The Annals of Applied Probability
1997, Vol. 7, No. 4, 1106–1117

INEQUALITIES FOR THE PROBABILITY CONTENT OF A ROTATED ELLIPSE AND RELATED STOCHASTIC DOMINATION RESULTS

By Thomas Mathew[1] and Kenneth Nordström[2]

University of Maryland Baltimore County and University of Helsinki

Let X_i and Y_i follow noncentral chi-square distributions with the same degrees of freedom ν_i and noncentrality parameters δ_i^2 and μ_i^2, respectively, for $i = 1, \ldots, n$, and let the X_i's be independent and the Y_i's independent. A necessary and sufficient condition is obtained under which $\sum_{i=1}^n \lambda_i X_i$ is stochastically smaller than $\sum_{i=1}^n \lambda_i Y_i$ for all nonnegative real numbers $\lambda_1 \geq \cdots \geq \lambda_n$. Reformulating this as a result in geometric probability, solutions are obtained, in particular, to the problems of monotonicity and location of extrema of the probability content of a rotated ellipse under the standard bivariate Gaussian distribution. This complements results obtained by Hall, Kanter and Perlman who considered the behavior of the probability content of a square under rotation. More generally, it is shown that the vector of partial sums $(X_1, X_1 + X_2, \ldots, X_1 + \cdots + X_n)$ is stochastically smaller than $(Y_1, Y_1 + Y_2, \ldots, Y_1 + \cdots + Y_n)$ if and only if $\sum_{i=1}^n \lambda_i X_i$ is stochastically smaller than $\sum_{i=1}^n \lambda_i Y_i$ for all nonnegative real numbers $\lambda_1 \geq \cdots \geq \lambda_n$.

1. Introduction. This paper originates with a problem which in its simplest form can be stated as follows. Let $\chi_\nu^2(\delta^2)$ denote the noncentral chi-square distribution with $\nu > 0$ degrees of freedom and δ^2 as noncentrality parameter. Suppose that X_1 and X_2 are independent random variables distributed, respectively, as $\chi_1^2(\delta_1^2)$ and $\chi_1^2(\delta_2^2)$, and suppose that Y_1 and Y_2 are independent and distributed as $\chi_1^2(\delta_1^2 + \delta_2^2)$ and $\chi_1^2(0)$, respectively. Is it true that

$$\lambda X_1 + X_2 \leq_{\mathscr{D}} \lambda Y_1 + Y_2 \quad \text{for all } \lambda \geq 1, \tag{1.1}$$

or equivalently, that

$$\lambda X_1 + (1-\lambda) X_2 \leq_{\mathscr{D}} \lambda Y_1 + (1-\lambda) Y_2 \quad \text{for all } \lambda \in [1/2, 1], \tag{1.2}$$

where "$\leq_{\mathscr{D}}$" denotes ordinary stochastic domination? [For real-valued random variables U and V, $U \leq_{\mathscr{D}} V$ if $\mathbb{P}\{U \leq a\} \geq \mathbb{P}\{V \leq a\}$ holds for all $a \in \mathbb{R}$.] Heuristically, since $\mathbb{P}\{\chi_\nu^2(\delta^2) \leq a\}$ is decreasing as a function of the noncentrality parameter δ^2 (for fixed a), it follows that $X_1 \leq_{\mathscr{D}} Y_1$, and therefore

---

Received November 1996; revised April 1997.

[1]Research supported by U.S. Air Force Office of Scientific Research Grant AFOSR F49620-95-1-0335.

[2]Research initiated while a Senior Fulbright Scholar in residence at the Department of Mathematics and Statistics, University of Maryland Baltimore County. Also supported by grants from the Ella and Georg Ehrnrooth Foundation and the Finnish Society of Sciences and Letters.

AMS 1991 *subject classifications*. Primary 60D05, 60E15; secondary 62H99.

Key words and phrases. Coupling, majorization, noncentral chi-square distribution, Poisson mixture representation, Schur-convex function, stochastic ordering.

one would expect, for example, inequality (1.2) to hold at least for values of λ sufficiently close to unity. [Similarly, the converse inequality should hold in (1.2) for small enough nonnegative λ.]

Suppose, more generally, that X_i is distributed as $\chi^2_{\nu_i}(\delta_i^2)$ for $i = 1, \dots, n$, with the X_i's independent, and suppose that Y_i is distributed as $\chi^2_{\nu_i}(\mu_i^2)$ for $i = 1, \dots, n$, with the Y_i's independent. Under what condition on the noncentrality parameters can one have

$$\sum_{i=1}^{n} \lambda_i X_i \leq_{\mathscr{D}} \sum_{i=1}^{n} \lambda_i Y_i, \tag{1.3}$$

the λ_i's being any nonnegative scalars satisfying $\lambda_1 \geq \cdots \geq \lambda_n$? The original motivation for considering this problem comes from Mathew, Sharma and Nordström (1995), where a special case of (1.3) is encountered in the context of constructing confidence regions in a multivariate calibration problem.

The preceding problems can alternatively be recast as problems in geometric probability. This reformulation, which we find particularly intriguing, is perhaps best illustrated using (1.1). Let X_1, X_2, Y_1 and Y_2 be as in (1.1), and let

$$\delta = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \quad \text{and} \quad \mu = \begin{pmatrix} \sqrt{\delta_1^2 + \delta_2^2} \\ 0 \end{pmatrix}, \tag{1.4}$$

assuming, without loss of generality, that $\delta_1, \delta_2 \geq 0$. Given $\lambda > 1$, define $Q(\delta) = \lambda X_1 + X_2$ and $Q(\mu) = \lambda Y_1 + Y_2$, and let $Z = (Z_1, Z_2)'$ be standard bivariate Gaussian. Then

$$Q(\delta) =_{\mathscr{D}} \lambda(Z_1 - \delta_1)^2 + (Z_2 - \delta_2)^2 \tag{1.5}$$

and

$$Q(\mu) =_{\mathscr{D}} \lambda\left(Z_1 - \sqrt{\delta_1^2 + \delta_2^2}\right)^2 + Z_2^2, \tag{1.6}$$

and the inequality in (1.1) requires that

$$\mathbb{P}\{Q(\delta) \leq a\} \geq \mathbb{P}\{Q(\mu) \leq a\} \tag{1.7}$$

hold for all $a \in \mathbb{R}$.

Let $\|\cdot\|$ and $\|\cdot\|_\Lambda$ denote, respectively, the standard Euclidean norm and the weighted norm corresponding to the matrix $\Lambda = \operatorname{diag}(\lambda, 1)$, and define

$$\mathscr{E}_\eta = \{z \in \mathbb{R}^2 \colon \|z - \eta\|_\Lambda^2 \leq a\}, \qquad \eta \in \mathbb{R}^2,\ a > 0. \tag{1.8}$$

Then inequality (1.7) takes the form

$$(2\pi)^{-1} \int_{\mathscr{E}_\delta} \exp(-\|z\|^2/2)\, dz \geq (2\pi)^{-1} \int_{\mathscr{E}_\mu} \exp(-\|z\|^2/2)\, dz. \tag{1.9}$$

Clearly, $\mathscr{E}_\delta$ and $\mathscr{E}_\mu$ are ellipses with principal semiaxes of lengths $(a/\lambda)^{1/2}$ and $a^{1/2}$, and with centers δ and μ lying on the circle centered at the origin and with radius $\|\delta\|$ $(= \|\mu\|)$. In view of (1.9) it is seen that, if true, (1.7) [and hence

(1.1)] requires that the probability content of the ellipse $\mathscr{C}_\delta$ be larger than that of $\mathscr{C}_\mu$, regardless of the choice of $a \in \mathbb{R}$. In other words, the probability content under the standard bivariate Gaussian distribution should increase when the ellipse is rotated by an angle of θ ($0 \le \theta \le \pi/2$) by rotating its center from the position μ to position δ while keeping its principal axes parallel to the coordinate axes (see Figure 1), or equivalently (in view of spherical symmetry),

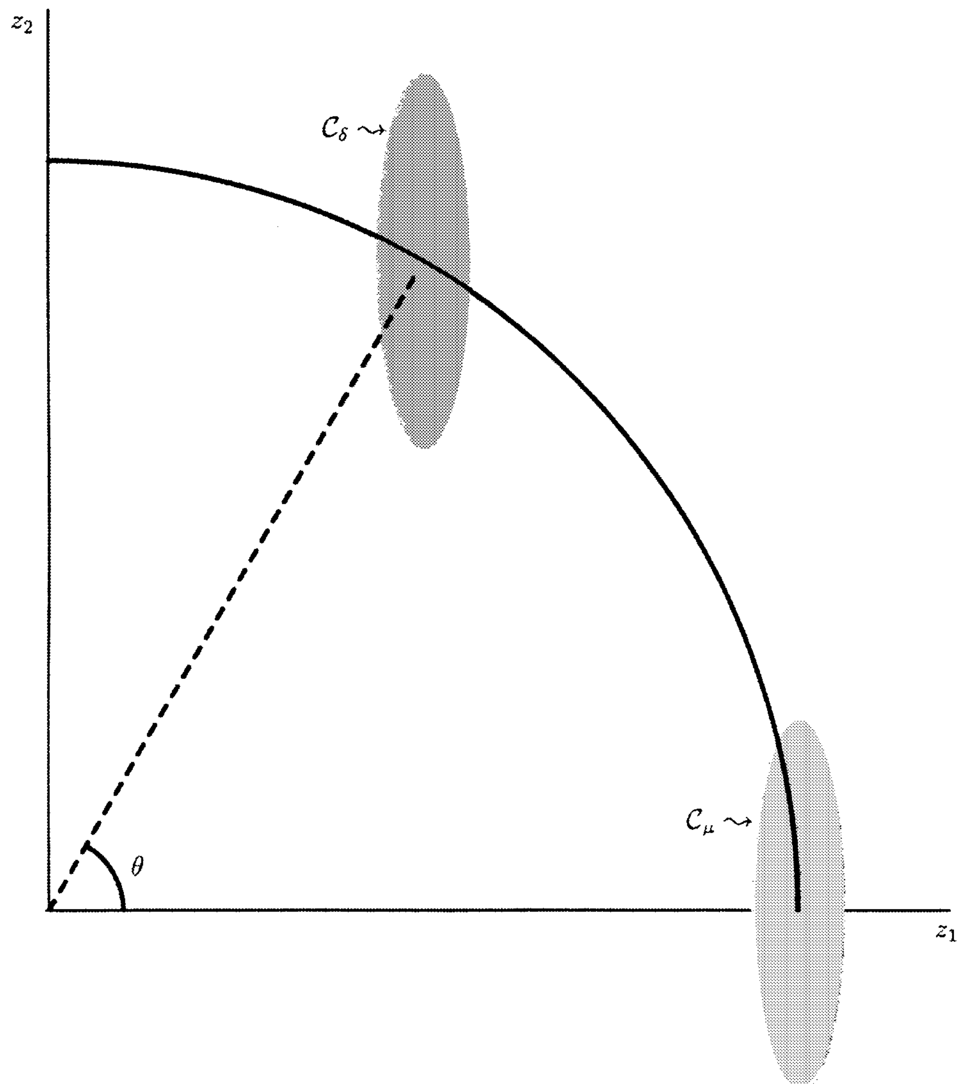

FIG. 1. *Rotating the ellipse* $\mathscr{C}_\mu$ *by rotating its center* $\mu = (\sqrt{\delta_1^2 + \delta_2^2}, 0)'$ *counterclockwise through an angle of* θ *along the boundary of the ball of radius* $\|\mu\|$ *to position* $\delta = (\delta_1, \delta_2)'$ ($\theta = \arccos(\delta_1/\sqrt{\delta_1^2 + \delta_2^2})$).

by keeping its center fixed at μ and rotating it through an angle of θ about its center.

More generally, one may consider the problems of monotonicity and location of extrema of the probability content of an ellipse under rotation. It is interesting to note that this is the exact analogue of a problem considered by Hall, Kanter and Perlman (1980), who studied the behavior of the probability content of a square under rotation. Indeed, from a general log concavity result for Laplace transforms, they deduced that the probability content is a maximum when one of the diagonals of the square (or the extension of the diagonal) passes through the origin. They also showed that the probability content decreases monotonically when the square is rotated from such a position, achieving its minimum after a rotation through an angle of $\pi/4$. Although the present problem is strikingly similar, there is a noteworthy difference which appears to call for different methods of proof. Namely, while a square, centered at the origin, is invariant under exchange of coordinates, this clearly is not the case with an ellipse. This invariance, together with the spherical symmetry of the standard Gaussian distribution, played a crucial role in Hall, Kanter and Perlman (1980).

In the next section, a result is given, which, for $n = 2$ and $\nu_1 = \nu_2 = 1$, provides a necessary and sufficient condition for the stochastic domination (1.3) to hold (Theorem 1). This result yields, in particular, a solution to the problem about rotating an ellipse outlined above. Suppose that instead of (1.4), μ is any vector in $\mathbb{R}^2$, that is, that

$$\delta = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \quad \text{and} \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}. \tag{1.10}$$

Given a vector μ in $\mathbb{R}^2$ and any nonnegative scalars $\lambda_1 \geq \lambda_2$, a characterization is in fact obtained of the region where the center δ of the ellipse $\mathscr{C}_\delta$, defined in terms of $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2)$, must lie such that the probability content of $\mathscr{C}_\delta$ exceeds that of $\mathscr{C}_\mu$. A further refinement is also obtained, characterizing subregions in terms of permissible families of ellipses such that $\mathscr{C}_\delta$ is assigned more probability mass than $\mathscr{C}_\mu$ (Theorem 2). The extension to ellipsoids in $\mathbb{R}^n$ is straightforward and can be obtained from the general necessary and sufficient condition for the stochastic domination (1.3) (Theorem 3).

Let X_i and Y_i, $i = 1, \dots, n$, be as in (1.3), and define for $k = 1, \dots, n$, $\mathscr{S}_k(X) = \sum_{i=1}^k X_i$ and $\mathscr{S}_k(Y) = \sum_{i=1}^k Y_i$, the partial sums of the random vectors $X = (X_1, \dots, X_n)'$ and $Y = (Y_1, \dots, Y_n)'$. Then (1.3) can be rewritten as (defining $\lambda_{n+1} = 0$)

$$\sum_{k=1}^n (\lambda_k - \lambda_{k+1}) \mathscr{S}_k(X) \leq_{\mathscr{D}} \sum_{k=1}^n (\lambda_k - \lambda_{k+1}) \mathscr{S}_k(Y). \tag{1.11}$$

Since the expressions in (1.11) are increasing functionals of the random vectors $\mathscr{S}(X) = (\mathscr{S}_1(X), \dots, \mathscr{S}_n(X))'$ and $\mathscr{S}(Y) = (\mathscr{S}_1(Y), \dots, \mathscr{S}_n(Y))'$, one may, more generally, ask for conditions under which

$$\mathscr{S}(X) \leq_{\mathscr{D}} \mathscr{S}(Y). \tag{1.12}$$

[For random vectors U and V, $U \leq_{\mathscr{D}} V$ if $\mathbb{E}\phi(U) \leq \mathbb{E}\phi(V)$ holds for all bounded increasing functionals ϕ; see, e.g., Kamae, Krengel and O'Brien (1977) and Section 3 for equivalent conditions and further details.] A necessary and sufficient condition is given also for the stochastic domination (1.12) (Theorem 4). Interestingly, the condition turns out to be exactly the same as the condition for (1.3). Although one would, in general, expect the multivariate stochastic domination property (1.12) to be substantially stronger than the univariate domination (1.3), it thus transpires that requiring (1.3) to hold for all nonnegative λ_i's satisfying $\lambda_1 \geq \cdots \geq \lambda_n$ is indeed enough to imply the multivariate stochastic domination (1.12) between the vectors of partial sums $\mathscr{S}(X)$ and $\mathscr{S}(Y)$.

The proofs of the results involve coupling constructions, that is, constructions of pointwise ordered random variables (vectors), are briefly outlined in Section 3, and are given in full in Mathew and Nordström (1996). That coupling should occur in this context is, of course, no surprise, in view of the well-known result by Strassen (1965).

2. Results.

The following result provides (for $n = 2$ and $\nu_1 = \nu_2 = 1$) a necessary and sufficient condition for the stochastic domination (1.3) to hold, and shows, in particular, the validity of inequality (1.1) [and (1.2)].

THEOREM 1. *Let X_i and Y_i be distributed, respectively, as $\chi_1^2(\delta_i^2)$ and $\chi_1^2(\mu_i^2)$, $i = 1, 2$, with X_1 and X_2 independent and Y_1 and Y_2 independent. Then*

$$\lambda_1 X_1 + \lambda_2 X_2 \leq_{\mathscr{D}} \lambda_1 Y_1 + \lambda_2 Y_2 \tag{2.1}$$

holds for all nonnegative $\lambda_1 \geq \lambda_2$ if and only if

$$\delta_1^2 \leq \mu_1^2 \tag{2.2}$$

and

$$\delta_1^2 + \delta_2^2 \leq \mu_1^2 + \mu_2^2 \tag{2.3}$$

are satisfied.

Let $\delta = (\delta_1, \delta_2)'$ and $\mu = (\mu_1, \mu_2)'$. Given $\mu \in \mathbb{R}_+^2$, the positive quadrant, let $\mathscr{B}_{\|\mu\|}^+$ denote the part of the ball $\mathscr{B}_{\|\mu\|}$ centered at the origin and of radius $\|\mu\|$ which lies within $\mathbb{R}_+^2$, and define the regions (cf. Figure 2)

$$\mathscr{R}_1 = \left\{\delta \in \mathscr{B}_{\|\mu\|}^+ \colon \delta_1 \leq \mu_1,\ \delta_2 \leq \mu_2\right\}, \tag{2.4}$$

$$\mathscr{R}_2 = \left\{\delta \in \mathscr{B}_{\|\mu\|}^+ \colon \delta_1 \leq \mu_1,\ \delta_2 > \mu_2\right\} \tag{2.5}$$

and

$$\mathscr{R}_3 = \left\{\delta \in \mathscr{B}_{\|\mu\|}^+ \colon \delta_1 > \mu_1,\ \delta_2 \leq \mu_2\right\}. \tag{2.6}$$

Restricting ourselves, without loss of generality (in view of spherical symmetry), to the positive quadrant $\mathbb{R}_+^2$ ($\delta_i,\ \mu_i \geq 0$), the inequality (2.1) holds if and only if $\delta \in \mathscr{R}_1 \cup \mathscr{R}_2$, in view of Theorem 1 above. However, the assumption

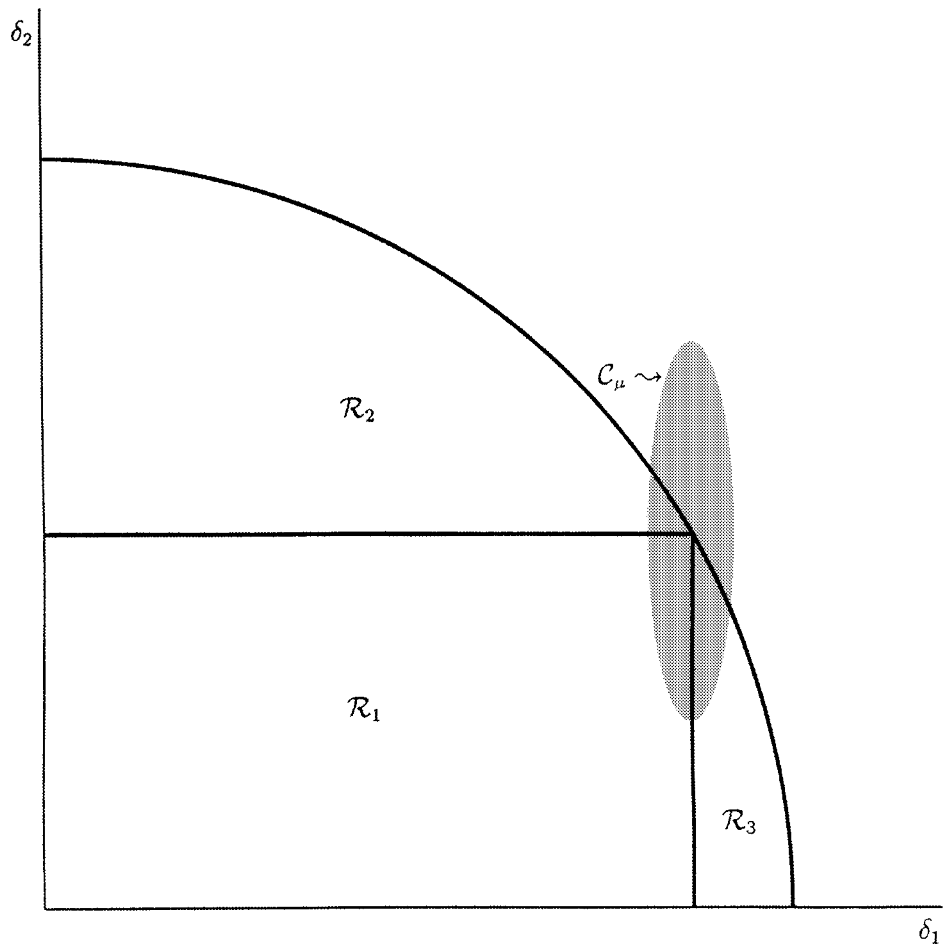

FIG. 2. *Given an ellipse $\mathscr{C}_\mu$ with center $\mu = (\mu_1, \mu_2)'$ $(\mu_1, \mu_2 \geq 0)$, the part of the ball of radius $\|\mu\|$ which lies in the positive quadrant splits into three disjoint regions $\mathscr{R}_1$, $\mathscr{R}_2$ and $\mathscr{R}_3$.*

$\lambda_1 \geq \lambda_2$ is needed only when $\delta \in \mathscr{R}_2$. This is, indeed, a consequence of the following result, which shows that the subregions of $\mathscr{B}^+_{\|\mu\|}$ can be characterized in terms of permissible families of ellipses for which the probability content of $\mathscr{C}_\delta$ exceeds that of $\mathscr{C}_\mu$. Given $\eta \in \mathbb{R}^2_+$, $\mathscr{C}_\eta$ denotes here an ellipse from the homothetic family (obtained by varying $a > 0$) defined by (1.8), with specified $\Lambda = \text{diag}(\lambda_1, \lambda_2)$.

THEOREM 2. *Let $\mu \in \mathbb{R}^2_+$ and $a > 0$ be given and suppose that $\delta \in \mathbb{R}^2_+$. Under the standard bivariate Gaussian distribution, the probability inequality*

$$\mathbb{P}\{\mathscr{C}_\delta\} \geq \mathbb{P}\{\mathscr{C}_\mu\} \tag{2.7}$$

holds:

(i) *for all ellipses $\mathscr{E}_\delta$ and $\mathscr{E}_\mu$ with major principal axis in the direction of the vertical coordinate axis if and only if $\delta \in \mathscr{R}_1 \cup \mathscr{R}_2$;*

(ii) *for all ellipses $\mathscr{E}_\delta$ and $\mathscr{E}_\mu$ with major principal axis in the direction of the horizontal coordinate axis if and only if $\delta \in \mathscr{R}_1 \cup \mathscr{R}_3$;*

(iii) *for all ellipses $\mathscr{E}_\delta$ and $\mathscr{E}_\mu$ if and only if $\delta \in \mathscr{R}_1$;*

where the regions $\mathscr{R}_1$, $\mathscr{R}_2$ and $\mathscr{R}_3$ are defined by (2.4)–(2.6).

From (i) it follows, in particular, that when $\lambda_1 > \lambda_2$ the probability content of $\mathscr{E}_\mu$ increases monotonically when its center moves from $(\|\mu\|, 0)'$ to $(0, \|\mu\|)'$ along the boundary $\partial\mathscr{B}_{\|\mu\|}$ in $\mathbb{R}^2_+$, the minimum and maximum being achieved (for fixed $\|\mu\|$) at these points, respectively. This provides an affirmative answer to the question about the probability content of an ellipse under rotation, outlined in Section 1.

As pointed out in Section 1, a rotation of an ellipse by an angle of θ can, from the point of view of probability content (under the standard Gaussian distribution), be thought of either as a rotation of the center while keeping the principal axes parallel to the coordinate axes (see Figure 1), or as a rotation of the ellipse through an angle of θ about its center while keeping the center fixed. The case of rotating an ellipse with center, say, on the horizontal coordinate axis, but with principal semiaxes not in generic position, that is, not parallel to the coordinate axes, is thus covered as well under the present setup. Indeed, in view of spherical symmetry, one simply rotates the coordinate axes into generic position relative to the ellipse, and the above results apply to such an ellipse in the new coordinate system.

It is interesting to note that inequality (2.7) can, in fact, be inferred for the entire region $\mathscr{R}_1 \cup \mathscr{R}_2$ (or $\mathscr{R}_1 \cup \mathscr{R}_3$) from a knowledge of the behavior of the probability content of $\mathscr{E}_\mu$ when $\mu = (\mu_1, \mu_2)'$ moves along the boundary $\partial\mathscr{B}_{\|\mu\|}$ only. Indeed, from a well-known convolution inequality due to T.W. Anderson (1955), it follows that the probability content of $\mathscr{E}_\mu$ increases monotonically when the center μ is pulled toward the origin along a ray passing through μ. Therefore the probability content of $\mathscr{E}_\delta$ exceeds that of $\mathscr{E}_\mu$ when $\delta \in \mathbb{R}^2_+$ can be obtained from $\mu \in \mathbb{R}^2_+$ by means of a rotation counterclockwise along $\partial\mathscr{B}_{\|\mu\|} \cap \mathbb{R}^2_+$, followed by a contraction toward the origin. This corresponds to the subregion of $\mathscr{R}_1 \cup \mathscr{R}_2$ lying above the ray passing through μ (cf. Figure 2). However, from Anderson's theorem we are not able to infer inequality (2.7) for the entire region $\mathscr{R}_1 \cup \mathscr{R}_2$. This is because Anderson's theorem is for convex centrally symmetric sets, while the ellipse exhibits even greater symmetry; namely, it is invariant under coordinate-wise sign changes (axial symmetry). But applying a variant of Anderson's inequality for such sets [see, e.g., Theorem 2.1 in Jogdeo (1977)] shows that the probability content of $\mathscr{E}_\mu$ is, in fact, monotonically decreasing in μ_1 and μ_2 separately, not only along rays emanating from the origin. Thus inequality (2.7) follows also for the part of $\mathscr{R}_1 \cup \mathscr{R}_2$

lying below the ray passing through μ (cf. Figure 2). (We are indebted to a referee for drawing our attention to this latter fact.)

The preceding results extend straightforwardly to n noncentral chi-square random variables and to ellipsoids in $\mathbb{R}^n$, $n \geq 2$. We content ourselves here with formulating the extension of Theorem 1 to the case $n \geq 2$ and $\nu_i > 0$, $i = 1, \dots, n$ [cf. (1.3)].

THEOREM 3. *Let X_i and Y_i be distributed, respectively, as $\chi^2_{\nu_i}(\delta_i^2)$ and $\chi^2_{\nu_i}(\mu_i^2)$, $i = 1, \dots, n$, with $X_1, \dots, X_n$ independent and $Y_1, \dots, Y_n$ independent. Then*

$$\sum_{i=1}^{n} \lambda_i X_i \leq_{\mathscr{D}} \sum_{i=1}^{n} \lambda_i Y_i \tag{2.8}$$

holds for all nonnegative λ_i's satisfying $\lambda_1 \geq \cdots \geq \lambda_n$ if and only if

$$\sum_{i=1}^{k} \delta_i^2 \leq \sum_{i=1}^{k} \mu_i^2 \quad \textit{for all } k = 1, \dots, n. \tag{2.9}$$

At the end of Section 1, the inequality (2.8) was rewritten in a form which suggested that multivariate stochastic domination could perhaps hold between the vectors of partial sums formed from the X_i's and the Y_i's; see (1.11) and (1.12). The following result shows that the same condition (2.9) is in fact necessary and sufficient also for such a multivariate stochastic domination property.

THEOREM 4. *Let X_i and Y_i, $i = 1, \dots, n$, be as in Theorem 3, let $X = (X_1, \dots, X_n)'$ and $Y = (Y_1, \dots, Y_n)'$, and let $\mathscr{S}(X) = (\mathscr{S}_1(X), \dots, \mathscr{S}_n(X))'$ and $\mathscr{S}(Y) = (\mathscr{S}_1(Y), \dots, \mathscr{S}_n(Y))'$, where $\mathscr{S}_k(X) = \sum_{i=1}^k X_i$ and $\mathscr{S}_k(Y) = \sum_{i=1}^k Y_i$, $k = 1, \dots, n$. Then*

$$\mathscr{S}(X) \leq_{\mathscr{D}} \mathscr{S}(Y) \tag{2.10}$$

holds if and only if condition (2.9) *is satisfied.*

It should be pointed out that, in the special case when all the degrees of freedom are equal, the fact that (2.9) implies (2.10) can also be obtained using Theorem 2 in Boland, Proschan and Tong (1992) [cf. also Proposition 2.13 in Shaked, Shanthikumar and Tong (1995)].

The multivariate stochastic dominance relation (2.10) is clearly equivalent to $\phi[\mathscr{S}(X)] \leq_{\mathscr{D}} \phi[\mathscr{S}(Y)]$ for all increasing functionals ϕ. Upon rewriting (2.8) in the form (1.11), one would expect the multivariate stochastic dominance relation (2.10) to be a substantially stronger domination property than (2.8). However, from Theorem 3 and Theorem 4 we conclude that (2.10) holds if and only if (2.8) holds for all nonnegative λ_i's satisfying $\lambda_1 \geq \cdots \geq \lambda_n$. Thus we have a situation where requiring stochastic dominance to hold only for a subclass of increasing functionals, namely, the class of linear functionals of the form $l'\mathscr{S}(X)$ and $l'\mathscr{S}(Y)$ with nonnegative coordinates l_i, $i = 1, \dots, n$, is enough to imply $\phi[\mathscr{S}(X)] \le_{\mathscr{D}} \phi[\mathscr{S}(Y)]$ for the class of all increasing functionals ϕ [cf. (1.11)].

REMARK 1. The condition (2.9) resembles weak majorization [see Marshall and Olkin (1979), Chapter 1]. The difference is that, unlike majorization and weak majorization, condition (2.9) is not stated in terms of the ordered components of the vectors of noncentrality parameters

$$\tilde{\delta} = (\delta_1^2, \dots, \delta_n^2)' \quad \text{and} \quad \tilde{\mu} = (\mu_1^2, \dots, \mu_n^2)'. \tag{2.11}$$

Note that while Schur-convex functions are necessarily permutation invariant, the probability $\mathbb{P}\{\sum_{i=1}^n \lambda_i X_i \ge a\}$, as a function of $\tilde{\delta}$, fails to be so, in general. However, suppose we denote the ordered components of $\tilde{\delta}$ as $\delta_{(1)}^2 \ge \delta_{(2)}^2 \ge \dots \ge \delta_{(n)}^2$, and write $\tilde{\delta}_{\downarrow} = (\delta_{(1)}^2, \delta_{(2)}^2, \dots, \delta_{(n)}^2)'$. Now if $X_i \sim \chi_{\nu_i}^2(\delta_{(i)}^2)$, $i = 1, \dots, n$, then Theorem 3 states that the probability $\mathbb{P}\{\sum_{i=1}^n \lambda_i X_i \ge a\}$, as a function of $\tilde{\delta}_{\downarrow}$, is increasing and Schur-convex for every $a > 0$ and $\lambda_1 \ge \dots \ge \lambda_n \ge 0$ [cf. Marshall and Olkin (1979), page 59].

3. Proofs.

We begin by recalling some auxiliary results and properties, which are needed in the proofs. The theorems in the previous section follow straightforwardly from these auxiliary results (A.1), (A.2) and (A.3), as well as from Lemma 1 stated below. Hence we only give a brief outline of the proofs here, referring to Mathew and Nordström (1996) for full details.

The first result is a standard result concerning the noncentral chi-square distribution.

A.1. For a noncentral chi-square distribution $\chi_\nu^2(\delta^2)$,

$$\chi_\nu^2(\delta^2) \sim \chi_{\nu+2N}^2 \quad \text{where } N \sim \text{Poisson}(\delta^2/2),$$

that is, it is representable as a Poisson mixture of central chi-squares.

For random vectors U and V, taking values in a common Euclidean space, stochastic domination (ordering) is defined by $U \le_{\mathscr{D}} V$ if $\mathbb{E}\phi(U) \le \mathbb{E}\phi(V)$ holds for all bounded functionals ϕ which are increasing w.r.t. the componentwise vector ordering. Alternatively, $U \le_{\mathscr{D}} V$ is characterized by $\mathbb{P}\{U \in \mathscr{A}\} \le \mathbb{P}\{V \in \mathscr{A}\}$ holding for all Borel sets $\mathscr{A}$ with increasing indicator $1_{\mathscr{A}}(\cdot)$ [see, e.g., Kamae, Krengel and O'Brien (1977)].

The following simple result will also be used repeatedly.

A.2. Let U, V and T be random vectors, U and V taking values in a common Euclidean space. If, for some determinations of the conditional distributions,

$$\mathbb{P}\{U \in \mathscr{A} | T = t\} \le \mathbb{P}\{V \in \mathscr{A} | T = t\} \tag{3.1}$$

holds for all Borel sets $\mathscr{A}$ with increasing indicator $1_{\mathscr{A}}(\cdot)$ and for all $t \in \text{supp}(T)$, the support of T, then (3.1) holds unconditionally. That is, we have $U \le_{\mathscr{D}} V$.

A further characterization of stochastic dominance is the characterization in terms of coupling, which forms the backbone of our proofs. We shall now state this characterization.

A.3. For random vectors U and V, taking values in a common Euclidean space, $U \leq_{\mathscr{D}} V$ holds if and only if there exist distributional copies $\widehat{U} =_{\mathscr{D}} U$ and $\widehat{V} =_{\mathscr{D}} V$, realized on a common probability space, such that $\mathbb{P}\{\widehat{U} \leq \widehat{V}\} = 1$.

When $U \leq_{\mathscr{D}} V$, Strassen (1965) proved the existence of such a coupling of U and V, but a constructive proof is known only when U and V are univariate; see, for example, Kamae, Krengel and O'Brien (1977) or Szekli (1995), Chapter 2 for further details.

We shall first state a lemma that is used to prove Theorem 1 as well as Theorem 4. The lemma actually establishes Theorem 4 and the equivalence of (2.10) and (2.8) for the special case $n = 2$ and follows directly from (A.1), (A.2) and (A.3).

LEMMA 1. *Let U_i and V_i be distributed, respectively, as $\chi^2_{\nu_i}(\varepsilon_i^2)$ and $\chi^2_{\nu_i}(\xi_i^2)$, $i = 1, 2$, with U_1 and U_2 independent and V_1 and V_2 independent.*

(i) *Suppose that $\varepsilon_1^2 \leq \xi_1^2$ and that $\varepsilon_1^2 + \varepsilon_2^2 \leq \xi_1^2 + \xi_2^2$. Then there exist independent Poisson random variables N_i, $i = 1, 2, 3, 4$, and central chi-square random variables Z_1, Z_2, W_1 and W_2 such that Z_1 and Z_2 are independent and W_1 and W_2 are independent, and such that, conditionally given $N_i = n_i$, $i = 1, 2, 3, 4$,*

$$\begin{pmatrix} U_1 \\ U_1 + U_2 \end{pmatrix} =_{\mathscr{D}} \begin{pmatrix} Z_1 \\ Z_1 + Z_2 \end{pmatrix}, \qquad \begin{pmatrix} V_1 \\ V_1 + V_2 \end{pmatrix} =_{\mathscr{D}} \begin{pmatrix} W_1 \\ W_1 + W_2 \end{pmatrix}, \tag{3.2}$$

where the distributions of Z_1, Z_2, W_1 and W_2 depend on the conditioning values n_i of N_i $(i = 1, 2, 3, 4)$. Furthermore,

$$Z_1 \leq W_1 \quad \text{and} \quad Z_1 + Z_2 \leq W_1 + W_2. \tag{3.3}$$

(ii) *The bivariate stochastic dominance*

$$\begin{pmatrix} U_1 \\ U_1 + U_2 \end{pmatrix} \leq_{\mathscr{D}} \begin{pmatrix} V_1 \\ V_1 + V_2 \end{pmatrix}$$

holds if and only if $\varepsilon_1^2 \leq \xi_1^2$ and $\varepsilon_1^2 + \varepsilon_2^2 \leq \xi_1^2 + \xi_2^2$ are satisfied.

PROOF OF THEOREM 1. By virtue of spherical symmetry, it can be assumed that $\delta_i,\ \mu_i \geq 0$, $i = 1, 2$. The necessity of (2.2) and (2.3) follows directly by taking $\lambda_2 = 0$ and $\lambda_1 = \lambda_2 > 0$, respectively, in (2.1). To prove sufficiency, note that if (2.2) and (2.3) hold, then

$$\begin{pmatrix} X_1 \\ X_1 + X_2 \end{pmatrix} \leq_{\mathscr{D}} \begin{pmatrix} Y_1 \\ Y_1 + Y_2 \end{pmatrix},$$

by Lemma 1(ii). This implies $l_1X_1 + l_2(X_1 + X_2) \leq_{\mathscr{D}} l_1Y_1 + l_2(Y_1 + Y_2)$ for all $l_1, l_2 \geq 0$. In other words, (2.1) holds for all nonnegative $\lambda_1 \geq \lambda_2$. □

PROOF OF THEOREM 2. Given $\mu \in \mathbb{R}^2_+$, and with δ restricted to $\mathbb{R}^2_+$, it follows from Theorem 1 that (2.1) holds for all $\lambda_1 \geq \lambda_2 \geq 0$ if and only if $\delta \in \mathscr{R}_1 \cup \mathscr{R}_2$. But the probability inequality (2.7) is equivalent to (2.1) [cf. (1.9)], and with $(a/\lambda_1)^{1/2}$ and $(a/\lambda_2)^{1/2}$ being, respectively, the length of the horizontal and the vertical principal semiaxes of the ellipses, the claim in (i) follows. Parts (ii) and (iii) can be similarly established. □

We shall now briefly outline the proof of Theorem 4.

PROOF OF THEOREM 4. Suppose that $\mathscr{S}(X) \leq_{\mathscr{D}} \mathscr{S}(Y)$. Then, clearly, $\sum_{i=1}^k X_i \leq_{\mathscr{D}} \sum_{i=1}^k Y_i$ for $k = 1, \ldots, n$. Hence the conditions (2.9) must hold.

Conversely, suppose that (2.9) holds. The proof that $\mathscr{S}(X) \leq_{\mathscr{D}} \mathscr{S}(Y)$ involves a coupling construction similar to that in part (i) of Lemma 1 together with the result A.3. We provide an outline only of the argument here, referring to Mathew and Nordström (1996) for full details. As in (2.11), let $\tilde{\delta}$ and $\tilde{\mu}$ denote the vectors of noncentrality parameters, and define $\tilde{\delta} \preceq \tilde{\mu}$ if conditions (2.9) hold. The relation "$\preceq$" defines a partial vector ordering, and it is possible to exhibit an ascending chain of vectors, all of which lie between $\tilde{\delta}$ and $\tilde{\mu}$ (relative to "$\preceq$"), and such that any two adjacent vectors in the chain differ in two coordinates at most. Indeed, define

$$\mu_{k*}^2 = \sum_{i=1}^{k} \mu_i^2 - \sum_{i=1}^{k-1} \delta_i^2, \qquad k = 2, \ldots, n,$$

and observe that, by construction,

$$\delta_j^2 \leq \mu_{j*}^2 \quad \text{and} \quad \delta_{k-1}^2 + \mu_{k*}^2 = \mu_{(k-1)*}^2 + \mu_k^2,$$

$j = 1, 2, \ldots, n$; $k = 2, \ldots, n$, where $\mu_{1*}^2 = \mu_1^2$. Let

$$\tilde{\delta}_{(j)} = (\delta_1^2, \ldots, \delta_{j-1}^2, \mu_{j*}^2, \mu_{j+1}^2, \ldots, \mu_n^2)', \qquad j = 2, \ldots, n.$$

Then we have

$$\tilde{\delta} \preceq \tilde{\delta}_{(n)} \preceq \tilde{\delta}_{(n-1)} \preceq \cdots \preceq \tilde{\delta}_{(2)} \preceq \tilde{\mu}. \tag{3.4}$$

Such a chain of vectors is also constructed in Boland, Proschan and Tong (1992), who use it in the proof of their Theorem 2. The proof of Theorem 4 is complete if we can show that, given any two adjacent vectors in the chain (3.4), the random vector of partial sums corresponding to the larger vector of noncentrality parameters is stochastically larger. This can be established using Lemma 1, (A.2) and (A.3). □

PROOF OF THEOREM 3. Theorem 3 follows from Theorem 4 using arguments similar to those in the proof of Theorem 1. □

Acknowledgments. We are grateful to two anonymous referees for helpful comments toward revision. In particular, one of the referees drew our attention to the relevance of the result in Jogdeo (1977).

REFERENCES

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

BOLAND, P. J., PROSCHAN, F. and TONG, Y. L. (1992). A stochastic ordering of partial sums of independent random variables and of some random processes. *J. Appl. Probab.* 29 645–654.

HALL, R. L., KANTER, M. and PERLMAN, M. D. (1980). Inequalities for the probability content of a rotated square and related convolutions. *Ann. Probab.* 8 802–813.

JOGDEO, K. (1977). Association and probability inequalities. *Ann. Statist.* 5 495–504.

KAMAE, T., KRENGEL, U. and O'BRIEN, G. L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* 5 899–912.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

MATHEW, T. and NORDSTRÖM, K. (1996). Inequalities for the probability content of a rotated ellipse and related stochastic domination results. Technical report, Dept. Mathematics and Statistics, Univ. Maryland Baltimore County.

MATHEW, T., SHARMA, M. K. and NORDSTRÖM, K. (1995). Conservative multiple use confidence regions in multivariate calibration. Unpublished manuscript.

SHAKED, M., SHANTHIKUMAR, J. G. and TONG, Y. L. (1995). Parametric Schur convexity and arrangement monotonicity properties of partial sums. *J. Multivariate Anal.* 53 293–310.

STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* 36 423–439.

SZEKLI, R. (1995). *Stochastic Ordering and Dependence in Applied Probability*. *Lecture Notes in Statist.* 97. Springer, New York.

DEPARTMENT OF MATHEMATICS
AND STATISTICS
UNIVERSITY OF MARYLAND BALTIMORE COUNTY
1000 HILLTOP CIRCLE
BALTIMORE, MARYLAND 21250
E-MAIL: mathew@umbc2.umbc.edu

DEPARTMENT OF STATISTICS
UNIVERSITY OF HELSINKI
P.O. BOX 54
FIN-00014 HELSINKI
FINLAND
E-MAIL: nordstro@noppa.helsinki.fi