considered by Anderson. Our results were motivated by a design problem which will be discussed later.

Let $\mathbf{X} = (X_1, \cdots, X_p)$ have the p-variate normal distribution with mean zero and covariance matrix Σ. For $d \in R$ (the real line) set $P[\Sigma, d] = P[X_1 \leqq d, \cdots, X_p \leqq d]$. Let I and J denote the identity matrix of order p and a $p \times p$ matrix

Received September, 1976; revised April, 1977.

[1] Research supported by NSF Grant MCS 72-04364 A04.
[2] Research supported by NSF Grant ENG 75-10487.

AMS 1970 *subject classifications*. Primary 62F07, 62K05.
Key words and phrases. Schur-concavity, total positivity, optimal design, ranking and selection.

with all elements equal to 1 respectively and define $\Sigma_{\alpha^2} = I + \alpha^2 J$ (covariance matrices of this form appear in multivariate analysis under the name of intraclass or equicorrelation pattern). For any fixed $\alpha \in R$ we obtain that the difference

$$h(d) = P[\Sigma_{\alpha^2}, d] - P[I, d] \tag{1.3}$$

has one sign change from positive to negative over $d \in R$, i.e., there exists a $d_0 \in R$ (depending on α and p) such that $(d - d_0)h(d) \leqq 0$ for all $d \in R$.

Our application requires the following generalization of this result. For $1 \leqq n \leqq p$ let A_n be a $p \times p$ matrix defined by $A_n = (a_{ij})$ where $a_{ij} = 1$ for $1 \leqq i, j \leqq n$, $a_{ij} = 0$ otherwise, and define $\Sigma_{\alpha^2,\beta^2} = I + \beta^2 J + \alpha^2 A_n$. In particular $\Sigma_{0,\beta^2} = I + \beta^2 J = \Sigma_{\beta^2}$ by our previous definition. For any fixed $\alpha, \beta \in R$ and $1 \leqq n \leqq p$ we obtain that the difference

$$g(d) = P[\Sigma_{\alpha^2,\beta^2}, d] - P[\Sigma_{\beta^2}, d] \tag{1.4}$$

has one sign change at d_0, from positive to negative as a function of $d \in R$, where $d_0 \in R$ depends on α^2, β^2, n and p.

In the case $n \leqq (p+1)/2$, $\beta^2 = 1$ we have $d_0 \leqq 0$, and by expressing the probabilities in (1.4) as integrals (denoting $m = p - n$) we obtain for all $d \geqq 0$

$$\int\int \Phi^n(d + x + \alpha y)\Phi^m(d + x)\, d\Phi(x)\, d\Phi(y) \leqq \int \Phi^{n+m}(d + x)\, d\Phi(x) \tag{1.5}$$

where $\Phi(x) = 1/(2\pi)^{\frac{1}{2}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$, $m + 1 \geqq n \geqq 1$, and all integrals are over R. In fact we can show that in this case the left-hand side of (1.5) is decreasing in $|\alpha|$, for any $d \geqq 0$.

We consider the application of (1.5) to a design problem for an analysis of covariance model. Suppose l observations are to be taken on each of k treatments and the responses $\{Y_{ij} \mid 1 \leqq i \leqq k,\ 1 \leqq j \leqq l\}$ satisfy the analysis of covariance model

$$Y_{ij} = \mu_i + \beta X_{ij} + \varepsilon_{ij}\,; \qquad 1 \leqq i \leqq k\,, \quad 1 \leqq j \leqq l \tag{1.6}$$

where $\{\mu_i\}$ are k unknown treatment effects, the $\{X_{ij}\}$ are the values of kl known concomitant variables, β is the common unknown slope of the k regressions and the ε_{ij} are independent normal variables with zero means and common known variance, σ^2. In certain applications the experimenter may have some control over the values of the concomitant variables $\{X_{ij}\}$. In particular we have in mind the situation in which the values of the X-variables for the kl subjects participating in the experiments are given, but their allocation to the k treatment groups is under the experimenter's control. If the purpose of the experiment is to test equality of the $\{\mu_i\}$ then it can be shown that the power of the F test is maximized when the $\{X_{ij}\}$ are allocated so that

$$\bar{X}_{1\cdot} = \bar{X}_{2\cdot} = \cdots = \bar{X}_{k\cdot} \tag{1.7}$$

where $\bar{X}_{i\cdot} = \sum_{j=1}^{l} X_{ij}/l$. If a given set of $\{X_{ij}\}$ values is to be used in the experiment then it may be impossible to achieve allocation (1.7) exactly, however a good approximation may be attainable if the experiment is sufficiently large. By invoking (1.5) we show that if the purpose of the analysis is to select the best treatment, i.e., the treatment with the largest μ_i, and if the "natural" selection procedure based on the BLUE's of the μ_i's is employed then the minimum probability of a correct selection outside a suitably chosen indifference zone is maximized by the same design (1.7). The main tool used to reduce this problem to (1.5) is Schur convexity as described in Marshall and Olkin (1974).

2. Inequalities for the normal distribution. Set $\phi(t) = 1/(2\pi)^{\frac{1}{2}} e^{-t^2/2}$ and $\Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt$.

THEOREM 2.1. *For any integer $n \geqq 1$ and $\alpha \in R$ the function h defined by $h(z) = \int \Phi^n(z + \alpha y)\, d\Phi(y) - \Phi^n(z)$ has exactly one sign change. More precisely, there exists $z_0 \in R$ such that $(z - z_0)h(z) \leqq 0$ for all $z \in R$.*

REMARK 2.1. A direct calculation shows that function $h(\cdot)$ of Theorem 2.1 is identical with $h(\cdot)$ defined by (1.3).

PROOF OF THEOREM 2.1. We first show that $h(z)$ must exhibit an odd number of sign changes by proving: (a) $h(z) > 0$ for $z < 0$; and (b) $h(z) < 0$ for z sufficiently large. Part (a) follows from the fact that for $z < 0$

$$\Phi^n(z) \leqq \{\Phi(z/(1 + \alpha^2)^{\frac{1}{2}})\}^n = \{\int \Phi(z + \alpha y)\, d\Phi(y)\}^n \leqq \int \Phi^n(z + \alpha y)\, d\Phi(y) .$$

Part (b) can be proved as follows: $h(z) \leqq \int \Phi(z + \alpha y)\, d\Phi(y) - \Phi^n(z) = 1 - \Phi^n(z) - [1 - \Phi(z/(1 + \alpha^2)^{\frac{1}{2}})] \leqq n[1 - \Phi(z)] - [1 - \Phi(z/(1 + \alpha^2)^{\frac{1}{2}})]$. Substituting the asymptotic approximation $1 - \Phi(x) \sim \phi(x)/x$ into the last two terms gives the result. The remainder of the proof consists of showing that $h(z)$ can have no more than two sign changes. Since $\Phi^n(z) = \int \Phi^n(z + \alpha y)\, d\delta_0(y)$ where δ_0 is the probability measure degenerate at zero, it follows that $h(z)$ can be uniformly approximated (in z) by the sequence $h_m(z) = \int \Phi^n(z + \alpha y)[\phi(y) - u_m(y)]\, dy$ as $m \to \infty$ where $u_m(y)$ is the uniform density on $[-1/m, 1/m]$. Hence it suffices to show that $h_m(z)$ has at most two sign changes for m sufficiently large. The function $f_m(y) = \phi(y) - u_m(y)$ has two sign changes, being negative in $[-1/m, 1/m]$ and positive outside this interval. The theory of total positivity (see Karlin (1968), Chapter 1, Section 3 and Chapter 5) implies that provided we can demonstrate that the kernel $K(z, y) = \Phi^n(z + \alpha y)$ is *totally positive* of order 3 (TP_3) or *sign regular* of order 3 (SR_3), then by the "variation diminishing" character of such kernels the number of sign changes the transform $h_m(z) = \int K(z, y) f_m(y)\, dy$ can have is bounded by the number of sign changes of $f_m(y)$. Since $f_m(y)$ has exactly two sign changes as noted above, the transform $h_m(z)$ has at most two sign changes and the proof is complete. (The sign regularity of $\Phi^n(z + \alpha y)$ which is required in the above proof is demonstrated in the Appendix.)

Theorem 2.2 below describes the behavior of the function $g(d)$ of (1.4) expressed in integral form.

THEOREM 2.2. *For any $n, m \geqq 1$, $\alpha, \beta \in R$ the function $g(d)$ defined by*

$$g(d) = \iint \Phi^n(d + \beta x + \alpha y)\Phi^m(d + \beta x)\, d\Phi(x)\, d\Phi(y) - \int \Phi^{n+m}(d + \beta x)\, d\Phi(x)$$

has one sign change. More precisely there exists $d_0 \in R$ (depending on n, m, α, β) such that $(d - d_0)g(d) \leqq 0$ for all $d \in R$.

PROOF. Set $f(z) = 1/\beta\Phi^m(z)h(z)$ where h is defined in Theorem 2.1. Then $g(d) = \int f(z)\phi((z - d)/\beta)\, dz$. The function f has one sign change by Theorem 2.1 and the desired result follows from the total positivity of the kernel $K(z, d) = \phi((z - d)/\beta)$.

REMARK 2.2. We can now derive inequality (1.5) provided that $g(0) \leqq 0$ in the case $m + 1 \geqq n \geqq 1$, $\beta^2 = 1$ since it follows that $g(d) \leqq 0$ by Theorem 2.2 in this case. In order to prove $g(0) \leqq 0$ set

$$H(\alpha, d) = \iint \Phi^n(d + x + \alpha y)\Phi^m(d + x)\, d\Phi(x)\, d\Phi(y) - \int \Phi^{n+m}(d + x)\, d\Phi(x) .$$

Note that $H(\alpha, d)$ is an even function in α for any fixed d and we can restrict attention to $\alpha > 0$.

Differentiation of $H(\alpha, 0)$ and a substitution yield

$$\frac{dH(\alpha, 0)}{d\alpha} = \frac{n}{2\alpha^2} \iint (u - v)[\Phi^{m-n+1}(v) - \Phi^{m-n+1}(u)] \times \Phi^{n-1}(u)\Phi^{n-1}(v)\phi\left(\frac{u - v}{\alpha}\right) d\Phi(u)\, d\Phi(v) .$$

It follows that $dH(\alpha, 0)/d\alpha = 0$ for $m + 1 = n$ and $dH(\alpha, 0)/d\alpha < 0$ for $m + 1 > n$ so that for $m + 1 \geqq n \geqq 1$, $H(\alpha, 0) \leqq H(0,0) = 0$. For $\beta^2 = 1$ and a given α we have $g(0) = H(\alpha, 0)$ implying $g(0) \leqq 0$ in the case $m + 1 \geqq n \geqq 1$ and $\beta^2 = 1$ and this completes the proof of (1.5).

A stronger result than (1.5), namely that the left-hand side of (1.5) is decreasing in $|\alpha|$ is now indicated. Invoking total positivity arguments similar to those employed above it can be shown that for any fixed $\alpha > 0$ the function $dH(\alpha,d)/d\alpha$ of d has at most one sign change which must be from positive to negative. It follows that for $d > 0$, $dH(\alpha, d)/d\alpha < 0$ for all $\alpha > 0$ when $m + 1 \geqq n > 1$ since $dH(\alpha, 0)/d\alpha \leqq 0$ and hence $H(\alpha, d)$ is decreasing in $|\alpha|$ by the symmetry of $H(\alpha, d)$.

3. An application to a design problem. Our results will now be applied to a design problem in selecting the best treatment in an analysis of covariance model. Suppose k treatments are studied by applying each to l subjects and observing the response Y_{ij} for the jth subject under treatment i and that $Y_{ij} = \mu_i + \beta X_{ij} + \varepsilon_{ij}$ where the μ_i's are unknown treatment effects, the X_{ij} are known values of concomitant variables with common unknown slope β and the ε_{ij} are independent $N(0, \sigma^2)$ variables. Let $\mu_{(i)} \leqq \cdots \leqq \mu_{(k)}$ denote the ranked effects and suppose the treatment with the largest effect, $\mu_{(k)}$ is considered the best treatment. The BLUE of μ_i is $\hat{\mu}_i = \bar{Y}_{i\cdot} - \hat{\beta}\bar{X}_{i\cdot}$ where $\bar{Y}_{i\cdot} = \sum_{j=1}^{l} Y_{ij}/l$, $\bar{X}_{i\cdot} = \sum_{j=1}^{l} X_{ij}/l$ and $\hat{\beta} = E_{XY}/E_{XX}$ where $E_{XY} = \sum_{i=1}^{k} \sum_{j=1}^{l} (X_{ij} - \bar{X}_{i\cdot})Y_{ij}$ and $E_{XX} = \sum_{i=1}^{k} \sum_{j=1}^{l} (X_{ij} - \bar{X}_{i\cdot})^2$. We consider the procedure which selects the treatment producing the largest $\hat{\mu}_i$ as the best treatment. Let $\Omega(\delta^*) = \{\boldsymbol{\mu} \mid \mu_{(k)} - \mu_{(k-1)} \geqq \delta^*\}$ for some $\delta^* > 0$ denote those $\boldsymbol{\mu}$ configurations where the best treatment is at least δ^* superior to the remaining ones. Let $P[CS \mid \boldsymbol{\mu}, \mathbf{X}]$ denote the probability of making a correct selection of the best treatment using the rule above given $\boldsymbol{\mu}$ is the true configuration of the treatment effects and $\mathbf{X} = \{X_{ij}\}$ are the values of the concomitant variables. It is easy to see that $\inf_{\boldsymbol{\mu} \in \Omega(\delta^*)} P[CS \mid \boldsymbol{\mu}, \mathbf{X}]$ is attained at $\boldsymbol{\mu}_0 = (0, \cdots, 0, \delta^*)$ and from now on we restrict our attention to this so-called least favorable configuration (see Bechhofer (1954)). Let $\hat{\mu}_{(i)}$ and $X_{(i)j}$ be the BLUE and jth covariate associated with the ith best treatment (which is unknown) $i = 1, \cdots, k$. We have

$$\begin{aligned} P[CS \mid \boldsymbol{\mu}_0, \mathbf{X}] &= P[\hat{\mu}_{(k)} > \hat{\mu}_{(j)}, 1 \leqq j \leqq k-1] \\ &= P[Z_j \geqq -\delta^*/\sigma, 1 \leqq j \leqq k-1] \end{aligned}$$

where $\mathbf{Z} = (Z_1, \cdots, Z_{k-1})$ is $(k-1)$ variate normal with zero mean vector and covariance matrix $\Sigma = (\sigma_{ij})$ where

$$\begin{aligned} \sigma_{ij} &= 2/l + (\bar{X}_{(k)\cdot} - \bar{X}_{(j)\cdot})^2/E_{XX}\,, & i = j \\ &= 1/l + (\bar{X}_{(k)\cdot} - \bar{X}_{(j)\cdot})(\bar{X}_{(k)\cdot} - \bar{X}_{(i)\cdot})/E_{XX}\,, & i \neq j\,. \end{aligned}$$

This covariance matrix and hence the probability of correct selection depends on $\bar{X}_{(k)\cdot}$, the mean of the covariates in the best treatment. Obviously the assignment of one of the values $\bar{X}_{1\cdot}, \cdots, \bar{X}_{k\cdot}$ to $\bar{X}_{(k)\cdot}$ is unknown since we do not know which of the k treatments is the best and hence the value of $\bar{X}_{(k)\cdot}$ is unknown. We therefore consider the minimum of the probability of correct selection over the k possible assignments of $\bar{X}_{1\cdot}, \cdots, \bar{X}_{k\cdot}$ to $\bar{X}_{(k)\cdot}$. This minimum can be expressed in integral form as follows:

$$\begin{aligned} (3.1) \qquad & \min_{\bar{X}_{(k)\cdot} = \bar{X}_{j\cdot}; 1 \leqq j \leqq k} P[CS \mid \boldsymbol{\mu}_0, \mathbf{X}] \\ &= \min_{1 \leqq j \leqq k} \int\int \prod_{i=1; i \neq j}^{k} \Phi(d + x + (b_j - b_i)y)\, d\Phi(x)\, d\Phi(y) \end{aligned}$$

where $d = l^{\frac{1}{2}}\delta^*/\sigma$ and $b_i = \bar{X}_{i\cdot} l^{\frac{1}{2}}/(E_{XX})^{\frac{1}{2}}$ $i = 1, \cdots, k$. For any given kl values for the X variables we seek the optimal allocation of the k treatment groups, i.e., the allocation which maximizes (3.1).

The following definitions and results will be required for the proof of Theorem 3.1 below.

DEFINITION. If $\mathbf{a}$ and $\mathbf{b}$ are p dimensional vectors then $\mathbf{a}$ is said to majorize $\mathbf{b}$ (written $\mathbf{a} \succ \mathbf{b}$) if upon reordering components to achieve $a_1 \geqq a_2 \geqq \cdots \geqq a_p$ and $b_1 \geqq b_2 \geqq \cdots \geqq b_p$ then $\sum_{j=1}^{k} a_j \geqq \sum_{j=1}^{k} b_j$ for $1 \leqq k \leqq p$ with equality for $k = p$. Real valued functions Ψ for which $\mathbf{a} \succ \mathbf{b}$ implies $\Psi(\mathbf{a}) \leqq \Psi(\mathbf{b})$ are called Schur-concave.

It is well known (see for example Marshall and Olkin (1974)) that $\Psi(\mathbf{a}) = \prod_{i=1}^{p} f(a_i)$ is Schur-concave if f is log-concave. From this it follows readily that the function

$$\Psi(\mathbf{a}) = (a_1, \cdots, a_p) = \int\int \prod_{i=1}^{p} \Phi(d + x + a_i y)\, d\Phi(x)\, d\Phi(y)$$

is Schur-concave.

THEOREM 3.1. *For any given* $X = \{X_{ij}\}$ *the expression* (3.1) *attains its maximum when* $\bar{X}_{1\cdot} = \bar{X}_{2\cdot} = \cdots = \bar{X}_{k\cdot}$. *The value of the maximum is independent of the values of* $\bar{X}_{i\cdot}$, *and thus any allocation of subjects to treatments which satisfies* $\bar{X}_{1\cdot} = \bar{X}_{2\cdot} = \cdots = \bar{X}_{k\cdot}$ *is optimal for the selection problem, and achieves the same probability of correct selection.*

PROOF. Assume without loss of generality $b_1 \leqq b_2 \leqq \cdots \leqq b_k$ and choose $j_0 = (k+1)/2$ or $(k+2)/2$ as k is odd or even. Setting $(a_1, \cdots, a_{k-1}) = (b_{j_0} - b_1, \cdots, b_{j_0} - b_{j_0-1}, b_{j_0} - b_{j_0+1}, \cdots, b_{j_0} - b_k)$ it will suffice to prove

$$(3.2)\qquad \Psi(a_1, \cdots, a_{k-1}) = \int\int \prod_{i=1}^{k-1} \Phi(d + x + a_i y)\, d\Phi(x)\, d\Phi(y) \leqq \int \Phi^{k-1}(d + x)\, d\Phi(x)\,.$$

Our choice of j_0 implies that we can write $k - 1 = m + n$ where $n = m$ or $n = m + 1$ as k is odd or even and we have $a_i \geqq 0$ for $i = 1, \cdots, n$ and $a_i \leqq 0$ otherwise. Let $v_n(c)$ denote a vector of length $n + m$ with the value c in the first n position and zeros in the remaining positions and set $\alpha = 1/n \sum_{i=1}^{k-1} a_i$. We have $\mathbf{a} \succ v_n(\alpha)$ and by the Schur-concavity of $\Psi(a_1, \cdots, a_{k-1})$ we obtain

$$\Psi(a_1, \cdots, a_{k-1}) \leqq \Psi\,(v_n(\alpha)) = \int\int \Phi^n(d + x + \alpha y)\Phi^m(d + x)\, d\Phi(x)\, d\Phi(y)\,.$$

Inequality (3.2) follows from inequality (1.5).

REMARK 3.1. Our result shows that the experimenter should design the experiment so that $\bar{X}_{1\cdot} = \bar{X}_{2\cdot} = \cdots = \bar{X}_{k\cdot}$. If this cannot be attained the continuity of the probability of correct selection as a function of $\mathbf{X} = \{X_{ij}\}$ suggests that an approximation of this design may be satisfactory. In such a case the probability of correct selection given in (3.1) depends not only on $\bar{X}_{1\cdot}, \cdots, \bar{X}_{k\cdot}$ but also on E_{XX} and an inspection of (3.1) shows that the approximation is improved by increasing the value of E_{XX}. It is therefore worthwhile to note that for any given kl values for the X variables, fixing $\sum_{ij} X_{ij}^2$, E_{XX} is a concave function of $\bar{X}_{1\cdot}, \cdots, \bar{X}_{k\cdot}$ which is maximized when $\bar{X}_{1\cdot} = \cdots = \bar{X}_{k\cdot}$.

4. Appendix. We prove here that the kernel $K(x, y) = \Phi^n(x - y)$ is strictly totally positive of order 3. This implies that $\Phi^n(x + ay)$ is strictly sign regular of order 3 as required for the proof of Theorem 2.1. (See Karlin (1968) Chapter 2 for definitions.)

For $n = 1$ it is well known that $\Phi(x - y)$ is STP of all orders. Since the product of STP_2 kernels is STP_2 (Karlin (1968) page 157), $\Phi^n(x - y)$ is STP_2 for all $n \geqq 1$ and invoking Theorem 2 of Karlin (1957) it remains to prove that for $n \geqq 2$ we have

$$(4.1)\qquad \begin{vmatrix} \Phi^n(w_1) & \frac{d}{dw}\Phi^n(w_1) & \frac{d^2}{dw^2}\Phi^n(w_1) \\ \Phi^n(w_2) & \frac{d}{dw}\Phi^n(w_2) & \frac{d^2}{dw^2}\Phi^n(w_2) \\ \Phi^n(w_3) & \frac{d}{dw}\Phi^n(w_3) & \frac{d^2}{dw^2}\Phi^n(w_3) \end{vmatrix} < 0$$

for all $w_1 < w_2 < w_3$. Some calculations show that the determinant in (4.1) can be written as

$$n^2\Phi^{n-2}(w_1)\Phi^{n-2}(w_2)\Phi^{n-1}(w_3) \times \left\{ \Phi(w_1)\Phi(w_2)\Phi(w_3) \begin{vmatrix} \Phi(w_1) & \frac{d}{dw}\Phi(w_1) & \frac{d^2}{dw^2}\Phi(w_1) \\ \Phi(w_2) & \frac{d}{dw}\Phi(w_2) & \frac{d^2}{dw^2}\Phi(w_2) \\ \Phi(w_3) & \frac{d}{dw}\Phi(w_3) & \frac{d^2}{dw^2}\Phi(w_3) \end{vmatrix} \right. \tag{4.2}$$

$$+ (n-1) \begin{vmatrix} \Phi(w_1) & \frac{d}{dw}\Phi(w_1) \\ \Phi(w_2) & \frac{d}{dw}\Phi(w_2) \end{vmatrix} \begin{vmatrix} \Phi(w_1) & \frac{d}{dw}\Phi(w_1) \\ \Phi(w_3) & \frac{d}{dw}\Phi(w_3) \end{vmatrix}$$

$$\left. \times \begin{vmatrix} \Phi(w_2) & \frac{d}{dw}\Phi(w_2) \\ \Phi(w_3) & \frac{d}{dw}\Phi(w_3) \end{vmatrix} \right\}.$$

All the determinants in (4.2) are negative since $\Phi(x - y)$ is strictly totally positive and (4.1) follows.

REFERENCES

[1] ANDERSON, T. W. (1955). The integral of a symmetric unimodel function. *Proc. Amer. Math. Soc.* **6** 170–176.

[2] BECHHOFER, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist.* **25** 16–39.

[3] DAS GUPTA, S., EATON, M. L., OLKIN, I., PERLMAN, M., SAVAGE, L. J. and SOBEL, M. (1972). Inequalities on the probability content of convex regions for elliptically contoured distributions. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 241–265, Univ. of Calif. Press.

[4] KARLIN, S. (1957). Pólya type distribution type distributions, II. *Ann. Math. Statist.* **27** 281–308.

[5] KARLIN, S. (1968). *Total Positivity*. Stanford Univ. Press.

[6] MARSHALL, A. and OLKIN, I. (1974). Majorization in multivariate distributions. *Ann. Statist.* **2** 1189–1200.

[7] SLEPIAN, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell System Tech. J.* **41** 463–501.

DEPARTMENT OF STATISTICS
THE HEBREW UNIVERSITY OF JERUSALEM
JERUSALEM, ISRAEL

DEPARTMENT OF OPERATIONS RESEARCH
CORNELL UNIVERSITY, UPSON HALL
ITHACA, N.Y. 14853