MULTIVARIATE CHEBYSHEV INEQUALITIES[1]

By Albert W. Marshall and Ingram Olkin[2]

Stanford University; Michigan State University and Stanford University

1. Summary and Introduction. If X is a random variable with $EX^2 = \sigma^2$, then by Chebyshev's inequality,

$$P\{ | X | \geqq \epsilon\} \leqq \sigma^2/\epsilon^2. \tag{1.1}$$

If in addition $EX = 0$, one obtains a corresponding one-sided inequality

$$P\{X \geqq \epsilon\} \leqq \sigma^2/(\epsilon^2 + \sigma^2) \tag{1.2}$$

(see, e.g., [8] p. 198). In each case a distribution for X is known that results in equality, so that the bounds are sharp. By a change of variable we can take $\epsilon = 1$.

There are many possible multivariate extensions of (1.1) and (1.2). Those providing bounds for $P\{\max_{1 \leqq j \leqq k} | X_j | \geqq 1\}$ and $P\{\max_{1 \leqq j \leqq k} X_j \geqq 1\}$ have been investigated in [3, 5, 9] and [4], respectively. We consider here various inequalities involving (i) the minimum component or (ii) the product of the components of a random vector. Derivations and proofs of sharpness for these two classes of extensions show remarkable similarities. Some of each type occur as special cases of a general theorem in Section 3.

Bounds are given under various assumptions concerning variances, covariances and independence.

Notation. We denote the vector $(1, \cdots, 1)$ by e and $(0, \cdots, 0)$ by 0; the dimensionality will be clear from the context. If $x = (x_1, \cdots, x_k)$ and $y = (y_1, \cdots, y_k)$, we write $x \geqq y (x > y)$ to mean $x_j \geqq y_j (x_j > y_j), j = 1, 2, \cdots, k$. If $\Sigma = (\sigma_{ij}): k \times k$ is a moment matrix, for convenience we write $\sigma_{jj} = \sigma_j^2$, $j = 1, \cdots, k$. Unless otherwise stated, we assume that Σ is positive definite.

2. On Proving Sharpness. Chebyshev inequalities are usually proved by defining a non-negative function f on R^k (k-dimensional Euclidean space) such that $f(x) \geqq 1$ for all $x \in T \subset R^k$. Then if X is a k-dimensional random vector,

$$Ef(X) = \int_{\{X \varepsilon T\}} f(X)\, dP + \int_{\{X \notin T\}} f(X)\, dP \geqq \int_{\{X \varepsilon T\}} f(X)\, dP \geqq P\{X \varepsilon T\}. \tag{2.1}$$

Ordinarily, one states the inequality with some hypotheses $\mathcal{H}$ on the distribution of X (e.g., $EX'X = \Sigma$) that permit an explicit determination of the bound $Ef(X)$.

We call such an inequality sharp, if for every $\epsilon > 0$ and every value of $Ef(Z)$

Received September 1, 1959; revised July 12, 1960.

[1] This work was sponsored in part by the Office of Ordnance Research at Michigan State University and the Office of Naval Research at Stanford University.

[2] Now at the University of Minnesota.

possible under $\mathcal{H}$, there exists a random vector Z satisfying $\mathcal{H}$, with $P\{Z \varepsilon T\} \geqq Ef(Z) - \epsilon$.

Except in Section 6, the sharpness of (2.1) will follow as a consequence of the stronger result that there exists a random vector (satisfying $\mathcal{H}$) for which equality is attained.

If one is to prove (2.1) sharp by exhibiting a distribution for X attaining equality, then that distribution must assign probability only to points $x \varepsilon T$ for which $f(x) = 1$ and to points $x \not\varepsilon T$ for which $f(x) = 0$. Hence, to obtain a distribution for X achieving equality in (2.1), we begin by considering distributions that assign probability only to the rows of a matrix $\begin{pmatrix} C \\ W \end{pmatrix}$ with

$$f(c^{(i)}) = 0, \qquad f(w^{(j)}) = 1, \qquad i = 1, \cdots, m; \qquad j = 1, \cdots, n, \tag{2.2}$$

where $c^{(i)}$ is the ith row of $C: m \times k$ and $w^{(j)}$ is the jth row of $W: n \times k$. Since $f(x) \geqq 1$ for $x \varepsilon T$, (2.2) implies $c^{(i)} \not\varepsilon T$ for all i, but we still must specify that

$$w^{(j)} \varepsilon T, \qquad \text{for all } j. \tag{2.3}$$

Conditions (2.2) and (2.3) may be sufficient to define both C and W (e.g., see [4, 5]). However, if f is a quadratic form that is not positive definite (p.d.) but only positive semi-definite (p.s.d.), then $\{x: f(x) = 0\}$ is not finite and (2.2) will not define C. This means that when p.s.d. functions are used, there is no clear-cut way to find the spectrum of a distribution attaining equality.

If $P\{c^{(i)}\} = p_i$, $i = 1, 2, \cdots, m$ and $P\{w^{(j)}\} = q_j$, $j = 1, 2, \cdots, n$ then attainment of equality in (2.1) means that

$$\sum q_q = q = Ef(X), \qquad \sum p_j = 1 - q. \tag{2.4}$$

Most inequalities considered in this paper are stated with the hypotheses $EX'X = \Sigma$, i.e.,

$$C'D_pC + W'D_qW = \Sigma, \tag{2.5}$$

and sometimes also with $EX = 0$, namely,

$$eD_pC + eD_qW = 0, \tag{2.6}$$

where $D_p = \operatorname{diag}(p_1, \cdots, p_m)$ and $D_q = \operatorname{diag}(q_1, \cdots, q_n)$.

One can try to solve these equations subject to conditions (2.3) and (2.4) with the realization that (2.2) is then satisfied. Then by (2.1), such a solution satisfies $c^{(i)} \not\varepsilon T$ for all i. The above requirements may not be sufficient to define the various parameters, in which case the example attaining equality is not unique.

If T is symmetric about the origin, then we lose no generality in assuming that the distribution attaining equality is symmetric about the origin, since (2.5) and the probability assigned to T are left unchanged if C, W, D_p, and D_q

are replaced by

$$\begin{pmatrix} C \\ -C \end{pmatrix}, \quad \begin{pmatrix} W \\ -W \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} D_p & 0 \\ 0 & D_p \end{pmatrix}, \text{ and } \frac{1}{2}\begin{pmatrix} D_q & 0 \\ 0 & D_q \end{pmatrix},$$

respectively, in which case (2.6) is automatically satisfied.

3. Bounds Involving Convex Sets. If we wish the bound $Ef(X)$ to be in terms of the first and second moments then $f(x)$ must be quadratic, possibly with linear terms, i.e., $f(x) = (x - \alpha) A(x - \alpha)'$. A bound is then obtained by minimizing $Ef(X)$ subject to the conditions $f(x) \geqq 0$, $f(x) \geqq 1$ for $x \varepsilon T$. If the complement of T is bounded, then clearly these conditions are satisfied only for p.d. A. However, if T is either convex or the union of two convex sets, a minimizing A cannot be p.d. For if A is p.d., then by (2.2) $C = 0{:}\,1 \times k$. Furthermore, $\{x{:}\,f(x) \leqq 1\}$ is strictly convex (an ellipsoid) so $\{f(x) = 1\} \cap T$ has at most two points and $W{:}\,1 \times k$ or $2 \times k$. However, a three point distribution is not in general sufficient to fulfill all the conditions $EX'X = \Sigma$.

The following theorem gives conditions when a minimizing A has rank 1, so that $A = a'a$, for some $a{:}\ 1 \times k$, and the above procedure leads to sharp inequalities.

3.1. *A General Theorem.*

THEOREM 3.1. *Let* $X = (X_1, \cdots, X_k)$ *be a random vector with* $EX = 0$, $EX'X = \Sigma$. *Let* $T = T_+ \cup \{x{:}\ -x \varepsilon T_+\}$, *where* $T_+ \subseteq R^k$ *is a closed, convex set.*

(i) If $\mathfrak{A} = \{a \varepsilon R^k{:}\ (ax') \geqq 1$ *for all* $x \varepsilon T_+\}$, *then*

$$P\{X \varepsilon T\} \leqq \inf_{a \varepsilon \mathfrak{A}} a\Sigma a', \tag{3.1}$$

$$P\{X \varepsilon T_+\} \leqq \inf_{a \varepsilon \mathfrak{A}} (a\Sigma a')/(1 + a\Sigma a'). \tag{3.2}$$

(ii) *Equality in* (3.1) *can be attained whenever the bound* $\leqq 1$; *equality in* (3.2) *can always be attained.*

REMARK. If $0 \not\varepsilon T$, then $\mathfrak{A}$ is non-empty, since T and $\{0\}$ have a separating hyperplane. If $0 \varepsilon T$, then the bound one is sharp for both T and T_+, and we henceforth assume that $0 \not\varepsilon T$.

PROOF OF (i). If $a \varepsilon \mathfrak{A}$, then (3.1) and (3.2) follow from (2.1) with $f(x) = (ax')^2$ and $f(x) = (ax' + a\Sigma a')^2/(1 + a\Sigma a')^2$, respectively.

Note that the hypothesis $EX = 0$ is not required for (3.1).

To prove (ii) we need the following lemmas. We write

$$q \equiv q(a) = a\Sigma a', \qquad q^* \equiv q^*(a) = q/(1 + q), \qquad w \equiv w(a) = a\Sigma/q.$$

LEMMA 3.2. $\Sigma - qw'w$ *is* p.s.d. *and* $\Sigma - q^*w'w$ *is* p.d. (*Recall that* Σ *was assumed to be* p.d.).

PROOF. From Cauchy's inequality, for all $x \varepsilon R^k$,

$$x\Sigma x' \geqq (xw')^2/(w\Sigma^{-1}w') = q(xw')^2 \geqq q^*(xw')^2.$$

If $x \neq 0$, then strict inequality must hold in one of the two inequalities.||

LEMMA 3.3. *There exists an* $a_0 \varepsilon \mathcal{A}$ *with* $\inf a\Sigma a' = a'\Sigma a_0'$. *For such an* a_0, $w_0 = w(a_0) \varepsilon T_+$.

PROOF. Since we can obtain $\Sigma = I$ by a change of variables, it is clear that there exists (uniquely) an $a_0 \varepsilon \mathcal{A}$ with $\inf a\Sigma a' = a_0\Sigma a_0'$.

To show that $w_0 \varepsilon T_+$, assume the contrary so that there is a hyperplane separating w_0 and T_+, i.e., there is a vector v and a number α with $vw_0' < \alpha$, $vt' \geqq \alpha$ for all $t \varepsilon T_+$. If we replace v by $v + (1 - \alpha)a_0$, we can assume $\alpha = 1$. Since $vw_0' < 1$ implies $a_0\Sigma a_0' > a_0\Sigma v'$, for sufficiently small ϵ,

$$\epsilon(v\Sigma v' - 2a_0\Sigma v' + a_0\Sigma a_0') < 2(a_0\Sigma a_0' - a_0\Sigma v').$$

This is equivalent to $u\Sigma u' < a_0\Sigma a_0'$, where $u = \epsilon v + (1 - \epsilon)a_0$. But $u \varepsilon \mathcal{A}$, which is a contradiction. Hence $w_0 \varepsilon T_+$. ||

REMARK. One can also obtain $a_0\Sigma a_0'$ by computing $1/[\inf_{t\varepsilon T_+} t\Sigma^{-1}t']$; since for $t \varepsilon T_+$, $(a_0\Sigma a_0')(t\Sigma^{-1}t') \geqq (at')^2 \geqq 1$, which implies that $a_0\Sigma a_0' = 1/(w_0\Sigma^{-1}w_0') \geqq 1/(t\Sigma^{-1}t')$ for all $t \varepsilon T_+$.

PROOF OF (ii). For convenience the subscripts on $q_0 = q(a_0)$, $q_0^* = q^*(a_0)$, and $w_0 = w(a_0)$ will be omitted.

We first prove that (3.1) is sharp. Choose $r \geqq k$ and let $M: r \times k$ be such that $M'M = \Sigma - qw'w$. Choose $D = \operatorname{diag}(p_1, \cdots, p_r)$, such that $p_i > 0$, $\Sigma p_i = 1 - q$, and define $C = D^{-\frac{1}{2}}M$. Consider a random vector Z with $P\{Z = c^{(i)}\} = P\{Z = -c^{(i)}\} = p_i/2$, $(i = 1, \cdots, r)$, $P\{Z = w\} = P\{Z = -w\} = q/2$, where $c^{(i)}$ is the ith row of C. Then $EZ = 0$, $EZ'Z = C'DC + qw'w = \Sigma$.

By Lemma 3.3, $w \varepsilon T$ so $P\{Z \varepsilon T\} \geqq q$, but $P\{Z \varepsilon T\} \leqq q$ by (3.1). Hence equality is attained whenever the random variable X has the same distribution as Z.

We next prove that (3.2) is sharp. By Lemma 3.2, there exists a non-singular matrix $M: k \times k$ such that $M'M = \Sigma - q^*w'w$. Choose an orthogonal matrix $\Gamma: k \times k$ which rotates $- q^*wM^{-1}$ to the positive orthant, i.e., $-q^*wM^{-1}\Gamma > 0$. Define $D = \operatorname{diag}(p_1, \cdots, p_k)$ and C by $eD^{\frac{1}{2}} = [(p_1)^{\frac{1}{2}}, \cdots, (p_k)^{\frac{1}{2}}] = -q^*wM^{-1}\Gamma$, $C = D^{-\frac{1}{2}}\Gamma'M$. Consider a random vector Z with $P\{Z = c^{(i)}\} = p_i$, $(i = 1, \cdots, k)$, $P\{Z = w\} = q^*$, where $c^{(i)}$ is the ith row of C. Then

$$EZ = eDC + wq^* = (-q^*wM^{-1}\Gamma)(\Gamma'M) + wq^* = 0,$$

$$EZ'Z = C'DC + q^*w'w = M'M + q^*w'w = \Sigma.$$

Let us verify that $\sum p_j = 1 - q^* = 1/(1 + q)$. Noting that $w\Sigma^{-1}w' = 1/q$,

$$\sum p_i = eDe' = (q^*)^2 w(\Sigma - q^*w'w)^{-1}w' = 1/(1 + q).$$

By Lemma 3.3, $w \varepsilon T_+$, so that $P\{Z \varepsilon T_+\} \geqq q^*$. Hence by (3.2), $P\{Z \varepsilon T_+\} = q^*$, and equality in (3.2) is attained whenever X has the same distribution as Z.||

REMARK. Suppose T_+ is not convex. The following example shows that (ii) need no longer be true even when $\mathcal{A}$ is non-empty.

Let $k = 2$, $T_+ = \{x: x \geqq 0, x_1^2 + x_2^2 \geqq 1\}$, and let $T = T_+ \cup \{x: -x \,\varepsilon\, T_+\}$. $P\{X \,\varepsilon\, T\} \leqq \sigma_1^2 + \sigma_2^2$ follows from (2.1) with $f(x) = x_1^2 + x_2^2$. Now $ax' \geqq 1$ on T_+ if and only if $a_1 \geqq 1$, $a_2 \geqq 1$. But $a\Sigma a' > \sigma_1^2 + \sigma_2^2$ whenever $\sigma_{12} > 0$ and $a_1 \geqq 1$, $a_2 \geqq 1$.

3.2. *Bounds Involving the Minimum Component.*

THEOREM 3.4. *If* $X = (X_1, \cdots, X_k)$ *is a random vector with* $EX = 0$, *and if* $EX'X = \Sigma$, *then*

$$(3.3) \quad P\{X \,\varepsilon\, T\} \equiv P\{\min_j X_j \geqq 1 \text{ or } \min_j (-X_j) \geqq 1\} \leqq \min 1/(e\Sigma_s^{-1}e'),$$

$$(3.4) \quad P\{X \,\varepsilon\, T_+\} \equiv P\{\min_j X_j \geqq 1\} \leqq \min 1/(1 + e\Sigma_s^{-1}e'),$$

where the minimum is taken over all principal submatrices Σ_s *of* Σ *such that* $e\Sigma_s^{-1} > 0$.

Equality can be attained in (3.3) *whenever the bound* $\leqq 1$; *equality can always be attained in* (3.4).

There always exist principal submatrices Σ_s of Σ such that $e\Sigma_s^{-1} > 0$ (e.g., if Σ_s is 1×1) so that (3.3) and (3.4) always provide a bound.

PROOF. The theorem follows from Theorem 3.1 if we show that the bound of (3.3) is the minimum of $a\Sigma a'$ for $a \,\varepsilon\, \mathfrak{A} = \{a: a \geqq 0, ae' \geqq 1\}$.

Suppose the minimum occurs at an a whose non-vanishing co-ordinates are $\dot{a} > 0$; clearly $\dot{a}e' = 1$. Since the minimum over $\dot{a}$ does not occur on the boundary (where some component is zero), the minimizing $\dot{a}$ must satisfy $2\Sigma_s\dot{a} + \lambda e' = 0$, obtained by differentiating $\dot{a}\Sigma_s\dot{a}' + \lambda(\dot{a}e' - 1$. Using $\dot{a}e' = 1$, we obtain

$$\dot{a} = e\Sigma_s^{-1}/e\Sigma_s^{-1}e' > 0.$$

Inequalities (3.3) and (3.4) can also be obtained from (2.1) with $f(x) = (e\Sigma_s^{-1}\dot{x}')^2/(e\Sigma_s^{1-}e')^2$, and $f(x) = (e\Sigma_s^{-1}\dot{x}' + 1)^2/(e\Sigma_s^{-1}e' + 1)^2$, respectively, where $E\dot{x}'\dot{x} = \Sigma_s$.||

REMARK. If $\Sigma_1 = (\Sigma_{ij})$, $i, j = 1, 2$, and $\Sigma_2 = \Sigma_{11}$, then $e\Sigma_1^{-1}e' \geqq e\Sigma_2^{-1}e'$. Thus in order to find the bound of (3.3) or (3.4), one need not investigate all submatrices Σ_s of Σ for which $e\Sigma_s^{-1} > 0$.

Some special cases of interest for which the bounds of Theorem 3.4 can be written more explicitly are given in the following examples.

EXAMPLE 1. If $k = 2$, $\sigma_1^2 \leqq \sigma_2^2$,

$$e\Sigma_s^{-1}e' = \begin{cases} 1/\sigma_1^2, & \text{if } \sigma_1^2 \leqq \sigma_{12}, \\ (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})/(\sigma_1^2\sigma_2^2 - \sigma_{12}^2) & \text{if } \sigma_1^2 > \sigma_{12}. \end{cases}$$

EXAMPLE 2. If $\sigma_i^2 = \sigma^2$, $\sigma_{ij} = \sigma^2\rho$ $(i \neq j)$, then $e\Sigma^{-1} > 0$ and $e\Sigma^{-1}e' = k/[\sigma^2(1 + (k - 1)\rho)]$.

EXAMPLE 3. Let $\Sigma = n(D_p - p'p)$, where $D_p = \operatorname{diag}(p_1, \cdots, p_k)$, $p = (p_1, \cdots, p_k)$, $u = \sum p_i < 1$, $v = \sum p_i^{-1}$. It is easily verified that $e\Sigma^{-1} > 0$ and $e\Sigma^{-1}e' = [v + k^2/(1 - u)]/n$. If $X = (X_1, \cdots, X_k, n - \sum_1^k X_i)$ has a multinomial distribution with parameters $p_1, \cdots, p_k, 1 - u$, the covariance matrix of X is singular, but the covariance matrix of $(X_1, \cdots, X_k)$ is Σ.

EXAMPLE 4. A special form of Green's matrix $A = (a_{ij})$, $a_{ij} = a_{ji} = u_i v_j (i \leqq j)$ is given by $\Sigma = (\sigma_{ij})$ with $\sigma_{ii} = \sigma^2$, $\sigma_{ij} = \sigma^2 \prod_{m=i}^{j-1} \alpha_m$, $i < j$, and is p.d. if $\alpha_i^2 < 1$ for all i. In this case Σ^{-1} has all elements zero except on the main, super- and sub-diagonals. It can be verified that $e\Sigma^{-1} > 0$, and

$$e\Sigma^{-1}e' = \frac{1}{\sigma^2}\left(\frac{1}{1+\alpha_1} + \sum_{j=1}^{k-2} \frac{1-\alpha_j\alpha_{j+1}}{(1+\alpha_j)(1+\alpha_{j+1})} + \frac{1}{1+\alpha_{k-1}}\right).$$

In the above examples where Σ has the form $\sigma^2 R$, one can replace Σ by DRD where $D = \operatorname{diag}(\sigma_1, \sigma_2, \cdots, \sigma_k)$. Then it may no longer be that $e\Sigma^{-1} > 0$, and examples can be obtained where various submatrices Σ_s lead to the best bound. However, if $\rho = 0$ in Example 2 or $\alpha_i = 0$ in Example 4, then $\Sigma = \operatorname{diag}(\sigma_1^2, \cdots, \sigma_k^2)$ so that $e\Sigma^{-1} > 0$ and $e\Sigma^{-1}e' = \Sigma_1^k \sigma_j^{-2}$.

EXAMPLE 5. Let $Y_1, Y_2, \cdots, Y_k$ be uncorrelated random variables with $EY_j = 0$, $EY_j^2 = \tau_j^2$, $j = 1, 2, \cdots, k$, and suppose that $X_i = \sum_1^i Y_j$, $i = 1, 2, \cdots, k$, are partial sums. Then $EX_i = 0$, $i = 1, 2, \cdots, k$, and $EX'X = \Sigma = (\sigma_{ij})$ with $\sigma_{ij} = \sigma_i^2$, $i \leqq j$, where $\sigma_i^2 = \sum_1^i \tau_j^2$, so that $\sigma_1^2 \leqq \sigma_2^2 \leqq \cdots \leqq \sigma_k^2$. In this case, $e\Sigma_s^{-1} > 0$ only for Σ_s: 1×1, and $\min 1/(e\Sigma_s^{-1}e') = \sigma_1^2$.

PROOF. If $U = (u_{ij})$ is upper triangular, $u_{ij} = 0$, $i > j$, $u_{ij} = \tau_i$, $i \leqq j$, then $\Sigma = UU'$ and $e\Sigma^{-1} = (\tau_1^{-2}, 0, \cdots, 0)$. All principal submatrices of Σ are of the same form as Σ, so that $e\Sigma_s^{-1} > 0$ only when Σ_s is 1×1. ||

3.3. *Bounds for products of random variables.*

THEOREM 3.5. *If $X = (X_1, \cdots, X_k)$ is a random vector with $EX = 0$ and $EX'X = \Sigma$, then*

$$(3.5)\qquad P\{X \in T\} \equiv P\{\,|\prod X_i| \geqq 1 \text{ and } X > 0 \text{ or } X < 0\} \leqq \min_{a \in \mathfrak{A}} a\Sigma a',$$

$$(3.6)\qquad P\{X \in T_+\} \equiv P\{|\prod X_j| \geqq 1, X > 0\} \leqq \min_{a \in \mathfrak{A}} (a\Sigma a')/(1 + a\Sigma a'),$$

where $\mathfrak{A}$ is given in Theorem 3.1.

There is a unique solution a^* of

$$(3.7)\qquad a\Sigma = (a\Sigma a')a_{-1}/k,$$

with $a^* > 0$ and $\prod a_j^* = k^{-k}$, where for any vector v, $v_{-1} = (v_1^{-1}, \cdots, v_k^{-1})$. Furthermore $a^* \in \mathfrak{A}$ and $\min a\Sigma a' = a^*\Sigma a^{*\prime}$.

Equality can be attained in (3.5) *whenever the bound* $\leqq 1$; *equality can always be attained in* (3.6).

PROOF. Inequalities (3.5), (3.6), and the fact that equality can be attained all follow from Theorem 3.1.

By the remark following the proof of Lemma 3.3,

$$\inf_{\mathfrak{A}} a\Sigma a' = a^*\Sigma a^{*\prime} = 1/(w\Sigma^{-1}w') \geqq 1/(t\Sigma^{-1}t')$$

for all $t \in T_+$, where $w = a^*\Sigma/(a^*\Sigma a^{*\prime}) \in T_+$. The minimum of $t\Sigma^{-1}t'$, $t \in T_+$ occurs at $t = w$, and w must be a boundary point of T_+, so that $\prod w_i = 1$, $w > 0$. Minimization of $t\Sigma^{-1}t'$ via Lagrange's multiplier yields $2w\Sigma^{-1} = \lambda w_{-1}$.

Post multiplication by w' yields λ and the equation $kw\Sigma^{-1} = (w\Sigma^{-1}w')w_{-1}$. Since $a^* = w\Sigma^{-1}/(w\Sigma^{-1}w')$, we obtain $ka^* = w_{-1}$ so the minimizing a^* must be a solution of $a_{-1} = kw = ka\Sigma/a\Sigma a'$. Furthermore, $1 = \prod w_i = 1/\prod(ka_i^*)$ so $\prod a_i^* = k^{-k}$.

To show uniqueness, suppose there is another solution u of (3.7) with $\prod u_j = k^{-k}$, $u > 0$; then $u\Sigma u' \geqq a^*\Sigma a^{*\prime}$. Post multiplication by $\Sigma^{-1}u_{-1}'$ in (3.7) yields $u_{-1}\Sigma^{-1}u_{-1}' = k^2/(u\Sigma u')$. Using the fact that the geometric mean is dominated by the arithmetic mean and then applying Cauchy's inequality we obtain

$$k = k\prod\left(\frac{a_j^*}{u_j}\right)^{1/k} \leqq u_{-1}\,a^{*\prime} \leqq [(u_{-1}\,\Sigma^{-1}u_{-1}')(a^*\,\Sigma\,a^{*\prime})]^{\frac{1}{2}} = \left[k^2\,\frac{a^*\,\Sigma\,a^{*\prime}}{u\Sigma u'}\right]^{\frac{1}{2}} \leqq k.$$

Hence we have equality, so that $u = a^*$. ||

Corollary 3.6. *Equation* (3.7) *has one and only one solution in each orthant, subject to* $|\prod a_j| = c > 0$.

Proof. One can replace the positive orthant in the arguments of Theorem 3.5 by any other orthant. ||

We now consider two special cases for which the bounds can be given explicitly.

Example 1. If $k = 2$, then $\min a\Sigma a' = (\sigma_1\sigma_2 + \sigma_{12})/2$.

Example 2. If the column sums of Σ are all equal, then e/k is a solution of (3.7) and $\min a\Sigma a' = (e\Sigma e')/k^2$, which is equal to $\sigma^2[1 + (k-1)\rho]/k$, when $\sigma_{ii} = \sigma^2$, $\sigma_{ij} = \sigma^2\rho$.

4. Some Related Bounds.

Theorem 4.1. *If X is a random vector with $EX'X = \Sigma$, where* $\min \sigma_j^2 = \sigma_1^2$, *then*

$$P\{\min_j |X_j| \geqq 1\} \leqq \sigma_1^2, \tag{4.1}$$

$$P\{|\prod_1^k X_j| \geqq 1\} \leqq \prod_1^k \sigma_j^{2/k}, \tag{4.2}$$

$$P\{\prod_1^k X_j \geqq 1\} \leqq \prod_1^k \sigma_j^{2/k}. \tag{4.3}$$

Proof. Since $\{\min |X_j| \geqq 1\} \subset \{|X_1| \geqq 1\}$, (4.1) follows from (1.1). Successive application of (1.1) and Hölder's inequality yields

$$P\{\prod |X_j|^{1/k} \geqq 1\} \leqq E|\prod X_j|^{2/k} \leqq [\prod EX_j^2]^{1/k},$$

which is (4.2). The relation $\{\prod X_i \geqq 1\} \subset \{|\prod X_i| \geqq 1\}$ and (4.2) give (4.3). ||

We now consider the question of sharpness. Suppose that $\sigma_1^2 = \cdots = \sigma_k^2 = \sigma^2$, in which case all three inequalities can be proved by (2.1) with $f(x) = \sum_1^k x_j^2/k$. In order to satisfy (2.2) and (2.3), C: $1 \times k$ must be the zero vector, and W: $n \times k$ must be a matrix with $w_{ij} = \pm 1$.

Matrices $H = (h_{ij})$: $m \times m$ with $h_{ij} = \pm 1$ and $HH' = mI$ are called Hadamard matrices. Various sufficient conditions for their existence can be found in

[1, 6, 7]; e.g., they exist if $m = 4(\pi^r + 1)$ where π is an odd prime, r is a positive integer. A necessary condition for their existence is that $m = 2$ or $m = 4t$ for some positive integer t. If H is a Hadamard matrix, so is HD where $D = \text{diag}\,(\pm 1, \cdots, \pm 1)$. Hence we can assume that the first row of H is e. In this case all other rows $w^{(j)}$ have an equal number of positive and negative entries because $w^{(1)}w^{(j)\prime} = 0$.

From (2.5) and the fact that C consists of the zero vector we know that attainment of equality depends on the solution of $W'DW = \Sigma$. Our use below of Hadamard matrices for W stems from the fact that matrices Σ of a certain class are diagonalized by Hadamard matrices with first row e.

THEOREM 4.2. *Let* $\sigma_{ii} = \sigma^2 \leqq 1$, $\sigma_{ij} = \sigma^2\rho$, $(i \neq j)$. (i) *Equality can be attained in* (4.1) *and* (4.2) *if a Hadamard matrix of order* k *exists or* $\rho \geqq 0$. (ii) *Otherwise equality may not be attainable.*

PROOF OF (i). Any Hadamard matrix W (with first row e) of order k will diagonalize Σ; i.e.,

$$\Sigma = W'D_qW, \tag{4.4}$$

where

$D_q = \text{diag}\,(q_1, \cdots, q_k) = \text{diag}\,\{[1 + (k-1)\rho], (1-\rho), \cdots, (1-\rho)\}\sigma^2/k$.

The characteristic roots of Σ are $kq_i > 0$.

Consider the random vector Z with

$$P\{Z = 0\} = 1 - \sigma^2, \qquad P\{Z = w^{(i)}\} = P\{Z = -w^{(i)}\} = q_i/2, \qquad (i = 1, 2, \cdots, k), \tag{4.5}$$

where $w^{(i)}$ is the ith row of W. Clearly $\sum q_i = \sigma^2$, $EZ'Z = \Sigma$, and $w^{(j)} \varepsilon T$ when $T = \{\min |x_j| \geqq 1\}$ or $\{|\prod x_j| \geqq 1\}$.

If $\rho \geqq 0$, let $\Sigma^* = \sigma^2[(1-\rho)I + \rho e'e]$: $m \times m$ where $m \geqq k$ is such that a Hadamard matrix of order m exists. Σ^* is p.d. and

$$\sigma^2 = P\{\min_{1 \leqq j \leqq m} |Z_j| \geqq 1\} \leqq P\{\min_{1 \leqq j \leqq k} |Z_j| \geqq 1\} \leqq \sigma^2,$$

$$\sigma^2 = P\{\prod_1^k Z_j \geqq 1\} = P\{|\prod_1^k Z_j| \geqq 1\},$$

where the distribution of $Z = (Z_1, \cdots, Z_m)$ is given as in (4.5).

PROOF OF (ii). By $\{\min |X_j| \geqq 1\} \subseteq \{|\prod X_j| \geqq 1\}$, it is sufficient to prove that (4.2) is not necessarily sharp. Since $w_{ij} = \pm 1$, a random vector Z for which equality is attained when $k = 3$ must have a distribution of the form $P\{\pm(1, 1, 1)\} \equiv P\{(1, 1, 1) \text{ or } (-1, -1, -1)\} = p_i$, $P\{\pm(-1, 1, 1)\} = p_2$, $P\{\pm(1, -1, 1)\} = p_3$, $P\{\pm(1, 1, -1,)\} = p_4$, $P\{0\} = 1 - \sigma^2$. The conditions $EZ'Z = \Sigma$ require $p_1 = \sigma^2(1 + 3\rho)/4$, $p_2 = p_3 = p_4 = \sigma^2(1-\rho)/4$. Σ is p.d. whenever $-\frac{1}{2} < \rho < 1$ and no distribution attaining equality in (4.2) exists when $-\frac{1}{2} < \rho < -\frac{1}{3}$. ||

THEOREM 4.3. *Let* $k > 2$ *and* $\sigma_{ii} = \sigma^2 \leqq 1$, $\sigma_{ij} = \sigma^2\rho$, $(i \neq j)$. *Equality can be* *attained in* (4.3) *whenever there exists a Hadamard matrix of order* k; *otherwise, equality may not be attainable.*

REMARK. Since $\{X_1X_2 \geqq 1\} = \{|X_1X_2| \geqq 1, \text{sign } X_1 = \text{sign } X_2\}$, an improvement of (4.3) for $k = 2$ and this Σ is given by (3.5) and Example 2 of Section 3.3.

PROOF. A distribution attaining equality in (4.3) is given by (4.5). We need only show that $w^{(j)} \varepsilon \{\prod x_j \geqq 1\}$. The first row $w^{(1)}$ of W is e, and all other rows of W must have an equal number of positive and negative entries. Because k is a multiple of 4, this means $w^{(j)}$ has an even number of negative entries.

Equality cannot be attained in (4.3) if it cannot be attained in (4.2). ||

5. Bounds when only variances are known.

THEOREM 5.1. *If* $X = (X_1, \cdots, X_k)$ *is a random vector with* $EX = 0$ *and* $EX_j^2 = \sigma_j^2$, *where* $\sigma_1^2 \leqq \sigma_j^2$, $j = 1, 2, \cdots, k$, *then*

$$P\{\min_j X_j \geqq 1 \quad or \quad \min_j (-X_j) \geqq 1\} \leqq \sigma_1^2, \tag{5.1}$$

$$P\{\min_j X_j \geqq 1\} \leqq \sigma_1^2/(1 + \sigma_1^2), \tag{5.2}$$

$$P\{|\prod X_j| \geqq 1 \quad and \quad X > 0 \quad or \quad X < 0\} \leqq \prod \sigma_j^{2/k}, \tag{5.3}$$

$$P\{|\prod X_j| \geqq 1 \quad and \quad X > 0\} \leqq \prod \sigma_j^{2/k}/(1 + \prod \sigma_j^{2/k}), \tag{5.4}$$

$$P\{\min_j |X_j| \geqq 1\} \leqq \sigma_1^2, \tag{5.5}$$

$$P\{|\prod X_j| \geqq 1\} \leqq \prod \sigma_j^{2/k}, \tag{5.6}$$

$$P\{\prod X_j \geqq 1\} \leqq \prod \sigma_j^{2/k}. \tag{5.7}$$

REMARK. The hypothesis $EX = 0$ is required only for (5.2) and (5.4).

PROOF. If $T \subseteq T^* \subseteq R^k$ and $P\{X \varepsilon T^*\} \leqq p$, then trivially $P\{X \varepsilon T\} \leqq p$. In this manner, inequalities (5.1) and (5.5) follow from (1.1); (5.2) follows from (1.2); and (5.3) follows from (4.2). Inequalities (5.4), (5.6) and (5.7) follow respectively from (5.3) and Theorem 3.5, (3.11), (4.2) and (4.3). ||

THEOREM 5.2. *Equality can be attained in* (5.1)-(5.6); *equality can be attained in* (5.7) *if* $k > 1$, *whenever the bound* $\leqq 1$.

PROOF. Equality in each of (5.1)-(5.7) is achieved by one of the following distributions after a change of variable.

(i) $P\{Y = e\} = P\{Y = -e\} = \sigma^2/2, \qquad P\{Y = 0\} = 1 - \sigma^2,$

(ii) $P\{Z = e\} = \sigma^2/(1 + \sigma^2), \qquad P\{Z = -\sigma^2 e\} = 1/(1 + \sigma^2),$

(iii) $P\{U = w^{(j)}\} = \sigma^2/(2k - 2), j = 1, \cdots, k,$

$P\{U = e\} = \sigma^2(k - 2)/(2k - 2), P\{U = 0\} = 1 - \sigma^2,$
where $w^{(j)}$ is the jth row of $(2I - e'e): k \times k$.

Equality is achieved in (5.1) and (5.5) if $X_j = (\sigma_j/\sigma)Y_j$, in (5.2) if $X_j = (\sigma_j/\sigma) Z_j$, where $\sigma = \sigma_1$.

Define $\sigma^2 = \prod_1^k \sigma_j^{2/k}$. Equality is achieved in (5.3), (5.6) and for k even in (5.7) if $X_j = (\sigma_j/\sigma) Y_j$, in (5.4) if $X_j = (\sigma_j/\sigma) Z_j$, in (5.7) for k odd if $X_j = (\sigma_j/\sigma) U_j$. ||

6. Analogs of Kolmogorov's Inequality. The following theorem restates some of the previously proven inequalities with the hypotheses strengthened so that they become, in a sense, analogs of Kolmogorov's inequality. Of course, no added hypothesis can destroy the validity of an inequality, but it may destroy sharpness by permitting a better bound. For the following inequalities, we show that this is not the case.

THEOREM 6.1. *If $Y_1, \cdots, Y_k$ are mutually independent random variables with $E(Y_j) = 0$ and $E(Y_j^2) = \sigma_j^2$, $j = 1, \cdots, k$, and $X_j = \sum_1^j Y_i$, then*

$$P\{\min_j X_j \geqq 1 \quad or \quad \min_j (-X_j) \geqq 1\} \leqq \sigma_1^2, \tag{6.1}$$

$$P\{\min_j X_j \geqq 1\} \leqq \sigma_1^2/(1 + \sigma_1^2), \tag{6.2}$$

$$P\{\min_j | X_j | \geqq 1\} \leqq \sigma_1^2. \tag{6.3}$$

PROOF. (6.1), (6.2), (6.3) follow from (3.5), (3.6), and (4.1), respectively. Direct proofs are immediate since $\{\min X_j \geqq 1 \text{ or } \min -X_j \geqq 1\} \subseteq \{\min | X_j | \geqq 1\} \subseteq \{| X_1 | \geqq 1\}$, and $\{\min X_j \geqq 1\} \subseteq \{X_1 \geqq 1\}$. ||

We now show that the above inequalities are sharp. Inequalities (6.1) and (6.2) are the only inequalities that we prove sharp without showing that equality is attainable. (See Section 2 for a clarification of this distinction.) Indeed, we show that unless $\sigma_2^2 = \cdots = \sigma_k^2 = 0$, equality cannot be attained in (6.1) and (6.2) so that the probabilities of these inequalities are strictly less than the given bounds.

THEOREM 6.2. *Inequalities* (6.1), (6.2) *and* (6.3) *are sharp. Equality in* (6.1) *and* (6.2) *can be attained only if $\sigma_2 = \cdots = \sigma_k = 0$. Equality in* (6.3) *can be attained whenever the bound* $\leqq 1$.

PROOF.

Case of (6.1). Let $0 < \delta < 1$ and $Z = (Z_1, \cdots, Z_k)$ be a random vector with mutually independent components such that

$$P\{Z_1 = 1\} = P\{Z_1 = -1\} = \sigma_1^2/2, \qquad P\{Z_1 = 0\} = 1 - \sigma_1^2,$$

$$P\{Z_j = \sigma_j/\delta\} = P\{Z_j = -\sigma_j/\delta\} = \delta^2, \qquad P\{Z_j = 0\} = 1 - \delta^2, \qquad j = 2, \cdots, k.$$

Then

$$P\left\{\min_j \sum_1^j Z_i \geqq 1 \quad \text{or} \quad \min_j \left(-\sum_1^j Z_i\right) \geqq 1\right\}$$
$$> P\{Z_1 = 1 \quad \text{or } Z_1 = -1\} \prod_2^k P\{Z_j = 0\} = \sigma_1^2(1 - \delta^2)^{k-1},$$

which approaches σ_1^2 as $\delta \to 0$. Since $EZ_j = 0$ and $EZ_j^2 = \sigma_j^2$, $j = 1, \cdots, k$, (6.1) is sharp.

To attain equality in (6.1), i.e., in

$$P\{\min_j X_j \geqq 1 \quad \text{or} \quad \min_j (-X_j) \geqq 1\} \leqq P\{| X_1 | \geqq 1\} \leqq \sigma_1^2, \tag{6.4}$$

it must be that equality is attained in the right hand inequality. By (2.2), this means that if the random vector Z attains equality, $P\{Z_1 = 1 \text{ or } -1 \text{ or } 0\} = 1$, and since $EZ_1 = 0$, $P\{Z_1 = 1\} = P\{Z_1 = -1\} = \sigma_1^2$, $P\{Z_1 = 0\} = 1 - \sigma_1^2$. Suppose that i is the smallest index ($i > 1$) for which $\sigma_i^2 > 0$. Then $P\{Z_j = 0\} = 1, j = 2, \cdots, i - 1$, but Z_i must assume some value $v \neq 0$ with positive probability. If $v > 0$ and if $Z_1 = -1, Z_i = v$, then $(Z_1, \cdots, Z_k) \varepsilon T$ because $Z_1 \ngeqq 1$ and $-(Z_1 + \cdots + Z_i) \ngeqq 1$. But $P\{Z_1 = -1, Z_i = v\} = P\{Z_1 = -1\} P\{Z_i = v\} > 0$. This means equality is not attained in the left hand inequality of (6.4). A similar argument holds when $v < 0$.

Case of (6.2). Let Z be a random vector with mutually independent components such that $P\{Z_1 = 1\} = \sigma_1^2/(1 + \sigma_1^2)$, $P\{Z_1 = -\sigma_1^2\} = 1/(1 + \sigma_1^2)$, $P\{Z_j = \delta\} = \sigma_j^2/(\delta^2 + \sigma_j^2)$, $P\{Z_j = -\sigma_j^2/\delta\} = \delta^2/(\delta^2 + \sigma_j^2)$. Then if $\delta > 0$,

$$P\left\{\min_j (Z_1 + \cdots + Z_j) \geqq 1\right\} \geqq \frac{\sigma_1^2}{1 + \sigma_1^2} \prod_{j=2}^{k} \frac{\sigma_j^2}{\delta^2 + \sigma_j^2},$$

which approaches $[\sigma_1^2/(1 + \sigma_1^2)]$ as $\delta \to 0$, so that (6.2) is sharp.

The argument that equality cannot be attained in (6.2) is essentially the same as for (6.1). In this case a random vector Z attaining equality requires $P\{Z_1 = 1\} = \sigma_1^2/(1 + \sigma_1^2)$ and $P\{Z_1 = -\sigma_1^2\} = 1/(1 + \sigma_1^2)$.

Case of (6.3). Let $Z_j, j = 1, \cdots, k$ be mutually independent random variables such that $P\{Z_j = \pm 2^{j-1}\} = \sigma_j^2/2^{2i-1}$, $P\{Z_j = 0\} = 1 - \sigma_j^2/2^{2j-2}$, $(j = 1, \cdots, k)$, then $EZ_j = 0$, $EZ_j^2 = \sigma_j^2$ for all j, and $P\{\min_j | Z_1 + \cdots + Z_j | \geqq 1\} = P\{| Z_1 | \geqq 1\} = \sigma_1^2$. Hence, equality is attained in (6.3) whenever Y has the same distribution as Z.

Since $\{\min_j X_j \geqq 1 \text{ or } \min_j(-X_j) \geqq 1\} \subseteq \{\min_j | X_j | \geqq 1\}$, sharpness of (6.1) implies sharpness of (6.3), but does not imply that equality can be attained in (6.3) as we have just proved.

The inequalities of Theorem 6.1 can be obtained by specializing previous results to the case that $Y_1, \cdots, Y_k$ are uncorrelated, then adding the hypotheses that $Y_1, \cdots, Y_k$ are independent. The same procedure when applied to Lal's inequality [2] yields a bound for $P\{| Y_1 | \geqq 1 \text{ or } | Y_1 + Y_2 | \geqq 1\}$ with Y_1, Y_2 independent; in curious contrast to our case, it is not sharp since Kolmogorov's inequality provides a better bound.

7. Some Extensions. There are a number of methods by which the results can be extended. We mention only a few and give some partial results.

7.1. *Extensions to Stochastic Processes.* If $\{X_t, t \varepsilon T\}$ is a real stochastic process with $EX_t = 0$ for all $t \varepsilon T$, then

$$P\{\inf_{t\varepsilon T} | X_t | \geqq 1\} \leqq \inf_{t\varepsilon T} EX_t^2 \equiv p, \tag{7.1}$$

$$P\{\inf_{t\varepsilon T} X_t \geqq 1 \quad \text{or} \quad \sup_{t\varepsilon T} X_t \leqq -1\} \leqq p. \tag{7.2}$$

$$P\{\inf_{t\varepsilon T} X_t \geqq 1\} \leqq p/(1 + p) \tag{7.3}$$

whenever the probabilities are defined. The proof is analogous to that of Theorem 6.1.

In view of Theorem 3.4, we can hope sometimes to improve (7.2) and (7.3) if the covariance function of the process is known. General results are not easily obtained. We content ourselves with a single example for which (7.2) and (7.3) can be improved and concentrate on showing that no improvement is possible if the process is a martingale.

THEOREM 7.1. *If T is not finite and $\{X_t, t \varepsilon T\}$ is a process with $EX_t = 0$ and $EX_t^2 = \sigma^2$, $EX_sX_t = \sigma^2\rho$, $(s \neq t)$, for all $s, t \varepsilon T$, (where $0 \leqq \rho \leqq 1$), then*

$$P\{\inf_{t\varepsilon T} X_t \geqq 1 \quad \textit{or} \quad \sup_{t\varepsilon T} X_t \leqq -1\} \leqq \sigma^2\rho, \tag{7.4}$$

$$P\{\inf_{t\varepsilon T} X_t \geqq 1\} \leqq \sigma^2\rho/(1 + \sigma^2\rho) \tag{7.5}$$

whenever the probabilities are defined.

PROOF. This follows from Example 2 of Section 3.1. ||

THEOREM 7.2. *If $\{X_t, t \varepsilon T \equiv [0, \tau]\}$ is a martingale with $EX_t = 0$ and $EX_t^2 = \sigma^2(t)$, then*

$$P\{\inf_{t\varepsilon T} X_t \geqq 1 \quad \textit{or} \quad \sup_{t\varepsilon T} X_t \leqq -1\} \leqq \sigma^2(0), \tag{7.6}$$

$$P\{\inf_{t\varepsilon T} X_t \geqq 1\} \leqq \sigma^2(0)/[1 + \sigma^2(0)] \tag{7.7}$$

whenever the probabilities are defined.

Equality is attainable in both of these inequalities if $\sigma^2(\cdot)$ is right continuous.

REMARK. Inequalities (7.6) and (7.7) remain true if we replace the condition that the process is a martingale by the condition that the process has covariance $EX_sX_t = \sigma^2(s)$, $s \leqq t$ (of course, $\sigma^2(\cdot)$ must be non-decreasing). This is the case, e.g., if the process has orthogonal increments, and with this replacement the theorem would generalize Example 5 of Section 3.2. We have not chosen to weaken the conditions of the theorem because to do so would weaken the result that equality is attainable.

PROOF. (7.6) and (7.7) are immediate consequences of (7.2) and (7.3). We now indicate how a martingale $\{Z_t, t \varepsilon [0, \tau]\}$ can be defined attaining equality in (7.6). Let θ be a random time with its distribution to be chosen later. Let the sample functions Z_t be zero for $t < \theta$, and for $t \geqq \theta \neq 0$, let $Z_t = \alpha$ or $-\alpha$ witn equal probability. If $\theta = 0$, let Z_t be constant at 1 or -1 with equal probability.

Obviously Z_t is a martingale and $EZ_t = 0$. $EZ_t^2 = P\{\theta = 0\} + \alpha^2 P\{0 < \theta \leqq t\}$ and this is $\sigma^2(t)$ if θ has the distribution $P\{\theta = 0\} = \sigma^2(0)$, $P\{\theta \leqq t\} = [\sigma^2(t) - \sigma^2(0)]/\alpha^2 + \sigma^2(0)$. Choosing α so that $P\{\theta \leqq \tau\} = 1$ gives

$$\alpha = \left[\frac{\sigma^2(\tau) - \sigma^2(0)}{1 - \sigma^2(0)}\right]^{\frac{1}{2}}$$

which is less than one. Hence

$$P\{\inf_{t \varepsilon [0,\tau]} Z_t \geqq 1 \quad \text{or} \quad \sup_{t \varepsilon [0,\tau]} Z_t \leqq -1\} = P\{\theta = 0\} = \sigma^2(0),$$

so that equality is attained in (7.6).

To attain equality in (7.7), again let θ be random in $[0, \tau]$ and let $Z_t \equiv 1$ if $\theta = 0$. If $\theta \neq 0$, let $Z_t = -\sigma^2(0)$ for $t \leqq \theta$ and for $t > \theta$, let Z_t be $-\sigma^2(0) + \beta$ or $-\sigma^2(0) - \beta$ with equal probability. If $\beta = \{[1 + \sigma^2(0)][\sigma^2(\tau) - \sigma^2(0)]\}^{\frac{1}{2}}$ and θ has distribution

$$P\{\theta = 0\} = \sigma^2(0)/[1 + \sigma^2(0)], \qquad P\{\theta \leqq t\} = \frac{\sigma^2(0)}{1 + \sigma^2(0)} + \frac{\sigma^2(t) - \sigma^2(0)}{\beta^2} \quad (t \leqq \tau)$$

it is easily verified that the process Z_t has the desired properties. ||

7.2. *Extensions Through Transformations.* Let X be a random variable with $EX = 0$, $EX^2 = \sigma^2$. By means of the linear transformation $y = \eta x + \mu$, $\eta > 0$, one can obtain Chebyshev's inequality in its usual generality from (1.1) with $\epsilon = 1$.

Multivariate Chebyshev-type inequalities with hypotheses concerning means and covariances can be extended similarly by linear transformations, and, in fact, the possibilities are much greater than in the univariate case.

Let X be a random vector with $EX = 0$, $EX'X = \Sigma$, and suppose that one has the inequality

$$P\{X \varepsilon T\} \leqq p(\Sigma). \tag{7.8}$$

If H is a non-singular matrix, then using the transformation $y = xH^{-1} + \mu$ one obtains

$$P\{Y \varepsilon S\} \leqq p(H'\Pi H), \tag{7.9}$$

whenever Y is a random vector with $EY = \mu$, $EY'Y = \Pi$, and $S = \{y: (y - \mu)H \varepsilon T\}$.

Clearly, (7.9) is sharp whenever (7.8) is sharp.

Non-linear transformations may also be useful, e.g., with $Y_j = X_j^2$, $j = 1, \cdots, k$, the results of Section 5 yield corresponding results for positive random variables in terms of their expectations.

7.3. *Bounds for Subsets.* It is immediate that if $T_2 \subseteq T_1$ and (i) $P(T_1) \leqq p$, then (ii) $P(T_2) \leqq p$. Obviously, if (ii) is sharp, then (i) is sharp, and it is perhaps surprising that in many cases (ii) is sharp. Examples of this are (4.1), (5.1), and (5.2).

As a further application, let us consider inequality (3.4), but suppose that some entries of the covariance matrix Σ are unknown. Then we can consider subvectors $(X_{i_1}, \cdots, X_{i_n})$ of $(X_1, \cdots, X_k)$ for which the corresponding covariance matrix is known and apply (3.4), together with

$$P\{\min_{1 \leqq j \leqq k} X_j \geqq 1\} \leqq P\{\min_{1 \leqq j \leqq n} X_{i_j} \geqq 1\}.$$

Whether this procedure (which can also be applied to (3.3)) leads to sharp inequalities is not known, but in Section 5 we proved sharpness when only the diagonal elements of Σ are known. This procedure can be used whenever at least one diagonal element of Σ is known.

8. Acknowledgment. The authors are indeed grateful to H. Chernoff, S. Karlin, and John Pratt for suggesting proofs of Lemma 3.3, and for many valuable comments and discussions.

REFERENCES

[1] R. C. Bose and S. S. Shrikhande, "A note on a result in the theory of code construction," *Information and Control*, Vol. 2 (1959), pp. 183–194.

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

[3] Albert W. Marshall, "On the growth of stochastic processes," Technical Report No. 32, Department of Math., Univ. of Washington, 1958, pp. 1–77.

[4] Albert W. Marshall and Ingram Olkin, "A one-sided inequality of the Chebyshev type," *Ann. Math. Stat.*, Vol. 31 (1960), pp. 488–491.

[5] Ingram Olkin and John W. Pratt, "A multivariate Tchebycheff inequality," *Ann. Math. Stat.* Vol. 29 (1958), pp. 226–234.

[6] R. E. A. C. Paley, "On orthogonal matrices," *J. Math. and Phys.*, Vol. 12 (1933), pp. 311–320.

[7] R. L. Plackett and J. P. Burman, "The design of optimum multifactorial experiments," *Biometrika*, Vol. 33 (1946), pp. 305–325.

[8] J. V. Uspensky, *Introduction to Mathematical Probability*, McGraw-Hill, New York, 1937.

[9] P. Whittle, "A multivariate generalization of Tchebichev's inequality," *Quart. J. Math.*, Oxford 2nd Ser., Vol. 9 (1958), pp. 232–240.