SOME MULTIVARIATE CHEBYSHEV INEQUALITIES WITH EXTENSIONS TO CONTINUOUS PARAMETER PROCESSES[1]

By Z. W. Birnbaum and Albert W. Marshall

University of Washington; University of Washington and Stanford University

0. Summary. In this paper we obtain some multivariate generalizations of Chebyshev's inequality, two of which are extended to continuous parameter stochastic processes. The extensions are obtained in a natural way by taking into account separability and letting the number of variables approach infinity.

Particular attention is paid to the question of sharpness. To show that the bound of the inequality cannot be improved, examples are given in a number of cases that attain equality.

1. Introduction. We begin by discussing a model for the various generalizations of Chebyshev's inequality, and for a standard proof that we shall use. Examination of this proof will enable us to make some general comments concerning the problems of deriving inequalities and of proving sharpness.

Let $(\Omega, \mathcal{B}, P)$ be a probability space, and let $(\mathfrak{X}, \mathfrak{A})$ be a measurable space. For each $i \varepsilon I$, an arbitrary index set, let $\mathfrak{C}_i \subset \mathfrak{A}$ and let $\mathfrak{F}_i$ be a class of random variables on $(\Omega, \mathcal{B})$ taking values in $(\mathfrak{X}, \mathfrak{A})$ such that $X \varepsilon \mathfrak{F}_i$ whenever $Y \varepsilon \mathfrak{F}_i$ has the same distribution as X. Chebyshev's inequality and its generalizations are of the following form:

$$(1.1) \quad X \varepsilon \mathfrak{F}_i \text{ implies } P\{X \varepsilon A\} \leqq \Phi_i(A) \quad \text{for all} \quad A \varepsilon \mathfrak{C}_i \quad \text{and} \quad \text{all} \quad i \varepsilon I,$$

where for each $i \varepsilon I$, Φ_i is a non-negative function on $\mathfrak{C}_i$.

For the usual Chebyshev inequality, $\mathfrak{X}$ is the real line, $\mathfrak{A}$ is the Borel sets, $I = (-\infty, \infty) \times [0, \infty)$, $\mathfrak{F}_{(\mu,\sigma^2)}$ is the set of all real-valued random variables X with expectation μ and variance σ^2, $\mathfrak{C}_{(\mu,\sigma^2)}$ consists of all sets of the form $A_\epsilon = (\mu - \epsilon, \mu + \epsilon)^c$ (E^c denotes the complement of the set E), and

$$\Phi_{(\mu,\sigma^2)}(A_\epsilon) = \sigma^2/\epsilon^2.$$

In Sections 2 and 3, $\mathfrak{X}$ will be Euclidean n-space R^n for some n, and $\mathfrak{A}$ will again be the Borel sets.

Inequalities of the type (1.1) can very often be proved as follows: for each $i \varepsilon I$, one defines a function f_i on $\mathfrak{C}_i \times \mathfrak{X}$ to R such that, for each $A \varepsilon \mathfrak{C}_i$,

1. $f_i(A, \cdot)$ is measurable,
2. $\int_\Omega f_i(A, X)\, dP$ is independent of $X \varepsilon \mathfrak{F}_i$,
3. $\int_{\{X \not\varepsilon A\}} f_i(A, X)\, dP \geqq 0$ for all $X \varepsilon \mathfrak{F}_i$,
4. $\int_{\{X \varepsilon A\}} f_i(A, X)\, dP \geqq P\{X \varepsilon A\}$ for all $X \varepsilon \mathfrak{F}_i$.

Then

Received April 19, 1960; revised March 31, 1961.

[1] Research sponsored by the Office of Naval Research at the University of Washington and at Stanford University.

$$\Phi_i(A) = \int_\Omega f_i(A, X)\, dP \geqq \int_{\{X \varepsilon A\}} f_i(A, X)\, dP \geqq P\{X \varepsilon A\} \quad \text{for all} \quad X \varepsilon \mathfrak{F}_i .$$

Often conditions 3. and 4. are replaced by the stronger conditions

3′. $f_i(A, X) \geqq 0$ for all $x \varepsilon \mathfrak{X}$,
4′. $f_i(A, X) \geqq 1$ for all $x \varepsilon A$.

This replacement will be made in Section 3, but not in Section 2.

The above model has been presented with various degrees of clarity and generality by a number of authors. It is used quite extensively by Fréchet [4], who credits Cantelli with its first presentation.

An inequality of the form (1.1) is said to be sharp if for all i in I and A in $\mathcal{C}_i$ there is a sequence $\{X_k\}_{k=0}^\infty$ of elements of $\mathfrak{F}_i$ such that

$$\lim_{k\to\infty} P\{X_k \varepsilon A\} = \begin{cases} \Phi_i(A) & \text{if} \quad \Phi_i(A) \leqq 1 \\ 1 & \text{if} \quad \Phi_i(A) > 1. \end{cases}$$

By examining a proof of the kind described above where conditions 3′. and 4′. are satisfied, it is often possible to find an example for which equality holds in (1.1) and thereby demonstrate sharpness. If X is a random variable in $\mathfrak{F}_i$ for which equality holds for arbitrary but fixed i in I and A in $\mathcal{C}_i$, then equality must hold in 3′. and 4′. Hence (neglecting sets of zero probability) it must be that X assumes in $A(A^c)$ only values x for which $f_i(A, x) = 1 (f_i(A, x) = 0)$. Using this determination of the values that X may assume with positive probability together with the requirement $X \varepsilon \mathfrak{F}_i$, one can often find a distribution for X if one exists.

When deriving an inequality of the form (1.1), there are certain procedures one may use to find the functions f_i. For example, if the bound is to involve only second moments of the random variables, then f_i must be a quadratic form. This together with conditions 3′. and 4′. may so severely limit the possible candidates for f_i that one can write f_i as a function involving only a few unknown parameters. Their values can sometimes be found by the requirement that they minimize the bound Φ_i. Alternatively, one can begin by using the method of the preceding paragraph to find (in terms of the unknown parameters) the values that a random variable X may assume with positive probability to achieve equality in (1.1). The requirement that $X \varepsilon \mathfrak{F}_i$ may then determine the unknown parameters.

2. A generalization of Kolmogorov's inequality.

THEOREM 2.1. *Let $X_1, X_2, \cdots, X_n$ be random variables such that*

$$E(|X_k| \mid X_1, \cdots, X_{k-1}) \geqq \psi_k |X_{k-1}| \qquad \text{a.e.}^2,$$

where $\psi_k \geqq 0$, $k = 2, 3, \cdots, n$. Let

[2] Even though we usually neglect to mention the underlying probability space $(\Omega, \mathcal{B}, P)$, we use the abbreviation a.e. to mean "almost everywhere with respect to P."

$$a_k > 0, \quad b_k = \max (a_k, a_{k+1}\psi_{k+1}, a_{k+2}\psi_{k+1}\psi_{k+2}, \cdots, a_n \prod_{i=k+1}^{n} \psi_i),$$

$k = 1, 2, \cdots, n$, $b_{n+1} = 0$, *and let* $X_0 = 0$. *If* $r \geqq 1$ *is such that* $E|X_k|^r < \infty$, $k = 1, 2, \cdots, n$, *then*

$$(2.1) \quad P\{\max_{1 \leqq k \leqq n} a_k|X_k| \geqq 1\} \leqq \sum_{k=1}^{n} (b_k^r - \psi_{k+1}^r b_{k+1}^r) E|X_k|^r$$
$$= \sum_{k=1}^{n} b_k^r (E|X_k|^r - \psi_k^r E|X_{k-1}|^r).$$

REMARKS. Several known generalizations of Kolmogorov's inequality follow from this theorem by setting $\psi_k = 1$, $X_k = Y_1 + Y_2 + \cdots + Y_k$, $k = 1, 2, \cdots, n$, and further specializing the assumptions. In particular, assuming $E(Y_1) = 0, E(Y_k \mid Y_1, \cdots, Y_{k-1}) = 0$ a.e., $k = 2, \cdots, n, r = 2, a_1 = \cdots = a_n = 1/\epsilon$, one obtains an inequality given by Loève [7, p. 386] and by Doob [3, p. 315]; assuming $Y_1, \cdots, Y_n$ mutually independent, $E(Y_1) = 0$, $r \geqq 1$, $a_1 = \cdots = a_n = 1/\epsilon$, one obtains an inequality given by Loève [7, p. 263]; and assuming $Y_1, \cdots, Y_n$ mutually independent, $E(Y_1) = 0$, $r = 2$, $a_1 \geqq a_2 \geqq \cdots \geqq a_n > 0$, one obtains a result due to Hájek and Rényi [5].

PROOF. Since $E(|X_k| \mid X_1, \cdots, X_{k-1}) \geqq \psi_k|X_{k-1}|$ a.e. implies that

$$E(|X_k^r| \mid X_1^r, \cdots, X_{k-1}^r) \geqq \psi_k^r|X_{k-1}^r| \qquad \text{a.e.},$$

where $X_j^r = (\operatorname{sign} X_j)|X_j|^r$, we can take $r = 1$ without loss of generality.

Let $A_k = \{a_i|X_i| < 1, \ i = 1, 2, \cdots, k-1, \ a_k|X_k| \geqq 1\}$, $k = 1, \cdots, n$. Then if $j > k$ (we denote the characteristic function of a set C by χ_C),

$$\int_{A_k} |X_j| \, dP = E\{\chi_{A_k} E[|X_j| \mid X_1, \cdots, X_{j-1}]\} \geqq E\{\chi_{A_k} \psi_j |X_{j-1}|\}$$
$$= \psi_j \int_{A_k} |X_{j-1}| \, dP,$$

and by induction it follows that

$$\int_{A_k} |X_j| \, dP \geqq \left(\prod_{i=k+1}^{j} \psi_i\right) \int_{A_k} |X_k| \, dP.$$

Since $\sum_{j=k}^{n} (b_j - \psi_{j+1}b_{j+1})(\prod_{i=k+1}^{j} \psi_i) = b_k \geqq a_k$, and since $b_k \geqq \psi_{k+1}b_{k+1}$, $k = 1, 2, \cdots, n$,

$$\sum_{j=1}^{n} b_j[E|X_j| - \psi_j E|X_{j-1}|] = \sum_{j=1}^{n} (b_j - \psi_{j+1}b_{j+1}) E|X_j|$$
$$\geqq \sum_{k=1}^{n} \sum_{j=1}^{n} (b_j - \psi_{j+1} b_{j+1}) \int_{A_k} |X_j| \, dP \geqq \sum_{k=1}^{n} \sum_{j=k}^{n} (b_j - \psi_{j+1} b_{j+1}) \int_{A_k} |X_j| \, dP$$
$$\geqq \sum_{k=1}^{n} \sum_{j=k}^{n} (b_j - \psi_{j+1} b_{j+1}) \left(\prod_{i=k+1}^{j} \psi_i\right) \int_{A_k} |X_k| \, dP$$

$$\geqq \sum_{k=1}^{n} \sum_{j=k}^{n} (b_j - \psi_{j+1}\, b_{j+1}) \left(\prod_{i=k+1}^{j} \psi_i \right) a_k^{-1} P(A_k) \geqq \sum_{k=1}^{n} P(A_k)$$
$$= P\left\{ \max_{1 \leqq k \leqq n} a_k |X_k| \geqq 1 \right\}.$$

The proof is now complete, and we list some special cases.

If $E(X_k \mid X_1, \cdots, X_{k-1}) = X_{k-1}$ a.e., then $\psi_k = 1$, $k = 2, 3, \cdots, n$, and (2.1) becomes

$$P\{ \max_{1 \leqq k \leqq n} a_k |X_k| \geqq 1 \} \leqq \sum_{k=1}^{n} (b_k^r - b_{k+1}^r) E|X_k|^r = \sum_{k=1}^{n} b_k^r (E|X_k|^r - E|X_{k-1}|^r). \tag{2.2}$$

If $\psi_k \neq 0$, $k = 2, 3, \cdots, n$, then with the change of variables

$$X_1 = X_1', \; X_k = X_k' \left(\prod_{i=1}^{k-1} \psi_i \right)^{-1}, \qquad k = 2, \cdots, n,$$

in (2.2), one obtains (2.1) after removing the primes.

If $\{X_1, \cdots, X_n\}$ is a semi-martingale, then so is $\{X_1^+, \cdots, X_n^+\}$ where $X_k^+ = \max(X_k, 0)$ (see, e.g., [3], p. 295). In this case it follows from Theorem 2.1 that

$$P\{ \max_{1 \leqq k \leqq n} a_k X_k \geqq 1 \} = P\{ \max_{1 \leqq k \leqq n} a_k X_k^+ \geqq 1 \} \leqq \sum_{k=1}^{n} (b_k^r - b_{k+1}^r) E(X_k^+)^r. \tag{2.3}$$

With $a_1 \geqq a_2 \geqq \cdots \geqq a_n > 0$ and $r = 1$, this inequality has been given by Chow [2]. It follows from (2.3) that if $\{X_1, \cdots, X_n\}$ is a semi-martingale

$$P\{ \max_{1 \leqq k \leqq n} a_k X_k \geqq 1 \} \leqq \sum_{k=1}^{n} (b_k^r - b_{k+1}^r) E|X_k|^r. \tag{2.4}$$

If $\psi_k = 0$, $k = 2, 3, \cdots, n$, then $a_k = b_k$, $k = 1, 2, \cdots, n$, and (2.1) becomes

$$P\{ \max_{1 \leqq k \leqq n} a_k |X_k| \geqq 1 \} \leqq \sum_{k=1}^{n} a_k^r E|X_k|^r. \tag{2.5}$$

This inequality was obtained by Olkin and Pratt [8, p. 234] with $r = 2$ and the additional assumption that $X_1, \cdots, X_n$ are uncorrelated. With $r = 2$ it also appears as a special case of Theorem 3.1.

If $a_k \leqq a_n(\prod_{j=k+1}^{n} \psi_j)$, $k = 1, 2, \cdots, n$, then again $a_k = b_k$, $k = 1, 2, \cdots, n$ and we obtain from (2.1)

$$P\{ \max_{1 \leqq k \leqq n} a_k |X_k| \geqq 1 \} \leqq a_n^r E|X_n|^r. \tag{2.6}$$

With $n = r = 2$, we obtain the following from Theorem 2.1. Let X_1 and X_2 be random variables such that $E(X_i) = 0$, $E(X_i^2) = \sigma_i^2 < \infty$, $i = 1, 2$, and $E(X_1 X_2) = \sigma_1 \sigma_2 \rho$. If the regression of X_2 on X_1 is linear, then for every positive

a_1 and a_2,

$$(2.7)\quad P\{a_1|X_1| \geqq 1 \quad \text{or} \quad a_2|X_2| \geqq 1\} \leqq \begin{cases} a_1^2\sigma_1^2 + a_2^2\sigma_2^2(1-\rho^2) & \text{if} \quad a_1^2\sigma_1^2 \geqq a_2^2\sigma_2^2\rho^2 \\ a_2^2\sigma_2^2 & \text{if} \quad a_1^2\sigma_1^2 \leqq a_2^2\sigma_2^2\rho^2. \end{cases}$$

To obtain this, we have used the relation $\xi\sigma_1 = \rho\sigma_2$ where $E(X_2 \mid X_1) = \xi X_1$ a.e.

THEOREM 2.2. *Equality can be achieved in* (2.1), *so that* (2.1) *is sharp.*

REMARK. Actually we prove slightly more than this. Even though the hypotheses of Theorem 2.1 are strengthened by assuming that $E(X_1) = m$ and $E(X_k \mid X_1, \cdots, X_{k-1}) = \xi_k X_{k-1}$ a.e. (in which case we take $\psi_k = |\xi_k|$), $k = 2, 3, \cdots, n$, equality can be attained in (2.1) so long as $b_1^r E|X_1|^r \geqq b_1|m|$. If the hypothesis $E(X_1) = m$ is added to Theorem 2.1 and $b_1^r E|X_1|^r < b_1|m|$, then (2.1) is no longer sharp, and a better bound for the case $n = 1$, $r = 2$ has been obtained by Selberg [10].

PROOF. We introduce the notations $E(X_1) = m$, $\mu_0^r = 0$, $E|X_k|^r = \mu_k^r$, $k = 1, \cdots, n$. Since $r \geqq 1$ and $E(|X_k| \mid X_1, \cdots, X_{k-1}) \geqq \psi_k|X_{k-1}|$ a.e., it follows from Hölder's inequality that

$$E|X_k|^r = E[E\{|X_k|^r \mid X_1, \cdots, X_{k-1}\}] \geqq E[E\{|X_k| \mid X_1, \cdots, X_{k-1}\}]^r$$
$$\geqq \psi_k^r E|X_{k-1}|^r, \qquad k = 2, 3, \cdots, n.$$

Hence $b_k^r(\mu_k^r - \psi_k^r\mu_{k-1}^r) \geqq 0$, $k = 2, 3, \cdots, n$.

Now suppose that $b_1^r\mu_1^r \geqq b_1|m|$ and that $\sum_{k=1}^n b_k^r(\mu_k^r - \psi_k^r\mu_{k-1}^r) \leqq 1$, and consider a random vector $Z = (Z_1, \cdots, Z_n)$ with the following distribution (where $\psi_k = |\xi_k|$, $k = 2, 3, \cdots, n$):

$z = (z_1, \cdots, z_n)$	$P(Z = z)$
$b_1^{-1}\left(1, \xi_2, \xi_2\xi_3, \cdots, \prod_{i=2}^n \xi_i\right)$	$\frac{1}{2}(b_1^r\mu_1^r + b_1 m)$
$-b_1^{-1}\left(1, \xi_2, \xi_2\xi_3, \cdots, \prod_{i=2}^n \xi_i\right)$	$\frac{1}{2}(b_1^r\mu_1^r - b_1 m)$
$\pm b_2^{-1}\left(0, 1, \xi_3, \cdots, \prod_{i=3}^n \xi_i\right)$	$\frac{1}{2}(b_2^r(\mu_2^r - \psi_2^r\mu_1^r))$
----	----
$\pm b_k^{-1}\left(0, \cdots, 0, 1, \xi_{k+1}, \cdots, \prod_{i=k+1}^n \xi_i\right)$	$\frac{1}{2}(b_k^r(\mu_k^r - \psi_k^r\mu_{k-1}^r))$
----	----
$\pm b_n^{-1}(0, \cdots, 0, 1)$	$\frac{1}{2}(b_n^r(\mu_n^r - \psi_n^r\mu_{n-1}^r))$
$(0, \cdots, 0)$	$1 - \sum_{k=1}^n b_k^r(\mu_k^r - \psi_k^r\,\mu_{k-1}^r).$

Then $Z_1, Z_2, \cdots, Z_n$ satisfy the conditions of Theorem 2.1; in fact they satisfy the stronger assumptions given in the above remark.

Suppose that for some k, $b_k|Z_k| \geqq 1$. Then for some $j \geqq k$,

$$b_k = a_j(\textstyle\prod_{i=k+1}^{j} \psi_i), \quad \text{and on} \quad \{b_k|Z_k| \geqq 1\}, \qquad Z_j = Z_k(\textstyle\prod_{i=k+1}^{j} \xi_i).$$

Hence

$$b_k|Z_k| = a_j|Z_j| \geqq 1 \quad \text{so that} \quad \max_{1 \leqq k \leqq n} b_k|Z_k| \geqq 1$$

$$\text{implies} \quad \max_{1 \leqq k \leqq n} a_k|Z_k| \geqq 1, \qquad \text{and}$$

$$P\{\max_{1 \leqq k \leqq n} a_k|Z_k| \geqq 1\} = \textstyle\sum_{k=1}^{n} b_k^r(\mu_k^r - \psi_k^r\mu_{k-1}^r).$$

Thus the random vector Z attains equality in (2.1).

Next suppose that $b_1^r\mu_1^r \geqq b_1|m|$ but that $\sum_{k=1}^{n} b_k^r(\mu_k^r - \psi_k^r\mu_{k-1}^r) > 1$, i.e., the bound of (2.1) exceeds unity. Choose $c_1, c_2, \cdots, c_n$ such that $0 < c_k \leqq b_k$, $k = 1, 2, \cdots, n$, $\sum_{k=1}^{n} c_k^r(\mu_k^r - \psi_k^r\mu_{k-1}^r) = 1$, and such that $c_1^r\mu_1^r \geqq c_1|m|$. Then in the distribution of Z, replace b_k by c_k, $k = 1, 2, \cdots, n$. For a random vector defined in this manner,

$$P\{\max_{1 \leqq k \leqq n} a_k|Z_k| \geqq 1\} = P\{\max_{1 \leqq k \leqq n} b_k|Z_k| \geqq 1\} \geqq P\{\max_{1 \leqq k \leqq n} c_k|Z_k| \geqq 1\} = 1,$$

and the bound of unity is attained.

3. Generalizations of Berge's inequality. We consider now multivariate generalizations of Chebyshev's inequality providing bounds for

$$P\{\max_{1 \leqq i \leqq n} a_i|X_i| \geqq 1\}$$

under assumptions regarding second moments.

In 1919, Karl Pearson [9] published a generalization of Chebyshev's inequality providing an upper bound in terms of second moments for the probability that a two-dimensional random vector falls outside a given ellipse. His results may be described as follows: Let $\mathfrak{F}_\sigma$ be the class of random vectors $X = (X_1, X_2)$ with $E(X_i) = 0$, $E(X_i^2) = \sigma_i^2$, $i = 1, 2$, and $E(X_1X_2) = \sigma_{12}$; let $\mathfrak{C}$ be the class of sets $A_e \subset R^2$ of the form $A_e = \{x = (x_1, x_2) : f_e(x) < 1\}$, where the boundary of A_e is the ellipse $f_e(x) = e_1x_1^2 + e_2x_2^2 + e_3x_1x_2 = 1$. Since f_e is positive definite and not less than one on A_e^c, it follows that $X \in \mathfrak{F}_\sigma$ and $A_e \in \mathfrak{C}$ implies

$$e_1\sigma_1^2 + e_2\sigma_2^2 + e_3\sigma_{12} = \int_\Omega f_e(X)\,dP \geqq \int_{\{X \notin A_e\}} f_e(X)\,dP \geqq P\{X \notin A_e\}.$$

If $A = \{(x_1, x_2) : a_i|x_i| < 1, i = 1, 2\}$, where a_1 and a_2 are positive, it is trivial that

$$(3.1) \quad P\{a_1|X_1| \geqq 1 \text{ or } a_2|X_2| \geqq 1\} = P\{X \notin A\} \leqq \inf_{A \supset A_e \in \mathfrak{C}} \int_\Omega f_e(X)\,dP.$$

In a special case, P. O. Berge [1] computed this bound and obtained the following inequality: If $X = (X_1, X_2) \in \mathfrak{F}_\sigma$, then for all $k > 0$,

$$(3.2) \qquad P\{|X_1| \geqq k\sigma_1 \quad \text{or} \quad |X_2| \geqq k\sigma_2\} \leqq [1 + (1 - \rho^2)^{\frac{1}{2}}]/k^2,$$

where $\rho = \sigma_{12}/(\sigma_1\sigma_2)$. Whenever $\sigma = (\sigma_1^2, \sigma_2^2, \sigma_{12})$ is such that the covariance matrix is positive definite (i.e., $\sigma_1^2 > 0$, $\sigma_1^2\sigma_2^2 > \sigma_{12}^2$) and the bound is not greater than one, Berge gave an example attaining equality to show that (3.2) is sharp. The bound of (3.1) was computed in general by D. N. Lal [6] and follows from (3.4) with $n = 2$, $\nu = 1$.

We describe now a natural generalization of Berge's result to higher dimensions. Use a prime to denote transpose and let $\mathfrak{F}_\Lambda$ be the class of random vectors $X = (X_1, X_2, \cdots, X_n)'$ taking values in R^n with moment matrix $\Lambda = (E(X_iX_j))$. Replace the functions f_e above by quadratic forms $F_M(x) = x'Mx$ where $x = (x_1, \cdots, x_n)' \varepsilon R^n$ and M is an $n \times n$ positive definite matrix. If $A_M = \{x : x'Mx < 1\}$, it follows as before that for $X \varepsilon \mathfrak{F}_\Lambda$,

$$\Phi(M, \Lambda) = \int_\Omega X'MX\,dP \geqq \int_{\{X \notin A_M\}} X'MX\,dP \geqq P\{X \not\varepsilon A_M\}.$$

If $A = \{x : a_i|x_i| < 1, i = 1, 2, \cdots, n\}$, we obtain

$$P\{X \not\varepsilon A\} \leqq \inf_{A_M \subset A} \Phi(M, \Lambda), \tag{3.3}$$

which is the desired inequality. Unfortunately the bound is not easily computed.

This generalization of Berge's result has been recently investigated by Olkin and Pratt [8] and by Whittle [11]. They consider the set $\mathfrak{A}$ of positive definite $n \times n$ matrices M for which $A_M = \{x : x'Mx < 1\} \subset A$ and prove that there is a unique element M^* of this set such that $\inf_{\mathfrak{A}} \int_\Omega F_M(X)\,dP = \int_\Omega F_{M^*}(X)\,dP$ where $F_{M^*}(x) = xM^*x$. They have not succeeded in obtaining M^* but were able to characterize it as the solution of a certain matrix equation. Using this result they prove that the inequality (3.3) is sharp.

It is possible to obtain many inequalities related to (3.3) that are not sharp, since, $M \varepsilon \mathfrak{A}$ implies $P\{X \not\varepsilon A\} \leqq \int_\Omega X'MX\,dP$. An inequality of this kind was given by Lal [6] and a better one by Olkin and Pratt [8].

Extension of (3.1) to n dimensions are obtained by generalizing the sets $A = \{x \varepsilon R^2 : a_1|x_1| < 1, a_2|x_2| < 1\}$ to n dimensions, and the sets $\mathfrak{F}_\sigma$ to sets of n-dimensional random vectors. Clearly both of these generalizations may be accomplished in many ways other than those used to obtain (3.3). In the remainder of this section, we obtain an extension of (3.1) which differs from (3.3) in that only certain terms of the covariance matrix are assumed known.

THEOREM 3.1. *Let ν be an integer in the interval $0 \leqq \nu \leqq n - 1$, and let $r_1, \cdots, r_\nu$ be integers such that $r_1 = 1$ and $r_{k-1} \leqq r_k \leqq k$, $k = 2, 3, \cdots, \nu$. If $X = (X_1, \cdots, X_n)'$ is a random vector with $E(X_i^2) = \sigma_i^2 < \infty$, $i = 1, 2, \cdots, n$ and $E(X_{r_i}X_{i+1}) = \varphi_i < \sigma_{r_i}\sigma_{i+1}$, $i = 1, 2, \cdots, \nu$ and if $\epsilon_i > 0$, $i = 1, 2, \cdots, n$, then*

$$P\{|X_i| < \epsilon_i, i = 1, 2, \cdots, n\} \geqq 1 - \sum_{i=1}^{n} \frac{\sigma_i^2}{\epsilon_i^2} + \sum_{i=1}^{\nu} \frac{c_i - d_i^{\frac{1}{2}}}{2} \tag{3.4}$$

where

$$c_i = \frac{\sigma_{r_i}^2}{\epsilon_{r_i}^2} + \frac{\sigma_{i+1}^2}{\epsilon_{i+1}^2} \quad \textit{and} \quad d_i = c_i^2 - 4\frac{\varphi_i^2}{\epsilon_{r_i}\epsilon_{i+1}}, \ i = 1, 2, \cdots, \nu$$

(*if* $\nu = 0$, *regard the empty sum of the bound as zero*).

REMARK. Inequality (3.4) is applicable whenever all the moments $E(X_i^2)$ are known. It utilizes all knowledge of second moments whenever (possibly after a permutation of the random variables X_i) every 3×3 principle minor of the matrix $(E(X_iX_j))$ has at least one unknown entry.

PROOF. We begin by assuming that $\epsilon_i = 1$. For $0 \leqq i \leqq n$, let

$$\alpha_i = \begin{cases} \dfrac{c_i - d_i^{\frac{1}{2}}}{2\varphi_i} & \text{if } 1 \leqq i \leqq \nu \text{ and } \varphi_i \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

Since $\sigma_{r_i}^2\sigma_{i+1}^2 > \varphi_i^2$ for $1 \leqq i \leqq \nu$, $d_i \neq 0$ and

$$d_i^{\frac{1}{2}} = [(c_i + 2|\varphi_i|)(c_i - 2|\varphi_i|)]^{\frac{1}{2}} > c_i - 2|\varphi_i|$$

so that $|\alpha_i| < 1$, $0 \leqq i \leqq n$.

For $k = 1, 2, \cdots, n - 1$, let

$$F_{k+1}(x) = F_{k+1}(x_1, \cdots, x_n) = \sum_{i=0}^{k} (x_{i+1} - \alpha_i x_{r_i})^2/(1 - \alpha_i^2).$$

If $\alpha_k \neq 0$,

$$F_{k+1}(x) = F_k(x) - x_{r_k}^2 + (x_{r_k} - \alpha_k x_{k+1})^2/(1 - \alpha_k^2) + x_{k+1}^2;$$

using this and the relation $F_{k+1}(x) \geqq F_k(x)$, it is easily established by induction that $F_k(x) \geqq x_i^2$, $1 \leqq i \leqq k$. Hence

$$F_n(x) \geqq 0 \quad \text{and} \quad F_n(x) \geqq 1 \quad \text{for} \quad x \notin A = \{x : x_i < 1, i = 1, 2, \cdots, n\}.$$

From this and the relations

$$\frac{\alpha_i}{1 - \alpha_i^2} = \frac{\varphi_i}{d_i^{\frac{1}{2}}}, \qquad \frac{\alpha_i^2}{1 - \alpha_i^2} = \frac{c_i - d_i^{\frac{1}{2}}}{2d_i^{\frac{1}{2}}}$$

we obtain

$$\begin{aligned} E[F_n(X)] &= \sum_{i=1}^{\nu} \frac{(c_i - d_i^{\frac{1}{2}})c_i}{2d_i^{\frac{1}{2}}} - 2\sum_{i=1}^{\nu} \frac{\varphi_i^2}{d_i^{\frac{1}{2}}} + \sum_{i=1}^{n} \sigma_i^2 \\ &= \sum_{i=1}^{\nu} \frac{d_i^{\frac{1}{2}} - c_i}{2} + \sum_{i=1}^{n} \sigma_i^2 \geqq \int_{\{X \notin A\}} F_n(X)\, dP \geqq P\{X \notin A\}. \end{aligned} \tag{3.6}$$

Inequality (3.4) follows from (3.6) after the change of variables $X_i^* = \epsilon_i X_i$, $i = 1, 2, \cdots, n$ is made and the asterisks removed, so that the proof is complete.

If in (3.4) we take $\nu = n - 1$ and $r_i = i$, we obtain

$$P\{|X_i| < \epsilon_i, i = 1, 2, \cdots, n\} \tag{3.7}$$
$$\geqq 1 - \frac{1}{2}\left\{\frac{\sigma_1^2}{\epsilon_1^2} + \frac{\sigma_n^2}{\epsilon_n^2} + \sum_{i=1}^{n-1}\left[\left(\frac{\sigma_i^2}{\epsilon_i^2} + \frac{\sigma_{i+1}^2}{\epsilon_{i+1}^2}\right)^2 - \frac{4\varphi_i^2}{\epsilon_i^2\epsilon_{i+1}^2}\right]^{\frac{1}{2}}\right\}$$

where $\varphi_i = E(X_iX_{i+1})$, $i = 1, 2, \cdots, n - 1$. With $\nu = n - 1$ and $r_i = 1$, (3.4) becomes

$$P\{| X_i | < \epsilon_i, i = 1, 2, \cdots, n\} \tag{3.8}$$
$$\geqq 1 - \frac{1}{2}\left\{\sum_{i=2}^{n}\frac{\sigma_i^2}{\epsilon_i^2} - (n - 3)\frac{\sigma_1^2}{\epsilon_1^2} + \sum_{i=1}^{n-1}\left[\left(\frac{\sigma_1^2}{\epsilon_1^2} + \frac{\sigma_{i+1}^2}{\epsilon_{i+1}^2}\right)^2 - \frac{4\varphi_i^2}{\epsilon_1^2\epsilon_{i+1}^2}\right]^{\frac{1}{2}}\right\}$$

where $\varphi_i = E(X_1X_{i+1})$, $i = 1, 2, \cdots, n - 1$.

Note that with $\nu = 0$, (3.4) becomes (2.4) with $r = 2$.

We investigate the sharpness of (3.4) only in the special cases (3.7) and (3.8), and only under the additional hypotheses that $E(X_i) = 0$, $i = 1, 2, \cdots, n$.

Consider first (3.7); i.e., assume $\nu = n - 1$ and $r_i = i$. If $Z = (Z_1, \cdots, Z_n)'$ is a random vector for which equality holds in (3.7) then equality must hold throughout (3.6) when X is replaced by Z. This means

$$F_n(Z) = 0 \text{ a.e. on } \{Z \varepsilon A\}, \tag{3.9}$$

and

$$F_n(Z) = 1 \text{ a.e. on } \{Z \not\varepsilon A\}. \tag{3.10}$$

Since $\{F_n(x) \leqq 1\}$ is strictly convex, $F_n(Z) \geqq 1$ on $\{Z \not\varepsilon A\}$ implies that there is at most one root of the equation $F_n(x) = 1$ on each plane $x_i = \pm 1$. Hence there are at most $2n$ roots not in A of the equation $F_n(x) = 1$. It is easily verified that these roots are plus and minus the columns $b^{(1)}, \cdots, b^{(n)}$ of the Green's matrix $B = (b_{ij})$ where $b_{ij} = b_{ji} = \beta_j/\beta_i$ $(i \leqq j)$, $\beta_1 = 1$ and $\beta_k = \prod_{m=1}^{k-1} \alpha_m$, $k > 1$.

In order that (3.9) be satisfied, Z must with probability one assume in A only the value $(0, \cdots, 0)'$; in order that (3.10) be satisfied, Z must with probability one assume in A^c only the values $\pm b^{(i)}$. Thus Z must have a distribution of the form

$$P\{Z = b^{(i)}\} = p_i/2, \qquad P\{Z = -b^{(i)}\} = p_i^*/2, \tag{3.11}$$
$$P\{Z = 0\} = 1 - \sum_{i=1}^{n}\left(\frac{p_i}{2} + \frac{p_i^*}{2}\right).$$

If $u = (u_1, u_2, \cdots, u_n)'$ where $u_i = (p_i - p_i^*)/2$, then $E(Z) = B'u$. Since $|B| = \prod_{i=1}^{n}(1 - \alpha_i^2)$ and $|\alpha_i| < 1$, B is positive definite and $E(Z) = (0, \cdots, 0)'$ if and only if $u_i = 0$, i.e., $p_i = p_i^*$, $i = 1, 2, \cdots, n$.

Now consider the equations

$$E(Z_i^2) = \sigma_i^2, i = 1, 2, \cdots, n, E(Z_iZ_{i+1}) = \varphi_i, i = 1, 2, \cdots, n - 1. \tag{3.12}$$

From (3.12) we obtain $\sigma_i^2 + \sigma_{i+1}^2 - \alpha_i\varphi_i = \varphi_i/\alpha_i$, and since we require $|\alpha_i| < 1$, (3.12) is consistent with (3.5). It also follows from (3.12) that

$$(3.13)\quad p_k = \begin{cases} \dfrac{\sigma_1^2 - \alpha_1^2\sigma_2^2}{1 - \alpha_1^4}, & k = 1, \\[2ex] \dfrac{\sigma_k^2 - 2\alpha_{k-1}^2\sigma_{k-1}^2 + \alpha_{k-1}^4\sigma_k^2}{2(1 - \alpha_{k-1}^4)} + \dfrac{\sigma_k^2 - 2\alpha_k^2\sigma_{k+1}^2 + \alpha_k^4\sigma_k^2}{2(1 - \alpha_k^4)}, & 2 \leqq k \leqq n-1, \\[2ex] \dfrac{\sigma_n^2 - \alpha_{n-1}^2\sigma_{n-1}^2}{1 - \alpha_{n-1}^4}, & k = n. \end{cases}$$

The expressions for p_1 and p_n are easily verified; the other p_k are obtained by computing $(\sigma_k^2 - \alpha_{k-1}^2\sigma_{k-1}^2) - \alpha_{k-1}^2(\sigma_{k-1}^2 - \alpha_{k-1}^2\sigma_k^2)$ and $(\sigma_k^2 - \alpha_k^2\sigma_{k+1}^2) - \alpha_k^2(\sigma_{k+1}^2 - \alpha_k^2\sigma_k^2)$ in terms of the p_i.

Since (3.13) is the solution of (3.12) it follows that (3.11) together with (3.13) does provide an example satisfying the hypotheses of (3.7) providing that $p_k \geqq 0$ for all k. Furthermore,

$$\sum_{i=1}^n p_i = \sum_{i=1}^n \sigma_i^2 - \sum_{i=1}^{n-1} \alpha_i\varphi_i = \frac{1}{2}\left\{\sigma_1^2 + \sigma_n^2 + \sum_{i=1}^{n-1} [(\sigma_i^2 + \sigma_{i+1}^2)^2 - 4\varphi_i^2]^{\frac{1}{2}}\right\},$$

so that the example attains equality in (3.7).

It follows from Schwarz's inequality that $\sigma_1/\sigma_2 \geqq |\alpha_1|$ and this implies $p_1 \geqq 0$; similarly, $p_n \geqq 0$. For $2 \leqq k \leqq n - 1$, $p_k \geqq 0$ if $\sigma_k^2 - 2\alpha_{k-1}^2\sigma_{k-1}^2 + \alpha_{k-1}^4\sigma_k^2 \geqq 0$ and $\sigma_k^2 - 2\alpha_k^2\sigma_{k+1}^2 + \alpha_k^4\sigma_k^2 \geqq 0$. That is, $p_k \geqq 0$ if

$$(3.14)\qquad 1 + \frac{\sigma_k^2}{\sigma_{k-1}^2} \geqq \frac{2\varphi_{k-1}^2}{\sigma_{k-1}^2\sigma_k^2} \quad \text{and} \quad 1 + \frac{\sigma_k^2}{\sigma_{k+1}^2} \geqq \frac{2\varphi_k^2}{\sigma_k^2\sigma_{k+1}^2}.$$

These conditions are satisfied, e.g., for $\sigma_k^2 = \sigma^2$, $k = 1, 2, \cdots, n$ or for $\varphi_k = 0$, $k = 1, 2, \cdots, n - 1$.

With $n = 3$, the covariance matrix

$$(3.15)\qquad \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{7}}{8} & \frac{7}{16} \\[1ex] \frac{\sqrt{7}}{8} & \frac{1}{4} & \frac{\sqrt{7}}{8} \\[1ex] \frac{7}{16} & \frac{\sqrt{7}}{8} & \frac{1}{2} \end{bmatrix}$$

provides an example of $p_2 < 0$, since both conditions of (3.14) are violated. Thus we cannot claim sharpness for (3.7) under all conditions.

If we write $F_n(x)$ in the form $F_n(x) = x'Mx$, then $B^{-1} = M$. In view of Theorem 3.7 of [8], this is as expected.

To investigate the sharpness (3.8), let $b^{(1)}, b^{(2)}, \cdots, b^{(n)}$ be the columns of the matrix $B = (b_{ij})$ where $b_{ii} = 1$, $b_{ij} = \alpha_{i-1}\alpha_{j-1}(i \neq j)$ and α_k is given by (3.5) with $r_k = 1, k = 1, 2, \cdots, n-1$ and $\alpha_0 = 1$. Suppose that $Z = (Z_1, Z_2, \cdots, Z_n)'$ has the distribution

$$P\{Z = b^{(1)}\} = P\{Z = -b^{(1)}\} = \frac{1}{2}\left[\sigma_1^2 - \sum_{i=1}^{n-1} \frac{\alpha_i^2(\sigma_{i+1}^2 - \alpha_i^2\sigma_1^2)}{1-\alpha_i^4}\right] = \frac{p_1}{2},$$

$$P\{Z = b^{(k)}\} = P\{Z = -b^{(k)}\} = \frac{\sigma_k^2 - \alpha_{k-1}^2\sigma_1^2}{2(1-\alpha_{k-1}^4)} = \frac{p_k}{2}, k = 2, 3, \cdots, n,$$

$$P\{Z = 0\} = 1 - \sum_{k=1}^{n} p_k.$$

Using the relations

$$1 + \alpha_i^2 = \frac{c_i(c_i - d_i^{\frac{1}{2}})}{2\varphi_i^2}, \qquad \frac{\alpha_i^2}{1+\alpha_i^2} = \frac{c_i(c_i - d_i^{\frac{1}{2}}) - 2\varphi_i^2}{c_i(c_i - d_i^{\frac{1}{2}})},$$

one verifies that $\sum_{k=1}^{n} p_k$ is the bound of (3.8). It is straightforward to verify that $E(Z_i^2) = \sigma_i^2, i = 1, 2, \cdots, n$ and that $E(Z_iZ_{i+1}) = \varphi_i, i = 1, 2, \cdots, n-1$. Since $P\{\max_{1\leq i\leq n} |Z_i| \geqq 1\} = \sum_{k=1}^{n} p_k$, equality is attained in (3.8) whenever X has the same distribution as Z. Of course the example is valid only if $p_k \geqq 0$, $k = 1, 2, \cdots, n$. From Schwarz's inequality, it follows that $\sigma_k^2 - \alpha_{k-1}^2\sigma_1^2 \geqq 0$ so that $p_k \geqq 0, k = 2, 3, \cdots, n$. However, if the first and second rows and columns of (3.15) are interchanged, an example is obtained for which $p_1 < 0$. Thus, as in the case of (3.7), we cannot claim that (3.8) is sharp under all conditions.

This example can be obtained by arguments similar to those used in investigating sharpness of (3.7). Both examples can be obtained using the results of [8].

4. A lemma on separability. In the remainder of this paper, results of the preceding sections are used to obtain some inequalities of the Chebyshev type for continuous parameter stochastic processes. Separability of the processes will of course be required (the term "separable" will be used to mean "separable relative to the class of all closed subsets of the extended real line", although a weaker separability would suffice). From now on, the underlying probability space $(\Omega, \mathcal{B}, P)$ will be such that P is complete.

If $\{X_t, t \geqq 0\}$ is a separable process and S is a countable set satisfying the definition of separability and containing the points $0, \tau$, then $\{\sup_{t\varepsilon[0,\tau]} |X_t| < 1\}$ is measurable and $P\{\sup_{t\varepsilon[0,\tau]} |X_t| < 1\} = P\{\sup_{t\varepsilon S\cap[0,\tau]} |X_t| < 1\}$. However for a positive function f on $[0, \infty)$ it is not clear that $\{\sup_{t\varepsilon[0,\tau]} [|X_t|/f(t)] < 1\}$ is measurable and the following lemma is required.

LEMMA 4.1. *Let $\{X_t, t \geqq 0\}$ be a separable process, let f be a positive function on $[0, \infty)$ having at most countably many discontinuities, and let $\tau > 0$. If S is a countable set dense in $[0, \infty)$ satisfying the definition of separability and containing the set of discontinuities of f as well as 0 and τ, then $\{\omega: \sup_{t\varepsilon[0,\tau]} [|X_t(\omega)|/f(t)] < 1\}$ is measurable and*

$$P\left\{\sup_{t\varepsilon[0,\tau]} \frac{|X_t|}{f(t)} < 1\right\} = \lim_{k\to\infty} P\left\{|X_t| \leqq \frac{(k-1)f(t)}{k} \quad \text{for all} \quad t\varepsilon S\cap[0,\tau]\right\}. \tag{4.1}$$

PROOF. Let $\{t_j\}_{j=1}^{\infty}$ be an ordering of $S \cap [0, \tau]$ with the property that $\sup_{t \varepsilon [0,\tau]} \inf_{1 \leqq k \leqq n} |t - t_k| \to 0$ as $n \to \infty$. Let $\{s_{0,n}, s_{1,n}, \cdots, s_{n,n}, s_{n+1,n}\} = \{0, t_1, t_2, \cdots, t_n, \tau\}$ where $0 = s_{0,n} \leqq s_{1,n} < \cdots < s_{n,n} \leqq s_{n+1,n} = \tau$. Let $a_{k,n} = \sup_{t \varepsilon (s_{k-1,n}, s_{k,n})} f(t)$, $k = 1, 2, \cdots, n + 1$, $n = 1, 2, \cdots$, and let

$$f_n(\cdot) = \sum_{k=1}^{n+1} a_{k,n}\, \chi_{(s_{k-1,n}, s_{k,n})}(\cdot) + \sum_{k=0}^{n+1} f(s_{k,n})\, \chi_{\{s_{n,n}\}}(\cdot),$$

where for any set E, χ_E represents its characteristic function. By considering separately the case that $t \varepsilon S \cap [0, \tau]$ and the case that t is a continuity point of f, it is easily shown that $\lim_{n\to\infty} f_n(t) = f(t)$ for all $t \varepsilon [0, \tau]$.

Let $A_n = \{|X_t| \leqq f_n(t)$ for all $t \varepsilon S \cap [0, \tau]\}$, and let $B_n = \{|X_t| \leqq f_n(t)$ for all $t \varepsilon [0, \tau]\}$. Since $\{X_t, t \geqq 0\}$ is separable and P is complete, it can be shown that for all n, B_n is measurable and $P(A_n) = P(B_n)$. Since $f_n \geqq f_{n+1} \geqq f$, it follows that $A_n \supset A_{n+1}$ and $B_n \supset B_{n+1}$ for all n. Hence

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} P(B_n) = P\left(\bigcap_{n=1}^{\infty} B_n\right). \tag{4.2}$$

But

$$\begin{aligned} \bigcap_{n=1}^{\infty} A_n &= \{|X_t| \leqq f(t) \text{for all } t \,\varepsilon\, S \cap [0, \tau]\} \text{ and} \\ \bigcap_{n=1}^{\infty} B_n &= \{|X_t| \leqq f(t) \text{for all } t \,\varepsilon\, [0, \tau]\}. \end{aligned} \tag{4.3}$$

Now let $C_k = \{X_t \leqq [(k - 1)/k]f(t)$ for all $t \varepsilon [0, \tau]\}$, $k = 1, 2, \cdots$; applying (4.2) and (4.3) with $f(t)$ replaced by $[(k - 1)/k]f(t)$, it follows that C_k is measurable and $P(C_k) = P\{|X_t| \leqq [(k - 1)/k]f(t)$ for all $t \varepsilon S \cap [0, \tau]\}$. Since $C_k \subset C_{k+1}$ for all k, $P\{\sup_{t \varepsilon [0,\tau]}(|X_t|/f(t)) < 1\} = \lim_{k\to\infty} P(C_k) = \lim_{k\to\infty} P\{|X_t| \leqq [(k - 1)/k]f(t)$ for all $t \varepsilon S \cap [0, \tau]\}$, as was to be proved.

We remark that the assumptions of the above lemma are not sufficient to imply that the set $\{|X_t| < f(t)$ for all $t \varepsilon [0, \tau]\}$ is measurable. However $\{\sup_{t \varepsilon [0,\tau]} (|X_t|/f(t)) < 1\} \subset \{|X_t| < f(t)$ for all $t \varepsilon [0, \tau]\}$ so that if this latter set is measurable and if $P\{\sup_{t \varepsilon [0,\tau]} (|X_t|/f(t)) < 1\} \geqq \Phi$, then $P\{|X_t| < f(t)$ for all $t \varepsilon [0, \tau]\} \geqq \Phi$.

5. An inequality for semi-martingales. In this section we apply (2.3) to obtain an inequality for semi-martingales, and give an example to demonstrate sharpness.

THEOREM 5.1. *If $\{X_t, t \geqq 0\}$ is a separable semi-martingale such that $E|X_t| = \mu(t) < \infty$ for all $t \leqq \tau$, and if f is a non-decreasing positive function on $[0, \tau]$ such that the Riemann-Stieltjes integral in the following bound exists, then*

$$P\left\{\sup_{t \varepsilon [0,\tau]} \frac{X_t}{f(t)} \geqq 1\right\} \leqq \frac{\mu(0)}{f(0)} + \int_0^{\tau} \frac{d\mu(t)}{f(t)}. \tag{5.1}$$

PROOF. Define S, $s_{0,n}s_{1,n}, \cdots, s_{n+1,n}$ as in the proof of Lemma 4.1, and let

$f_k = [(k-1)/k]f$. Since $X_{s_{0,n}}, X_{s_{1,n}}, \cdots, X_{s_{n+1,n}}$ satisfy the conditions of (2.4),

$$P\left\{\max_{0 \leq i \leq n+1} \frac{X_{s_{i,n}}}{f_k(s_{i,n})} > 1\right\} \leqq \frac{\mu(0)}{f_k(0)} + \sum_{i=1}^{n+1} \frac{\mu(s_{i,n}) - \mu(s_{i-1,n})}{f_k(s_{i,n})}.$$

Since $\lim_{n\to\infty} \sup_{i=1,2,\cdots,n+1} (s_{i,n} - s_{i-1,n}) = 0$ and since the integral exists,

$$\lim_{n\to\infty} \sum_{i=1}^{n+1} \frac{\mu(s_{i,n}) - \mu(s_{i-1,n})}{f_k(s_{i,n})} = \int_0^\tau \frac{d\mu(t)}{f_k(t)}.$$

Then since

$$\lim_{n\to\infty} P\left\{\max_{0 \leq i \leq n+1} \frac{X_{s_{i,n}}}{f_k(s_{i,n})} > 1\right\} = P\left\{\sup_{t \varepsilon S \cap [0,\tau]} \frac{X_s}{f_k(t)} > 1\right\},$$

we obtain

$$P\left\{\sup_{t \varepsilon S \cap [0,\tau]} \frac{X_t}{f_k(t)} \leqq 1\right\} \geqq 1 - \frac{\mu(0)}{f_k(0)} - \int_0^\tau \frac{d\mu(t)}{f_k(t)}.$$

Hence by Lemma 4.1,

$$P\left\{\sup_{t \varepsilon [0,\tau]} \frac{X_t}{f(t)} < 1\right\} \geqq 1 - \frac{\mu(0)}{f(0)} - \int_0^\tau \frac{d\mu(t)}{f(t)} \text{ as claimed.}$$

If $\{X_t, t \geqq 0\}$ is a martingale and $r \geqq 1$, $\{|X_t|^r, t \geqq 0\}$ is a semi-martingale and it follows from Theorem 5.1 that if $\mu^r(t) = E\,|X_t|^r$,

$$P\left\{\sup_{t \varepsilon [0,\tau]} \frac{|X_t|}{f(t)} \geqq 1\right\} \leqq \frac{\mu^r(0)}{f^r(0)} + \int_0^\tau \frac{d\mu^r(t)}{f^r(t)}. \tag{5.2}$$

The restriction of Theorem 5.1 that f be monotone is not necessary; in any case, $g(t) = \inf_{t \leqq s \leqq \tau} f(s)$ is monotone and $g(t) \leqq f(t)$ on $[0, \tau]$ so that

$$P\left\{\sup_{t \varepsilon [0,\tau]} \frac{X_t}{f(t)} \geqq 1\right\} \leqq \frac{\mu(0)}{g(0)} + \int_0^\tau \frac{d\mu(t)}{g(t)}. \tag{5.3}$$

One can prove (5.3) sharp by replacing f by g in the example of the following theorem.

THEOREM 5.2. *Equality can be attained in* (5.1) *whenever the bound does not exceed one.*

PROOF. Let ω be a random variable such that

$$P\{\omega \leqq \omega_0\} = [\mu(0+)/f(0)] + \alpha(\omega_0)$$

where $\alpha(\omega)$ is the Lebesgue-Stieltjes integral $\int_{(0,\omega]} [d\mu(t+)/f(t)]$ (we denote $\lim_{s\,t} \mu(s)$ by $\mu(t+)$ and similarly define $\mu(t-)$). Let

$$\eta(t) = [\mu(t) - \mu(t-)]/[\mu(t+) - \mu(t-)]$$

unless μ is continuous at t, in which case $\eta(t) = 1$ (define $\mu(0-) = \mu(0)$). The process $\{Z_t, 0 \leqq t \leqq \tau\}$ defined on $[0, \tau]$ by

$$Z_t(\omega) = \begin{cases} 0, & t < \omega, \\ \eta(\omega) f(\omega), & t = \omega, \\ f(\omega), & \tau \geqq t > \omega, \end{cases}$$

is a semi-martingale since it has non-decreasing sample functions. Furthermore,

$$E|Z_t| = \frac{\mu(0+)}{f(0)}|Z_t(0)| + \int_{(0,t)} |Z_t(\omega)|\frac{d\mu(\omega+)}{f(\omega)} + |Z_t(t)|\frac{\mu(t+)-\mu(t-)}{f(t)} = \mu(t)$$

so that the process satisfies the conditions of Theorem 5.1.

The existence of the Riemann-Stieltjes integral $\int_0^\tau [d\mu(t)/f(t)]$ implies that f and μ have no common discontinuity points, so that $\sup_{t\epsilon[0,\tau]} [Z_t(\omega)/f(t)] = 1$ whenever $\omega \leqq \tau$. But $P\{\omega \leqq \tau\}$ is the bound of (5.1) so that the process attains equality, and the proof is complete.

It is possible to modify the Z_t process of Theorem 5.2 to obtain a martingale attaining equality in (5.2). Where the sample function $Z_t(\omega)$ jumps to some value, say v, the modification jumps to v or $-v$ with probabilities chosen so that the martingale condition is satisfied.

6. An inequality for a class of second-order processes. We now apply (3.7) to obtain an inequality for second-order processes satisfying certain regularity conditions.

The inequality, together with an ingeneous heuristic derivation, has already been given by Whittle [12]. Using (3.7) as a starting point we give a more straightforward proof, and, in the stationary case, we show that the inequality is sharp by defining a process attaining equality.

The procedures used to obtain an inequality for processes from (3.7) might also be used with (3.8) as a starting point. If this is tried, only a trivial bound is obtained.

THEOREM 6.1. *Let $\{X_t, t \geqq 0\}$ be a separable stochastic process with $E(X_t) = 0$ for all t, and let f be a positive function on $[0, \infty)$ with at most countably many discontinuities. For non-negative s and t, let $\sigma(s, t) = E(X_sX_t)$, $\sigma^2(t) = \sigma(t, t)$, and $g(s, t) = \sigma(s, t)/[f(s)f(t)]$. If g has continuous third partial derivatives, then*

$$P\left\{\sup_{t\epsilon[0,\tau]} \frac{|X_t|}{f(t)} \geqq 1\right\} \leqq \frac{1}{2}[g(0,0)+g(\tau,\tau)] + \int_0^\tau \left[g(t,t)\frac{\partial^2}{\partial x \partial y}g(x,y)|_{x=y=t}\right]^{\frac{1}{2}} dt. \tag{6.1}$$

PROOF. Since $g(\cdot, \cdot)$ is symmetric, it follows that for all non-negative s and t,

$$\begin{aligned} g_1(t) &= \left.\frac{\partial g(x,y)}{\partial x}\right|_{x=y=t} = \left.\frac{\partial g(x,y)}{\partial y}\right|_{x=y=t} = g_2(t) \\ g_{1,2}(t) &= \left.\frac{\partial^2 g(x,y)}{\partial x \partial y}\right|_{x=y=t} = \left.\frac{\partial^2 g(x,y)}{\partial y \partial x}\right|_{x=y=t} = g_{2,1}(t) \\ g_{1,1}(t) &= \left.\frac{\partial^2 g(x,y)}{\partial x^2}\right|_{x=y=t} = \left.\frac{\partial^2 g(x,y)}{\partial y^2}\right|_{x=y=t} = g_{2,2}(t). \end{aligned} \tag{6.2}$$

Since the third partial derivatives of g are continuous, it follows from Taylor's theorem that

$$g(t, t) + g(t + \Delta, t + \Delta) = 2g(t, t) + \Delta[g_1(t) + g_2(t)] + \tfrac{1}{2}\Delta^2[g_{1,1}(t) + 2g_{1,2}(t) + g_{2,2}(t)] + o(\Delta^2),$$

and

$$g(t, t + \Delta) = g(t, t) + \Delta g_2(t) + \tfrac{1}{2}\Delta^2 g_{2,2}(t) + o(\Delta^2)$$

for all non-negative Δ and t. Making use of (6.2) we obtain

$$\begin{aligned}(6.3)\quad &\{[g(t, t) + g(t + \Delta, t + \Delta)]^2 - 4g^2(t, t + \Delta)\}^{\frac{1}{2}} \\ &= 2\Delta\{g(t, t)g_{1,2}(t, t)[1 + o(\Delta^2)/\Delta^2]\}^{\frac{1}{2}} = 2\Delta[g(t, t)g_{1,2}(t, t)]^{\frac{1}{2}}[1 + o(\Delta)/\Delta], \\ &\qquad \text{for all } t, \Delta \geqq 0.\end{aligned}$$

Define $S, s_{0,n}, s_{1,n}, \cdots, s_{n+1,n}$ as in the proof of Lemma 4.1, and let $f_k = [(k - 1)/k]f$. Applying (3.7) and (6.3) we obtain

$$\begin{aligned}(6.4)\quad &\left\{P \max_{0 \leqq i \leqq n+1} \frac{|X_{s_{i,n}}|}{f_k(s_{i,n})} > 1\right\} \leqq \frac{k^2}{2(k-1)^2}\Big\{g(0,0) + g(\tau,\tau) \\ &\quad + \sum_{i=0}^{n} [[g(s_{i,n}, s_{i,n}) + g(s_{i+1,n} s_{i+1,n})]^2 - 4g^2(s_{i,n} s_{i+1,n})]^{\frac{1}{2}}\Big\} \\ &= \frac{k^2}{2(k-1)^2}\Big\{g(0,0) + g(\tau,\tau) + 2\sum_{i=1}^{n} (s_{i+1,n} - s_{i,n})[g(s_{i,n}, s_{i,n})g_{12}(s_{i,n})]^{\frac{1}{2}} \\ &\qquad \cdot \left(1 + \frac{o(s_{i+1,n} - s_{i,n})}{s_{i+1,n} - s_{i,n}}\right)\Big\}.\end{aligned}$$

The limit on the right side of (6.4) as $n \to \infty$ is

$$M_k = \frac{k^2}{(k-1)^2}\left\{\frac{1}{2}[g(0,0) + g(\tau,\tau)] + \int_0^\tau [g(t, t)g_{1,2}(t, t)]^{\frac{1}{2}} dt\right\},$$

and the limit of the left side of (6.4) as $n \to \infty$ is $P\{\sup_{t \varepsilon S \cap [0,\tau]} [|X_t|/f_k(t)] > 1\}$. From Lemma 4.1 it follows that

$$P\left\{\sup_{t \varepsilon [0,\tau]} \frac{|X_t|}{f(t)} \geqq 1\right\} = \lim_{k\to\infty} P\left\{\sup_{t \varepsilon S \cap [0,\tau]} \frac{|X_t|}{f_k(t)} > 1\right\} \leq \lim_{k\to\infty} M_k,$$

and the proof is complete.

COROLLARY 6.2. *Retain the hypotheses and notation of Theorem* 6.1 *and suppose that there is a real function h such that* $g(x, y) = h(y - x)$ *for all non-negative x and y. If* $H(\cdot) = \{1 - [h^2(\cdot)/h^2(0)]\}^{\frac{1}{2}}$ *has a derivative* $H'(0)$ *at the origin, then* $h'(0) = 0$ *and*

$$(6.5)\quad P\left\{\sup_{t \varepsilon [0,\tau]} \frac{|X_t|}{f(t)} \geqq 1\right\} \leqq h(0)[1 + \tau H'(0)] = h(0) + \tau[-h(0)h''(0)]^{\frac{1}{2}}.$$

PROOF. The second bound of (6.5) follows directly from (6.1). If

$$t_i = \frac{(i-1)\tau}{n-1}, \frac{\sigma_i^2}{\epsilon_i^2} = q(t_i, t_i) = h(0), i = 1, 2, \cdots, n, \frac{\varphi_i}{\epsilon_i \epsilon_{i+1}} = g(t_i t_{i+1}) = h\left(\frac{\tau}{n-1}\right),$$

$i = 1, 2, \cdots, n - 1$, then by applying (3.7) and passing to the limit on n, one obtains the first bound of (6.5). This bound can be rigorously established by showing that it is equal to the second bound. By computing $H'(t)$ and using the fact that $H'(0)$ exists, one can show that $h'(0) = 0$. The hypotheses of Theorem 6.1 imply that H' is continuous at the origin, so that by using a Taylor series expansion of $h^2(t)$, one obtains

$$H'(0) = - \lim_{t \to 0} \frac{h'(t)}{[h^2(0) - h^2(t)]^{\frac{1}{2}}} = \left[- \frac{h''(0)}{h(0)} \right]^{\frac{1}{2}},$$

and the desired result follows.

THEOREM 6.3. *Equality can be achieved in* (6.5) *whenever the bound does not exceed one.*

PROOF. The bound of (6.5) depends on $\{X_t, t \geqq 0\}$ only through $h(0)$ and $H'(0)$. To prove the theorem, we show that for all possible values of these parameters, there is a process $\{Z_t, 0 \leqq t \leqq \tau\}$ attaining equality in (6.5) with $E(Z_t) \equiv 0$ and with

$$h_Z(\Delta) = E(Z_t Z_{t+\Delta})/[f(t)f(t + \Delta)], \qquad H_Z(\Delta) = \{1 - [h_Z^2(\Delta)/h_Z^2(0)]\}^{\frac{1}{2}}$$

satisfying $h_Z(0) = h(0)$, $H_Z'(0) = H'(0)$.

Let $\Omega = [0, \tau] \times \{-1, 1\} \cup \{(0, 0)\}$, $\mathfrak{B}$ be the Borel subsets of Ω, and let P be the probability measure defined on $\mathfrak{B}$ by

$$P\{(0, 1)\} = P\{(0, -1)\} = P\{(\tau, 1)\} = P\{(\tau, -1)\} = h(0)/4,$$

$$P\{(\theta, \delta) : 0 < \theta < a, \delta = 1\} = P\{(\theta, \delta) : 0 < \theta < a, \delta = -1\}$$

$$= \tfrac{1}{2} a h(0) H'(0), 0 < a \leqq \tau,$$

$$P\{(0, 0)\} = 1 - h(0)[1 + \tau H'(0)].$$

Define the process $\{Z_t, 0 \leqq t \leqq \tau\}$ on $(\Omega, \mathfrak{B}, P)$ by

$$Z_t(\theta, \delta) = f(t)\delta \exp[- |t - \theta| H'(0)].$$

Then $E(Z_t) \equiv 0$ by symmetry, and

$$\begin{aligned} h_Z(\Delta) = {} & \tfrac{1}{2} h(0) \{\exp[- |t - 0| H'(0) - |t + \Delta - 0| H'(0)] \\ & + \exp[- |t - \tau| H'(0) - |t + \Delta - \tau| H'(0)]\} \\ & + \textstyle\int_0^\tau \exp[- |t - \theta| H'(0)] \exp[- |t + \Delta - \theta| H'(0)] h(0) H'(0)\, d\theta \\ & = h(0) \exp\{- \Delta H'(0)\}[1 + \Delta H'(0)]. \end{aligned}$$

Thus $h_Z(0) = h(0)$. Direct computation of $H_Z'(0) = [- h_Z''(0)/h(0)]^{\frac{1}{2}}$ yields $H_Z'(0) = H'(0)$. Thus the process $\{Z_t, 0 \leqq t \leqq \tau\}$ satisfies the conditions of Corollary 6.2. Since

$$P\left\{ \sup_{t \in [0, \tau]} \frac{|Z_t|}{f(t)} \geqq 1 \right\} = P\{(\theta, \delta) : 0 \leqq \theta \leqq \tau\} = h(0)[1 + \tau H'(0)],$$

the proof is complete.

In order to apply (6.5), one does not need to know the function h but only $h(0)$ and $H'(0)$; presumably a better bound could be given if h were known. The preceding example shows that if h is of the form $h(\Delta) = \sigma^2(1 + a\Delta)e^{-a\Delta}$ ($a \geqq 0$, $\Delta \geqq 0$), then no such improvement is possible.

Acknowledgments. The authors are most grateful to Robert M. Blumenthal and to the referee for a number of valuable suggestions and comments.

REFERENCES

[1] P. O. Berge, "A note of a form of Tchebycheff's theorem for two variables," *Biometrika*, Vol. 29 (1937), pp. 405–406.
[2] Y. S. Chow, "A martingale inequality and the law of large numbers," *Proc. Amer. Math. Soc.*, Vol. 11 (1960), pp. 107–111.
[3] J. L. Doob, *Stochastic Processes*. John Wiley and Sons, New York, 1953.
[4] Maurice Fréchet, *Recherches Théoriques Modernes sur la Théorie des Probabilitiés*, Book 1, Gauthier-Villars, Paris, 1937.
[5] J. Hájek and A. Rényi, "Generalization of an inequality of Kolmogorov," *Acta Math. Acad. Sci. Hungar.*, Vol. 6 (1955), pp. 281–283.
[6] D. N. Lal, "A note on a form of Tchebycheff's inequality for two or more variables," *Sankhyā*, Vol. 15 (1955), pp. 317–320.
[7] Michel Loève, *Probability Theory*. New York, D. Van Nostrand Co., 1955.
[8] Ingram Olkin and John W. Pratt, "A multivariate Tchebycheff inequality," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 226–234.
[9] Karl Pearson, "On generalized Tchebycheff theorems in the mathematical theory of statistics," *Biometrika*, Vol. 12 (1919), pp. 284–296.
[10] Henrik L. Selberg, "Zwei Ungleichungen zur ergänzung des Tchebycheffschen lemmas," *Skandinavisk Aktuarietidskrift*, Vol. 23 (1940), pp. 121–125.
[11] P. Whittle, "A multivariate generalization of Tchebichev's inequality," *Quart. J. Math.*, Oxford 2nd Ser., Vol. 9 (1958), pp. 232–240.
[12] P. Whittle, "Continuous generalizations of Tchebichev's inequality," *Teoriya veroyatnosteĭ i ee premeneniya*, Vol. 3 (1958), pp. 386–394.