THIS REPORT HAS BEEN DELIMITED
AND CLEARED FOR PUBLIC RELEASE
UNDER DOD DIRECTIVE 5200.20 AND
NO RESTRICTIONS ARE IMPOSED UPON
ITS USE AND DISCLOSURE.

DISTRIBUTION STATEMENT A

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

"NOTICE: When Government or other drawings, specifications or other data are used for any purpose other than in connection wi: a definitely related Government procurement operation, the U.S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulat furnished, or in any way supplied the said drawings, specifications or other data is not to be regarded by implication or otherwise a in any manner licensing the holder or any other person or corpor tion, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related there:

MULTIVARIATE CHEBYSHEV TYPE INEQUALITIES

by

ALBERT W. MARSHALL and INGRAM OLKIN

TECHNICAL REPORT NO. 52

December 3, 1959

PREPARED UNDER CONTRACT N6onr-25140

(NR 342-022)

OFFICE OF NAVAL RESEARCH

This work was sponsored by the Army, Navy, and Air Force through the Joint Services Advisory Committee for Research Groups in Applied Mathematics and Statistics by Contract No. N6onr-25140 (NR 342-002).

APPLIED MATHEMATICS AND STATISTICS LABORATORIES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

MULTIVARIATE CHEBYSHEV TYPE INEQUALITIES*

by

Albert W. Marshall and Ingram Olkin

Stanford University; Michigan State University and Stanford University

1. Summary and Introduction.

Exact distribution-theoretic results in the field of multivariate analysis are too often unknown. In order to make specific numerical probabilistic statements, one often must resort to the use of approximations or electronic computers or both.

In a number of situations the use of bounds provided by multivariate Chebyshev type inequalities may be more appropriate. These bounds have the advantage of being distribution free and may be applied even if unknowns preclude the legitimate application of distribution theory.

In this paper, inequalities involving the minimum component and the product of components of a random vector are investigated. If we are interested in whether all of k variances exceed some preassigned value, or in estimating the reliability of a system of k components whose performance critically depends on the smallest value obtained by some characteristic of the components, then the former kind of inequality may be useful. The second kind of inequality may be also useful in reliability, which is sometimes measured by the probability that a product of indicator variables exceeds a critical value. Another area of application for the second kind of inequality arises from the fact that the likelihood ratio statistic for testing certain hypotheses (e.g. , independence of sets of variates) in multivariate analysis is a product of variables and its distribution is known only in a very few cases.

---

* 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.

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

$$P\{|X| \geq \epsilon\} \leq \sigma^2/\epsilon^2 . \tag{1.1}$$

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

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

(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. Note that a change of variable permits the choice of ϵ to be unity with no loss in generality.

There are many possible multivariate extensions of (1.1) and (1.2). Those providing bounds for $P\{\max_{1 \leq j \leq k} |X_j| \geq 1\}$ and $P\{\max_{1 \leq j \leq k} X_j \geq 1\}$ have been investigated in [3,5,9] and [4], respectively. We consider here various inequalities (are considered) involving (1) the minimum component or (2) the product of components of a random vector. Derivations and proofs of sharpness for these ~~two~~ 2 classes of inequalities 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 $(1,\ldots,1)$ by e and $(0,\ldots,0)$ by 0; the dimensionality will be clear from the context. If $x = (x_1,\ldots,x_k)$ and $y = (y_1,\ldots,y_k)$, we write $x \geq y$ $(x > y)$ to mean $x_j \geq y_j$ $(x_j > y_j)$, $j = 1,2,\ldots,k$. If $\Sigma = (\sigma_{ij}) : k \times k$ is a moment matrix, for convenience we write $\sigma_{jj} = \sigma_j^2$, $j = 1,\ldots,k$.

Unless otherwise stated, we assume that Σ is positive definite.

2. On Proving Sharpness.

In many cases Chebyshev type inequalities can be proved by defining a non-negative function f on R^k (k-dimensional Euclidean space) such that $f(x) \geq 1$ for all $x \in T \subset R^k$. Then if X is a k-dimensional random vector,

$$(2.1)\qquad Ef(X) = \int_{\{X \in T\}} f(X)dP + \int_{\{X \notin T\}} f(X)dP \geq \int_{\{X \in T\}} f(X)dP \geq P\{X \in T\} .$$

The bound of (2.1) depends on the distribution of X. Ordinarily, one states the inequality with some further hypotheses $\mathcal{H}$, e.g., $EX'X = \Sigma$, in order to obtain a more explicit determination of the bound $Ef(X)$.

We call such an inequality sharp, if for every $\epsilon > 0$ and every value of $Ef(Z)$ possible under $\mathcal{H}$ there exists a random vector Z satisfying $\mathcal{H}$, with

$$P\{Z \in T\} \geq Ef(Z) - \epsilon ,$$

in which case no better bound can be given without stronger hypotheses.

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 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 $x \in T$ for which $f(x) = 1$ and to points $x \notin T$ 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

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

where $c^{(i)}$ is the i-th row of $C: m \times k$ and $w^{(j)}$ is the j-th row of $W: n \times k$. Since $f(x) \geq 1$ for $x \in T$, (2.2) implies $c^{(i)} \notin T$ for all i, but we still must specify that

$$(2.3) \qquad w^{(j)} \in T \quad \text{for all } j \ .$$

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 but only positive semi-definite, then $\{x: f(x) = 0\}$ is not finite and (2.2) will not define C. In this paper most proofs of sharpness are complicated by the fact that positive semi-definite functions are used.

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

$$(2.4) \qquad \Sigma q_j = q = Ef(x) \ , \quad \Sigma p_i = 1-q \ .$$

Let $D_p = \text{diag}(p_1,\ldots,p_m)$ and $D_q = \text{diag}(q_1,\ldots,q_n)$.

Most inequalities considered in this paper are stated with the hypotheses $E\,X'X = \Sigma$ and sometimes also with $E\,X = 0$. The condition $E\,X = 0$ is equivalent to

$$e\,D_p\,C + e\,D_q\,W = 0\,, \tag{2.5}$$

and the condition $E\,X'X = \Sigma$ is equivalent to

$$C'\,D_p\,C + W'\,D_q\,W = \Sigma\,. \tag{2.6}$$

One can try to solve equations (2.5) and (2.6) (or equation (2.6) alone when $E\,X = 0$ is not a hypothesis for (2.1)) subject to conditions (2.4) with the realization that (2.2) and (2.3) must be satisfied. These 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, it may be convenient to replace C, W, D_p and D_q by

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

respectively, in which case (2.5) is automatically satisfied and (2.6) is unchanged.

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 $f(x) \geq 0$, $f(x) \geq 1$ for $x \in T$. If the complement of T is bounded then clearly these are satisfied only for positive definite A . However, if T is either convex or the union of two convex sets, a minimizing A cannot be positive definite. For if A is positive definite, then by (2.2) $C = 0: 1 \times k$. Furthermore, $\{x: f(x) \leq 1\}$ 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 (in which case A has a representation $A = a'a$, $a: 1 \times k$) and the above procedure leads to sharp inequalities.

3.1. A General Theorem.

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

(i) If

$$(3.1) \qquad \mathcal{Q} = \{a \in R^k: ax' \geq 1 \text{ for all } x \in T_+\} ,$$

then

(3.2) $$P\{X \in T\} \leq \inf_{a \in \mathcal{Q}} a \Sigma a' ,$$

(3.3) $$P\{X \in T_+\} \leq \inf_{a \in \mathcal{Q}} \frac{a \Sigma a'}{1 + a \Sigma a'} .$$

(ii) Equality in (3.2) can be attained whenever the bound is less than one; equality in (3.3) can always be attained.

Remark. Note that if the origin is not in T, then $\mathcal{Q}$ is non-empty, since T and 0 have a separating hyperplane. If the origin is in T, then the bound one is sharp.

Proof of (i). If $a \in \mathcal{Q}$, then (3.2) and (3.3) 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 (3.2).

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

$$q = q(a) = a \Sigma a' , \quad q^* = \frac{q}{1+q} , \quad w \equiv w(a) = \frac{1}{q} a \Sigma.$$

Lemma 3.2. If Σ is positive definite, then for all $a \in \mathcal{Q}$, $\Sigma - qw'w$ is positive semi-definite and $\Sigma - q^*w'w$ is positive definite.

Proof. $w \Sigma^{-1} w' = 1/q$, and since $aw' = 1$, it follows from Cauchy's inequality that for all $x \in R^k$

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

If $x \neq 0$, then strict inequality must hold in one of the two inequalities, so that $x(\Sigma - qw'w)x' \geq 0$, $x(\Sigma - q^*w'w)x' > 0$. ‖

Lemma 3.3. If $\inf_{\mathcal{A}} a \Sigma a' = 1/q_o > 0$, then there exists $a_o \in \mathcal{A}$ such that

$$\text{(3.4)} \qquad \inf_{\mathcal{A}} a \Sigma a' = a_o \Sigma a_o' ,$$

and such that $w_o \equiv w(a_o) \in T_+$.

Proof. We show that the minimization can be viewed in terms of a game in which T_+ is one of the strategy spaces.

Suppose for the moment that T_+ is bounded and consider the game $G = (S, T_+, g)$ where $S = \{s: s \Sigma s' \leq 1\}$ and $g(s,t) = st'$. It follows [1, Theorem 2.4.3, p. 50] that G has a value v and there exist pure strategies $s_o \in S$, $t_o \in T_+$ such that

$$st_o' \leq s_o t_o' \leq s_o t'$$

for all $s \in S$, $t \in T_+$. Since $0 \in S$, $v \geq 0$. In accordance with the remark following the theorem, we assume $0 \notin T_+$. Then if $v = 0$, $0 \geq st_o'$ for all $s \in S$ implies $t_o = 0$, a contradiction, so $v > 0$. Clearly $s_o \Sigma s_o' = 1$.

We first show that $t_o = v s_o \Sigma$. Let $s^* = t_o \Sigma^{-1}/b$, where $b > 0$ is chosen so that $s^* \Sigma s^{*\prime} = 1$. By Cauchy's inequality, (3.5) $$(s^* \Sigma s_o')^2 \leq (s^* \Sigma s^{*\prime})(s_o \Sigma s_o') = 1 .$$

But

(3.6) $$b\, s^* \Sigma s_o' = s_o t_o' = v \geq s^* t_o' = b\, s^* \Sigma s^{*\prime} = b ,$$

so that $s^* \Sigma s_o' \geq 1$. Hence equality holds in (3.5) and $s^* = \alpha s_o$, where α is a scalar. Since $1 = s^* \Sigma s^{*\prime} = \alpha^2 s_o \Sigma s_o' = \alpha^2$, $\alpha = \pm 1$ and $s_o = \pm s^*$, and by (3.6) $b = \pm v$. But $b > 0$, $v > 0$, so that $s_o = s^*$. Hence, $t_o = v\, s_o \Sigma$.

Note that t_o has squared norm

$$t_o t_o' = v^2 \frac{s_o \Sigma^2 s_o'}{s_o \Sigma s_o'} \leq v^2 \rho^2 ,$$

where ρ^2 is the maximum characteristic root of Σ. If we replace T_+ by some larger set, then the value of the game cannot be increased and the best strategy for player II still has norm less than or equal to $v\rho$. From this it follows that the boundedness assumption on T_+ may be removed.

Let $a_o = s_o/v$; then $a_o \Sigma a_o' = 1/v^2$ and $a_o t' = s_o t'/v \geq 1$ whenever $t \in T_+$, so that $a_o \in \mathcal{Q}$. Using Cauchy's inequality, $t_o \Sigma^{-1} t_o' = v^2$ and $a_o t_o' \geq 1$, we obtain

$$v^2(a \Sigma a') = (t_o \Sigma^{-1} t_o')(a \Sigma a') \geq (a t_o')^2 \geq 1 .$$

Hence, $a \Sigma a' \geq 1/v^2 = a_o \Sigma a_o'$. Thus (3.4) is satisfied; furthermore $w_o = t_o = v s_o \Sigma = a_o \Sigma/q_o \in T_+$. ‖

Proof of (ii). For convenience the subscripts on $q_o = q(a_o)$, $q_o^* = q^*(a_o)$, and $w_o = w(a_o)$ will be omitted.

We first prove that (3.2) is sharp. Choose $r \geq k$ and let $M: r \times k$ be such that

$$M' M = \Sigma - q w'w .$$

Choose $D = \text{diag}(p_1,\ldots,p_r)$, such that $p_i > 0$, $\Sigma p_i = 1-q$, and define $C = D^{-1/2} M$. Consider a random vector Z with

$$P(Z = c^{(i)}) = P(Z = - c^{(i)}) = p_i/2 , \qquad i = 1,\ldots,r ,$$

$$P(Z = w) = P(Z = - w) = q/2 ,$$

where $c^{(i)}$ is the i-th row of C. Then

$$EZ = 0 , \quad EZ'Z = C' DC + q w'w = \Sigma .$$

By (2.2), $c^{(j)} \notin T$ if $a C' = 0$. But this holds since $aw' = 1$ and

$$a C' D C a' = a(\Sigma - q w'w)a' = 0 .$$

By Lemma 3.3, $w \in T$. Hence, $P\{Z \in T\} = 1$ and equality in (3.2) is attained whenever $P\{X = Z\} = 1$.

We next prove that (3.3) 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 $-q^* w M^{-1}$ to the positive orthant, i.e., $-q^* w M^{-1} \Gamma > 0$. Define $D = \mathrm{diag}(p_1,\ldots,p_k)$ and C by

$$c D^{1/2} = (\sqrt{p_1},\ldots,\sqrt{p_k}) = -q^* w M^{-1} \Gamma , \qquad C = \Gamma' M D^{-1/2} .$$

Consider a random vector Z with

$$P\{Z = c^{(i)}\} = p_i , \qquad i = 1,\ldots,k , \qquad P\{Z = w\} = q^* ,$$

where $c^{(i)}$ is the i-th row of C. Then

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

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

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

$$\Sigma p_i = e D e' = q^{*2} w M^{-1} M M'^{-1} w' = q^{*2} w(\Sigma - q^* w'w)^{-1} w'$$

$$= q^{*2} w\left(\Sigma^{-1} + \frac{q^* \Sigma^{-1} w'w \Sigma^{-1}}{1 - q^* w \Sigma^{-1} w'}\right) w' = 1/(1+q) .$$

By Lemma 3.2, $w \in T_+$. Since $(a C' + eq) D(C a' + qe') = 0$, $aC' = -eq$, and $c^{(i)} \notin T_+$. Hence, $P\{Z \in T_+\} = q^*$, and equality in (3.2) is attained whenever $P\{X = Z\} = 1$. ||

<u>Remark</u>. Suppose T_+ is not convex. The following example shows that (ii) need no longer be true even when $\mathcal{Q}$ is non-empty.

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

3.2. <u>Bounds Involving the Minimum Components.</u>

<u>Theorem 3.4</u>. If $X = (X_1, \ldots, X_k)$ is a random vector with $EX = 0$ and $EX'X = \Sigma$ (Σ positive definite), then

$$(3.7) \qquad P\{\min_{1 \leq j \leq k} X_j \geq 1 \text{ or } \min_{1 \leq j \leq k} (-X_j) \geq 1\} \leq \min_s \frac{1}{e \Sigma_s^{-1} e'} ,$$

$$(3.8) \qquad P\{\min_{1 \leq j \leq k} X_j \geq 1\} \leq \min_s \frac{1}{1 + e \Sigma^{-1} e'} ,$$

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

Proof. Let $X = (X_{i_1}, \ldots, X_{i_s})$ be such that $EX'X = \Sigma_s$. Consider the function

$$(3.9) \qquad f(x) = \frac{(e\,\Sigma_s^{-1}\,x')^2}{(e\,\Sigma_s^{-1}\,e')^2} .$$

Clearly $f(x) \geq 0$ for all x and $f(x) \geq 1$ for $x \in T = \{\min_j x_j \geq 1$ or $\min_j (-x_j) \geq 1\}$. Since $Ef(X) = 1/(e\,\Sigma_s^{-1}\,e')$, we obtain (3.7) from (2.1). To obtain (3.8), replace $f(x)$ by $(e\,\Sigma_s^{-1}\,x' + 1)^2/(e\,\Sigma_s^{-1}\,e' + 1)^2$ and T by $\{\min_j x_j \geq 1\}$ in the above proof. ∎

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 (3.7) and (3.8) always provide a bound.

Theorem 3.5. Equality in (3.7) and (3.8) can be attained whenever the bound is less than one.

Proof. The function $f(x)$ of (3.9) is of the form $(ax')^2$. In order to apply Theorem 3.1 we must show that the bound (3.7) is obtained by minimizing $q = a\,\Sigma\,a' = E(a\,X')^2$ subject to the restriction that $a \in \mathcal{A}$. Here $T_+ = \{\min_j x_j \geq 1\}$ and $\mathcal{A} = \{a\colon ax' \geq 1$ for $x \in T_+\} = \{a\colon a \geq 0,\ ae' \geq 1\}$. By Cauchy's inequality

$$a\,\Sigma\,a' \geq (ae')^2/\,(e\,\Sigma^{-1}\,e') ,$$

and $\min_{\mathcal{A}} a\,\Sigma\,a' = \min_{ae'=1} a\,\Sigma\,a' = 1/(e\,\Sigma^{-1}\,e')$ occurs at $a = e\,\Sigma^{-1}/\,(e\,\Sigma^{-1}\,e')$.

If $e\Sigma^{-1} \not> 0$, then $\min_{\mathcal{a}} a\Sigma a'$ occurs on the boundary of $\mathcal{a}$, i.e., where some components of a are zero.

Suppose $a_j = 0$ when $j \notin \{i_1,\ldots,i_s\}$, and let $\dot{a} = (a_{i_1},\ldots,a_{i_s})$, $\dot{x} = (x_{i_1},\ldots,x_{i_s})$. If $E\dot{X}'\dot{X} = \Sigma_s$,

$$a\Sigma a' = \dot{a}\Sigma_s\dot{a}' \geq (\dot{a}e')^2/(e\Sigma_s^{-1}e'),$$

with equality if $\dot{a} = e\Sigma_s^{-1}/(e\Sigma_s^{-1}e')$ in which case $Ef(X) = 1/(e\Sigma_s^{-1}e')$. Thus

$$\min_{\mathcal{a}} a\Sigma a' = \min_{s} 1/(e\Sigma_s^{-1}e') . \blacksquare$$

Remark. If Σ_2 is a submatrix of Σ_1, then $e\Sigma_1^{-1}e' \geq e\Sigma_2^{-1}e'$. Thus in order to find the bound of (3.7) or (3.8), one need not investigate all submatrices Σ_s of Σ for which $e\Sigma_s^{-1} > 0$.

To prove the inequality let

$$\Sigma_1 = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad \text{and} \quad \Sigma_2 = \Sigma_{11} .$$

Then $a\Sigma a' \geq 1/(e\Sigma_1^{-1}e')$ by Cauchy's inequality and the condition $ae' = 1$. But $a\Sigma a' = 1/(e\Sigma_{11}^{-1}e')$ when $a = (e\Sigma_{11}^{-1}/(e\Sigma_{11}^{-1}e'), 0)$.

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 $k = 2$, $\sigma_1^2 \le \sigma_2^2$,

$$\min_s \frac{1}{e\Sigma_s^{-1}e'} = \begin{cases} \sigma_1^2 & \text{if } \sigma_1^2 \le \sigma_{12}, \\ \dfrac{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} & \text{if } \sigma_1^2 > \sigma_{12}. \end{cases}$$

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

Proof. $\Sigma^{-1} = (I - \alpha e'e)/[\sigma^2(1-\rho)]$, where $\alpha = \rho[1 + (k-1)\rho]^{-1}$. Since the column sums of Σ^{-1} are all equal, and $e\Sigma^{-1}e' > 0$, it follows that $e\Sigma^{-1} > 0$. The computation of $e\Sigma^{-1}e'$ is immediate. ∎

Example 3. Let $\Sigma = n(D - p'p)$, where $D = \mathrm{diag}(p_1,\ldots,p_k)$, $p = (p_1,\ldots,p_k)$, $\sum_1^k p_i < 1$. It is easily verified that $\Sigma^{-1} = \frac{1}{n}(D^{-1} + \frac{e'e}{1-ep'})$, $e\Sigma^{-1} > 0$ and

$$e\Sigma^{-1}e' = \frac{1}{n}\left[\sum_1^k \frac{1}{p_i} + \frac{(k-1)^2}{1 - \sum_1^k p_i}\right].$$

If the random vector $X = (X_1,\ldots,X_{k+1})$, $\sum_1^{k+1} X_i = n$, has a multinomial distribution with parameters $p_1,\ldots,p_{k+1}$, the covariance matrix of X is singular, but the covariance matrix of $(X_1,\ldots,X_k)$ is Σ.

Example 4. A special form of Green's matrix $A = (a_{ij})$, $a_{ij} = a_{ji} = u_i v_j$ $(i \le j)$ is given by

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \alpha_1 & \alpha_1\alpha_2 & \cdots & & \prod_1^{k-1}\alpha_j \\ \alpha_1 & 1 & \alpha_2 & \cdots & & \prod_2^{k-1}\alpha_j \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \prod_1^{k-2}\alpha_j & \prod_2^{k-2}\alpha_j & & \cdots & 1 & \alpha_{k-1} \\ \prod_1^{k-1}\alpha_j & \prod_2^{k-1}\alpha_j & & \cdots & \alpha_{k-1} & 1 \end{pmatrix}$$

and is positive definite if $\alpha_j^2 < 1$ for all j. In this case Σ^{-1} has main diagonal

$$\frac{1}{\sigma^2}\left(\frac{1}{1-\alpha_1^2}, \frac{1-\alpha_1^2\alpha_2^2}{(1-\alpha_1^2)(1-\alpha_2^2)}, \ldots, \frac{1-\alpha_{k-2}^2\alpha_{k-1}^2}{(1-\alpha_{k-2}^2)(1-\alpha_{k-1}^2)}, \frac{1}{1-\alpha_{k-1}^2}\right),$$

super and sub diagonal $\frac{1}{\sigma^2}\left(-\frac{\alpha_1}{1-\alpha_1^2}, \ldots, -\frac{\alpha_{k-1}}{1-\alpha_{k-1}^2}\right)$, and all other elements are zero. It is easily verified that $e\,\Sigma^{-1} > 0$, and

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

which takes the form $[k(1-\alpha)+2\alpha]/[\sigma^2(1+\alpha)]$ if $\alpha_j = \alpha$ for all j. ||

In the above examples where Σ has the form $\sigma^2 R$, one can replace Σ by DRD where $D = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k)$. 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 $\Sigma = \mathrm{diag}(\sigma_1^2,\ldots,\sigma_k^2)$ so that $e\Sigma^{-1} > 0$ and $e\Sigma^{-1}e' = \sum_1^k \frac{1}{\sigma_j^2}$.

Example 5. Let $Y_1, Y_2,\ldots,Y_k$ be uncorrelated random variables with $EY_j = 0$, $EY_j^2 = \tau_j^2$, $j = 1,2,\ldots,k$, and suppose that $X_i = \sum_1^i Y_j$, $i = 1,2,\ldots,k$, are partial sums. Then $EX_i = 0$, $i = 1,2,\ldots,k$, and

$$EX'X = \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \cdots & \sigma_1^2 \\ \sigma_1^2 & \sigma_2^2 & \sigma_2^2 & \cdots & \sigma_2^2 \\ \sigma_1^2 & \sigma_2^2 & \sigma_3^2 & \cdots & \sigma_3^2 \\ & & \cdots & & \\ \sigma_1^2 & \sigma_2^2 & \cdots & \cdots & \sigma_k^2 \end{pmatrix},$$

where $\sigma_i^2 = \sum_1^i \tau_j^2$, so that $\sigma_1^2 \le \sigma_2^2 \le \cdots \le \sigma_k^2$. In this case, $e\Sigma_s^{-1} > 0$ only for $\Sigma_s : 1 \times 1$, and $\min_s 1/(e\Sigma_s^{-1}e') = \sigma_1^2$.

Proof: If T is the upper triangular matrix

$$T = \begin{pmatrix} \tau_1 & \tau_1 & \cdots & \tau_1 \\ & \tau_2 & \cdots & \tau_2 \\ 0 & & \ddots & \\ & & & \tau_k \end{pmatrix},$$

then $\Sigma = TT'$. Since T^{-1} has main diagonal $\{\frac{1}{\tau_1}, \frac{1}{\tau_2}, \ldots, \frac{1}{\tau_k}\}$, super diagonal $\{-\frac{1}{\tau_2}, -\frac{1}{\tau_3}, \ldots, -\frac{1}{\tau_k}\}$ and all other entries zero, $e\,T^{-1} = (\tau_1^{-1}, 0, \ldots, 0)$, and $e\,\Sigma^{-1} = (\tau_1^{-2}, 0, \ldots, 0)$. 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. Thus $\min \dfrac{1}{e\,\Sigma_s^{-1}\,e'} = \sigma_1^2$. ‖

3.3. Bounds for the product of random variables.

Theorem 3.6. If $X = (X_1, \ldots, X_k)$ is a random vector with $EX = 0$ and $EX'X = \Sigma$, then

$$P\{|\prod X_j| \geq 1 \text{ and } X > 0 \text{ or } X < 0\} \leq \min_{a \in \mathcal{A}} a\,\Sigma\,a' , \tag{3.10}$$

$$P\{|\prod X_j| \geq 1, X > 0\} \leq \min_{a \in \mathcal{A}} \frac{a\,\Sigma\,a'}{1 + a\,\Sigma\,a'} , \tag{3.11}$$

where $\mathcal{A}$ is given by (3.1).

There is a unique solution a^* of

$$a\,\Sigma = \frac{a\,\Sigma\,a'}{k} \left(\frac{1}{a}\right) , \tag{3.12}$$

$(\frac{1}{a}) = (\frac{1}{a_1}, \ldots, \frac{1}{a_k})$, with $a^* > 0$ and $\prod a_j^* = k^{-k}$. Furthermore, $\min_{a \in \mathcal{A}} a\,\Sigma\,a' = a^*\,\Sigma\,a^{*\prime}$.

Proof. Inequalities (3.10) and (3.11) follow from Theorem 3.1.

To prove the second part of the theorem, we begin by showing that

$\mathcal{Q} = \mathcal{Q}_1 = \{a: a \geq 0, \prod a_j \geq k^{-k}\}$ where $\mathcal{Q}$ is given by (3.1) with $T_+ = \{|\prod x_j| \geq 1, x > 0\}$. If $a \in \mathcal{Q}_1, \prod x_j \geq 1$ and $x > 0$, then using $\Sigma a_j x_j / k \geq \prod (a_j x_j)^{1/k}$ we obtain

$$ax' \geq ax'/(\prod x_j)^{1/k} \geq k \prod a_j^{1/k} \geq 1 ,$$

and $\mathcal{Q}_1 \subseteq \mathcal{Q}$. If $a \in \mathcal{Q}$ then $a \geq 0$; for if $a_i < 0, \prod x_j \geq 1$ and $x > 0$, then one can increase x_i and thereby decrease ax' while preserving $\prod x_j \geq 1$. Furthermore, $a \in \mathcal{Q}$ implies that $\prod a_j \geq k^{-k}$. For if we suppose the contrary, then since $x = \prod a_j^{1/k}(\frac{1}{a})$ satisfies $x > 0, \prod x_j \geq 1$, we obtain the contradiction $ax' = k \prod a_j^{1/k} < 1$. Thus $\mathcal{Q} \subseteq \mathcal{Q}_1$, and we conclude that $\mathcal{Q} = \mathcal{Q}_1$.

We wish to replace $\mathcal{Q}$ by a bounded set. Denote the minimum (maximum) characteristic root of Σ by ρ_m (ρ_M), choose $\varphi > 1$ and define $\beta = \varphi \rho_M / k \rho_m$, $\mathcal{Q}^* = \mathcal{Q} \cap \{a: aa' \leq \beta\}$. If $aa' \geq \beta$, then since $xx' \rho_m \leq x \Sigma x' \leq xx' \rho_M$ for all $x \in R^k$,

$$\text{(3.13)} \quad \frac{e}{k} \Sigma \frac{e'}{k} \leq \frac{1}{k} \rho_M < \beta \frac{a \Sigma a'}{aa'} \leq a \Sigma a' .$$

But $e/k \in \mathcal{Q}$ so that $\inf_{\mathcal{Q}} a \Sigma a' = \inf_{\mathcal{Q}^*} a \Sigma a'$.

Since $\mathcal{Q}^*$ is compact, there exists $a^* \in \mathcal{Q}^*$ for which $\inf_{\mathcal{Q}^*} a \Sigma a' = a^* \Sigma a^*$. By considering multiples of a^*, one obtains $\prod a_j^* = k^{-k}$.

The above arguments show that the problem reduces to that of finding

$\inf_{a \geq 0,\ aa' \leq \beta} a \Sigma a'$ subject to the condition $\prod a_j = k^{-k}$. This infimum does not occur on the boundary of $\{a \geq 0 , aa' \leq \beta\}$, because the condition $\prod a_j = k^{-k}$ does not allow $a_j = 0$, and strict inequality in (3.13) does not allow $aa' = \beta$. Thus we conclude via Lagrange's multiplier that

$$a \Sigma = \lambda \prod a_j \left(\frac{1}{a}\right) , \quad \prod a_j = k^{-k}$$

must have a simultaneous solution a_o such that $a_o > 0$, $a_o a_o' < \beta$. Furthermore the desired minimizer a^* is among such solutions. Post multiplication of the first equation above by a' yields $\lambda = a \Sigma a' / (k \prod a_j)$, so that (3.12) is obtained.

Suppose there is another solution u of (3.12) with $\prod u_j = k^{-k}$, $u > 0$; then $u \Sigma u' \geq a^* \Sigma a^{*\prime}$. Post multiplication by $\Sigma^{-1}(\frac{1}{u})'$ in (3.12) yields

$$\left(\frac{1}{u}\right) \Sigma^{-1} \left(\frac{1}{u}\right)' = \frac{k^2}{u \Sigma u'} .$$

Using the fact that the geometric mean is dominated by the arithmetic mean and then applying Cauchy's inequality we obtain

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

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

Corollary 3.7. Equation (3.12) 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 3.6 by any other orthant, and so conclude that (3.12) has a unique solution in every orthant. ‖

Theorem 3.8. Equality in (3.10) and (3.11) can be attained whenever the bound is less than one.

Proof. This is a special case of (ii), Theorem 3.1.

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

Example 1. If $k = 2$, then

$$(3.14) \qquad \min_{a} a \Sigma a' = (\sigma_1\sigma_2 + \sigma_{12})/2 .$$

Proof. In this case the direct solution of (3.12) together with $a > 0$ and $a_1a_2 = 2^{-2}$ yields $a_1 = \sqrt{\sigma_2/4\sigma_1}$, $a_2 = \sqrt{\sigma_1/4\sigma_2}$.

Example 2. If the column sums of Σ are all equal, then $\min_{a} a \Sigma a' = (e \Sigma e')/k^2$. In particular if $\Sigma = \sigma^2[(1-\rho)I + \rho e'e]$ then $\min_{a} a \Sigma a' = \sigma^2[1+(k-1)\rho]/k$.

Proof. Equality of the column sums of Σ means that

$$e\Sigma = \frac{e\,\Sigma\,e'}{k}\;e = \frac{e\,\Sigma\,e'}{k}\left(\frac{I}{e}\right) \quad ,$$

and hence e/k is the desired solution of (3.12).

4. Some Related Bounds

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

$$(4.1) \qquad P\left\{\min_j |X_j| \geq 1\right\} \leq \sigma_1^2 \quad ,$$

$$(4.2) \qquad P\left\{\left|\prod_1^k X_j\right| \geq 1\right\} \leq \prod_1^k \sigma_j^{2/k} \quad ,$$

$$(4.3) \qquad P\left\{\prod_1^k X_j \geq 1\right\} \leq \prod_1^k \sigma_j^{2/k} \quad .$$

<u>Proof.</u> Since $\left\{\min_j |X_j| \geq 1\right\} \subset \left\{|X_1| \geq 1\right\}$, (4.1) follows from (1.1). Successive application of (1.1) and Holder's inequality yields

$$P\left\{\left|\prod X_j\right| \geq 1\right\} = P\left\{\left|\prod X_j\right|^{1/k} \geq 1\right\} \leq E\left|\prod X_j\right|^{2/k}$$

$$\leq [\prod EX_j^2]^{1/k} = \prod \sigma_j^{2/k} \,,$$

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

We now consider the question of sharpness. Suppose that $\sigma_1^2 = \dots = \sigma_k^2 = \sigma^2$, 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), $C:1 \times k$ 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 [2, 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, then so is HD_ϵ, where $D_\epsilon = \text{diag}(\pm 1, \dots, \pm 1)$. Hence we can assume that the first row of H is e.

From (2.6) and the fact that C consists of the zero vector we know that the 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 whose first row is e.

<u>Theorem</u> 4.2. Let $\Sigma = \sigma^2[(1-\rho)I + \rho e'e]$.

(i) Equality can be attained in (4.1) and (4.2) if a Hadamard matrix of order k exists or if $\rho \geq 0$.

(ii) Otherwise equality may not be attainable.

<u>Proof of (i)</u>. Because of the form of Σ, any Hadamard matrix W of order k will diagonalize Σ; i.e.,

$$\Sigma = \frac{W'}{\sqrt{k}} D_\rho \frac{W}{\sqrt{k}} = W'D_q W , \tag{4.4}$$

where $D_q = \text{diag}(q_1, \ldots, q_k) = \text{diag}\left\{[1+(k-1)\rho], (1-\rho), \ldots, (1-\rho)\right\} \sigma^2/k$. The characteristic roots of Σ are $kq_i > 0$.

Consider the random vector Z with

$$\begin{aligned} &P\left\{Z = (0, \ldots, 0)\right\} = 1-\sigma^2 , \\ &P\left\{Z = w^{(i)}\right\} = P\left\{Z = -w^{(i)}\right\} = q_i/2 , \ i=1, 2, \ldots, k , \end{aligned} \tag{4.5}$$

where $w^{(i)}$ is the i^{th} row of W. Clearly $\Sigma q_i = \sigma^2$, $EZ'Z = \Sigma$, and $w^{(i)} \in T$ when $T = \left\{\min |x_j| \geq 1\right\}$ or $\left\{|\prod x_j| \geq 1\right\}$.

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

$$\sigma^2 = P\left\{\min_{1 \leq j \leq m} |Z_j| \geq 1\right\} \leq P\left\{\min_{1 \leq j \leq k} |Z_j| \geq 1\right\} \leq \sigma^2 ,$$

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

<u>Proof of (ii)</u>. By $\left\{\min |X_j| \geq 1\right\} \subseteq \left\{|\prod X_j| \geq 1\right\}$, 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

$$\begin{aligned} &P\left\{\pm(1, 1, 1)\right\} = p_1 , \ P\left\{\pm(-1, 1, 1)\right\} = p_2 , \ P\left\{\pm(1, -1, 1)\right\} = p_3 , \\ &P\left\{\pm(1, 1, -1)\right\} = p_4 , \ P\left\{(0, 0, 0)\right\} = 1-\sigma^2 . \end{aligned}$$

The condition $EZ'Z = \Sigma$ lead to the solution

$$p_1 = \sigma^2(1+3\rho)/4 \ , \ p_2 = p_3 = p_4 = \sigma^2(1-\rho)/4 \ .$$

Σ is positive definite whenever $-1/2 < \rho < 1$ and no distribution attaining equality in (4.2) exists when $-1/2 < \rho < -1/3$. ∎

Theorem 4.3. Let $k > 2$ and $\Sigma = \sigma^2[(1-\rho)I + \rho e'e]$. Equality can be attained in (4.3) whenever there exists a Hadamard matrix of order k; otherwise, equality may not be attainable.

Remark. Since $\{X_1X_2 \geq 1\} = \{|X_1X_2| \geq 1 \text{ , sign } X_1 = \text{sign } X_2\}$, an improvement of (4.3) for $k = 2$ is given by (3.10) and (3.14).

Proof. A distribution attaining equality in (4.3) is given by (4.5). We need only show that $w^{(j)} \in \{\prod x_j \geq 1\}$. The first row $w^{(1)}$ of W is e. All other rows of W must have an equal number of positive and negative entries because $w^{(1)}w^{(j)'} = 0$. Because k is a multiple of 4, this means that $w^{(j)}$ 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 $X = (X_1, \dots, X_k)$ is a random vector with $EX_j^2 = \sigma_j^2$, where $\sigma_1^2 \le \sigma_j^2$, $j = 1, 2, \dots, k$, then

$$(5.1) \qquad P\left\{\min_j X_j \ge 1 \text{ or } \min_j (-X_j) \ge 1\right\} \le \sigma_1^2 ,$$

$$(5.2) \qquad P\left\{\min_j X_j \ge 1\right\} \le \sigma_1^2/(1+\sigma_1^2) ,$$

$$(5.3) \qquad P\left\{|\prod X_j| \ge 1 \text{ and all } X_j \text{ are of the same sign}\right\} \le \prod \sigma_j^{2/k} ,$$

$$(5.4) \qquad P\left\{|\prod X_j| \ge 1 \text{ and all } X_j > 0\right\} \le \prod \sigma_j^{2/k}/(1+\prod \sigma_j^{2/k}) ,$$

$$(5.5) \qquad P\left\{\min_j |X_j| \ge 1\right\} \le \sigma_1^2 ,$$

$$(5.6) \qquad P\left\{|\prod X_j| \ge 1\right\} \le \prod \sigma_j^{2/k} ,$$

$$(5.7) \qquad P\left\{\prod X_j \ge 1\right\} \le \prod \sigma_j^{2/k} .$$

Proof. If $T \subseteq T^* \subseteq R^k$ and $P\{X \in T^*\} \le p$, then trivially $P\{X \in T\} \le p$. Inequalities (5.1) and (5.5) follow from (1.1), and (5.2) follows from (1.2) in this manner. Inequalities (5.3), (5.4), (5.6) and (5.7) follow respectively from (3.10), (3.11), (4.2) and (4.3). ||

Theorem 5.2. Equality in (5.1) - (5.7) can be attained.

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

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

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

(iii) $P\{U = w^{(j)}\} = \sigma^2/k$, $j = 1, \ldots, k$, $P\{U = 0\} = 1-\sigma^2$,

where $w^{(j)}$ is the j^{th} row of $(2I - ee'):k \times k$.

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

Define $\sigma^2 = \prod_1^k \sigma_j^{2/k}$ in (i) - (iii) . Equality is achieved in (5.3), (5.6) and for k even in (5.7) if $X_j = (\sigma_j/\sigma)\, Z_j$, in (5.4) if $X_j = (\sigma_j/\sigma)\, Y_j$, 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 $Y_1,\dots,Y_k$ are mutually independent random variables with $EY_j = 0$ and $EY_j^2 = \sigma_j^2$, $\sigma_1^2 \le \sigma_j^2$, $j = 1,\dots,k$, then

$$P\{\min_j (Y_1+\dots+Y_j) \ge 1 \text{ or } \min_j [-(Y_1+\dots+Y_j)] \ge 1\} \le \sigma_1^2 , \tag{6.1}$$

$$P\{\min_j (Y_1+\dots+Y_j) \ge 1\} \le \sigma_1^2/(1+\sigma_1^2) , \tag{6.2}$$

$$P\{\min_j |Y_1+\dots+Y_j| \ge 1\} \le \sigma_1^2 . \tag{6.3}$$

Proof. (6.1), (6.2) and (6.3) are special cases of (3.7), (3.8) and (4.1), respectively. Direct proofs are immediate since

$$\{\min(Y_1+\dots+Y_j) \ge 1 \text{ or } \min[-(Y_1+\dots+Y_j)] \ge 1\}$$

$$\subseteq \{\min|Y_1+\dots+Y_j| \ge 1\} \subseteq \{|Y_1| \ge 1\} ,$$

and

$$\{\min(Y_1+\dots+Y_j) \ge 1\} \subseteq \{Y_1 \ge 1\} . \blacksquare$$

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 always be attained.

Proof. Case of (6.1). Choose ϵ, $0 < \epsilon < \sigma_1^2$, and let $\delta = [1-(1-\epsilon/\sigma_1^2)^{1/(k-1)}]^{\frac{1}{2}}$. Let $Z = (Z_1,\ldots,Z_k)$ be a random vector with mutually independent components such that

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

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

Setting $T = \{\min_j (y_j+\cdots+y_j) \geq 1 \text{ or } \min_j [-(y_1+\cdots+y_j)] \geq 1\}$, we have

$$P(Z \in T) > P\{Z_1 = \pm 1\} \prod_{j=2}^{k} P\{Z_j = 0\} = \sigma_1^2 = \epsilon ,$$

which proves that (6.1) is sharp.

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

$$(6.4) \qquad P\{X \in T\} \le P\{|X_1| \ge 1\} \le \sigma_1^2 ,$$

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, $P\{Z_1 = \pm 1 \text{ or } 0\} = 1$, and since $EZ_1 = 0$,

$$P\{Z_1 = 1\} = P\{Z_1 = -1\} = \sigma_1^2 , \quad P\{Z_1 = 0\} = 1 - \sigma_1^2 .$$

Suppose that i is the smallest index $(i > 1)$ for which $\sigma_i^2 > 0$. Then $P\{Z_j = 0\} = 1$, $j = 2,\dots,i-1$, 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 $(Z_1,\dots,Z_k) \notin T$ because $Z_1 \not\geq 1$ and $-(Z_1+\dots+Z_i) \not\geq 1$. But $P\{Z_1 = -1 , Z_i = v\} = P\{Z_1 = -1\} P\{Z_i = v\} > 0$. This means equality is not attained at the left hand inequality of (6.4). A similar argument holds when $v < 0$.

<u>Case of (6.2)</u>. Choose $\epsilon > 0$ and let Z be a random vector with mutually independent components such that

$$P\{Z_1 = 1\} = \sigma_1^2/ (1 + \sigma_1^2) , \quad P\{Z_1 = -\sigma_1^2\} = 1/(1 + \sigma_1^2)$$

$$P\{Z_j = \delta\} = \sigma_j^2/ (\delta^2 + \sigma_j^2), \quad P\{Z_j = -\sigma_j^2/\delta\} = 1/(\delta^2 + \sigma_j^2) .$$

Then if $\delta > 0$,

$$P\{\min_j (Z_1 + \cdots + Z_j) \geq 1\} \geq \frac{\sigma_1^2}{1 + \sigma_1^2} \prod_{j=2}^{k} \frac{\sigma_j^2}{\delta^2 + \sigma_j^2} .$$

Clearly this is greater than $[\sigma_1^2/(1 + \sigma_1^2)] - \epsilon$ for δ sufficiently small, 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 $P\{Z_1 = -\sigma_1^2\} = 1/(1 + \sigma_1^2)$.

<u>Case of (6.3)</u>. Let Z_j , $j = 1,\ldots,k$ be mutually independent random variables such that

$$P\{Z_j = \pm 2^{j-1}\} = \sigma_j^2 / 2^{2j-1} , \quad P\{Z_j = 0\} = 1 - \sigma_j^2 / 2^{2j-2} , \quad j = 1,\ldots,k ,$$

then $EZ_j = 0$, $EZ_j^2 = \sigma_j^2$ for all j , and

$$P\{\min_j |Z_1 + \cdots + Z_j| \geq 1\} = P\{|Z_1| \geq 1\} = \sigma_1^2 .$$

Hence, equality is attained in (6.3) whenever $P\{Y = Z\} = 1$. ‖

By (6.4), 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.

One can strengthen the hypotheses of Lal's inequality [5, p. 229] (which provides a sharp upper bound for $P\{|X_1| \geq 1$ or $|X_2| \geq 1\}$

in terms of covariances) to obtain a bound for $P\{|Y_1| \geq 1$ or $|Y_1 + Y_2| \geq 1\}$, where Y_1 and Y_2 are independent. It is curious to note that while this specialization is a parallel to those of Theorem 6.1, 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 $\{X_t, t \in T\}$ is a real Stochastic process with $E(X_t) = 0$ for all $t \in T$, then

$$P\{\inf_{t \in T} |X_t| \geq 1\} \leq \inf_{t \in T} E(X_t^2) \equiv p, \tag{7.1}$$

$$P\{\inf_{t \in T} X_t \geq 1 \text{ or } \sup_{t \in T} X_t \leq -1\} \leq p, \tag{7.2}$$

$$P\{\inf_{t \in T} X_t \geq 1\} \leq p/(1+p), \tag{7.3}$$

whenever the probabilities are defined. These inequalities are trivial consequences of (1.1) and (1.2) since

$$\{\inf_{t \in T} X_t \geq 1 \text{ or } \sup_{t \in T} X_t \leq -1\} \subseteq \{\inf_{t \in T} |X_t| \geq 1\} \subseteq \{|X_s| \geq 1\},$$

and

$$\{\inf_{t \in T} X_t \geq 1\} \subseteq \{X_s \geq 1\}$$

for all $s \in T$.

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 $\{X_t, t \in T\}$ is a process with $EX_t = 0$ and $E X_s^2 = \sigma^2$, $E X_s X_\tau = \sigma^2\rho$, $(s \neq \tau)$, for all $s,\tau \in T$, (where $0 \leq \rho \leq 1$), then

(7.4) $$P[\inf_{t \in T} X_t \geq 1 \text{ or } \sup_{t \in T} X_t \leq -1] \leq \sigma^2\rho ,$$

(7.5) $$P[\inf_{t \in T} X_t \geq 1] \leq \sigma^2\rho/(1 + \sigma^2\rho) ,$$

whenever the probabilities are defined.

Proof. This follows from Example 2 of Section 3.1.

Theorem 7.2. If $\{X_t, t \in [0,\tau]\}$ is a martingale with $E X_t = 0$ and $E X_t^2 = \sigma^2(t)$, then

(7.6) $$P[\inf_{t \in [0,\tau]} X_t \geq 1 \text{ or } \sup_{t \in [0,\tau]} X_t \leq -1] \leq \sigma^2(0) ,$$

(7.7) $$P[\inf_{t \in [0,\tau]} X_t \geq 1] \leq \sigma^2(0)/[1 + \sigma^2(0)] ,$$

whenever the probabilities are defined.

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

<u>Remark</u>. 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 $E X_s X_t = \sigma^2(s)$, $s \leq t$ (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.

<u>Proof</u>. (7.6) and (7.7) are immediate consequences of (7.2) and (7.3). We now define a martingale attaining equality in (7.6). Let

$$\varphi(\theta) = \begin{cases} 0 , & \theta < 0 , \\ \sigma^2(0) + \dfrac{\sigma^2(\theta) - \sigma^2(0)}{\alpha^2} , & 0 \leq \theta < \tau , \\ 1 , & \theta \geq \tau , \end{cases}$$

where $\alpha = \sqrt{\dfrac{\sigma^2(\tau) - \sigma^2(0)}{1 - \sigma^2(0)}}$, and let μ be the measure induced on the Borel subsets $\mathcal{B}^*$ of $[0,\tau]$ by the right continuous distribution function φ . Let $\Omega = [0,\tau] \times \{-1,1\}$, and let $\mathcal{B}$ be the Borel subsets of Ω . Define a probability measure P on $\mathcal{B}$ by $P(B) = \mu(B^*)/2$ for $B = B^* \times \{1\}$ or $B = B^* \times \{-1\}$ and $B^* \in \mathcal{B}^*$. Denote a point in Ω by (θ,δ) and let $\{Z_t, t \in [0,\tau]\}$ be the process defined on the probability space $(\Omega, \mathcal{B}, P)$ by

$$Z_t(\theta,\delta) = \begin{cases} \delta , & \theta = 0 , \\ \delta\alpha , & 0 < \theta \le t , \\ 0 , & t < \theta . \end{cases}$$

Then $E Z_t^2 = P(\theta = 0) + \alpha^2 \int_{(0,t]} d\varphi = \sigma^2(0) + \alpha^2[\varphi(t) - \varphi(0)] = \sigma^2(t)$.

Clearly $E Z_t = 0$ and the process is a martingale. Furthermore,

$P[\inf_{t \in [0,\tau]} Z_t \ge 1$ or $\sup_{t \in [0,\tau]} Z_t \le -1] = P(\theta = 0) = \sigma^2(0)$ so that

equality is attained in (7.6).

To obtain a process attaining equality in (7.7), replace the function φ above by

$$\varphi(\theta) = \begin{cases} 0 , & \theta < 0 , \\ \dfrac{\sigma^2(0)}{1 - \sigma^2(0)} + \dfrac{\sigma^2(\theta) - \sigma^2(0)}{\beta^2} , & 0 \le \theta < \tau , \\ 1 , & \theta \ge \tau , \end{cases}$$

where $\beta = \sqrt{[1 + \sigma^2(0)][\sigma^2(\tau) - \sigma^2(0)]}$, and repla $\{Z_t, t \in [0,\tau]\}$ by the process $\{Y_t, t \in [0,\tau]\}$, where

$$Y_t(\theta,\delta) = \begin{cases} 1 , & \theta = 0 , \\ -\sigma^2(0) , & 0 \le t \le \theta \ne 0 , \\ -\sigma^2(0) + \delta\beta , & 0 < \theta < t . \end{cases}$$

It is easily verified that the process $\{Y_t\ , t \in [0,\tau]\}$ is a martingale satisfying the conditions of the theorem and attaining equality in (7.7). ‖

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

$$(7.9) \qquad P(X \in T) \leq p(\Sigma) .$$

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

$$(7.10) \qquad P(Y \in S) \leq p(H' \Pi H) ,$$

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

Clearly, (7.10) is sharp whenever (7.9) is sharp.

Non-linear transformations may also be useful, e.g., with $Y_j = X_j^2$, $j = 1,\ldots,k$, 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) $P(T_1) \leq p$, then (ii) $P(T_2) \leq p$. 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 (3.8), but suppose that some entries of the covariance matrix Σ are unknown. Then we can consider subvectors $(X_{i_1},\ldots,X_{i_n})$ of $(X_1,\ldots,X_k)$ for which the corresponding covariance matrix is known and apply (3.8), together with

$$P[\min_{1 \leq j \leq n} X_{i_j} \geq 1] \geq P[\min_{1 \leq j \leq k} X_j \geq 1] .$$

Whether this procedure (which can also be applied to (3.7)) 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. Acknowledgement. The authors are indeed grateful to Herman Chernoff for suggesting the proof of Lemma 3.3, and for several valuable comments and discussions.

REFERENCES

[1] Blackwell, D. and Girshick, M.A., Theory of Games and Statistical Decisions, John Wiley and Sons, New York, 1954.

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

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

[4] Marshall, A.W. and Olkin, I. "A one-sided inequality of the Chebyshev type," Technical Report No. 48, Applied Math. and Statistics Laboratory, Stanford University, 1959, p.1-11.

[5] Olkin, I. and Pratt, J.W., "A multivariate Tchebycheff inequality," Ann. Math. Stat. Vol. 29 (1958), pp. 226-234.

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

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

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

[9] Whittle, P., "A multivariate generalization of Tchebychev's inequality," Quarterly Jour. of Math., Oxford 2-nd Ser., Vol. 9 (1958), pp. 232-240.