A MULTIVARIATE TCHEBYCHEFF INEQUALITY[1]

By Ingram Olkin[2] and John W. Pratt[3]

University of Chicago

0. Abstract. A multivariate Tchebycheff inequality is given, in terms of the covariances of the random variables in question, and it is shown that the inequality is sharp, i.e., the bound given can be achieved. This bound is obtained from the solution of a certain matrix equation and cannot be computed easily in general. Some properties of the solution are given, and the bound is given explicitly for some special cases. A less sharp but easily computed and useful bound is also given.

1. Introduction and outline. Tchebycheff's inequality states that if y is any real random variable with mean 0 and variance σ^2, then

$$P(|y| \geqq k\sigma) \leqq 1/k^2. \tag{1.1}$$

Berge [1] has generalized this result as follows. If y_1 and y_2 are any real random variables with means 0, variances σ_1^2 and σ_2^2 respectively, and correlation ρ, then

$$P(|y_1| \geqq k\sigma_1 \quad \text{or} \quad |y_2| \geqq k\sigma_2) \leqq \frac{1 + \sqrt{1 - \rho^2}}{k^2}. \tag{1.2}$$

Berge gives an example where the inequality is achieved.

Suppose $y = (y_1, \cdots, y_p)$ is a random vector with mean 0 and nonsingular covariance matrix Σ. We seek an upper bound, depending on Σ and $k_1, \cdots, k_p$, for $P(|y_i| \geqq k_i\sigma_i$ for some $i)$.

The problem can be reduced by letting $x_i = y_i/(k_i\sigma_i)$. Then $x = (x_i, \cdots, x_p)$ has mean 0 and covariance matrix $\Pi = K^{-1}RK^{-1}$, where $R = (\rho_{ij})$ is the correlation matrix of y (and of x), $\Pi_{ij} = \sigma_{ij}/(\sigma_i\sigma_j k_i k_j) = \rho_{ij}/(k_i k_j)$, and K is a diagonal matrix with diagonal elements $k_1, \cdots, k_p$. Furthermore, $|y_i| \geqq k_i\sigma_i$ if and only if $|x_i| \geqq 1$, so $P(|y_i| \geqq k_i\sigma_i$ for some $i) = P(|x_i| \geqq 1$ for some $i)$.

Suppose A is a $p \times p$ matrix such that

$$xAx' \geqq 1 \quad \text{if} \quad |x_i| = 1 \quad \text{for some } i. \tag{1.3}$$

Then, looking at scalar multiples of x, we see that

$$xAx' \geqq 1 \quad \text{if} \quad |x_i| \geqq 1 \quad \text{for some } i, \tag{1.4}$$

Received February 18, 1957; revised June 11, 1957.

[1] This research was sponsored by the Office of Ordnance Research, U.S. Army, and the Statistics Branch, Office of Naval Research.

[2] Work done while on leave from Michigan State University.

[3] Now at Harvard University.

and that

$$(1.5) \qquad xAx' \geqq 0 \quad \text{for all } x,$$

i.e., A is positive definite. Therefore

Lemma 1.1. *If A satisfies* (1.3), *then*

$$(1.6) \quad P(|y_i| \geqq k_i\sigma_i \text{ for some } i) = P(|x_i| \geqq 1 \text{ for some } i) \leqq E(xAx') = \operatorname{tr} A\Pi,$$

where tr *denotes trace.*

Each A satisfying (1.3) therefore gives an upper bound for

$$P(|x_i| \geqq 1 \text{ for some } i).$$

The smallest bound obtainable in this way is the minimum of tr $A\Pi$ over all A satisfying (1.3). The set $\mathfrak{A}$ of all such matrices A is obviously convex, closed, and bounded from below, and tr $A\Pi$ is linear in A, so this minimum is achieved at an extreme point of $\mathfrak{A}$. In Theorem 3.3 it is shown that A is an extreme point of $\mathfrak{A}$ if and only if A^{-1} is positive definite and has 1's on the main diagonal. Furthermore, there is a unique extreme point of $\mathfrak{A}$ minimizing tr $A\Pi$, namely that extreme point A such that $A\Pi A$ is diagonal (Theorem 3.5). The bound thus obtained is the best possible, inasmuch as, if it is less than 1, there is a distribution for x (with mean 0 and covariance matrix Π) under which it is achieved, and otherwise there is a distribution for x under which

$$P(|x_i| \geqq 1 \text{ for some } i) = 1$$

(Theorem 3.7).

The minimizing matrix is easy to compute explicitly only in some special cases (Sec. 5). In the case $p = 2$, $k_1 = k_2 = k$, Berge lets $A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}^{-1}$, shows that A satisfies (1.3), and minimizes tr $A\Pi$ with respect to a. Following this lead, in Sec. 2 we let $A = [(1 - a)I + ae'e]^{-1}$, where $e = (1, \cdots, 1)$, show that A satisfies (1.3) for $1 > a > -1/(p - 1)$, and minimize tr $A\Pi$ with respect to a, obtaining the bound in Theorem 2.3. Though the minimum over such A is in general, except in the case $p = 2$, not the minimum over all A satisfying (1.3), it provides a useful and easily computed bound. Lal [3] considers a matrix similar in form to that of Sec. 2. However, this does not lead to the best bound, as Lal asserts, and indeed his bound is not as tight as that given in Theorem 2.3 unless $p = 2$ or $\rho_{ij} = 0$ for all $i \neq j$.

2. A multivariate inequality. We will now carry out the program of the last paragraph.

Lemma 2.1. $A = [(1 - a)I + ae'e]^{-1}$ *satisfies* (1.3) *if* $1 > a > -1/(p - 1)$.

PROOF. $A = [(1 - a)I + ae'e]^{-1} = (I - \alpha e'e)/(1 - a)$, *where* $\alpha = a/[1 + (p - 1)a]$. $x[I - \alpha e'e]x' = \sum x_i^2 - \alpha(\sum x_i)^2$

$$\geqq \begin{cases} \sum x_i^2 & \text{if } 0 \geqq a \geqq -1/(p-1), \quad \text{i.e.,} \quad \alpha \leqq 0; \\ (1 - p\alpha)\sum x_i^2 & \text{if } 0 \leqq a < 1, \quad \text{i.e.,} \quad 0 \leqq \alpha < 1/p \end{cases}$$

(The second case follows from $(\sum x_i)^2 \leqq p\sum x_i^2$.) The right-hand side becomes infinite with $\sum x_i^2$, so the minimum over all $(p - 1)$ – vectors z of

$$(1, z)(I - \alpha e'e)(1, z)'$$

occurs at a finite z. Differentiating

$$(1, z)(I - \alpha e'e)(1, z)' = 1 + \sum z_i^2 - \alpha(1 + \sum z_i)^2$$

with respect to each z_i we find that the minimizing z must satisfy $2z_i - 2\alpha(1 + \sum z_j) = 0$ for all i, or $z - \alpha ze'e - \alpha e = 0$. (Here e has $p - 1$ coordinates.) It follows that all z_i are equal, and that $\sum z_i = (p - 1)a$, so $z = ae$. Therefore the minimum over z of $(1, z)(I - \alpha e'e)(1, z)'$ is $1 - a$, and thus the minimum over z of

$$(1, z)A(1, z)'$$

is 1. The lemma follows. (See also Lemma 5.1.) || (This symbol will be used to indicate the end of a proof.)

LEMMA 2.2. *tr* $[(1 - a)I + ae'e]^{-1}\Pi$ *is minimized for* $1 > a > -1/(p - 1)$ *by*

$$a = \frac{t - \sqrt{u(pt - u)/(p - 1)}}{u - (p - 1)t}, \tag{2.1}$$

where $t = \operatorname{tr} \Pi = \sum \Pi_{ii} = \sum 1/k_i^2$ *and* $u = e\Pi e' = \sum \Pi_{ij} = \sum \rho_{ij}/(k_i k_j)$.

PROOF. $\operatorname{tr} [(1 - a)I + ae'e]^{-1}\Pi = \operatorname{tr} (I - \alpha e'e)\Pi/(1 - a) = (t - \alpha u)/(1 - a)$. The derivative of this quantity with respect to a has zeros at

$$a = \frac{t \pm \sqrt{u(pt - u)/(p - 1)}}{u - (p - 1)t}.$$

The condition $1 > a > -1/(p - 1)$ is satisfied if and only if

$$\mp\sqrt{u(pt - u)/(p - 1)}$$

is between $u/(p - 1)$ and $(pt - u)$. The upper sign is impossible because

$$u/(p - 1) \qquad \text{and} \qquad (pt - u)$$

are both positive. The lower sign is possible because $\sqrt{u(pt - u)/(p - 1)}$ is the geometric mean of $u/(p - 1)$ and $(pt - u)$. The extremum is a minimum since $(t - \alpha u)/(1 - a) \to \infty$ as $a \to 1$ or $a \to -1/(p - 1)$. ||

Substituting (2.1) in (1.6) and simplifying, we obtain, by Lemmas 1.1, 2.1, and 2.2, THEOREM 2.3. *$P(|y_i| \geqq k_i\sigma_i$ for some $i) = P(|x_i| \geqq 1$ for some $i)$*

$$\leqq \frac{p-1}{p} t - \frac{p-2}{p^2} u + \frac{2}{p^2} \sqrt{u(pt-u)(p-1)} \cdot$$
$$= [\sqrt{u} + \sqrt{(pt-u)(p-1)}]^2/p^2.$$

In the case $p = 2$, we obtain

$$P(|y_1| \geqq k_1\sigma_1 \text{ or } |y_2| \geqq k_2\sigma_2) \leqq \frac{1}{2k_1^2k_2^2}[k_1^2 + k_2^2 + \sqrt{(k_1^2+k_2^2)^2 - 4\rho^2k_1^2k_2^2}],$$

which is Lal's equation (B), and is to be compared with Berge's result, (1.2).

3. The sharpest inequality. In this section we seek the tightest bound obtainable from Lemma 1.1, and show that it is sharp, following the outline in the next-to-last paragraph of Sec. 1. What we seek, then, is the minimum of tr $A\Pi$ for A satisfying (1.3), i.e., for $A \varepsilon \mathcal{A}$. As remarked before, the minimum occurs at an extreme point of $\mathcal{A}$. We start by characterizing, in Lemma 3.2, the matrices in $\mathcal{A}$, and, in Theorem 3.3, the extreme points of $\mathcal{A}$. We use the following lemma, which has some independent interest.

LEMMA 3.1. *If A is positive definite, the minimum of xAx' for $x_1 = 1$ is $1/b_{11}$ and occurs at $(1, b/b_{11})$, and only there, where*

$$B = \begin{pmatrix} b_{11} & b \\ b' & B_{22} \end{pmatrix} = A^{-1} = \begin{pmatrix} a_{11} & a \\ a' & A_{22} \end{pmatrix}^{-1}.$$

PROOF. It is easily checked that

$$b_{11} = (a_{11} - aA_{22}^{-1}a')^{-1}, \qquad b = -b_{11}aA_{22}^{-1}, \qquad B_{22} = A_{22}^{-1} + A_{22}^{-1}a'b_{11}aA_{22}^{-1}.$$

"Completing the square," we have

$$\begin{aligned}(1, z)A(1, z)' &= a_{11} + 2az' + zA_{22}z' \\ &= a_{11} - aA_{22}^{-1}a' + (z + aA_{22}^{-1})A_{22}(z + aA_{22}^{-1})' \\ &= b_{11}^{-1} + (z - b_{11}^{-1}b)A_{22}(z - b_{11}^{-1}b)'.\end{aligned}$$

Since A_{22} is positive definite, the lemma follows. Alternatively, $(1, z)A(1, z)'$ could be differentiated with respect to each coordinate of z, as in the proof of Lemma 2.1. ||

It follows from this lemma and (1.5) that

LEMMA 3.2. *$A \varepsilon \mathcal{A}$ if and only if $B = A^{-1}$ is positive definite and $b_{ii} \leqq 1$, $i = 1, \cdots, p$.*

THEOREM 3.3. *A is extreme in $\mathcal{A}$ if and only if $B = A^{-1}$ is positive definite and $b_{ii} = 1, i = 1, \cdots, p$.*

PROOF. (i) Suppose B is positive definite and all $b_{ii} = 1$. Then, by Lemma 3.2, $A \varepsilon \mathcal{A}$. Suppose $A = (A_1 + A_2)/2$, $A_1 \varepsilon \mathcal{A}$, $A_2 \varepsilon \mathcal{A}$. For each i, by Lemma

3.1,

$$1 = 1/b_{ii} = \min_{x_i=1} xAx' \geqq \tfrac{1}{2}\,[\min_{x_i=1} xA_1 x' + \min_{x_i=1} xA_2 x'],$$

$$\min_{x_i=1} xA_1 x' \geqq 1. \qquad \min_{x_i=1} xA_2 x' \geqq 1.$$

It follows that

$$\min_{x_i=1} xA_1 x' = 1 = \min_{x_i=1} xA_2 x',$$

and the minima occur at the same point. This implies, by Lemma 3.1, that the ith row of A_1^{-1} equals the ith row of A_2^{-1}. As this is true for each i, $A_1 = A_2$. Therefore A is extreme in $\mathfrak{A}$, which proves the "if".

(ii) If B is not positive definite, $A \in \mathfrak{A}$, by Lemma 3.2. Suppose B is positive definite but $b_{ii} < 1$ for some i, say $b_{11} < 1$. Let

$$B(\delta) = B + \begin{pmatrix} \delta & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} b_{11} + \delta & b \\ b' & B_{22} \end{pmatrix}.$$

By Lemma 3.2, $B^{-1}(\delta) \in \mathfrak{A}$ for δ small enough. If we can choose $\delta_1 \neq \delta_2$ such that $B^{-1}(\delta_1) \in \mathfrak{A}$, $\quad B^{-1}(\delta_2) \in \mathfrak{A}$, and

$$A = B^{-1} = \theta B^{-1}(\delta_1) + (1 - \theta) B^{-1}(\delta_2) \tag{3.1}$$

for some θ, $0 < \theta < 1$, we will have shown that A is not extreme in $\mathfrak{A}$.

According to the first sentence of the proof of Lemma 3.1, with A and B interchanged, $B^{-1}(\delta)$ is a linear function of its upper left element $a_{11}(\delta)$, so (3.1) is equivalent to

$$a_{11} = a_{11}(0) = \theta a_{11}(\delta_1) + (1 - \theta) a_{11}(\delta_2).$$

Furthermore,

$$a_{11}(\delta) = \frac{1}{b_{11} + \delta - bB_{22}^{-1} b'} = \frac{1}{\delta + 1/a_{11}} = \frac{a_{11}}{1 + \delta a_{11}}.$$

Therefore (3.1) is equivalent to

$$\frac{\theta \delta_1}{1 + \delta_1 a_{11}} + \frac{(1 - \theta)\delta_2}{1 + \delta_2 a_{11}} = 0,$$

and it is clear that δ_1 and δ_2 can be chosen as desired. ||

This reduces the problem to that of minimizing $\operatorname{tr} B^{-1}\Pi$ for $B \in \mathfrak{B}$, where $\mathfrak{B}$ is the set of positive definite matrices with ones on the main diagonal. We will now show that $\operatorname{tr} B^{-1}\Pi$ is minimized at a unique interior point $\bar{B}$ of $\mathfrak{B}$, (Theorem 3.4), and characterize $\bar{B}$ (Theorem 3.5).

THEOREM 3.4. $\operatorname{tr} B^{-1}\Pi$ *is a strictly convex function of* B *for* $B \in \mathfrak{B}$, *and has a unique minimum, which occurs at an interior point* $\bar{B}$ *of* $\mathfrak{B}$.

PROOF. Let $B(t)$ be a straight line in $\mathfrak{B}$. Then dB/dt is a symmetric matrix,

$d^2B/dt^2 = 0$, and

$$\frac{d}{dt}\operatorname{tr} B^{-1}\Pi = -\operatorname{tr} B^{-1}\left(\frac{dB}{dt}\right)B^{-1}\Pi,$$

$$\frac{d^2}{dt^2}\operatorname{tr} B^{-1}\Pi = 2\operatorname{tr} B^{-1}\left(\frac{dB}{dt}\right)B^{-1}\left(\frac{dB}{dt}\right)B^{-1}\Pi > 0.$$

This proves the strict convexity. The rest follows, since $\mathcal{B}$ is convex and bounded, and $\operatorname{tr} B^{-1}\Pi \to \infty$ as B approaches the boundary of $\mathcal{B}$. The latter follows from the fact that

$$\operatorname{tr} B^{-1}\Pi \geqq (\operatorname{tr} B^{-1})(\text{smallest eigenvalue of } \Pi). \;\|$$

THEOREM 3.5. *$\bar{B}$ is the unique point of $\mathcal{B}$ such that $\bar{B}^{-1}\Pi\bar{B}^{-1}$, or equivalently $\bar{B}\Pi^{-1}\bar{B}$, is diagonal.*

PROOF. By Theorem 3.4, $\bar{B}$ is the unique point of $\mathcal{B}$ for which

$$\frac{d}{db_{ij}}\operatorname{tr} B^{-1}\Pi = \operatorname{tr} B^{-1}\left(\frac{dB}{db_{ij}}\right)B^{-1}\Pi = \operatorname{tr}\left(\frac{dB}{db_{ij}}\right)B^{-1}\Pi B^{-1} = 2c_{ij} = 0$$

for $i \neq j$, where $C = B^{-1}\Pi B^{-1}$, and dB/db_{ij} is a matrix with all elements zero except the (i, j)-th and (j, i)-th, which are one. ||

We note that $B^{-1}\Pi B^{-1} = C$ if and only if

$$B = \Pi^{1/2}(\Pi^{1/2}C\Pi^{1/2})^{-1/2}\Pi^{1/2} = C^{-1/2}(C^{1/2}\Pi C^{1/2})^{1/2}C^{-1/2}$$

By Theorems 3.3, 3.4, and 3.5, the tightest inequality obtainable from Lemma 1.1 is

THEOREM 3.6. *$P(|y_i| \geqq k_i\sigma_i$ for some $i) = P(|x_i| \geqq 1$ for some $i)$*

$$\leqq \operatorname{tr} \bar{B}^{-1}\Pi = \operatorname{tr} \bar{B}^{-1}\Pi\bar{B}^{-1},$$

where $\bar{B}$ is the unique positive definite matrix having ones on the main diagonal such that $\bar{B}\Pi^{-1}\bar{B}$ is diagonal.

We note that $\operatorname{tr} \bar{B}^{-1}\Pi = \operatorname{tr} (\bar{B}^{-1}\Pi\bar{B}^{-1})\bar{B} = \operatorname{tr} \bar{B}^{-1}\Pi\bar{B}^{-1}$, since $\bar{B}^{-1}\Pi\bar{B}^{-1}$ is diagonal and $\bar{B}$ has ones on the diagonal.

According to the following theorem, the bound given in Theorem 3.6 is the smallest possible bound except when the smallest possible bound is the trivial bound 1.

THEOREM 3.7. *Let $\Theta = \bar{B}^{-1}\Pi\bar{B}^{-1}$ and $\theta_1, \cdots, \theta_p$ be its diagonal elements. Then*

$$\operatorname{tr} \bar{B}^{-1}\Pi = \operatorname{tr} \bar{B}^{-1}\Pi\bar{B}^{-1} = \operatorname{tr}\Theta = \sum \theta_i .$$

If $\sum \theta_i \leqq 1$, equality holds in Theorem 3.6 if and only if

$$P(x = b^i) = P(x = -b^i) = \theta_i/2, \qquad i = 1, \cdots, p, \tag{3.2}$$
$$P(x = 0) = 1 - \sum \theta_i ,$$

where $b^1, \cdots, b^p$ *are the rows of* $\bar{B}$. *If* $\sum \theta_i > 1$, $P(|x_i| \geqq 1$ *for some* $i) = 1$ *if*

$$P(x = \sqrt{\textstyle\sum \theta_i}\, b^i) = P(x = -\sqrt{\textstyle\sum \theta_i}\, b^i) = \theta_i/(2 \textstyle\sum \theta_i), \qquad i = 1, \cdots, p. \tag{3.3}$$

PROOF. If $\sum \theta_i \leqq 1$, (3.2) is a distribution for x, and if x has this distribution, equality holds in Theorem 3.6. If x has the distribution (3.3) and $\sum \theta_i > 1$, then, with probability one, $|x_i| \geqq \sqrt{\sum \theta_i} > 1$ for some i. In either case, x has mean 0 and covariance matrix

$$E(x'x) = \sum \theta_i b^{i\prime} b^i = \bar{B}\Theta\bar{B} = \Pi.$$

This proves the "if".

It remains to prove the "only if". Suppose $\sum \theta_i \leqq 1$ and equality holds in Theorem 3.6. Then, by the relation of (1.6) to (1.4) and (1.5), with probability one,

$$x\bar{B}^{-1}x' = 1 \quad \text{if} \quad |x_i| \geqq 1 \quad \text{for some } i,$$

and

$$x\bar{B}^{-1}x' = 0 \quad \text{otherwise}.$$

It follows, by Lemma 3.1, that the distribution of x is concentrated at 0 and $\pm b^1, \cdots, \pm b^p$. Then

$$E(x) = \sum [P(x = b^i) - P(x = -b^i)]b^i.$$

But $E(x) = 0$ and $b^1, \cdots, b^p$ are linearly independent, since they are the rows of a non-singular matrix, so $P(x = b^i) = P(x = -b^i)$ for all i. Then

$$E(x'x) = \sum 2P(x = b^i)b^{i\prime}b^i = \bar{B}D\bar{B},$$

where D is a diagonal matrix with diagonal elements

$$2P(x = b^1), \cdots, 2P(x = b^p).$$

But

$$E(x'x) = \Pi, \quad \text{so} \quad D = \bar{B}^{-1}\Pi\bar{B}^{-1} = \Theta,$$

and (3.2) follows.||

4. On the solution of $\bar{B}\Theta\bar{B} = \Pi$. From $\Pi = \bar{B}\Theta\bar{B}$, we find that

$$\Pi_{ij} = \textstyle\sum_\alpha \bar{b}_{i\alpha}\theta_\alpha\bar{b}_{\alpha j},$$

and for $i = j$ we have the system of equations

$$1/k_i^2 = \sum_\alpha \bar{b}_{i\alpha}^2\theta_\alpha, \qquad i = 1, \cdots, p.$$

If we write $\bar{B} \times \bar{B} = (\bar{b}_{ij}^2)$, then

$$(\theta_1, \cdots, \theta_p) = (k_1^{-2}, \cdots, k_p^{-2})(\bar{B} \times \bar{B})^{-1}.$$

Thus given $\bar{B}$ and $k_1, \cdots, k_p$, we can solve for Θ and Π. The matrix $B \times B$ is the Hadamard product, and is positive definite if B is ([2], p. 143). Given $k_1, \cdots, k_p$, $\bar{B}$ results from some Π if and only if $\bar{B} \in \mathcal{B}$ and

$$(k_1^{-2}, \cdots, k_p^{-2})(\bar{B} \times \bar{B})^{-1}$$

has positive elements. The following example shows that this last condition is not automatically satisfied.

$$B = \begin{pmatrix} 1 & .8 & .8 \\ .8 & 1 & .5 \\ .8 & .5 & 1 \end{pmatrix}, \qquad |B \times B|(B \times B)^{-1} = \begin{pmatrix} .9375 & -.4800 & -.4800 \\ -.4800 & .5904 & .1596 \\ -.4800 & .1596 & .5904 \end{pmatrix},$$

$$k_1 = \cdots = k_p = 1.$$

Every $\bar{B} \in \mathcal{B}$ results from some $k_1, \cdots, k_p$ and Π, e.g., for

$$(k_1^{-2}, \cdots, k_p^{-2}) = (1, \cdots, 1)\bar{B} \times \bar{B}.$$

This section began with a procedure for determining Π from $\bar{B}$ by standard matrix operations. It appears that $\bar{B}$ cannot be obtained from Π by standard matrix operations except in special cases. We now give two properties of the solution (Theorems 4.1 and 4.2).

THEOREM 4.1. *If P is a permutation matrix and $P\Pi P = \Pi$, then $P\bar{B}P = \bar{B}$.*

PROOF.

$$(P\bar{B}P)\Pi^{-1}(P\bar{B}P) = P\bar{B}\Pi^{-1}\bar{B}P = P\Theta^{-1}P = \Theta^{-1}.$$

$P\bar{B}P \in \mathcal{B}$, so by the uniqueness in Theorem 3.5, $P\bar{B}P = \bar{B}$.||

THEOREM 4.2. *If* $\Pi = \begin{pmatrix} \Pi_1 & 0 \\ 0 & \Pi_2 \end{pmatrix}$, *then* $\bar{B} = \begin{pmatrix} \bar{B}_1 & 0 \\ 0 & \bar{B}_2 \end{pmatrix}$, *where* $\bar{B}_i$ minimizes

$$\operatorname{tr} \bar{B}_i \Pi_{ii}^{-} \quad in \quad \mathcal{B}_i, \qquad i = 1, 2.$$

PROOF. If $\bar{B}_i\Pi_i^{-1}\bar{B}_i$ is diagonal, $i = 1, 2$, then $\bar{B}\Pi^{-1}\bar{B}$ is diagonal, and by the uniqueness of $\bar{B}$, the conclusion follows. ||

5. Special cases.

THEOREM 5.1. *If $\Pi^{1/2}$ has equal diagonal elements, say, d, then*

$$\bar{B} = \Pi^{1/2}/d, \qquad \Theta = d^2 I$$

and

$$P(|y_i| \geqq k_i\sigma_i \textit{ for some } i) = P(|x_i| \geqq 1 \textit{ for some } i) \leqq \operatorname{tr} \bar{B}^{-1}\Pi = d^2 p.$$

This follows from Theorem 3.5. (The result for singular Π is an easy consequence of the result for non-singular Π.)

We note that $\Pi^{1/2}$ has equal diagonal elements if the group of permutation matrices P such that $P\Pi P = \Pi$ is transitive, i.e., every coordinate of x can be carried into every other one by a permutation of coordinates which preserves the covariances, i.e., $k_1 = \cdots = k_p$, and every coordinate of y can be carried into every other by a permutation of coordinates which preserves the correlations. This follows from the fact that $P\Pi^{1/2}P = \Pi^{1/2}$ if $P\Pi P = P$, since then $(P\Pi^{1/2}P)^2 = P\Pi P = \Pi$.

$\bar{B} = (1 - a)I + ae'e$, i.e., the inequality of Sec. 2 is the best possible, if and only if the elements of Π are

$$\Pi_{ii} = 1/k_i^2,$$
$$\Pi_{ij} = \rho_{ij}/k_i k_j = \frac{a}{1+a}\left[k_i^{-2} + k_j^{-2} + \frac{a(1-a)}{1+(p-1)a^2}\sum k_\alpha^{-2}\right], \tag{5.1}$$

in which case

$$P(|y_i| \geqq k_i\sigma_i \text{ for some } i) \leqq \operatorname{tr} \bar{B}^{-1}\Pi = \sum k_i^{-2}/[1 + (p-1)a^2]. \tag{5.2}$$

In the case $p = 2$, Π is always of this form and (5.2) yields (2.6).

If $k_1 = \cdots = k_p = k$, and $\Pi_{ii} = 1/k^2$, $\Pi_{ij} = \rho/k^2$, then Π is of the form (5.1) and

$$P(|y_i| \geqq k\sigma_i \text{ for some } i) \leqq \operatorname{tr} \bar{B}^{-1}\Pi$$
$$= \frac{p}{k^2[1+(p-1)a^2]} = \frac{[(p-1)\sqrt{1-\rho} + \sqrt{1+(p-1)\rho}]^2}{pk^2}. \tag{5.3}$$

This could also be obtained from Theorem 2.3, or from Theorem 5.1.

$$\Pi^{1/2} = \frac{\sqrt{1-\rho}}{k} I + \frac{[\sqrt{1+p(-1)p} - \sqrt{1-\rho}]}{kp} e'e.$$

For special values of p and ρ we obtain in addition to Berge's result (1.2), the following inequalities.

(i) $\rho = 1$: $P(|y_i| \geqq k\sigma_i \text{ for some } i) \leqq 1/k^2$, which amounts to the univariate Tchebycheff inequality.

(ii) $\rho = 0$: For p *uncorrelated* random variables,

$$P(|y_i| \geqq k_i\sigma_i \text{ for some } i) \leqq \sum k_i^{-2},$$

whereas for p *independent* random variables, the univariate Tchebycheff inequality yields the bound $1 - \prod_{i=1}^{p} (1 - k_i^{-2})$.

(iii) $\rho = -1/(p-1)$: $P(|y_i| \leqq k\sigma_i \text{ for some } i) \leqq (p-1)/k^2$.

Acknowledgements. We wish to acknowledge the very helpful comments and suggestions of W. Hoeffding and L. J. Savage.

REFERENCES

[1] P. O. Berge, "A note on a form of Tchebycheff's theorem for two variables," *Biometrika*, Vol. 29 (1937), pp. 405–406.

[2] P. R. Halmos, "Finite dimensional vector spaces," *Annals of Mathematics Studies*, No. 7, Princeton University Press, Princeton, 1947.

[3] D. N. Lal, "A note on a form of Tchebycheff's inequality for two or more variables," *Sankhyā*, Vol. 15 (1955), pp. 317–320.