INEQUALITIES ON THE PROBABILITY CONTENT OF CONVEX REGIONS FOR ELLIPTICALLY CONTOURED DISTRIBUTIONS

S. DAS GUPTA
UNIVERSITY OF MINNESOTA

M. L. EATON
UNIVERSITY OF CHICAGO

I. OLKIN
STANFORD UNIVERSITY

M. PERLMAN
UNIVERSITY OF MINNESOTA

L. J. SAVAGE
YALE UNIVERSITY

M. SOBEL
UNIVERSITY OF MINNESOTA

We dedicate this paper to the memory of L. J. Savage.

1. Introduction

Let $x = (x_1, \cdots, x_p)$ be a p dimensional random vector,

$$P(\Sigma) \equiv P_\Sigma\{|x_1| \leqq h_1, \cdots, |x_p| \leqq h_p\} = \int_{-h}^{h} |\Sigma|^{-1/2} f(x\Sigma^{-1}x')\,dx, \tag{1.1}$$

and

$$P^+(\Sigma) \equiv P_\Sigma\{x_1 \leqq \ell_1, \cdots, x_p \leqq \ell_p\} = \int_{-\infty}^{\ell} |\Sigma|^{-1/2} f(x\Sigma^{-1}x')\,dx, \tag{1.2}$$

where $h = (h_1, \cdots, h_p)$ and $\ell = (\ell_1, \cdots, \ell_p)$ are constant vectors, $h_i \geqq 0$, $i = 1, \cdots, p$, and $\Sigma = (\sigma_{ij})$ is a positive definite matrix. We call a density (with respect to Lebesgue measure) of the form

$$|\Sigma|^{-1/2} f(x\Sigma^{-1}x'), \tag{1.3}$$

This work was supported in part by the National Science Foundation Grant No. 17172 at Stanford University, Grants No. 11021, 9593, and 21074 at the University of Minnesota and Grant No. 25911 at the University of Chicago; by the Army, Navy, Air Force and NASA under a contract (N00014-67-4-0097-0014, Task Number NR-042-242) administered by the Office of Naval Research at Yale University, by the Army under a contract (DA-ARO-D31-124-70-G-102) at the University of Minnesota.

where $\int_0^\infty r^{p-1} f(r^2)\, dr < \infty$, elliptically contoured; it has also been called a spherical distribution by Lord [19], since the transformation $y = x\Sigma^{-1/2}$ yields a density which is uniform on spheres, or equivalently, since the distribution has spherical symmetry in the Euclidean geometry defined by the distance $\rho(x, y) = [(x - y)\Sigma^{-1}(x - y)']^{1/2}$. The normal distribution is clearly a special case and other examples are given below.

Inequalities for $P(\Sigma)$ and $P^+(\Sigma)$ have been obtained in a series of papers, and with a number of variants. However, in almost all instances the results are based on the normal distribution. The first inequality of which we know for $P^+(\Sigma)$ is due to Slepian [28].

THEOREM 1.1. *If* $f(z) = (2\pi)^{-p/2} e^{-z/2}$, $\Sigma = (\sigma_{ij})$, *and* $\Gamma = (\gamma_{ij})$, *with* $\sigma_{ij} \leqq \gamma_{ij}$ *and* $\sigma_{ii} = \gamma_{ii}$, *then* $P^+(\Sigma) \leqq P^+(\Gamma)$.

Slepian's elegant proof is based on a property of the normal distribution

$$\frac{\partial \varphi}{\partial \sigma_{ii}} = \frac{1}{2}\frac{\partial^2 \varphi}{\partial x_i^2}, \qquad \frac{\partial \varphi}{\partial \sigma_{ij}} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j}, \qquad i \neq j, \tag{1.4}$$

where $\varphi(x, \Sigma) = |\Sigma|^{-1/2} \exp\{-\frac{1}{2}(x\Sigma^{-1}x')\}$. Because the proof depends so heavily on (1.4), it is not easily adaptable to the more general class of elliptically contoured distributions. An alternative geometrical proof for the normal distribution is given by Chartres [2]. However, normality really does not play an essential role, and his proof can be modified to apply to elliptically contoured distributions. A general proof using a reflection-inclusion argument is given in Section 5. As a by product in the study of dependence, Lehmann [18] obtains the inequality $P^+(\Sigma) \geqq P^+(I)$ for the bivariate normal distribution with $\sigma_{11} = \sigma_{22} = 1$, and $\sigma_{12} \geqq 0$.

Inequalities for $P(\Sigma)$ perhaps originate with special results of Dunnett and Sobel [7] and of Dunn [5], in which it is shown that $P(\Sigma) \geqq P(I)$ for special forms of Σ (with $\sigma_{ii} = 1$) or for special values of p. The most general result for the normal distribution obtained by Šidák [25], [26] is that $P(\Sigma_\lambda)$ is a monotone increasing function of λ, $0 \leqq \lambda \leqq 1$, where

$$\Sigma_\lambda = \begin{pmatrix} \Sigma_{11} & \lambda\Sigma_{12} \\ \lambda\Sigma_{21} & \sigma_{22} \end{pmatrix}, \qquad \Sigma_{11} : p - 1 \times p - 1. \tag{1.5}$$

This he proved by using a conditional argument together with an inequality of Anderson [1] on the integral of a unimodal symmetric distribution. Jogdeo [13] provides a simpler proof of the same result by combining property (1.4) and the same inequality of Anderson. Scott [23] provides an alternative proof for the special case $P(\Sigma) \geqq P(I)$ using a conditional argument. (His proof contains a flaw which is discussed in Section 3.1.)

When the distribution is normal, Chover [3] uses a geometric argument to prove both the one sided and two sided inequalities. His proof is of interest since it treats both cases simultaneously, although the result is not completely general because the covariance matrices must satisfy certain conditions.

Our main result is an extension of Šidák's result to general elliptically contoured densities (Section 2). We give a stronger version dealing with a convex symmetric set, which permits several extensions (Section 3). In Section 3.1 we show how previous results for the normal distribution are interrelated. Complementary inequalities and reversals are given in Sections 3.2 and 3.3. As a consequence of our main theorem, an inequality for the probability of a convex symmetric set when one covariance dominates another is given in Section 3.4. Counterexamples in Section 3.5 show that certain assumptions cannot be weakened.

One motivation for seeking bounds for P_Σ stems from the study of simultaneous confidence bounds. Inequalities for Studentized variates have been obtained by Dunnett and Sobel [7], Halperin [11], Khatri [15], and Šidák [26], [27]. Section 4 provides some extensions of these results.

Property (1.4) for $p = 2$ apparently is an old result. A proof for general p was provided by Plackett [21]. The class of elliptically contoured distributions has been studied in some detail by Kelker [14], who obtains a number of characterizations. In Section 6 we discuss several characterizations, and incidentally give a direct and simple proof that (1.4) also characterizes the multivariate normal distribution. This property was independently obtained by another method by Patil and Boswell [20].

2. Main theorem

In this section we prove our main two sided inequality for elliptically contoured distributions.

THEOREM 2.1. *Let*

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \sigma_{pp} \end{pmatrix} \tag{2.1}$$

be a $p \times p$ positive definite matrix with $\Sigma_{11}: (p-1) \times (p-1)$, and let $x = (x_1, \cdots, x_p)$ be a random vector with density function $|\Sigma_\lambda|^{-1/2} f(x\Sigma_\lambda^{-1}x')$, where

$$\Sigma_\lambda = \begin{pmatrix} \Sigma_{11} & \lambda\Sigma_{12} \\ \lambda\Sigma_{21} & \sigma_{pp} \end{pmatrix}, \qquad 0 \leqq \lambda \leqq 1. \tag{2.2}$$

If C is a convex symmetric set in R^{p-1}, then $P_\lambda\{(x_1, \cdots, x_{p-1}) \in C, |x_p| \leqq h\}$ is nondecreasing in λ.

The proof depends heavily on an extension of Anderson's [1] inequality for the integral of a symmetric unimodal function. Anderson proves that if C is a convex set, symmetric about the origin, and if $f(x) \geqq 0$ is a real valued function satisfying $f(x) = f(-x)$, $\{x: f(x) \geqq u\} = K_u$ is convex for every u, and $\int_C f(x)\,dx < \infty$, then for any y,

$$u(k) \equiv \int_C f(x + ky)\,dx \tag{2.3}$$

is nonincreasing in k for $0 \leqq k \leqq 1$.

This result is strengthened by Sherman [24] as follows. Let $f(x)$ be a real valued function on R^{p-1}, and $\|f\|_* = \max\{\|f\|_1, \|f\|_\infty\}$. Let $\mathscr{C}$ be the closed (with respect to $\|\cdot\|_*$) convex cone generated by indicator functions of convex symmetric sets in R^{p-1}. Sherman proves that $f \in \mathscr{C}$, $g \in \mathscr{C}$ implies $f * g \in \mathscr{C}$, where $(f * g)(x) \equiv \int f(x - t)g(t)\,dt$. Since $h(kx)$ is nonincreasing in k, $0 \leqq k < \infty$, for $h \in \mathscr{C}$, Anderson's result follows.

REMARK 2.1. It is noted in Gnedenko and Kolmogorov ([10], p. 255), that the convolution of two univariate unimodal distributions need not be unimodal. However, the convolution of two univariate symmetric unimodal distributions is unimodal (Wintner [31]). Sherman's result is a multivariate generalization of Wintner's theorem.

PROOF OF THEOREM 2.1. Since C can be expressed as the decreasing intersection of convex symmetric polyhedrons, it suffices to consider the case where C is such a polyhedron. Furthermore, the matrix Σ_λ has a representation in which

$$\Sigma_{11} = MM', \qquad \Sigma_{21} = \sigma_{pp}^{1/2}(0, \cdots, 0, \lambda\rho)M', \ \rho \geqq 0. \tag{2.4}$$

This, in effect, is the usual transformation to canonical correlations. Because the transformation $(x_1, \cdots, x_{p-1}) \to (x_1, \cdots, x_{p-1})M^{-1}$, $x_p \to x_p/\sigma_{pp}^{1/2}$ leaves the hypotheses of the theorem unchanged, we can assume that

$$\Sigma_\lambda = \begin{pmatrix} I & \Lambda' \\ \Lambda & 1 \end{pmatrix}, \qquad \Lambda = (0, \cdots, 0, \lambda). \tag{2.5}$$

Also, $\Sigma_\lambda = T_\lambda T'_\lambda$, where

$$T_\lambda = \begin{pmatrix} I & 0 \\ \Lambda & (1 - \lambda^2)^{1/2} \end{pmatrix}. \tag{2.6}$$

Let $x = yT'_\lambda$, so that y has density $f(yy')$. Partition y as

$$y = (\dot{y}, y_{p-1}, y_p); \tag{2.7}$$

then

$$\begin{aligned} P_\lambda\{(x_1, \cdots, x_{p-1}) \in C, |x_p| \leqq h\} \\ &= P\{(\dot{y}, y_{p-1}) \in C, |\lambda y_{p-1} + (1 - \lambda^2)^{1/2} y_p| \leqq h\} \\ &= E[P\{(\dot{y}, y_{p-1}) \in C, |\lambda y_{p-1} + (1 - \lambda^2)^{1/2} y_p| \leqq h \mid \|y\|\}]. \end{aligned} \tag{2.8}$$

Alternatively, we can write the conditional probability in terms of indicator functions: for $\|y\|$ fixed, define

$$\xi(\lambda) = \int_S \chi_C(\dot{y}, y_{p-1})\, \chi_{[-h,h]}[\lambda y_{p-1} + (1 - \lambda^2)^{1/2} y_p]\, d\mu(y), \tag{2.9}$$

where S is the surface of the sphere with radius $r = \|y\|$, and μ is the uniform surface measure on S.

Let $f_\varepsilon^{(n)}$ be the approximate identity on R^n given by

$$f_\varepsilon^{(n)}(x) = (2\pi\varepsilon)^{-n/2} \exp\left\{-\frac{1}{2}\frac{\|x\|^2}{\varepsilon}\right\}. \tag{2.10}$$

If $\varphi_\varepsilon = \chi_C * f_\varepsilon^{(p-1)}$, then $\varphi_\varepsilon \in \mathscr{C}$, φ_ε is infinitely differentiable, bounded, has bounded derivatives, and as $\varepsilon \downarrow 0$, $\varphi_\varepsilon(x) \to \chi_C(x)$, unless $x \in \partial C$. Similarly, define

$$\psi_\varepsilon = \chi_{[-h,h]} * f_\varepsilon^{(1)}, \tag{2.11}$$

and let

$$\xi_\varepsilon(\lambda) = \int_S \varphi_\varepsilon(\dot{y}, y_{p-1})\psi_\varepsilon[\lambda y_{p-1} + (1-\lambda^2)^{1/2} y_p]\, d\mu(y). \tag{2.12}$$

By the bounded convergence theorem, $\xi_\varepsilon(\lambda) \to \xi(\lambda)$ as $\varepsilon \to 0$. (Note that $\mu(\partial C) = 0$, since C is polyhedral.)

We now assert that $\xi_\varepsilon(\lambda)$ is nondecreasing in λ, and hence $\xi(\lambda)$ is nondecreasing in λ, thereby completing the proof. To prove this assertion make an orthogonal transformation

$$\dot{y} \to \dot{y},\ (y_{p-1}, y_p) \to (u, v) = (y_{p-1}, y_p)\begin{pmatrix} \lambda & (1-\lambda^2)^{1/2} \\ (1-\lambda^2)^{1/2} & -\lambda \end{pmatrix}, \tag{2.13}$$

so that

$$\xi_\varepsilon(\lambda) = \int_S \varphi_\varepsilon(\dot{y}, \lambda u + (1-\lambda^2)^{1/2} v)\psi_\varepsilon(u)\, d\mu(\dot{y}, u, v). \tag{2.14}$$

With the notation $D_k f(z) \equiv \partial f(z_1, \cdots, z_n)/\partial z_k$, we obtain

$$\begin{aligned} \frac{d\xi_\varepsilon(\lambda)}{d\lambda} &= \int_S D_{p-1}\varphi_\varepsilon(\dot{y}, y_{p-1})\big(u - \lambda v(1-\lambda^2)^{-1/2}\big)\psi_\varepsilon(u)\, d\mu(\dot{y}, u, v) \\ &= \int_S D_{p-1}\varphi_\varepsilon(\dot{y}, y_{p-1})\big(y_p(1-\lambda^2)^{-1/2}\big) \\ &\qquad \psi_\varepsilon\big(\lambda y_{p-1} + (1-\lambda^2)^{1/2} y_p\big)\, d\mu(y). \end{aligned} \tag{2.15}$$

Note that $\psi_\varepsilon\big(\lambda y_{p-1} + (1-\lambda^2)^{1/2} y_p\big)$ is an even function of y and $D_{p-1}\varphi_\varepsilon(\dot{y}, y_{p-1})$ is an odd function of y. We now invoke the divergence theorem as follows. Suppose $x = (x_1, \cdots, x_p)$, $B = \{\|x\| \leqq r\}$, and $g(x)$ is an odd function. Then (under appropriate regularity conditions)

$$\int_S x_i g(x)\, d\mu(x) = \int_B \frac{\partial g(x)}{\partial x_i}\, dx. \tag{2.16}$$

Applying this to (2.15), we obtain

$$\begin{aligned}(2.17)\quad \frac{d\xi_\varepsilon(\lambda)}{d\lambda} &= \int_B D_{p-1}\varphi_\varepsilon(\dot{y}, y_{p-1})\psi'_\varepsilon[\lambda y_{p-1} + (1-\lambda^2)^{1/2}y_p]\,d\dot{y}\,dy_{p-1}\,dy_p \\ &= \int_B D_{p-1}\varphi_\varepsilon[\dot{y}, \lambda u + (1-\lambda^2)^{1/2}v]\psi'_\varepsilon(u)\,d\dot{y}\,du\,dv \\ &= \int_{-1}^{1} \psi'_\varepsilon(u)\left\{\int_{\|\dot{y}\|^2+v^2\leqq 1-u^2} D_{p-1}\varphi_\varepsilon[\dot{y}, \lambda u + (1-\lambda^2)^{1/2}v]\,d\dot{y}\,dv\right\}du \\ &\equiv \int_{-1}^{1} \psi'_\varepsilon(u)\,s(u)\,du.\end{aligned}$$

We now show that $d\xi_\varepsilon(\lambda)/d\lambda \geqq 0$ by showing that for each fixed u, $\psi'_\varepsilon(u)s(u) \geqq 0$, which completes the proof.

As a consequence of Wintner's result [31] $\psi_\varepsilon(u)$ is nonincreasing in $|u|$. Hence,

$$(2.18)\qquad \psi'_\varepsilon(u) \geqq 0 \text{ if } u \leqq 0, \qquad \psi'_\varepsilon(u) \leqq 0 \text{ if } u \geqq 0.$$

Next, let $w = (1-\lambda^2)^{1/2}v$, so that

$$(2.19)\qquad s(u) = \frac{1}{(1-\lambda^2)^{1/2}}\int_\Omega D_{p-1}\varphi_\varepsilon(\dot{y}, w + \lambda u)\,d\dot{y}\,dw,$$

where $\Omega = \{(\dot{y}, w)\colon \|\dot{y}\|^2 + w^2/(1-\lambda^2) \leqq 1 - u^2\}$. Note that Ω is a convex symmetric set. From the Anderson-Sherman result, $\chi_\Omega * \varphi_\varepsilon \in \mathscr{C}$; therefore,

$$(2.20)\qquad \int_\Omega \varphi_\varepsilon(\dot{y}, w + ku)\,d\dot{y}\,dw$$

is nonincreasing in k, $0 \leqq k$, so that $us(u) \leqq 0$. But u and $\psi'_\varepsilon(u)$ also have opposite signs, and hence, $\psi'_\varepsilon(u)$ and $s(u)$ have the same sign. *Q.E.D.*

REMARK 2.2. This theorem is evidently an extension of the theorem by Šidák mentioned earlier. We have found the theorem to be very rich in implications. Just why that is cannot be said in a word, but numerous illustrations will occur throughout the paper.

A more geometric, and less algebraic, expression of the theorem may be helpful. Let C^* be a cylinder, that is, the Cartesian product of a $p-1$ dimensional symmetric convex set C and a one dimensional linear subspace orthogonal to it. Consider a vector v confined to a fixed circle with center at the origin and passing through the axis of C^*, and consider a slab S_v of thickness $2h$ with its plane of symmetry passing through the origin and orthogonal to v. As the angle between v and the axis of C^* increases from 0 to $\pi/2$, the probability of the intersection between C^* and S_v is nondecreasing. That is precisely equivalent to Theorem 2.1. Rather evidently, none of its content would be lost if the spherically symmetric measure were taken to be the uniform measure on the surface of the unit sphere. If the reader will try to visualize the geometric form of the theorem for that distribution and for $p = 3$, he will see that it is plausible though not obvious.

In the bivariate case there is a very elementary geometrical proof. Unfortunately, it is not clear how to extend it to higher dimensions; but because the proof is so suggestive we give it for the simplest case $P_\Sigma\{|x_1| \leqq h, |x_2| \leqq h\} \geqq P_{\Sigma_\lambda}\{|x_1| \leqq h, |x_2| \leqq h\}$, where

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \qquad \Sigma_\lambda = \begin{pmatrix} 1 & \lambda\rho \\ \lambda\rho & 1 \end{pmatrix}, \qquad 0 \leqq \lambda \leqq 1. \tag{2.21}$$

Although the proof can be extended to the general bivariate case, where $h_1 \neq h_2$, the essential ideas are exhibited here.

Making the two transformations $(x_1, x_2) = [y_1, y_1\rho + y_2(1 - \rho^2)^{1/2}]$ and $(x_1, x_2) = (y_1, y_1\lambda\rho + y_2(1 - \lambda^2\rho^2)^{1/2})$, we wish to show that $\int f(y_1^2 + y_2^2)\, dy_1 dy_2$ taken over the set $\{|y_1| \leqq h; |y_1\rho + y_2(1 - \rho^2)^{1/2}| \leqq h\}$ is larger than or equal to the same integral taken over the set $\{|y_1| \leqq h; |y_1\lambda\rho + y_2(1 - \lambda^2\rho^2)^{1/2}| \leqq h\}$. We can assume without loss of generality that $\rho \geqq 0$, for if $\rho < 0$, we can replace (x_1, x_2) by $(x_1, -x_2)$. Draw a circle with center at the origin with radius h, draw the tangent lines $|y_1| = h$, $|y_1\rho + y_2(1 - \rho^2)^{1/2}| = h$, $|y_1\lambda\rho + y_2(1 - \lambda^2\rho^2)^{1/2}| = h$, and draw a circle through the point a as in Figure 1.

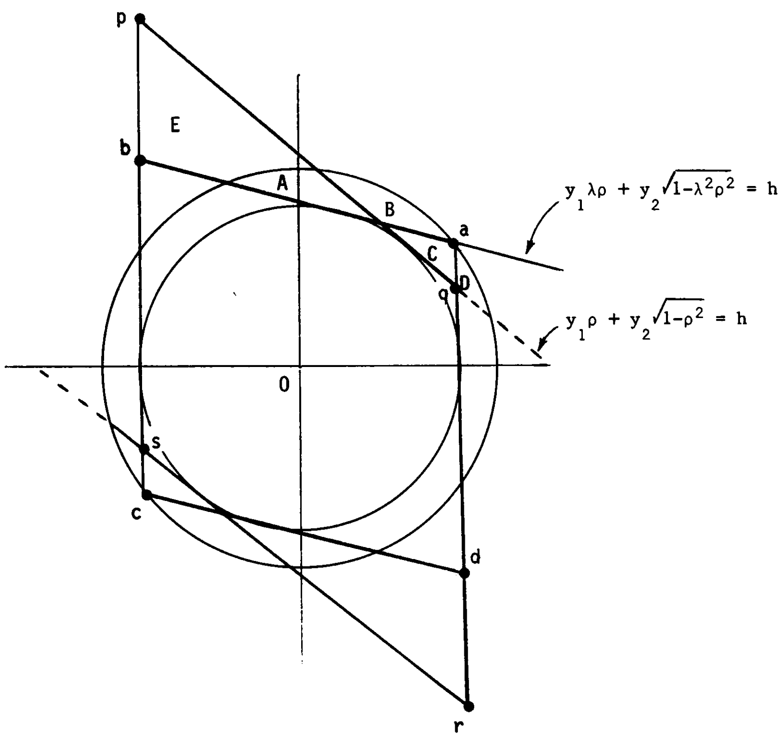

FIGURE 1

Because the tangent lines are equidistant from the origin we have by circular symmetry that $P\{A \cup B\} = P\{B \cup C \cup D\}$. Consequently, there is at least as much probability content in the parallelogram $(pqrs)$ as in the parallelogram $(abcd)$, which completes the proof. In higher dimensions this geometric argument would require consideration of many different cases, and this aspect becomes extremely complicated.

3. Related results and extensions

In this section we discuss the interrelations among results in this area (Section 3.1); a number of complementary and reversal inequalities are given in Section 3.2 and Section 3.3. The comparison of integrals from elliptically contoured distributions with covariance matrices Σ_1 and Σ_2 when $\Sigma_1 - \Sigma_2$ is positive semi-definite is provided in Section 3.4. Some counterexamples are given in Section 3.5.

3.1. *Some extensions and the normal case.* First notice that Theorem 2.1 has the following immediate corollary (Σ is a fixed positive definite matrix).

COROLLARY 3.1. *If $x = (x_1, \cdots, x_p)$ is a random vector with density function $|\Sigma_\lambda|^{-1/2} f(x\Sigma_\lambda^{-1}x')$, where $\lambda = (\lambda_1, \cdots, \lambda_p)$, $\Sigma_\lambda = (\sigma_{ij}(\lambda))$, $\sigma_{ii}(\lambda) = \sigma_{ii}$, $\sigma_{ij}(\lambda) = \lambda_i\lambda_j\sigma_{ij}$, $i \neq j$, $1 \leqq ij \leqq p$, then*

$$(3.1) \qquad P_{\lambda_1, \ldots, \lambda_p}\{|x_1| \leqq h_1, \cdots, |x_p| \leqq h_p\}$$

is nondecreasing in each λ_i, $0 \leqq \lambda_i \leqq 1$. In particular, $P(\Sigma) \geqq P(D_\sigma)$, where $D_\sigma = \operatorname{diag}(\sigma_{11}, \cdots, \sigma_{pp})$.

The inequality $P(\Sigma) \geqq P(D_\sigma)$ was proved by Dunn [5], [6] for the bivariate and trivariate normal distributions, and for general p when Σ has the special structure $\Sigma = D_\tau + \alpha'\alpha$, with $D_\tau = \operatorname{diag}(\tau_1, \cdots, \tau_p)$ and $\alpha = (\alpha_1, \cdots, \alpha_p)$. In the case of a normal distribution, Corollary 3.1 was proved by Šidák [26] and by Jogdeo [13]. Both proofs make use of Anderson's theorem [1]—Šidák uses a conditional argument and Jogdeo uses the differential identity (1.4) to obtain a short proof. A close examination of Jogdeo's argument shows in fact that in the case of the normal distribution, Anderson's theorem and Theorem 2.1 are equivalent. More precisely, the following two assertions concerning a symmetric set $C \subset R^{p-1}$ are equivalent.

ASSERTION 3.1. *For each $a = (a_1, \cdots, a_{p-1})$ and each $p - 1 \times p - 1$ positive definite matrix Ψ,*

$$(3.2) \qquad u(k) \equiv \int_C |\Psi|^{-1/2} \exp\{-\tfrac{1}{2}(y - ka)\Psi^{-1}(y - ka)'\}\, dy_1 \cdots dy_{p-1}$$

is a nonincreasing function of k, $0 \leqq k$.

ASSERTION 3.2. *For each $p \times p$ positive definite matrix Σ and each $\delta > 0$,*

$$(3.3) \qquad v(\lambda) = \int_{-\delta}^{\delta} \left[\int_C |\Sigma_\lambda|^{-1/2} \exp\{-\tfrac{1}{2}x\, \Sigma_\lambda^{-1}\, x'\}\, dx_1 \cdots dx_{p-1}\right] dx_p$$

is a nondecreasing function of λ, $0 \leqq \lambda \leqq 1$, where Σ_λ is defined in Theorem 2.1.

The equivalence of these assertions is based on the identity (1.4):

$$(3.4)\qquad \begin{aligned} v'(\lambda) &= \int_{-\delta}^{\delta}\int_C \frac{d}{d\lambda}\varphi(x, \Sigma_\lambda)\,dx_1\cdots dx_p \\ &= \int_{-\delta}^{\delta}\int_C \sum_{i=1}^{p-1}\sigma_{ip}\frac{\partial\varphi(x,\Sigma_\lambda)}{\partial\sigma_{ip}}\,dx_1\cdots dx_p \\ &= \int_{-\delta}^{\delta}\int_C \sum_{i=1}^{p-1}\sigma_{ip}\frac{\partial^2\varphi(x,\Sigma_\lambda)}{\partial x_i\partial x_p}\,dx_1\cdots dx_p \\ &= 2\int_C \sum_{i=1}^{p-1}\sigma_{ip}\frac{\partial\varphi(x^*,\Sigma_\lambda)}{\partial x_i}\,dx_1\cdots dx_{p-1}, \end{aligned}$$

where $x^* = (x_1, \cdots, x_{p-1}, \delta)$. By completing the square with $\Psi = \Sigma_{11} - \lambda^2\,\Sigma_{12}\,\Sigma_{21}/\sigma_{pp}$, $b = \sigma_{pp}^{-1/2}\exp\{-\frac{1}{2}\delta^2/\sigma_{pp}\}$, $a = \delta\Sigma_{21}/\sigma_{pp}$, $\dot{x} = (x_1, \cdots, x_{p-1})$, we obtain

$$(3.5)\qquad \begin{aligned} v'(\lambda) &= 2b\int_C \sum_{i=1}^{p-1}\sigma_{ip}\frac{\partial}{\partial x_i}\left[|\Psi|^{-1/2}\exp\{-\tfrac{1}{2}(\dot{x}-\lambda a)\Psi^{-1}(\dot{x}-\lambda a)'\right]\prod_1^{p-1}dx_j \\ &= \frac{-2b\sigma_{pp}}{\delta}\frac{d}{dk}\left\{\int_C |\Psi|^{-1/2}\exp\{-\tfrac{1}{2}(\dot{x}-ka)\Psi^{-1}(\dot{x}-ka)'\}\prod_1^{p-1}dx_j\right\}_{k=\lambda} \\ &= -2b\sigma_{pp}\frac{u'(\lambda)}{\delta}. \end{aligned}$$

Thus, $v'(\lambda) \geqq 0$ if and only if $u'(\lambda) \leqq 0$.

In Theorem 2.1 the vector x is partitioned into two parts of dimension $p - 1$ and 1, respectively. This suggests the conjecture: if $x = (\dot{x}, \ddot{x})$ is a random vector (with $\dot{x}$ of dimension p and $\ddot{x}$ of dimension q) having density $|\Sigma|^{-1/2} f(x\,\Sigma^{-1}\,x')$, and if $C_1 \subset R^p$ and $C_2 \subset R^q$ are convex symmetric sets, then

$$(3.6)\qquad P_\Sigma\{\dot{x}\in C_1, \ddot{x}\in C_2\} \geqq P_{\bar{\Sigma}}\{\dot{x}\in C_1, \ddot{x}\in C_2\},$$

where

$$(3.7)\qquad \Sigma = \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix},\qquad \bar{\Sigma} = \begin{pmatrix}\Sigma_{11} & 0\\ 0 & \Sigma_{22}\end{pmatrix}.$$

For the multivariate normal distribution—in which case the right side of (3.6) is simply $P_{\Sigma_{11}}\{\dot{x}\in C_1\}P_{\Sigma_{22}}\{\ddot{x}\in C_2\}$—Scott [23] used a conditioning argument in an ostensible proof of a simple case of (3.6). This argument was adopted by Das Gupta [4] and Khatri [16] in attempts to prove (3.6) in the general (normal) case. (Khatri's conclusion is reported and applied by Jensen [12], Corollary 1, p. 145.) Unfortunately, the conditioning argument contains a flaw. Implicit use is made of the "fact" that the conditional distribution of $y \sim N(0, I)$, given that y lies in an arbitrary subspace containing certain coordinate axes, is again multivariate normal. However, this is not the case; the condition that y lies in an arbitrary subspace containing certain axes is a condition on *angles*, and such conditional distributions are not normal. For example, if $(x, y) \sim N_2(0, I)$ has modified "polar" coordinates (R, Θ) where $-\infty < R < \infty$ and $\Theta \in [0, \pi)$, then the conditional distribution of R given Θ is not normally distributed, but has a "double χ^2" distribution. Thus (3.6) remains unsolved even for normally distributed variates.

Certain special cases of (3.6) can be obtained when Σ has a particular structure. Let

$$\Sigma(\lambda) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12}D_\lambda \\ D_\lambda \Sigma_{21} & D_\tau \end{pmatrix}, \tag{3.8}$$

where $D_\lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_q)$, $D_\tau = \operatorname{diag}(\tau_1, \cdots, \tau_q)$, and let $x = (\dot{x}, \ddot{x})$, where $\dot{x}$ is p dimensional and $\ddot{x}$ is q dimensional. If x has a density function $|\Sigma(\lambda)|^{-1/2} f(x\Sigma(\lambda)^{-1}x')$ and C is a convex, symmetric set in R^p, then as a consequence of Theorem 2.1

$$P_{\Sigma(\lambda)}\{\dot{x} \in C, |\ddot{x}_1| \leqq h_1, \cdots, |\ddot{x}_q| \leqq h_q\} \tag{3.9}$$

is nondecreasing in each λ_i, $0 \leqq \lambda_i \leqq 1$.

For the normal distribution Khatri [15] shows that (3.6) holds when $\Sigma_{12} = \operatorname{Cov}(\dot{x}, \ddot{x})$ has rank 1. Actually, Khatri's proof can be extended to the case where the mean of x is not zero (but is restricted).

THEOREM 3.1. *Let $x = (\dot{x}, \ddot{x})$ have the multivariate normal distribution*

$$N_{p+q}\left[(\mu_1, \mu_2), \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right], \tag{3.10}$$

where rank $(\Sigma_{12}) = 1$. Suppose there exists a scalar η such that $\eta_2\Sigma_{11} - \mu_1'\mu_1$ and $\eta^2\Sigma_{22} - \mu_2'\mu_2$ are positive semidefinite and $\eta^2\Sigma_{12} = \mu_1'\mu_2$. Then (3.6) *holds for any two convex symmetric sets $C_1 \subset R^p$ and $C_2 \subset R^q$.*

PROOF. Let $\alpha = \mu_1/\eta$, $\beta = \mu_2/\eta$ (if $\eta = 0$ choose any α, β such that $\alpha'\beta = \Sigma_{12}$) and choose the random variable z so that $(\dot{x}, \ddot{x}, z)$ has a $(p + q + 1)$ variate normal distribution with mean (μ_1, μ_2, η) and covariance matrix

$$\hat{\Sigma} = \begin{pmatrix} \Sigma_{11} & \alpha'\beta & \alpha' \\ \beta'\alpha & \Sigma_{22} & \beta' \\ \alpha & \beta & 1 \end{pmatrix}. \tag{3.11}$$

(The covariance matrix $\hat{\Sigma}$ is positive semidefinite by the hypotheses of the theorem.) The conditional distribution of $(\dot{x}, \ddot{x})$ given z is

$$N_{p+q}\left[(z\alpha, z\beta), \begin{pmatrix} \Sigma_{11} - \alpha'\alpha & 0 \\ 0 & \Sigma_{22} - \beta'\beta \end{pmatrix}\right], \tag{3.12}$$

so that

$$P\{\dot{x} \in C_1, \ddot{x} \in C_2 \mid z\} = P\{\dot{x} \in C_1 \mid z\} P\{\ddot{x} \in C_2 \mid z\}. \tag{3.13}$$

It follows from Anderson's theorem that both factors on the right of (3.13) are decreasing functions of $|z|$. Hence, they are similarly ordered, so integration of (3.13) with respect to z yields (3.6). *Q.E.D.*

In the bivariate case ($p = q = 1$) the hypotheses of Theorem 3.1 can be stated in simpler form, as follows. Let $x = (\dot{x}, \ddot{x})$ have the bivariate normal distribution

$$N_2\left[(\mu_1, \mu_2), \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right], \quad \rho \neq 0, \tag{3.14}$$

where $0 < \rho \leqq \mu_1/\mu_2 \leqq 1/\rho$ or $1/\rho \leqq \mu_1/\mu_2 \leqq \rho < 0$. Then for $h_1 > 0, h_2 > 0$,

$$P\{|\dot{x}| \leqq h_1, |\ddot{x}| \leqq h_2\} \geqq P\{|\dot{x}| \leqq h_1\}P\{|\ddot{x}| \leqq h_2\}. \tag{3.15}$$

REMARK 3.1. Theorem 3.1 appears to be the first such inequality for the case of nonzero means.

The right side of (3.6) suggests a comparison in the general elliptically contoured case between

$$P_\Sigma\{\dot{x} \in C_1, \ddot{x} \in C_2\} \text{ and } P_{\Sigma_{11}}\{\dot{x} \in C_1\}P_{\Sigma_{22}}\{\ddot{x} \in C_2\}. \tag{3.16}$$

Of course, when the distribution is normal these two expressions are equal. However, we see that in general no inequality exists even in the simple bivariate case. To see this, consider the difference

$$P_I\{\dot{x} \in C_1, \ddot{x} \in C_2\} - P_I\{\dot{x} \in C_1\}P_I\{\ddot{x} \in C_2\}. \tag{3.17}$$

When $p = q = 1$, consider the following figures. In Figure 2, put unit mass

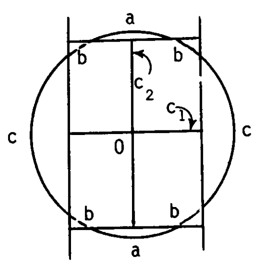

FIGURE 2

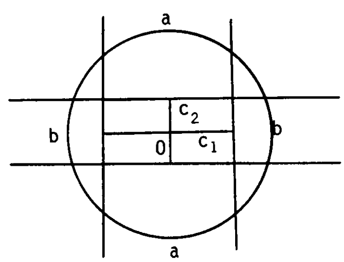

FIGURE 3

uniformly on the circle and let the arc lengths a, b, c represent the mass so that $2a + 4b + 2c = 1$. Then

$$P\{|x_1| \leqq c_1\} = 2(a + 2b), \qquad P\{|x_2| \leqq c_2\} = 2(c + 2b), \qquad P\{|x_1| \leqq c_1, |x_2| \leqq c_2\} = 4b, \tag{3.18}$$

from which we find that the difference (3.17) is negative. In Figure 3 put mass $\frac{1}{2}$ at the origin and mass $\frac{1}{2}$ uniformly on the circle. Then

$$P\{|x_1| \leqq c_1\} = \tfrac{1}{2} + 2a, \qquad P\{|x_2| \leqq c_2\} = \tfrac{1}{2} + 2b, \qquad P\{|x_1| \leqq c_1, |x_2| \leqq c_2\} = \tfrac{1}{2}. \tag{3.19}$$

For a and b sufficiently small, the difference (3.17) is positive. Although densities do not exist in these two examples, they may be approximated by distributions having densities.

The inequalities discussed thus far involve a Cartesian product of symmetric convex sets. We now drop the assumption of Cartesian product, and consider possible inequalities between $P_R\{x \in C\}$ and $P_I\{x \in C\}$, where R is a correlation matrix and C is a symmetric convex set in R^p. Even in a simple case where $C = \{x: \Sigma x_i^2 \leqq h^2\}$ and $x \sim N(0, I)$ there is no such inequality. Indeed $E_R(\Sigma x_i^2) = E_I(\Sigma x_i^2) = p$, so that we cannot have the same ordering between $P_R\{\Sigma x_i^2 \leqq h^2\}$ and $P_I\{\Sigma x_i^2 \leqq h^2\}$ for all h, unless $R = I$. When $p = 2$ and $\rho = 1$, the comparison is one between a random variable having a χ_2^2 distribution and a random variable having a $2\chi_1^2$ distribution.

If R_1 and R_2 are correlation matrices, the difference $R_1 - R_2$ is never positive semidefinite. However, if Σ_1 and Σ_2 are covariance matrices, then $\Sigma_1 - \Sigma_2$ may be positive definite. When this happens, we do obtain certain inequalities. (See Section 3.4.)

3.2. *Complementary inequalities.* When x has an elliptically contoured distribution the monotonicity property of

$$P_{\Sigma_\lambda}\{(x_1, \cdots, x_{p-1}) \in C, |x_p| \leqq c_p\}, \tag{3.20}$$

as given in Theorem 2.1 can be used to obtain a monotonicity property for

$$P_{\Sigma_\lambda}\{(x_1, \cdots, x_{p-1}) \notin C, |x_p| \geqq c_p\}. \tag{3.21}$$

To see this note that

$$P\{AB\} - P\{A\}P\{B\} = P\{A^cB^c\} - P\{A^c\}P\{B^c\}, \tag{3.22}$$

where A^c is the complement of A. Furthermore, if $x = (\dot{x}, \ddot{x})$ has a density $|\Sigma|^{-1/2}f(x\Sigma^{-1}x')$, then $\dot{x}$ has a density $|\Sigma_{11}|^{-1/2}g(\dot{x}\Sigma_{11}^{-1}\dot{x}')$, (see Section 6). Consequently, if

$$\Sigma_\lambda = \begin{pmatrix} \Sigma_{11} & \lambda\Sigma_{12} \\ \lambda\Sigma_{21} & \sigma_{pp} \end{pmatrix}, \tag{3.23}$$

and

$$A = x: (x_1, \cdots, x_{p-1}) \in C\}, B = \{x_p: |x_p| \leqq c_p\}, \tag{3.24}$$

then $P\{A\}$ and $P\{B\}$ are independent of λ, from which we obtain that $P\{AB\}$ and $P\{A^cB^c\}$ have the same monotonicity properties.

Turning to the normal case, (3.22) implies that if $x = (x_1, x_2)$ has a bivariate normal distribution $N(0, \Sigma)$, then

$$P_\Sigma\{|x_1| \geqq c_1, |x_2| \geqq c_2\} \geqq P_{\sigma_{11}}\{|x_1| \geqq c_1\}P_{\sigma_{22}}\{|x_2| \geqq c_2\}. \tag{3.25}$$

The extension to more than two dimensions is in general false (notwithstanding the similar appearance of Corollary 3.1), as noted by Šidák [27]. However, when $x \sim N(0, \Sigma)$ with $\Sigma = D_\tau + \alpha'\alpha$, $D_\tau = \text{diag}(\tau_1, \cdots, \tau_p)$, $\alpha = (\alpha_1, \cdots, \alpha_p)$, Khatri [15] shows that

$$(3.26)\qquad P_\Sigma\{|x_1| \geqq c_1, \cdots, |x_p| \geqq c_p\} \geqq \prod_1^p P_{\sigma_{ii}}\{|x_i| \geqq c_i\}.$$

Also by (3.22), if Σ_{12} is of rank 1, then for C_1, C_2 convex and symmetric,

$$(3.27)\qquad P_\Sigma\{\dot{x} \notin C_1, \ddot{x} \notin C_2\} \geqq P_{\Sigma_{11}}\{\dot{x} \notin C_1\}P_{\Sigma_{22}}\{\ddot{x} \notin C_2\}.$$

3.3. *Reversal inequalities.* By a reversal inequality we mean a bound in the opposite direction. In the general elliptically contoured case, reversal inequalities cannot be obtained. However, if the density has an additional monotonicity property, a reversal inequality can be found for the inequality $P(\Sigma) \geqq P(D_\sigma)$ (as in Corollary 3.1).

THEOREM 3.2. *Let x be a p dimensional random vector with density function $|R|^{-1/2}f(xR^{-1}x')$, where R is a correlation matrix. If f is monotone decreasing on $[0, \infty)$, then*

$$(3.28)\qquad P_R\{|x_1| \leqq c_1, \cdots, |x_p| \leqq c_p\} \leqq P_I\left\{|x_1| \leqq \frac{c_1}{\gamma_1}, \cdots, |x_p| \leqq \frac{c_p}{\gamma_p}\right\},$$

where $\gamma_k^2 = \det R_{(k)}/\det R_{(k-1)}$, and $R_{(k)} = (r_{ij})$, $1 \leqq i, j \leqq k$.

PROOF. Let $R = TT'$, where T is a lower triangular matrix and $y = xT'^{-1}$. Then y has a density $f(yy')$ and

$$(3.29)\qquad P_R\{|x_i| \leqq c_i, i = 1, \cdots, p\} = P_I\left\{\left|\sum_{j=1}^i t_{ij}y_j\right| \leqq c_i, i = 1, \cdots, p\right\},$$

which we may write as

$$(3.30)\qquad \int_{\Omega(\dot{y})} \prod_1^{p-1} dy_i \int_H f(y_1^2 + \cdots + y_p^2)\, dy_p,$$

where $\Omega(\dot{y}) = \{\dot{y} = (y_1, \cdots, y_{p-1}) : |\Sigma_{j=1}^{p-1} t_{ij}y_j| \leqq c_i, i = 1, \cdots, p-1\}$ is independent of y_p, and where $H = \{|t_{p1}, y_1 + \cdots + t_{pp}y_p| \leqq c_p\}$. If f is monotone nonincreasing, then as a consequence of Wintner's theorem [31]

$$(3.31)\qquad \int_H f(y_1^2 + \cdots + y_p^2)\, dy_p \leqq \int_{|t_{pp}y_p| \leqq c_p} f(y_1^2 + \cdots + y_p^2)\, dy_p.$$

By a repetitive argument,

$$(3.32)\qquad P_I\left\{\left|\sum_{j=1}^p t_{ij}y_j\right| \leqq c_i, i = 1, \cdots, p\right\} \leqq P_I\{|t_{ii}y_i| \leqq c_i, i = 1, \cdots, p\}.$$

The argument is completed by noting that $\det R_{(k)} = \Pi_1^k t_{ii}^2$, so that $t_{kk}^2 = \det R_{(k)}/\det R_{(k-1)}$. *Q.E.D.*

3.4. *Comparisons when one covariance matrix dominates another.* The following result was first proved by Anderson ([1], Corollary 3) in the multivariate normal case.

THEOREM 3.3. *Let the random vector y have a density* $|\Gamma|^{-1/2}g(y\Gamma^{-1}y')$. *If C is a convex, symmetric set and* $\Gamma_2 - \Gamma_1$ *is positive semidefinite, then*

$$P_{\Gamma=\Gamma_1}\{y \in C\} \geqq P_{\Gamma=\Gamma_2}\{y \in C\}. \tag{3.33}$$

For a general elliptically contoured distribution, Theorem 3.3 follows from Theorem 2.1 and

THEOREM 3.4. *Consider the two statements concerning a Lebesgue set* $C \subset R^{p-1}$:

(a) *For the p dimensional random vector* $x = (\dot{x}, x_p)$ *with density* $|\Sigma|^{-1/2}f(x\Sigma^{-1}x')$,

$$P_\Sigma\{\dot{x} \in C, |x_p| \leqq h\} \geqq P_{\bar{\Sigma}}\{\dot{x} \in C, |x_p| \leqq h\}, \tag{3.34}$$

where

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \sigma_{pp} \end{pmatrix}, \qquad \bar{\Sigma} = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \sigma_{pp} \end{pmatrix}, \qquad \Sigma_{11}: p - 1 \times p - 1. \tag{3.35}$$

(b) *For the* $(p - 1)$ *dimensional random vector y with density* $|\Gamma|^{-1/2}g(y\Gamma^{-1}y')$, (3.33) *holds.*

If the set C is such that (a) *holds for all f, all positive definite matrices* Σ, *and all sufficiently small h, then C is such that* (b) *holds for all g,* Γ_1 *and* Γ_2 *with* $\Gamma_2 - \Gamma_1$ *positive semidefinite.*

PROOF. Let g, Γ_1 and Γ_2 be given with $\int_0^\infty r^{p-2}g(r^2)\,dr < \infty$. There exists a smooth function γ with compact support such that

$$\int_0^\infty r^{p-2}|g(r^2) - \gamma(r^2)|\,dr \tag{3.36}$$

can be made arbitrarily small, so that

$$\int |\Gamma|^{-1/2}|g(y\Gamma^{-1}y') - \gamma(y\Gamma^{-1}y')|\,dy \tag{3.37}$$

can be made arbitrarily small simultaneously for all Γ. Thus, it suffices to assume that g is smooth with compact support. Since $\Gamma_2 - \Gamma_1$ is positive semidefinite, it may be written as the sum of positive semidefinite matrices of rank 1. Consequently, we may assume that $\Gamma_2 = \Gamma_1 + \alpha'\alpha$, where $\alpha = (\alpha_1, \cdots, \alpha_{p-1})$, and proceed inductively. Let

$$\Sigma = \begin{pmatrix} \Gamma_2 & \alpha' \\ \alpha & 1 \end{pmatrix}, \tag{3.38}$$

and choose $c > 0$ such that $c|\Sigma|^{-1/2}g(x\Sigma^{-1}x')$ is a density in R^p. That is, let $x = (\dot{x}, x_p)$ have density $c|\Sigma|^{-1/2}g(x\Sigma^{-1}x')$. By (a), with $f = cg$,

$$P_\Sigma\{\dot{x} \in C, |x_p| \leqq h\} \geqq P_{\bar{\Sigma}}\{\dot{x} \in C, |x_p| \leqq h\}. \tag{3.39}$$

Note that the marginal distribution of x_p is the same under Σ and $\bar{\Sigma}$, (see Section 6) and that $P\{|x_p| \leqq h\} > 0$. Dividing both sides of (3.39) by $P\{|x_p| \leqq h\}$ and letting $h \to 0$, we obtain

$$P_{\Sigma}\{\dot{x} \in C \mid x_p = 0\} \geqq P_{\bar{\Sigma}}\{\dot{x} \in C \mid x_p = 0\}. \tag{3.40}$$

But the conditional density of $\dot{x}$ given $x_p = 0$ under $\Sigma(\bar{\Sigma})$ is $|\Gamma_1|^{-1/2}g(\dot{x}\Gamma_1^{-1}\dot{x}')$ $[|\Gamma_2|^{-1/2}g(\dot{x}\Gamma_2^{-1}\dot{x}')]$. Hence

$$\int_C |\Gamma_1|^{-1/2}g(\dot{x}\Gamma_1^{-1}\dot{x}')\,d\dot{x} \geqq \int_C |\Gamma_2|^{-1/2}g(\dot{x}\Gamma_2^{-1}\dot{x}')\,d\dot{x}, \tag{3.41}$$

which proves (3.33). *Q.E.D.*

An alternative proof of Theorem 3.4 in a more general context is given by Fefferman, Jodeit and Perlman [9].

If the set C in Theorem 3.3 is an ellipsoid centered at the origin, the result can be proved by a direct inclusion argument. There exists a nonsingular matrix W and a diagonal matrix $D = \operatorname{diag}(d_1, \cdots, d_{p-1})$, with $0 < d_i \leqq 1$, $i = 1, \cdots, p-1$, such that $\Gamma_2 = WD^{-1}W'$, $\Gamma_1 = WW'$. If we let the random vector z have density $g(zz')$, then (3.33) becomes

$$P_I\{z \in K\} \geqq P_I\{z \in DK\}, \tag{3.42}$$

where $K = W^{-1}C$ is again an ellipsoid centered at the origin. Since the distribution of z is invariant under orthogonal transformations L, in order to prove (3.39) it suffices to produce an L such that

$$LDK \subseteqq K. \tag{3.43}$$

Now $K = \{x : xQx' \leqq \alpha\}$ and $DK = \{x : xD^{-1}QD^{-1}x' \leqq \alpha\}$, for some positive definite Q and $\alpha > 0$. The inclusion (3.43) holds if for some orthogonal L,

$$x(LD^{-1}QD^{-1}L' - Q)x' \geqq 0. \tag{3.44}$$

Let $\lambda_j(A)$ denote the jth largest characteristic root of the matrix A. From

$$\lambda_j(Q) = \lambda_j(DD^{-1}QD^{-1}D) \leqq \lambda_1(D^2)\lambda_j(D^{-1}QD^{-1}) \leqq \lambda_j(D^{-1}QD^{-1}), \tag{3.45}$$

it follows that there exists an orthogonal matrix L such that $LD^{-1}QD^{-1}L' - Q$ is positive semidefinite, which proves (3.43).

REMARK 3.2. This inclusion argument fails in three or more dimensions if C is an arbitrary convex symmetric set.

As an immediate consequence of Theorem 3.3 for the case of an ellipsoid C, note that if y has density $|\Gamma|^{-1/2}g(y\Gamma^{-1}y')$, if S is a random positive definite matrix independent of y, and if $\Gamma_2 - \Gamma_1$ is positive semidefinite, then for any $k > 0$

$$P_{\Gamma=\Gamma_1}\{yS^{-1}y' \leqq k\} \geqq P_{\Gamma=\Gamma_2}\{yS^{-1}y' \leqq k\}. \tag{3.46}$$

This result is of particular interest when y has a multivariate normal distribution and S has a Wishart distribution, since the statistic $yS^{-1}y'$ is proportional to the T^2 statistic.

3.5. *Some counterexamples.* The following counterexamples show that the results already given do not remain valid if C is taken to be the union of two convex symmetric sets (such a C is not convex but is star shaped with respect to the origin). This holds even in the normal case.

In two dimensions, choose $0 < a$ and define $C_1 = \{y_1 : |y_1| \leqq a\}$, $C_2 = \{y_2 : |y_2| \leqq a\}$, $C = C_1 \cup C_2$. Let $y = (y_1, y_2) \sim N(0, \Gamma(d))$ with

$$\Gamma(d) = \begin{pmatrix} 1 & 1 \\ 1 & 1 + d \end{pmatrix}. \tag{3.47}$$

Then for $d > 0$

$$P_{\Gamma(0)}\{y \in C\} < P_{\Gamma(d)}\{y \in C\}, \tag{3.48}$$

so that the conclusion of Theorem 3.3 is no longer valid.

By Theorem 3.4 it then follows that when (x_1, x_2, x_3) has a trivariate normal distribution with mean zero, it can happen that

$$P\{(x_1, x_2) \in C, |x_3| \leqq h\} < P\{(x_1, x_2) \in C\}P\{|x_3| < h\}, \tag{3.49}$$

so that the conclusion of Theorem 2.1 is no longer valid. Finally, from the equivalence of Assertions 3.1 and 3.2 in Section 3.1, it follows that the conclusion of Anderson's theorem is not necessarily valid even for a normal distribution if C is the union of two convex symmetric sets.

4. Statistical applications

The study of simultaneous confidence regions has prompted the development of inequalities for sample variances and for Studentized variates. We now present a number of inequalities which are obtained from our previous results.

Several results are consequences of the following theorem. Let $\Sigma^{(k)} = (\sigma_{ij}^{(k)})$, $1 \leqq k \leqq n$, be positive definite, $\lambda_{(k)} = (\lambda_1^{(k)}, \cdots, \lambda_p^{(k)})$, $0 \leqq \lambda_j^{(k)} \leqq 1$, and $\lambda = (\lambda_{(1)}, \cdots, \lambda_{(n)})$. Define

$$\Sigma_{\lambda(k)}^{(k)} = (\sigma_{ij}^{(k)}(\lambda_{(k)})), \tag{4.1}$$

where $\sigma_{ii}^{(k)}(\lambda_{(k)}) = \sigma_{ii}^{(k)}$, $\sigma_{ij}^{(k)}(\lambda_{(k)}) = \lambda_i^{(k)}\lambda_j^{(k)}\sigma_{ij}^{(k)}$, $i \neq j$, (see Corollary 3.1).

THEOREM 4.1. *If the columns of the $p \times n$ matrix $X = (x_1', \cdots, x_n')$ are independently distributed, where the kth column has the density function*

$$|\Sigma_{\lambda(k)}^{(k)}|^{-1/2} f_k(x(\Sigma_{\lambda(k)}^{(k)})^{-1}x'), \tag{4.2}$$

and $t > 0$, then

$$P_{\lambda}\left\{\sum_{\alpha=1}^{n} |x_{i\alpha}|^t \leqq a_i, i = 1, \cdots, p\right\} \tag{4.3}$$

is nondecreasing in each $\lambda_j^{(k)}$, $0 \leqq \lambda_j^{(k)} \leqq 1$.

PROOF. By conditioning on $x_2, \cdots, x_n$, we can express (4.3) as

$$E_{\lambda_{(2)},\cdots,\lambda_{(n)}}\left[P_{\lambda_{(1)}}\left\{|x_{i1}|^t \leqq a_i - \sum_{\alpha=2}^{n} |x_{i\alpha}|^t, i = 1, \cdots, p \,|\, x_2, \cdots, x_n\right\}\right]. \tag{4.4}$$

By Corollary 3.1, this is nondecreasing in each $\lambda_1^{(1)}, \cdots, \lambda_p^{(1)}$. The argument may be repeated for the remaining $\boldsymbol{\lambda}_{(k)}$. *Q.E.D.*

Of particular interest is the case $t = 2$, for then (4.3) becomes

$$P_{\boldsymbol{\lambda}}\{s_{ii} \leqq a_i, i = 1, \cdots, p\}, \tag{4.5}$$

where $(s_{ij}) = S = XX'$. If in addition, $f_k(z) = (2\pi)^{-p/2}e^{-z/2}$, $\Sigma^{(k)} = \Sigma$, and $\boldsymbol{\lambda}_{(k)} = (\lambda_1, \cdots, \lambda_p)$ for each k, then S has the Wishart distribution $W_p(n, \Sigma)$ with $E(S/n) = \Sigma$. As a consequence of Theorem 4.1, we obtain

COROLLARY 4.1. *If $S \sim W_p(n, \Sigma)$, then*

$$P_\Sigma\{s_{11} \leqq a_1, \cdots, s_{pp} \leqq a_p\} \geqq \prod_1^p P_{\sigma_{ii}}\{s_{ii} \leqq a_i\}. \tag{4.6}$$

Also, by a conditional argument similar to the proof of Theorem 4.1, but utilizing the inequality (3.26) rather than Corollary 3.1, we obtain

COROLLARY 4.2. *If $S \sim W_p(n, \Sigma)$ and $\Sigma = D_\tau + \alpha'\alpha$, then*

$$P_\Sigma\{s_{11} \geqq a_1, \cdots, s_{pp} \geqq a_p\} \geqq \prod_1^p P_{\sigma_{ii}}\{s_{ii} \geqq a_i\}. \tag{4.7}$$

Again by a similar conditional argument, but now utilizing Theorem 3.3, we have

COROLLARY 4.3. *If $S \sim W_p(n, \Sigma)$ and $\Sigma_2 - \Sigma_1$ is positive semidefinite, then*

$$P_{\Sigma_1}\{s_{11} \leqq a_1, \cdots, s_{pp} \leqq a_p\} \geqq P_{\Sigma_2}\{s_{11} \leqq a_1, \cdots, s_{pp} \leqq a_p\}. \tag{4.8}$$

Perhaps the most important result in terms of applications is the following.

THEOREM 4.2. *Let the p dimensional random vector y have a density function $|\Gamma|^{-1/2}f(y\Gamma^{-1}y')$; let the columns of X be independently and identically distributed with density $N_p(0, \Sigma)$, and let $S = XX'$. If $\Sigma = D_\tau + \alpha'\alpha = (\sigma_{ij})$ and $\Gamma = (\gamma_{ij})$ then*

$$P_{\Gamma,\Sigma}\{y_i^2 \leqq k_i s_{ii}, i = 1, \cdots, p\} \geqq P_{D_\gamma, D_\sigma}\{y_i^2 \leqq k_i s_{ii}, i = 1, \cdots, p\}, \tag{4.9}$$

where $D_\gamma = \operatorname{diag}(\gamma_{11}, \cdots, \gamma_{pp})$, $D_\sigma = \operatorname{diag}(\sigma_{11}, \cdots, \sigma_{pp})$.

PROOF. For fixed S, we have from Corollary 3.1 that

$$P_{\Gamma,\Sigma}(y_i^2 \leqq k_i s_{ii}, i = 1, \cdots, p \,|\, S\} \geqq P_{D_\gamma,\Sigma}\{y_i^2 \leqq k_i s_{ii}, i = 1, \cdots, p \,|\, S\}, \tag{4.10}$$

and hence (4.10) holds unconditionally. Now condition on y, and by Corollary 4.2,

$$P_{D_\gamma,\Sigma}\left\{s_{ii} \geqq \frac{y_i^2}{k_i}, i = 1, \cdots, p \,\middle|\, y\right\} \geqq P_{D_\gamma, D_\sigma}\left\{s_{ii} \geqq \frac{y_i^2}{k_i}, i = 1, \cdots, p \,\middle|\, y\right\}, \tag{4.11}$$

so that (4.11) holds unconditionally. *Q.E.D.*

This result was proved for the bivariate normal case when $p = 2$ by Halperin [11], and for general p by Šidák [27] when y is normal. As in Theorem 4.1, Theorem 4.2 can be extended to the case where the columns of X are not identically distributed.

Lastly, suppose the columns of the $p \times n$ matrix $X = (x'_1, \cdots, x'_n)$ are independently and identically distributed, each with density function $|\Sigma_\lambda|^{-1/2} f(x\Sigma_\lambda^{-1}x')$, (see Corollary 3.1 for notation). Let

$$(4.12) \qquad X = \begin{pmatrix} \dot{X} \\ \ddot{x} \end{pmatrix}, \qquad \dot{X} = (\dot{x}'_1, \cdots, \dot{x}'_n), \qquad \ddot{x} = (x_{p1}, \cdots, x_{pn}).$$

THEOREM 4.3. *Consider the regions* $\{\dot{X} \in C_1\}$ *and* $\{\ddot{x} \in C_2\}$. *If* C_1 *and* C_2 *are sets in* $R^{(p-1)n}$ *and* R^n, *respectively, which are convex and symmetric in each column of* $\dot{X}$ *and* $\ddot{x}$, *respectively, when the remaining columns are held fixed, then*

$$(4.13) \qquad P_{\lambda_1, \ldots, \lambda_p}\{\dot{X} \in C_1, \ddot{x} \in C_2\}$$

is monotone nondecreasing in each λ_i, $0 \leqq \lambda_i \leqq 1$.

PROOF. The proof is similar to that of Theorem 4.1. By conditioning on $x_2, \cdots, x_n$, we can express (4.13) as

$$(4.14) \qquad E[P_{\lambda_1, \cdots, \lambda_p}\{(\dot{x}_1, \cdots, \dot{x}_n) \in C_1, (x_{p1}, \cdots, x_{pn}) \in C_2 | x_2, \cdots, x_n\}].$$

For fixed $x_2, \cdots, x_n$, $\dot{X} \in C_1$ becomes $\dot{x}_1 \in C_1^* \equiv C_1^*(\dot{x}_2, \cdots, \dot{x}_n)$, where C_1^* is a set in R^{p-1} which, by hypothesis, is convex and symmetric; similarly, $\ddot{x} \in C_2$ becomes $|x_{p1}| \leqq h(x_{p2}, \cdots, x_{pn})$. By Theorem 2.1, the conditional probability is monotone nondecreasing in λ_1, and hence, maintains this property unconditionally. The result follows by repeating the argument with respect to other subscripts. *Q.E.D.*

REMARK 4.1. The proof is valid for a more general formulation of the theorem in which the columns of X are not identically distributed.

Also note that by specializing C_1 to $\Sigma_{\alpha=1}^n |x_{i\alpha}|^t \leqq a_i$, $i = 1, \cdots, p$, we obtain Theorem 4.1 for $t \geqq 1$.

As a consequence of Theorem 4.3, we have

COROLLARY 4.4. *If* $S \sim W_p(n\Sigma)$ *and* $\lambda_1(S_{11}) \geqq \cdots \geqq \lambda_{p-1}(S_{11})$ *are the ordered characteristic roots of* S_{11}, *where* $S_{11} = (s_{ij})$, $1 \leqq i, j \leqq p - 1$, *then*

$$(4.15) \qquad P_\Sigma\left\{\sum_1^m \lambda_i(S_{11}) \leqq k_1, s_{pp} \leqq k_2\right\} \geqq P_{\Sigma_{11}}\left\{\sum_1^m \lambda_i(S_{11}) \leqq k_1\right\} P_{\sigma_{pp}}\{s_{pp} \leqq k_2\}.$$

for $m = 1, 2, \cdots, p - 1$.

PROOF. The proof depends on the following characterization of the roots (von Neumann [30]): if $V: n \times n$ is positive semidefinite, then

$$\max_{\mathscr{U}} \operatorname{tr} UVU' = \sum_{1}^{m} \lambda_i(V), \qquad m = 1, \cdots, n, \tag{4.16}$$

where $\mathscr{U}$ is the set of $m \times n$, row orthogonal matrices, $m \leqq n$.

To complete the proof, we need only show that the set

$$C_1 = \{\dot{X}: \max_{\mathscr{U}} \operatorname{tr} U\dot{X}\dot{X}'U' \leqq k_1\} \tag{4.17}$$

is a convex symmetric set in each column of $\dot{X}$. To see this, note that

$$\max_{\mathscr{U}} \operatorname{tr} U\dot{X}\dot{X}'U' = \max_{\mathscr{U}} \left[\operatorname{tr} U\dot{x}_1'\dot{x}_1 U' + \sum_{2}^{n} \operatorname{tr} U\dot{x}_j'\dot{x}_j U' \right], \tag{4.18}$$

which for fixed $\dot{x}_2, \cdots, \dot{x}_n$, is a convex, symmetric function of $\dot{x}_1$. Hence, C_1 is a convex, symmetric set in $\dot{x}_1$, and, by permutation of subscripts, in each $\dot{x}_j$. Similarly, $C_2 = \{\ddot{x}: \Sigma_\alpha x_{p\alpha}^2 \leqq k_2\}$ is convex and symmetric in each $x_{p\alpha}$ when the remaining components are held fixed. *Q.E.D.*

5. The one sided inequality for $P^+(\Sigma)$

In this section, we present a proof based on a reflection-inclusion argument to provide a generalization of Slepian's theorem to elliptically contoured distributions. Chartres' proof [2], although stated for the normal distribution, can be modified to treat the general case. In fact, a slightly stronger inequality is obtained this way. Recall that if $x = (x_1, \cdots, x_p)$ has density $|\Sigma|^{-1/2} f(x\Sigma^{-1}x')$ and if the $\{\ell_i\}$ are given real numbers, we define

$$P^+(\Sigma) = P_\Sigma\{x_1 \leqq \ell_1, \cdots, x_p \leqq \ell_p\}. \tag{5.1}$$

THEOREM 5.1. *For any two positive definite (symmetric) $p \times p$ matrices $\Gamma = (\gamma_{ij})$ and $\Sigma = (\sigma_{ij})$ such that $\gamma_{ii} = \sigma_{ii}$, $1 \leqq i \leqq p$, and $\gamma_{ij} \geqq \sigma_{ij}$, $1 \leqq i < j \leqq p$,*

$$P^+(\Gamma) \geqq P^+(\Sigma). \tag{5.2}$$

Strict inequality holds in (5.2) *if $\gamma_{ij} > \sigma_{ij}$ for at least one pair $i < j$ and the support of f is unbounded, that is, $P_I\{\|x\| \leqq k\} < 1$ for all $k > 0$.*

The proof is based on a fundamental geometric lemma:

LEMMA 5.1. *Let $z = (z_1, \cdots, z_p)$ have density $f(zz')$, and let $a_1, \cdots, a_r, b$ be a set of $(1 \times p)$ vectors such that $a_1a_1' = bb'$ and $a_1a_j' \leqq ba_j'$ for $2 \leqq j \leqq r$. If*

$$\begin{aligned} E &= \{z \mid za_2' \leqq \ell_2, \cdots, za_r' \leqq \ell_r\}, \\ F &= \{z \mid za_1' \leqq \ell_1\}, \\ G &= \{z \mid zb' \leqq \ell_1\}, \end{aligned} \tag{5.3}$$

then $P[z \in E \cap F] \leqq P[z \in E \cap G]$. Strict inequality holds if $a_1a_j' < ba_j'$ for at least one $j \geqq 2$ and the support of f is unbounded.

PROOF. (The reader is urged to consider the case $p = r = 2$ in order to picture the underlying geometry.) Clearly, we may assume $a_1 \neq b$. Writing $P[H]$

for $P[z \in H]$, we have

$$P[E \cap G] - P[E \cap F] = P[E \cap F^c \cap G] - P[E \cap F \cap G^c]. \tag{5.4}$$

Since the distribution of z is orthogonally invariant, it suffices to produce an orthogonal transformation $T: R^p \to R^p$ such that

$$T(E \cap F \cap G^c) \subset (E \cap F^c \cap G). \tag{5.5}$$

Let T be a reflection through the $(p-1)$ dimensional subspace $\{z \mid (a_1 - b)z' = 0\}$, that is,

$$T: z \to zT = z - 2\frac{(uz')}{(uu')}u = z\left(I - 2\frac{u'u}{uu'}\right), \tag{5.6}$$

where $u = a_1 - b$. Note that if α and β lie on the same side of $\{z \mid uz' = 0\}$, that is, $(u\alpha')(u\beta') \geqq 0$, then

$$(\alpha T)\beta' = \alpha\beta' - 2\frac{(u\alpha')(u\beta')}{uu'} \leqq \alpha\beta'. \tag{5.7}$$

Also note that T is self-adjoint, that is, for any α and β

$$(\alpha T)\beta' = \alpha(\beta T)'. \tag{5.8}$$

Furthermore,

$$a_1T - b = (a_1 - b)\left[1 - \frac{2(a_1 - b)a_1'}{(a_1 - b)(a_1 - b)'}\right] = 0, \tag{5.9}$$

where we use the fact that $a_1a_1' = bb'$. Since $T^2 = I$, we have $bT = a_1$.

To deduce (5.5) suppose $z \in E \cap F \cap G^c$, so that $za_j' \leqq \ell_j$, $j \geqq 2$, $za_1' \leqq \ell_1$, $zb' > \ell_1$. Therefore, $uz' < 0$, and by hypothesis $ua_j' \leqq 0, j \geqq 2$, so by (5.7)

$$(zT)a_j' \leqq za_j' \leqq \ell_j; \tag{5.10}$$

hence, $zT \in E$. Also $zT \in F^c$, since by (5.8) and (5.9)

$$(zT)a_1' = z(a_1T)' = zb' > \ell_1. \tag{5.11}$$

Similarly, $zT \in G$, so $zT \in (E \cap F^c \cap G)$, proving (5.5). The statement concerning strict inequality follows from the fact that

$$(E \cap F^c \cap G) - T(E \cap F \cap G^c) \tag{5.12}$$

is a nonempty unbounded polyhedral wedge (consider the case $p = r = 2$). *Q.E.D.*

Proof of Theorem 5.1. First we show that the problem can be reduced to the case where $\sigma_{12} < \gamma_{12}$ and $\sigma_{ij} = \gamma_{ij}$ if $(i, j) \neq (1, 2)$. The set M of $p \times p$ positive definite (symmetric) matrices $\Delta = (\delta_{ij})$ having fixed diagonal elements $\delta_{ii} = \sigma_{ii}$, $1 \leqq i \leqq p$, is an open convex subset of $R^{p(p-1)/2}$ (with coordinates $\{\delta_{ij}: 1 \leqq i < j \leqq p\}$). Since $\Sigma \in M$ and $\Gamma \in M$, the closed line segment (in $R^{p(p-1)/2}$) between Σ and Γ lies totally inside M, and hence, is bounded away from the boundary of M. Therefore, since $\sigma_{ij} \leqq \gamma_{ij}$ for $1 \leqq i < j \leqq p$, there exists a polygonal path Π with vertices

$$\Delta^{(0)} = \Sigma, \Delta^{(1)}, \cdots, \Delta^{(n-1)}, \Delta^{(n)} = \Gamma \tag{5.13}$$

such that $\Pi \subset M$ and for each $k = 1, \cdots, n$, $\delta_{ij}^{(k-1)} \leqq \delta_{ij}^{(k)}$ with strict inequality holding for exactly one pair $i < j$ (depending on k). Proceeding stepwise, it thus suffices to treat the case where Σ and Γ differ in exactly one position, say $\sigma_{12} < \gamma_{12}$, (and its symmetric counterpart $\sigma_{21} < \gamma_{21}$).

Now write $\Sigma = AA'$ and $\Gamma = BB'$, where A and B are nonsingular $p \times p$ matrices, and partition

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} \equiv \begin{pmatrix} a_1 \\ A_2 \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} \equiv \begin{pmatrix} b_1 \\ B_2 \end{pmatrix}, \tag{5.14}$$

with A_2 and B_2: $(p-1) \times p$. By the assumption of the preceding paragraph $A_2A_2' = B_2B_2'$, so there exists a $p \times p$ orthogonal matrix L such that $B_2 = A_2L$; hence, $b_j = a_jL$ for $j \geqq 2$. Furthermore, by assumption, $a_1a_2' < b_1b_2'$ and $a_1a_j' = b_1b_j'$ for $j \neq 2$. These facts are now used to obtain (5.2).

Let z be a random vector with density $f(zz')$, so that

$$\begin{aligned} P^+(\Sigma) &= P\{za_1' \leqq \ell_1, \cdots, za_p' \leqq \ell_p\}, \\ P^+(\Gamma) &= P\{zb_1' \leqq \ell_1, \cdots, zb_p' \leqq \ell_p\}. \end{aligned} \tag{5.15}$$

However, the distribution of z is the same as that of zL, so

$$\begin{aligned} P^+(\Gamma) &= P\{zLb_1' \leqq \ell_1, zLb_2' \leqq \ell_2, \cdots, zLb_p' \leqq \ell_p\} \\ &= P\{zLb_1' \leqq \ell_1, za_2' \leqq \ell_2, \cdots, za_p' \leqq \ell_p\}. \end{aligned} \tag{5.16}$$

The proof is completed by applying Lemma 5.1 with $b = b_1L'$ and $r = p$. *Q.E.D*

REMARK 5.1. Under the assumptions of Theorem 5.1, x and $-x$ have the same distribution, so also

$$P_\Sigma\{x_1 \geqq \ell_1, \cdots, x_p \geqq \ell_p\} \leqq P_\Gamma\{x_1 \geqq \ell_1, \cdots, x_p \geqq \ell_p\}. \tag{5.17}$$

When $\sigma_{ij} = \gamma_{ij}$ for $(i, j) \neq (1, 2)$ and $\sigma_{12} < \gamma_{12}$, a modification of Chartres' argument [2] extends Theorem 5.1 as follows:

$$\begin{aligned} &P_\Sigma(x_1 \leqq \ell_1, x_2 \leqq \ell_2, (x_3, \cdots, x_p) \in F) \\ &\qquad \leqq P_\Gamma(x_1 \leqq \ell_1, x_2 \leqq \ell_2, (x_3, \cdots, x_p) \in F), \end{aligned} \tag{5.18}$$

where F is any measurable set in R^{p-2}. To see this, write $x^{(1)} = (x_1, x_2)$, $x^{(2)} = (x_3, \cdots, x_p)$, $\ell^{(1)} = (\ell_1, \ell_2)$ and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \qquad \Sigma_{11\cdot 2} \equiv \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \tag{5.19}$$

where Σ_{11} is 2×2. Setting $\tilde{\ell} = \tilde{\ell}(x^{(2)}) = \ell^{(1)} = x^{(2)}\Sigma_{22}^{-1}\Sigma_{21}$,

$$(5.20)\quad \begin{aligned} P_\Sigma\{x^{(1)} \leqq \ell^{(1)}, x^{(2)} \in F\} &= \int_F \int_{\{x^{(1)} \leqq \ell^{(1)}} fx\Sigma^{-1}x')\, dx \\ &= \int_F \left[\int_{\{w^{(1)} \leqq \tilde{\ell}(x^{(2)})\}} f(w^{(1)}\Sigma_{11\cdot 2}^{-1}w^{(1)\prime} + x^{(2)}\Sigma_{22}^{-1}x^{(2)\prime})\, dw^{(1)}\right] dx^{(2)}. \end{aligned}$$

Now, fix $x^{(2)}$ and note that $\Gamma_{12} = \Sigma_{12}$, $\Gamma_{22} = \Sigma_{22}$. Setting $k = x^{(2)}\Sigma_{22}^{-1}x^{(2)}$, to establish (5.18) it suffices to show that for each $\tilde{\ell}$,

$$(5.21)\quad \begin{aligned} &\int_{\{w^{(1)} \leqq \tilde{\ell}\}} f(w^{(1)}\Sigma_{11\cdot 2}^{-1}w^{(1)\prime} + k)\, dw^{(1)} \\ &\leqq \int_{\{w^{(1)} \leqq \tilde{\ell}\}} f(w^{(1)}\Gamma_{11\cdot 2}^{-1}w^{(1)\prime} + k)\, dw^{(1)}. \end{aligned}$$

The proof of (5.21) is accomplished by setting $\Sigma_{11\cdot 2} = AA' : 2 \times 2$, $\Gamma_{11\cdot 2} = BB' : 2 \times 2$ and arguing as in the proof of Theorem 5.1 for the case of $p = 2$. This argument shows that to prove Theorem 5.1, it suffices to establish the result for the case $p = 2$.

6. Some characterizations for elliptically contoured distributions

The family of elliptically contoured distributions has been considered in great detail by Kelker [14]. If x has a density $\mathscr{L}(x)$ given by

$$(6.1)\quad \mathscr{L}(x) = |\Sigma|^{-1/2}f(x\Sigma^{-1}x'),$$

and we partition $x = (\dot{x}, \ddot{x})$ and

$$(6.2)\quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

conformably, then

$$(6.3)\quad \mathscr{L}(\dot{x}) = |\Sigma_{11}|^{-1/2}g(\dot{x}\Sigma_{11}^{-1}\dot{x}'),$$

where

$$(6.4)\quad g(u) = \int |\Sigma_{22\cdot 1}|^{-1/2}f(u + \ddot{x}\Sigma_{22\cdot 1}^{-1}\ddot{x}')\, d\ddot{x} = \int f(u + \ddot{x}\ddot{x}')\, d\ddot{x},$$

and $\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$. This follows by noting that

$$(6.5)\quad x\Sigma^{-1}x' = \dot{x}\Sigma_{11}^{-1}\dot{x}' + (\ddot{x} - \dot{x}\Sigma_{11}^{-1}\Sigma_{12})\Sigma_{22\cdot 1}^{-1}(\ddot{x} - \dot{x}\Sigma_{11}^{-1}\Sigma_{12})'.$$

The characterization proved by Kelker is that if, for any marginal distribution, f and g are equal up to a constant, then x has a normal distribution. A similar characterization holds for conditional distributions.

It is easily verified that if x has finite second moments then $Ex = 0, Ex'x = c\Sigma$. Also, if $y = xA$, where $A : p \times k$ is of rank k, then

$$\mathscr{L}(y) = |A'\Sigma A|^{-1/2} g(y(A'\Sigma A)^{-1}y'). \tag{6.6}$$

We now obtain another characterization, namely, that the only elliptically contoured family, with finite second moments, which is closed under convolution is the normal distribution. More precisely:

THEOREM 6.1. *Let $x = (x_1, \cdots, x_p)$ and $y = (y_1, \cdots, y_p)$ be independently distributed, and $z = x + y$. If the components of x and y have finite second moments, $\mathscr{L}(x) = |A_1|^{-1/2} f(xA_1^{-1}x')$, $\mathscr{L}(y) = |A_2|^{-1/2} f(yA_2^{-1}y')$, and $\mathscr{L}(z) = |A_3|^{-1/2} f(zA_3^{-1}z')$, then $f(u) = \exp\{-\frac{1}{2}mu\}$, where m is a positive constant.*

PROOF. Note that $z\theta' = x\theta' + y\theta'$ for all θ. From (6.6),

$$\begin{aligned} \mathscr{L}(x\theta') &= \alpha^{-1} g\left(\frac{(x\theta')^2}{\alpha^2}\right), \qquad \mathscr{L}(y\theta') = \beta^{-1} g\left(\frac{(y\theta')^2}{\beta^2}\right), \\ \mathscr{L}(z\theta') &= \gamma^{-1} g\left(\frac{(z\theta')^2}{\gamma^2}\right), \end{aligned} \tag{6.7}$$

where $\alpha^2 = \theta A_1 \theta'$, $\beta^2 = \theta A_2 \theta'$, $\gamma^2 = \theta A_3 \theta'$. However, from the equality of second moments, $\gamma^2 = \alpha^2 + \beta^2$. Letting $\varphi(t) = E(\exp itx)$, it follows that $\varphi(\gamma t) = \varphi(\alpha t)\varphi(\beta t)$. But this is a characterizing property of the normal distribution due to Pólya [22]. Hence $x\theta'$ is normal for all θ, so that x is normal. *Q.E.D.*

REMARK 6.1. The result of Pólya [22] was also obtained by Vincze [29] and generalized to an infinite product by Laha, Lukacs, and Rényi [17]. Its most general version is given by Eaton [8] for the multivariate case, which also provides a proof of Theorem 6.1 as follows. Let

$$\tilde{x} = x_1^{-1/2}\Gamma_1, \qquad \tilde{y} = yA_2^{-1/2}\Gamma_2, \qquad \tilde{z} = zA_3^{-1/2}, \tag{6.8}$$

where $A^{1/2}$ is any square root of A, and Γ_1, Γ_2 are orthogonal matrices chosen so that $\Gamma_j A_j^{1/2} A_3^{-1/2} \equiv S_j$, $j = 1, 2$, are positive definite (symmetric). Consequently, $\mathscr{L}(\tilde{z}) = \mathscr{L}(\tilde{x}S_1 + \tilde{y}S_2)$, with $S_1^2 + S_2^2 = I$, from which the conclusion follows from Eaton ([8], Theorem 2).

It was noted previously in (1.4) that if x is normally distributed with density $\varphi(Ex = 0,\ Ex'x = \Sigma)$, then $\partial^2\varphi/dx_i dx_j = (1 - \frac{1}{2}\delta_{ij})\partial\varphi/\partial\sigma_{i,j}$. This result for $p = 2$ appears to be an old one; for general p Plackett [21] gives a proof using characteristic functions. The following provides both an alternative proof and a characterization of the normal distribution, namely, that the normal distribution is the only elliptically contoured distribution for which the differential equality holds.

THEOREM 6.2. *Let $\varphi_\Sigma(x) \equiv |\Sigma|^{-1/2} f(x\Sigma^{-1}x')$, then*

$$\frac{\partial\varphi}{\partial\sigma_{ij}} = \frac{\partial^2\varphi}{\partial x_i \partial x_j}(1 - \tfrac{1}{2}\delta_{ij}), \tag{6.9}$$

for all $i, j = 1, \cdots, p$, if and only if $f(u) = c\exp\{-\frac{1}{2}u\}$.

PROOF. Note that $(d\Sigma^{-1}) = -\Sigma^{-1}(d\Sigma)\Sigma^{-1}$,

$$\frac{\partial}{\partial\sigma_{ij}}(x\Sigma^{-1}x') = \frac{\partial}{\partial\sigma_{ij}}(tr\,\Sigma^{-1}x'x), \qquad |\Sigma|^{-1}\frac{\partial|\Sigma|}{\partial\sigma_{ij}} = \sigma^{ij}. \tag{6.10}$$

where $\Sigma^{-1} = (\sigma^{ij})$. A direct calculation then yields

$$\begin{aligned} \left(\frac{d\varphi}{\partial\sigma_{ij}}\right) &= -2|\Sigma|^{-1/2}(\Sigma^{-1}x'x\Sigma^{-1})f' - |\Sigma|^{-1/2}\Sigma^{-1}f, \\ \left(\frac{\partial^2\varphi}{\partial x_i \partial x_j}\right) &= 2|\Sigma|^{-1/2}\Sigma^{-1}f' + 4|\Sigma|^{-1/2}(\Sigma^{-1}x'x\Sigma^{-1})f'', \end{aligned} \tag{6.11}$$

where $f' = df(u)/du|_{u=x\Sigma^{-1}x'}$. Consequently, property (6.9) becomes, after simplification,

$$\Sigma^{-1}(f + 2f') + 2(\Sigma^{-1}x'x\Sigma^{-1})(f' + 2f'') = 0. \tag{6.12}$$

If $f(u) = c\exp\{-\frac{1}{2}u\}$, then it is immediate that $f' + 2f = 0$ and $f' + 2f'' = 0$, so that (6.12) holds.

If (6.12) holds, then because f, f', f'' are functions of $x\Sigma^{-1}x'$, we can choose two vectors x for which $x\Sigma^{-1}x'$ is constant. Since the first term in (6.12) remains fixed, the second term must be zero, from which it follows that $f + 2f' = 0$. Hence $f(u) = c\exp\{-\frac{1}{2}u\}$, where the constant c is determined so that f is a density. *Q.E.D.*

REFERENCES

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

[2] B. A. CHARTRES, "A geometrical proof of a theorem due to Slepian," *SIAM Review*, Vol. 5 (1963), pp. 335–341.

[3] J. CHOVER, "Certain convexity conditions on matrices with applications to Gaussian processes," *Duke Math J.*, Vol. 29 (1962), pp. 141–150.

[4] S. DAS GUPTA, "Some inequalities for multivariate normal distribution," *Calcutta Statist. Assoc. Bull.*, Vol. 18 (1968), pp. 179–180.

[5] O. J. DUNN, "Estimation of the means of dependent variables," *Ann. Math. Statist.*, Vol. 29 (1958), pp. 1095–1111.

[6] ———, "Confidence intervals for the means of dependent normally distributed variables," *J. Amer. Statist. Assoc.*, Vol. 54 (1959), pp. 613–621.

[7] C. W. DUNNETT and M. SOBEL, "Approximations to the probability integral and certain percentage points to a multivariate analogue of Student's *t*-distribution," *Biometrika*, Vol. 42 (1955), pp. 258–260.

[8] M. L. EATON, "Characterization of distributions by the identical distribution of linear forms," *J. Appl. Prob.*, Vol. 3 (1966), pp. 481–494.

[9] C. FEFFERMAN, M. JODEIT, JR., and M. D. PERLMAN, "A spherical surface measure inequality for convex sets," *Proc. Amer. Math. Soc.*, Vol. 30.

[10] B. V. GNEDENKO and A. N. KOLMOGOROV, *Limit Distributions for Sums of Independent Random Variables*, Cambridge, Addison Wesley Publishing Co., 1954.

[11] M. HALPERIN, "An inequality in a bivariate Student's "t" distribution," *J. Amer. Statist. Assoc.*, Vol. 62 (1967), pp. 603–606.

[12] D. R. Jensen, "The joint distribution of traces of Wishart matrices and some applications," *Ann. Math. Statist.*, Vol. 41 (1970), pp. 133–145.

[13] K. Jogdeo, "A simple proof of an inequality for multivariate normal probabilities of rectangles," *Ann. Math. Statist.*, Vol. 41 (1970), pp. 1357–1359.

[14] D. Kelker, "Distribution theory of spherical distributions and some characterization theorems," Michigan State University, Technical Report Rm-210, DK-1, 1968.

[15] C. G. Khatri, "On certain inequalities for normal distributions and their applications to simultaneous confidence bounds," *Ann. Math. Statist.*, Vol. 38 (1967), pp. 1853–1867.

[16] ———, "Further contributions to some inequalities for normal distributions and their applications to simultaneous confidence bounds," *Ann. Inst. Statist. Math.*, Vol. 22 (1970), pp. 451–458.

[17] R. G. Laha, E. Lukacs, and A. Rényi, "A generalization of a theorem of E. Vincze," *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, Vol. 9 (1964), pp. 237–239.

[18] E. Lehmann, "Some concepts of dependence," *Ann. Math. Statist.*, Vol. 37 (1966), pp. 1137–1153.

[19] R. D. Lord, "The use of Hankel transforms in statistics, I. General theory and examples," *Biometrika*, Vol. 41 (1954), pp. 44–45.

[20] G. P. Patil and M. T. Boswell, "A characteristic property of the multivariate normal density function and some of its applications," *Ann. Math. Statist.*, Vol. 41 (1970), pp. 1970–1977.

[21] R. L. Plackett, "A reduction formula for normal multivariate integrals," *Biometrika*, Vol. 41 (1954), pp. 351–360.

[22] G. Pólya, "Herleitung des Gauss'schen Fehlergesetzes aus einer Funktionalgleichung," *Math. Z.*, Vol. 18 (1923), pp. 96–108.

[23] A. J. Scott, "A note on conservative confidence regions for the mean of a multivariate normal," *Ann. Math. Statist.*, Vol. 38 (1967), pp. 278–280.

[24] S. Sherman, "A theorem on convex sets with applications," *Ann. Math. Statist.*, Vol. 26 (1955), pp. 763–766.

[25] Z. Šidák, "Rectangular confidence regions for the means of multivariate normal distributions," *J. Amer. Statist. Assoc.*, Vol. 62 (1967), pp. 626–633.

[26] ———, "On multivariate normal probabilities of rectangles: their dependence on correlations," *Ann. Math. Statist.*, Vol. 39 (1968), pp. 1425–1434.

[27] ———, "On probabilities of rectangles in multivariate student distributions: their dependence on correlations," *Ann. Math. Statist.*, Vol. 42 (1971), pp. 169–175.

[28] D. Slepian, "The one-sided barrier problem for Gaussian noise," *Bell System Tech. J.*, Vol. 41 (1962), pp. 463–501.

[29] E. Vincze, "Bemerkung zur Charakterizierung des Gauss'schen Fehlergesetzes," *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, Vol. 7 (1962), pp. 357–361.

[30] J. von Neumann, "Some matrix-inequalities and metrization of matric-space," *Tomsk. Univ. Rev.*, Vol. 1 (1937), pp. 286–300. (In *Collected Works*, Vol. IV, Oxford, Pergamon Press, 1962.)

[31] A. Wintner, *Asymptotic Distributions and Infinite Convolutions*, Ann Arbor, Edwards Brothers, 1938.