The Annals of Probability
1981, Vol. 9, No. 3, 413–428

A METHOD OF APPROXIMATING EXPECTATIONS OF FUNCTIONS OF SUMS OF INDEPENDENT RANDOM VARIABLES[1,2]

BY MICHAEL J. KLASS

University of California, Berkeley

Let $X_1, X_2, \cdots$ be a sequence of independent random variables with $S_n = \sum_{i=1}^n X_i$. Fix $\alpha > 0$. Let $\Phi(\cdot)$ be a continuous, strictly increasing function on $[0, \infty)$ such that $\Phi(0) = 0$ and $\Phi(cx) \le c^\alpha \Phi(x)$ for all $x > 0$ and all $c \ge 2$. Suppose a is a real number and J is a finite nonempty subset of the positive integers. In this paper we are interested in approximating $E \max_{j \in J} \Phi(\mid a + S_j \mid)$. We construct a number $b_J(a)$ from the one-dimensional distributions of the X's such that the ratio $E \max_{j \in J} \Phi(\mid a + S_j \mid)/\Phi(b_J(a))$ is bounded above and below by positive constants which depend only on α. Bounds for these constants are given.

0. Introduction. Let $X_1, X_2, \cdots X_n$ be independent random variables whose distributions are considered to be given, and $S_n = \sum_{i=1}^n X_i$. Von Bahr and Esseen (1965) have shown that when the variables have zero means and $1 \le \beta \le 2$, $E \mid S_n \mid^\beta \le 2\sum_{i=1}^n E \mid X_i \mid^\beta$. When the variables are further assumed to be identically distributed, Davis (1973) obtained a lower bound of the form $E \mid S_n \mid \ge c_1 t$ where c_1 is a constant and t is some function[3] of the truncated second moment[4] of X. Under an additional moment condition, he also proved that $E \mid S_n \mid^\beta \le c_2 t^\beta$, where c_2 is another constant and β a number between 1 and 2. More recently, Klass (1979) obtained uniformly accurate upper and lower bounds for $E \mid S_n \mid$ which are independent of X for all $n \ge 1$. For each n, the bounds were shown to be nearly best possible. They are best possible asymptotically. For $\beta \ge 2$, the order of magnitude of $E \mid S_n \mid^\beta$ was found by Brillinger (1962) in the identically distributed, mean zero case, and by Rosenthal (1970) in the nonidentically distributed case.

Suppose $\phi(\cdot)$ is any function which is strictly increasing and continuous on $[0, \infty)$ such that $\Phi(0) = 0$ and some $\alpha > 0$, $\Phi(cx) \le c^\alpha \phi(x)$ for all $x > 0$ and all $c \ge 2$. (Any increasing continuous function on $[0, \infty)$ satisfying $\Phi(\mid x + y \mid) \le \gamma(\Phi(\mid x \mid) + \Phi(\mid y \mid))$, $\gamma > 0$, is of this type and conversely.) Can the n-dimensional integral $E\Phi(\mid a + S_n \mid) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Phi(\mid a + x_1 + \cdots + x_n \mid)\, dF_1(x_1) \cdots dF_n(x_n)$ be approximated in terms of 1-dimensional X_i-integrals? (Here $F_i(\cdot)$ is the distribution function of X_i and a is any constant.) We construct a number b_n based on a, Φ, and $\{F_i: i = 1, 2, \cdots, n\}$ and use $\Phi(b_n)$ to approximate $E\Phi(\mid a + S_n \mid)$. We derive uniform bounds (depending only on α) for the ratio $R_n = E\Phi(\mid a + S_n \mid)/\Phi(b_n)$. In fact a more general result is obtained. Let J be any nonempty subset of $\{1, 2, \cdots, n\}$. A number b_J is constructed such that $R_J = E \max_{j \in J} \Phi(\mid a + S_j \mid)/\Phi(b_J)$ is also uniformly bounded away from 0 and ∞. Actually, the special case $a = 0$ already embodies the general one. Nevertheless, we find it convenient to isolate any constant terms.

Received December 11, 1979; revised March 26, 1980.

[1] Research supported by National Science Foundation under Grant Number MCS75-10376-A01.

[2] Earlier (unpublished) versions of this paper were references in Klass (1975, 1976, 1979) under this and related titles.

[3] t plays the role of v_n as defined in Remark 7.2 of this paper.

[4] Quantity of the form $EX^2 I(\mid X \mid \le b)$.

AMS 1970 *subject classifications*. Primary 60G50, 60E15; secondary 60J15.

Key words and phrases. Sums of independent random variables, expectations, truncated mean, truncated expectation, truncated second moment, tail Φ-moment, K-function, approximation of expectations, approximation of integrals.

We split the derivation into a number of sections. Section 1 contains a few technical lemmas that are needed in subsequent proofs. Section 2 introduces the quantities to be used in the construction of b_J. In Section 3, an upper bound is obtained for R_J. Owing to a certain pathology, this particular bound can be arbitrarily poor. Section 4 discusses this phenomenon and suggests a remedy: centering the X's at their medians. Lower bounds for $E\Phi(\mid a + S_n \mid)$ in terms of certain tail Φ-moments[5] and truncated second moments of $Y_i \equiv X_i - \text{med } X_i$ are presented in Section 5. These are based on certain inequalities for symmetric random variables. Truncated second moments and tail Φ-moments govern the order of magnitude of $E\Phi(\mid a + S_n \mid)$ except when both the extremal Y's can be neglected and the distribution of $a + S_n$ is sufficiently concentrated about some nonzero value. Then the median of $\Phi(\mid a + S_n \mid) \doteq \Phi(\mid \text{med}(a + S_n) \mid)$ is a good approximation for $E\Phi(\mid a + S_n \mid)$. The median (or some other nearby quantile) can be approximated in terms of quantities of the form $a + \sum_{i=1}^n \text{meds } X_i + \sum_{i=1}^n EY_i I(\mid Y_i \mid \le b)$, termed "truncated expectations" or "truncated means". Details appear in Section 6. Section 7 contains the main theorem together with several remarks. These remarks touch on (i) cases in which centering at medians is unnecessary, (ii) alternative construction of truncation points, and (iii) one-sided bounds using another approximation scheme applicable in special cases.

Given an appropriate generalization of the notion of the median of a random variable, these results can be extended almost without change to sums of independent random elements taking values in a Hilbert space. Moreover, they can be further extended to Banach spaces, subject to some modification of the approximating quantities. These results will appear in a forthcoming paper [9].

1. Preliminaries. We will be mainly concerned with the following two classes of functions:

(1.1) $\mathscr{F} \equiv \{f(\cdot): f(\cdot)$ is continuous and strictly increasing on $[0, \infty)$, and satisfies $f(0) = 0$, $f(x) = f(-x)$ and $f(\infty) \equiv \lim_{x\to\infty} f(x)\}$.

(1.2) $\mathscr{F}_\alpha \equiv \{\Phi \in \mathscr{F}: 0 < \Phi(cx) \le c^\alpha \Phi(x)$ for all $x > 0$ and all $c \ge 2\}$.

Whenever we have a sequence of independent random variables (rv's) $X_1, X_2, \cdots$, the symbol S_n will mean the nth partial sum $\sum_{i=1}^n X_i$. The median of smallest absolute value of an rv X is written med X and $a \vee b$ ($a \wedge b$) denotes the maximum (minimum) of the real numbers a and b, with $a^+ = a \vee 0$, $a^- = (-a) \vee 0$. $\mathscr{P}_\lambda$ denotes a Poisson rv with mean λ.

We require the use of certain lemmas whose proofs we omit. We state them here for later reference. The first three can be proved via variational arguments. They are used in Lemma 3.3, Theorem 5.1, and Lemma 6.1 respectively.

LEMMA 1.1. *Suppose $\alpha \ge 0$ and let $I_1, I_2, \cdots, I_n$ be independent indicator* rvs. *Let $\sum_{j=1}^n P(I_j = 1) \le q \le 1$. Then*

$$E\left(1 + \sum_{j=1}^n I_j\right)^\alpha \le \begin{cases} E(1 + \mathscr{P}_q)^\alpha & \text{if } \alpha \ge 1 \\ 2^\alpha q + (1 - q) & \text{if } \alpha < 1. \end{cases}$$

LEMMA 1.2.

$$\inf\left\{\prod_{i=1}^{n-1} (1 - p_i): 0 \le p_i \le 1, 1 \le i \le n; \sum_{i=1}^n p_i \le q \le 1; \max_{1\le i\le n} p_i = p_n\right\}$$
$$= \left(1 - \frac{q}{n}\right)^{n-1} \downarrow e^{-q} \quad as \quad n \to \infty.$$

LEMMA 1.3. *Let $n \ge 1$. Then*

[5] Quantities of the form $E\Phi(\mid Y_i \mid) I(\mid Y_i \mid > b)$.

$$\inf\{\prod_{i=1}^n (1 - p_i): 0 \le p_i \le 1, 1 \le i \le n; \sum_{i=1}^n p_i \le q_i \le 1\} = 1 - q.$$

The next result is well known.

LEMMA 1.4. *Let μ, ν be positive σ-finite measures on $-\infty \le a \le b \le \infty$, such that for any $a \le x \le b$, $\mu[x, b] \le \nu[x, b]$. Let f be any nonnegative, nondecreasing function on $[a, b]$. Then for any $a \le c \le b$,*

$$\int_{[c,b]} f(x)\, d\mu(x) \le \int_{[c,b]} f(x)\, d\nu(x).$$

When f is absolutely continuous, the inequality follows from an integration by parts. However, since every nonnegative, nondecreasing function is the increasing limit of such absolutely continuous functions, application of monotone convergence establishes the result in general.

2. The approximating functions. Let $\vec{X} = (X_1, X_2, \cdots, X_n)$ be a vector of rv's. For $y > 0$ and any constant a, define

$$M_{\vec{X}}(y, a) \equiv \sup\{m: y \mid a + \textstyle\sum_{i=1}^n EX_iI(\mid X_i \mid \le m) \mid \ge m\}. \tag{2.1}$$

The set defining $M_{\vec{X}}(y, a)$ is nonempty. Application of dominated convergence shows that $\lim_{m\to\infty} E(\mid Y \mid /m)I(\mid Y \mid \le m) = 0$ for any rv Y. Hence $M_{\vec{X}}(y, a)$ is well-defined and finite. $M_{\vec{X}}(y, a)$ may be called a *truncated mean function* or simply truncated mean. We note that right-continuity of trucated expectations (expectations of truncated variables) yields

$$y \mid a + \textstyle\sum_{i=1}^n EX_iI(|X_i| \le m) \mid \le m \qquad \text{for} \quad m \ge M_{\vec{X}}(y, a). \tag{2.2}$$

Let $f \in \mathscr{F}$ (see (1.1)). As noted in the introduction, the approximation of $Ef(a + \sum_{i=1}^n X_i)$ depends not only on truncated means, but also on truncated second moments and tail f-moments. It is convenient to combine these last two items, adding their effects.

For each $y > 0$ let $K_{\vec{X}}(y)$ equal $\sum_{i=1}^n Ef(X_i)$ if this sum is 0 or ∞. Otherwise let $K_{\vec{X}}(y)$ be the unique positive real number such that

$$y \textstyle\sum_{i=1}^n EX^2I(|X_i| \le k) + y(k^2/f(k)) \sum_{i=1}^n Ef(X_i)I(|X_i| > k) \tag{2.3}$$

$$\begin{cases} > k^2 & \text{if} \quad 0 < k < K_{\vec{X}}(y) \\ = k^2 & \text{if} \quad k = K_{\vec{X}}(y) \\ < k^2 & \text{if} \quad k > K_{\vec{X}}(y) \end{cases}$$

$K_{\vec{X}}(\cdot)$ is the (generalized) *K-function* of $\vec{X}$ determined by f. The special case when $f(x) = |x|$ and $\vec{X} = X$, produces the K-function introduced in Klass (1976). To check that (2.3) actually holds for some $K_{\vec{X}}(y)$, divide by k^2 and note that for any rv X such that $0 < Ef(X) < \infty$,

$$\begin{aligned} g(k) &= E(X/k)^2I(|X| \le k) + E(f(X)/f(k))I(|X| > k) \\ &= E(|X| \wedge k)^2/k^2 + E(f(X) - f(k))^+/f(k) \end{aligned}$$

is the sum of two nonincreasing, continuous functions. The second function is strictly decreasing on $[0, \text{ess sup}|X|]$ while the first is strictly decreasing on $[\text{ess sup}|X|, \infty)$. Therefore $g(\cdot)$ is strictly decreasing and continuous on $(0, \infty)$. Moreover, it has range $(0, \infty)$. Thus, a unique, continuous, strictly increasing function$K_{\vec{X}}(\cdot)$ exists satisfying (2.3).

3. The upper bound. We devote this section to obtaining an upper bound for

$E\Phi(a + \sum_{i=1}^n X_i)$, with $\Phi \in \mathscr{F}_\alpha$ for some $\alpha > 0$. First we need an inequality giving an exponential bound for a certain type of tail probability.

LEMMA 3.1. *Let* $Y_1, Y_2, \ldots, Y_n$ *be independent* r.v.'s *with finite second moments and* $P(Y_i \le b) = 1$ *for* $1 \le i \le n$ *and some* $0 \le b \le \infty$. *Suppose* $y > 0$ *and* $\sum_{i=1}^n EY_i^2 \le v^2$. *Then*

$$P(\max_{1\le j\le n} \textstyle\sum_{i=1}^j (Y_i - EY_i) \ge y) \le \exp\{yb^{-1} - (yb^{-1} + v^2b^{-2})\log(1 + byv^{-2})\}.$$

PROOF. Let $x > 0$. A straightforward generalization of an inequality in Lemma 5.1 in Klass (1976) gives

$$P(\max_{1\le j\le n} \textstyle\sum_{i=1}^j (Y_i - EY_i) \le y) \ge \exp\{-xy + v^2b^{-2}(e^{xb} - 1 - xb)\}.$$

To obtain our lemma, simply put $x = b^{-1}\log(1 + \text{byv}^{-2})$. □

The next lemma utilizes the above bound on tail probabilities to derive an inequality for integrals involving sums of bounded centered variates.

LEMMA 3.2. *Let* $Y_1, Y_2, \ldots, Y_n$ *be independent* r.v.'s *such that* $\max_{1\le j\le n}|Y_i| \le b$ *with probability* 1 *and* $\sum_{i=1}^n EY_i^2 \le b^2$ *for some* $0 \le b \le \infty$. *Suppose* $\Phi \in \mathscr{F}_\alpha$ *for some* $\alpha > 0$. *Then*

$$E \max_{1\le j\le n} \Phi(\textstyle\sum_{i=1}^j (Y_i - EY_i)) \le C(\alpha)\Phi(b)$$

where $C(\alpha)$ *is a constant depending only on* α.

PROOF. If $b = 0$, the result is obvious. Observe that $\max_{1\le j\le n}\Phi(\sum_{i=1}^j (Y_i - EY_i)) \le \Phi(\infty)$. Hence if $b = \infty$ the result is also obvious. So suppose $0 < b < \infty$. Using Lemma 3.1 in the second inequality below,

$$\begin{aligned}
&P(\max_{l\le j\le n} |\textstyle\sum_{i=1}^j (Y_i - EY_i)| \ge y) \\
&\quad\le (P(\max_{1\le j\le n} \textstyle\sum_{i=1}^j (Y_i - EY_i) \ge y) + P(\max_{1\le j\le n} \textstyle\sum_{i=1}^j (EY_i - Y_i) \ge y)) \wedge 1 \\
&\quad\le (2\exp\{yb^{-1} - (yb^{-1} + 1)\log(1 + yb^{-1})\}) \wedge 1 \equiv \nu[y, \infty) \qquad \text{(say)}.
\end{aligned}$$

ν is a probability measure. Application of Lemma 1.4 yields

$$\begin{aligned}
E\Phi(\max_{1\le j\le n} |\textstyle\sum_{i=1}^j (Y_i - EY_i)|) &\le \int_0^\infty \Phi(y)\, d\nu(y) \\
&\le \Phi(4b)\nu[0, 4b) + \int_{4b}^\infty \Phi(y)\, d\nu(y) \\
&\le 4^\alpha\Phi(b)\nu[0, 4b) + \Phi(b)\int_{4b}^\infty (yb^{-1})^\alpha\, d\nu(y) \\
&\qquad \text{(since } \Phi \in \mathscr{F}_\alpha) \\
&= \Phi(b)\{4^\alpha + \int_{4b}^\infty ((yb^{-1})^\alpha - 4^\alpha)\, d\nu(y)\} \\
&= \Phi(b)\{4^\alpha + 2\alpha\int_4^\infty x^{\alpha-1}\exp[x - (1 + x)\log(1 + x)]\, dx\} \\
&= C(\alpha\Phi(b). \qquad \square
\end{aligned}$$

REMARK 3.1. Since $x - (1 + x)\log(1 + x) \le -x$ for $x \ge 4$, a crude upper bound for $C(\alpha)$ is

$$C(\alpha) \le 4^\alpha + 2\Gamma(\alpha + 1), \qquad \alpha > 0 \tag{3.1}$$

where $\Gamma(\cdot)$ is the gamma function.

The next result bounds the influence of the "big" X's.

LEMMA 3.3. *Let $\Phi \in \mathscr{F}_\alpha$ for some $\alpha > 0$. Let $\vec{X} = (X_1, X_2, \cdots, X_n)$ be a vector of independent* rv's *such that $\sum_{i=1}^n P(|X_i| > b) \le q$ for some $0 \le b \le \infty$ and $0 \le q \le 1$. Then*

$$E\Phi(\textstyle\sum_{i=1}^n |X_i| I(|X_i| > b)) \le \{(2^\alpha q + 1 - q) \vee E(1 + \mathscr{P}_q)^\alpha\} \sum_{i=1}^n E\Phi(X_i I(|X_i| > b)). \tag{3.2}$$

Thus if $b \ge K_{\vec{X}}(y)$ for some $y \ge 1$ (see (2.3)), then

$$E\Phi(\textstyle\sum_{i=1}^n |X_i| I(|X_i| > b)) \le y^{-1}\{(2^\alpha y^{-1} + 1 - y^{-1}) \vee E(1 + \mathscr{P}_{y^{-1}})^\alpha\}\Phi(b). \tag{3.3}$$

PROOF. Inequalities (3.2) and (3.3) are obvious if $b = \infty$. Hence suppose $0 \le b < \infty$. Write $X_i'' = X_i I(|X_i| > b)$. Since

$$\textstyle\sum_{i=1}^n |X_i''| \le (\sum_{i=1}^n I(|X_i| > b)) \max_{1\le j\le n} |X_j''|,$$

we have

$$\begin{aligned}\Phi(\textstyle\sum_{i=1}^n |X_i''|) &\le \textstyle\sum_{j=1}^n \Phi(\sum_{i=1}^n I(|X_i| > b)|X_j''|)\\ &= \textstyle\sum_{j=1}^n \Phi(|X_j''|[1 + \sum_{i=1, i\ne j}^n I(|X_i| > b)])\\ &\le \textstyle\sum_{j=1}^n [1 + \sum_{i=1, i\ne j}^n I(|X_i| > b)]^\alpha \Phi(|X_j''|).\end{aligned}$$

Taking expectations and using independence, we obtain

$$E\Phi(\textstyle\sum_{i=1}^n |X_i''|) \le E[1 + \sum_{i=1}^n I(|X_i| > b)]^\alpha \sum_{j=1}^n E\Phi(|X_j''|). \tag{3.4}$$

Application of Lemma 1.1 then yields (3.2). If $b \ge K_{\vec{X}}(y)$, (2.3) implies that

$$\textstyle\sum_{i=1}^n P(|X_i| > b) \le y^{-1} \quad \text{and} \quad \sum_{i=1}^n E\Phi(|X_j''|) \le \Phi(b)/y, \quad \text{even if} \quad b = 0.$$

Letting $q = y^{-1}$, (3.3) follows from (3.2). □

It is now easy to obtain a general upper bound.

THEOREM 3.1. *Let $\Phi \in \mathscr{F}_\alpha$ for some $\alpha > 0$ and $X_1, X_2, \cdots$ be a sequence of independent* rv's. *For $j \ge 1$, let $\vec{X}_j$ be the vector $(X_1, X_2, \cdots, X_j)$ and $S_j = \sum_{i=1}^j X_i$. Let J be a finite nonempty subset of the positive integers and for any real number a define*

$$K_j = K_{\vec{X}_j}(1), \qquad M_j(a) = M_{\vec{X}_j}(1, a), \qquad K_J = \max_{j\in J} K_j, \qquad M_J(a) = \max_{j\in j} M_j(a).$$

Suppose n is the largest element in J. Then

$$K_J = K_n \quad \textit{and} \quad E \max_{j\in J} \Phi(a + S_j) \le B(\alpha)\, \Phi(K_J \vee M_J(a))$$

where $B(\alpha)$ is a constant depending only on α.

PROOF. Fix a and J. Since K_j is nondecreasing in j, it follows immediately that $K_J = K_n$. Let $b = K_n \vee M_J(a)$. For $1 \le j \le n$, write

$$X_j' = X_j I(|X_j| \le b), X_j'' = X_j - X_j', \quad \text{and} \quad \bar{S}_j = \textstyle\sum_{i=1}^j (X_i' - EX_i').$$

We first observe that $X'_1, X'_2, \cdots, X'_n$ satisfy the condition in Lemma 3.2. For any $j \in J$, the triangle inequality and (2.2) yields

$$\Phi(a + S_j) \le \Phi(\mid a + \textstyle\sum_{i=1}^{j} EX'_i \mid + \sum_{i=1}^{j} \mid X''_i \mid + \mid \bar{S}_j \mid)$$
$$\le \Phi(b + \textstyle\sum_{i=1}^{n} \mid X''_i \mid + \max_{1 \le i \le n} \mid \bar{S}_i \mid)$$
$$\le 3^\alpha\{\Phi(b) + \Phi(\textstyle\sum_{i=1}^{n} \mid X''_i \mid) + \Phi(\max_{1 \le i \le n} \mid \bar{S}_i \mid \}$$

where the last inequality is a simple consequence of the fact that $\Phi \in \mathscr{F}_\alpha$. We now take max over $j \in J$ and then take expectations. Applying Lemma 3.2 and Lemma 3.3 to the result, we get

$$E \max_{j \in J} \Phi(a + S_j) \le 3^\alpha\{1 + C(\alpha) + 2^\alpha \vee E(1 + \mathscr{P}_1)^\alpha\}\Phi(b)$$
$$\equiv B(\alpha)\Phi(K_J \vee M_J(\alpha)). \qquad \square$$

REMARK 3.2. As a result of (3.1), an upper bound for $B(\alpha)$ is

$$B(\alpha) \le 3^\alpha\{1 + 4^\alpha + 2\Gamma(\alpha + 1) + 2^\alpha \vee E(1 + \mathscr{P}_1)^\alpha\}. \tag{3.5}$$

REMARK 3.3. In many interesting situations computation of $M_J(\alpha)$ is unnecessary because $K_n \ge M_J(\alpha)$. For instance, suppose $\Phi(x)/x$ is nondecreasing on $[0, \infty)$ (which happens whenever Φ is convex and increasing on $[0, \infty)$), $a = 0$ and the X's have zero means. Then for any $1 \le j \le n$ and $m > 0$,

$$\mid \textstyle\sum_{i=1}^{j} EX_iI(\mid X_i \mid \le m) \mid \le \sum_{i=1}^{j} E \mid X_i \mid I(\mid X_i \mid > m)$$
$$\le (m/\Phi(m)) \textstyle\sum_{i=1}^{j} E\Phi(X_i)I(\mid X_i \mid > m).$$

Now if $m > K_n$, the above quantity is less than m. Hence by (2.1), $M_j(0) \le K_n$. This being true for $1 \le j \le n$ implies that $M_J(0) \le K_n$.

4. The need to center the variables. Does the truncation point $b_n = K_n \vee M_n(a)$ generate a "good" approximation of $E\Phi(a + S_n)$? This is the key issue to be resolved. We perceive three potentially dominating aspects of the $(a + S_n)$ distribution:

(i) The *median* m_n of $(a + S_n)$ (or some other qth quantile where q is bounded away from 0 and 1). This may be thought of as the net trend of $a + S_1, a + S_2, \cdots, a + S_n$.

(ii) The *range* or dispersion r_n of the $(a + S_n)$-distribution as measured by, say, the median r_n of $\mid (a + S_n) - m_n \mid$ (or some nearby quantile). This may be thought of as the contribution to the expectation in question made by the size of the typical deviation of $a + S_n$ from its median. It is relevant when $r_n \ge 2 \mid m_n \mid$. Alternatively we may think of the range of $(a + S_n)$ as some appropriate quantile of $\mid S_n - \bar{S}_n \mid$, where S_n and $\bar{S}_n$ are independent and identically distributed (i.i.d.). This obviates the need to center $(a + S_n)$ to determine its range.

(iii) The maximal $\mid X_i \mid$.

Our b_n attempts to approximate the largest of the contributions made by (i), (ii) and (iii). To do so, it utilizes sums of (i′) truncated expectations, (ii′) truncated second moments, and (iii′) tail Φ-moments of the X_i's.

One important consideration has been neglected: the magnitude of b_n (and hence $\Phi(b_n)$) is sensitive to X_i's which are grossly improperly centered. For example, let W_1, W_2 be i.i.d. with $P(W_i = 1) = P(W_i = -1) = ½$. Choose any constant $c \ge 0$ and let $X_i = W_i + (-1)^i c$. Then $X_1 + X_2 = W_1 + W_2$ and $E\Phi(X_1 + X_2) = E\Phi(W_1 + W_2) = ½\, \Phi(2)$. As is easily checked, $b_n \sim c\sqrt{2}$ as $c \to \infty$. Hence the estimate given by (naive use of) Theorem 3.1 is asymptotic to $\Phi(c\sqrt{2})$ as $c \to \infty$. For unbounded Φ, this is arbitrarily larger than $½\, \Phi(2)$.

This problem does not arise when the sum of the second moments of the suitably truncated variables has the same order of magnitude as the sum of the variances of the same truncated variables. By transforming the variables, we can create such a situation.

Given any vector of constants $\vec{c} = (c_1, \cdots, c_n)$, $a + S_n = (a + \sum_{i=1}^n c_i) + \sum_{i=1}^n (X_i - c_i)$. Having selected $\vec{c}$, replace a by $a + \sum_{i=1}^n c_i$ and X_i by $X_i - c_i$ (thereby retaining independence) and then construct the associated truncation point $b_{\vec{c}}$. Using Theorem 3.1, we get

$$E\Phi(a + S_n) \leq B(\alpha) \inf_{\vec{c}} \Phi(b_{\vec{c}})$$

with a similar inequality holding for $E \max_{j\in J} \Phi(a + S_j)$. We make no attempt to construct the optimal $\vec{c}$, a difficult task at best. For our purposes it is adequate to center X_i at $c_i =$ med X_i. Thus we let $Y_i = X_i -$ med X_i, thereby constructing independent rv's whose distributions depend only on the corresponding X_i. They have the further property that for any $b > 0$, $Y_i' \equiv Y_i I(|Y_i| \leq b)$ satisfies $E(Y_i')^2 \leq 2 \operatorname{Var} Y_i'$ (whence $\sum_{i=1}^n E(Y_i')^2 \leq 2 \sum_{i=1}^n \operatorname{Var} Y_i'$). Let b_j^* denote the truncation point corresponding to $(a + \sum_{i=1}^j \text{med } X_i) + \sum_{i=1}^j Y_i$ and put $b_J^* = \max_{j\in J} b_j^*$. Theorem 3.1 then states that $E \max_{j\in J} \Phi(a + S_j)/\Phi(b_J^*)$ is bounded away from infinity above (uniformly in J and $\Phi \in \mathscr{F}_\alpha$). We will show that this ratio is also uniformly bounded away from zero below. Therefore the order of magnitude of (say) $E\Phi(a + S_n)$ is solely governed by the items in (i), (ii) and (iii) above provided we amend (iii) to read "maximal $|Y_i|$".

5. Lower bounds based on the *K*-function.

We need a lower bound to complement Theorem 3.1.

LEMMA 5.1. *Let $Y_1, Y_2, \cdots, Y_n$ be independent* rv's *symmetric about zero and $f \in \mathscr{F}$. Fix $b > 0$. Then*

$$Ef(Y_1 + Y_2 + \cdots + Y_n) \geq \tfrac{1}{2} \max\{E \max_{1\leq i\leq n} f(Y_i), Ef(\textstyle\sum_{i=1}^n Y_i I(|Y_i| \leq b))\}.$$

PROOF. We write $T_n = Y_1 + Y_2 + \cdots + Y_n$ for short. Let τ be the first i such that $|Y_i| = \max_{1\leq j\leq n} |Y_j|$, $1 \leq i \leq n$. The conditional distribution of $T_n - Y_\tau$ given Y_τ is symmetric about zero. Hence at least half the time it has the same sign as Y_τ, whence

$$|T_n| = |(T_n - Y_\tau) + Y_\tau| \geq |Y_\tau|$$

so that on such a set, $f(T_n) \geq f(Y_\tau)$. Therefore

$$Ef(T_n) \geq \tfrac{1}{2} Ef(Y_\tau) = \tfrac{1}{2} E \max_{1\leq i\leq n} f(Y_i).$$

To prove the other half of the result, we note that the conditional distribution of $\sum_{i=1}^n Y_i I(|Y_i| > b)$ given $\{Y_i I(|Y_i| \leqq b), 1 \leq i \leq n\}$ is symmetric about zero. A similar argument therefore yields

$$Ef(T_n) \geq \tfrac{1}{2} Ef(\textstyle\sum_{i=1}^n Y_i I(|Y_i| \leq b))$$

and this concludes the proof. □

REMARK 5. When f is also convex, an argument based on Jensen's inequality for conditional expectations shows that the constant factor ½ may be replaced by 1.

The idea of the next lemma is well-known. Variants of this theme appear in Chung (1974, page 48, exercise 11) and Davis (1973, Lemma 1).

LEMMA 5.2. *Let $Y_1, Y_2, \cdots, Y_n$ be independent mean zero* rv's *taking values in $[-b, b]$ for some $b > 0$. Let $v^2 = \sum_{i=1}^n EY_i^2 > 0$. Then for any $0 < \gamma < 1$,*

$$P(|\textstyle\sum_{i=1}^n Y_i| \geq \sqrt{\gamma}\, v) \geq (1 - \gamma)^2 v^2/(b^2 + 3v^2).$$

PROOF. Put $W = (\sum_{i=1}^n Y_i)^2$ and note that $EW = v^2$ and

$$\begin{aligned} EW^2 &= \sum_{i=1}^n EY_i^4 + 6\sum\sum_{1\le i<j\le n} EY_i^2Y_j^2 \\ &\le b^2 \sum_{i=1}^n EY_i^2 + 3\left(\sum_{i=1}^n EY_i^2\right)^2 \\ &= v^2(b^2+3v^2). \end{aligned}$$

Since

$$(1-\gamma)EW \le \int_{\{W\ge\gamma EW\}} W\,dP \le \{EW^2P(W\ge\gamma EW)\}^{1/2}$$

we may conclude that

$$\begin{aligned} P\left(\left|\sum_{i=1}^n Y_i\right| \ge \sqrt{\gamma}v\right) = P(W\ge\gamma EW) &\ge \{(1-\gamma)EW\}^2/EW^2 \\ &\ge (1-\gamma)^2v^2/(b^2+3v^2). \end{aligned}$$ □

THEOREM 5.1. *Let $\Phi \in \mathscr{F}_\alpha$ for some $\alpha > 0$. Let $\vec{Y} = (Y_1, Y_2, \cdots, Y_n)$ be a vector of independent* rv's *symmetric about zero. For fixed $y \ge 2$, let $b = K_{\vec{Y}}(y)$. Then, whether or not b is finite,*

$$E\Phi\left(\sum_{i=1}^n Y_i\right) \ge D(y,\alpha)\Phi(b)$$

where

$$D(y,\alpha) = 8(2y+3)^{-1}(\alpha+4)^{-2}\left[\frac{\alpha}{2y(\alpha+4)}\right]^{\alpha/2} \tag{5.1}$$

PROOF. Without loss of generality, we may suppose that $\Phi(b) > 0$ and $E\Phi(\sum_{i=1}^n Y_i) < \infty$. Therefore $E\Phi(Y_i) < \infty$ for $1 \le i \le n$ and so $0 < b < \infty$. For simplicity, we let

$$T_n = \sum_{i=1}^n Y_i,\ Y_i' = Y_iI(|Y_i| \le b),\quad \text{and}\quad Y_i'' = Y_i - Y_i' \quad \text{for}\quad 1 \le i \le n.$$

The definition of b implies that

$$yb^{-2}\sum_{i=1}^n E(Y_i')^2 + y(\Phi(b))^{-1}\sum_{i=1}^n E\Phi(Y_i'') = 1. \tag{5.2}$$

One of the two terms in the left-hand side must be $\ge \frac{1}{2}$. We consider both possibilities.

Case 1. Suppose

$$y\sum_{i=1}^n E\Phi(Y_i'') \ge \tfrac{1}{2}\Phi(b). \tag{5.3}$$

We first observe that (5.2) entails

$$\sum_{i=1}^n P(Y_i'' \ne 0) \le y^{-1} \le \tfrac{1}{2}. \tag{5.4}$$

Let

$$\tau = \begin{cases} 1\text{st } i, & 1 \le i \le n, \text{ such that } Y_i'' \ne 0 \\ \infty & \text{if no such } i \text{ exists.} \end{cases}$$

Lemma 5.1 implies that

$$\begin{aligned} E\Phi(T_n) &\ge \tfrac{1}{2}E\max_{1\le i\le n}\Phi(Y_i) \ge \tfrac{1}{2}\sum_{i=1}^n E\Phi(Y_i)I(\tau = i) \\ &= \tfrac{1}{2}\sum_{i=1}^n EI(\tau \ge i)\Phi(Y_i'') \\ &= \tfrac{1}{2}\sum_{i=1}^n P(\tau \ge i)E\Phi(Y_i'') \\ &\ge \tfrac{1}{2}P(\tau \ge n)\sum_{i=1}^n E\Phi(Y_i'') \\ &= \tfrac{1}{2}\prod_{i=1}^{n-1}\{1 - P(Y_i'' \ne 0)\}\sum_{j=1}^n E\Phi(Y_j''). \end{aligned} \tag{5.5}$$

We may reorder the Y_i's if necessary to ensure that $\max_{1\le i\le n} P(Y_i'' \neq 0) = P(Y_n'' \neq 0)$ without altering the value of $E \max_{1\le i\le n} \Phi(Y_i)$. In view of (5.4), Lemma 1.2 now yields

$$\prod_{i=1}^{n-1} \{1 - P(Y_i'' \neq 0)\} \geq [1-(ny)^{-1}]^{n-1} \geq e^{-1/y} \tag{5.6}$$

which together with (5.3) and (5.5) yields

$$E\Phi(T_n) \geq (4y)^{-1} e^{-1/y} \Phi(b) > D(y, \alpha)\Phi(b) \tag{5.7}$$

where the last inequality follows from elementary calculus.

Case 2. Next suppose

$$yb^{-2} \sum_{i=1}^n E(Y_i')^2 \geq ½. \tag{5.8}$$

Let $0 \le \gamma \le 1$ and $v^2 = \sum_{i=1}^n E(Y_i')^2$. Then $v^2 \geq b^2/(2y)$. Applying Lemma 5.2 in the second inequality below,

$$\begin{aligned} E\Phi(\textstyle\sum_{i=1}^n Y_i') &\geq \Phi(\sqrt{\gamma} v) P(|\textstyle\sum_{i=1}^n Y_i'| \geq \sqrt{\gamma} v) \\ &\geq \Phi(b(\gamma/2y)^{1/2})(1-\gamma)^2 v^2/(b^2+3v^2) \\ &\geq \Phi(b(\gamma/2y)^{1/2})(1-\gamma)^2 (b^2/2y)/\{b^2 + 3b^2/2y\} \\ &\geq (\gamma/2y)^{\alpha/2}\Phi(b)(1-\gamma)^2/(2y+3) \end{aligned} \tag{5.9}$$

where we use the fact that $\Phi \in \mathscr{F}_\alpha$ in the last inequality. The right-hand side of (5.9) is maximized when $\gamma = \alpha/(\alpha+4)$. Hence

$$E\Phi(\textstyle\sum_{i=1}^n Y_i') \geq (2y+3)^{-1} \left(\dfrac{4}{\alpha+4}\right)^2 \left[\dfrac{\alpha}{2y(\alpha+4)}\right]^{\alpha/2} \Phi(b). \tag{5.10}$$

Lemma 5.1 now finishes the proof of the theorem. □

In order to generalize the above result to rv's which are not necessarily symmetric, we need the following fact:

LEMMA 5.3. *Let* $\vec{Y} = (Y_1, Y_2, \cdots, Y_n)$ *be a vector of independent* rv's *with median zero. Let* $(\bar{Y}_1, \bar{Y}_2, \cdots, \bar{Y}_n)$ *be an independent copy of* $\vec{Y}$ *and define* $\vec{Y}^* = (Y_1 - \bar{Y}_1, Y_2 - \bar{Y}_2, \cdots, Y_n - \bar{Y}_n)$. *Then for any* $y > 0$ *and* $f \in \mathscr{F}$,

$$K_{\vec{Y}^*}(2y) \geq K_{\vec{Y}}(y).$$

PROOF. Consider independent rv's V and W such that med $W = 0$. For any symmetric function $h(\cdot)$, nondecreasing on $[0, \infty)$, it is easy to see that

$$Eh(V-W) \geq ½Eh(V).$$

For fixed $y > 0$, let $k = K_{\vec{Y}^*}(2y)$ and $Y_i^* = Y_i - \bar{Y}_i$, $1 \le i \le n$. We observe that for $f \in \mathscr{F}$,

$$\begin{aligned} h_k(x) &= x^2 I(|x| \le k) + (k^2/f(k)) f(x) I(|x| > k) \\ &= x^2 \wedge k^2 + (k^2/f(k))(f(x) - f(k))^+ \end{aligned}$$

is a symmetric function of x, nondecreasing on $[0, \infty)$. Thus

$$2Eh_k(Y_i^*) = 2Eh_k(Y_i - \bar{Y}_i) \geq Eh_k(Y_i), \qquad 1 \le i \le n.$$

The definition of k now implies that

$$k^2 = 2y \textstyle\sum_{i=1}^n Eh_k(Y_i^*) \geq y \sum_{i=1}^n Eh_k(Y_i)$$

which, in view of (2.3), shows that $k \geq K_{\vec{Y}}(y)$. □

By first centering the rv's at their medians and then applying a symmetrization argument, the following improvement of Theorem 5.1 is obtained.

THEOREM 5.2. *Let* $X_1, X_2, \cdots$ *be a sequence of independent* rv's *and* $\Phi \in \mathscr{F}_\alpha$ *for some* $\alpha > 0$. *Let* $Y_i = X_i - \text{med } X_i$ *for* $i \ge 1$ *and for an positive integer* j, *write* $S_j = X_1 + X_2 + \cdots + X_j$, $\vec{Y}_j = (Y_1, Y_2, \cdots, Y_j)$, *and* $K_j = K_{\vec{Y}_j}(1)$, *the K-function associated with* Φ *and* $\vec{Y}_j$. *Let* J *be any finite nonempty subset of the positive integers with largest element* n. *Then for any real* a,

$$E \max_{j\in J}\Phi(a + S_j) \ge G(\alpha)\Phi(K_n)$$

where

$$G(\alpha) = 7^{-1}4^{1-\alpha}(\alpha + 4)^{-2}\left(\frac{\alpha}{\alpha + 4}\right)^{\alpha/2} \tag{5.11}$$

PROOF. Let $(\bar{X}_1, \bar{X}_2, \cdots)$ be an independent copy of $(X_1, X_2, \cdots)$ and let $\bar{Y}_i = \bar{X}_i - \text{med } X_i$, $Y_i^* = Y_i - \bar{Y}_i = X_i - \bar{X}_i$, $\bar{S}_n = \bar{X}_1 + \bar{X}_2 + \cdots + \bar{X}_n$, and $\vec{Y}_n^* = (Y_1^*, Y_2^*, \cdots, Y_n^*)$. Theorem 5.1 yields

$$\begin{aligned} D(2, \alpha)\Phi(K_{\vec{Y}_n^*}(2)) &\le E\Phi(Y_1^* + Y_2^* + \cdots + Y_n^*) \\ &= E\Phi(a + S_n - (a + \bar{S}_n)) \\ &\le 2^\alpha\{E\Phi(a + S_n) + E\Phi(a + \bar{S}_n)\} \\ &= 2^{\alpha+1}E\Phi(a + S_n). \end{aligned}$$

Since Φ is nondecreasing, Lemma 5.3 yields

$$E \max_{j\in J}\Phi(a + S_j) \ge E\Phi(a + S_n) \ge 2^{-\alpha-1}D(2, \alpha)\Phi(K_n) = G(\alpha)\Phi(K_n). \qquad \square$$

6. Assessing trend. Let $W = a + \sum_{i=1}^n X_i$ and let $\Phi(\cdot) \in \mathscr{F}_\alpha$ for some $\alpha > 0$. Consider approximating $E\Phi(W)$ when the distribution of W is highly concentrated. A natural estimate is $\Phi(m)$, where m is a median (or nearby quantile) of W. Unfortunately, the complicated nature of convolutions typically prevents med W from being calculated. An approximation of med W is needed. Ideally, it should depend on one-dimensional quantities considered to be computable.

Classical results suggest an approach: Consider sums S_n of i.i.d. rv's with finite nonzero mean. Because these sums trend, their expectations ES_n and medians med S_n are asymptotically equivalent. Blind substitution of EW for med W, however, does not succeed in our more general context. The difficulty is posed by extremal X_i's which occur with low probability and therefore affect EW much more than med W. To avoid this contingency, we consider "truncated expectations".

Suppose it is possible to find $b > 0$ such that $m_b \equiv |a + \sum_{i=1}^n EX_iI(|X_i| \le b)|$ is nearly $|m|$. When W is highly concentrated (and the effect of badly centered X's can be neglected), it is reasonable to expect that b may be chosen almost equal to $|m|$ (which in turn is approximately m_b). This essentially defines b implicitly as $M_{\vec{X}}(1, a)$ and thereby provides an heuristic interpretation of the role played by $\Phi(M_{\vec{X}}(1, a))$ in the approximation of $E\Phi(W)$. In fact, for purposes of approximating $E\Phi(W)$, $M_{\vec{X}}(1, a)$ need only be close to the absolute value of med S_n (or nearby quantile) when $K_n < cM_{\vec{X}}(1, a)$ for some appropriate $c > 1$.

LEMMA 6.1. *Let* $\Phi \in \mathscr{F}_\alpha$ *for some* $\alpha > 0$. *Let* $\vec{X}_j = (X_1, X_2, \cdots, X_j)$ *be a vector of independent* rv's *and* $S_j = \sum_{i=1}^j X_i$ *for* $1 \le j \le n$. *Then for any real number* a *and* $1 \le j \le n$,

$$\Phi(M_{\vec{X}_j}(1, a)) \le \max\{\Phi(K_{\vec{X}_n}(2)), 2^{\alpha+2}E\Phi(a + S_j)\}.$$

PROOF. We fix a and j throughout the proof. To simplify notation, we define

$$M = M_{\vec{X}_j}(1, a), \qquad K = K_{\vec{X}_n}(2), \qquad X_i' = X_i I(|X_i| \le M), \qquad X_i'' = X_i - X_i',$$

$$T_j = \sum_{i=1}^{j} (X_i' - EX_i'), \qquad m = a + \sum_{i=1}^{j} EX_i'.$$

We assume $M > K$ since the lemma is immediate otherwise. First suppose that $X_1, X_2, \cdots, X_n$ are continuous rv's. Then $M = |m|$. We assume, without loss of generality, that $M = m$. Since $a + S_j = M + T_j + \sum_{i=1}^{j} X_i''$,

$$\begin{aligned} E\Phi(a + S_j) &\ge \Phi(\tfrac{1}{2}M)P(a + S_j \ge \tfrac{1}{2}M) \\ &\ge 2^{-\alpha}\Phi(M)P(\cap_{i=1}^{j} \{X_i'' = 0\}, T_j \ge -\tfrac{1}{2}M) \\ &\ge 2^{-\alpha}\Phi(M)\{P(\cap_{i=1}^{j} \{X_i'' = 0\}) - P(T_j < -\tfrac{1}{2}M)\}. \end{aligned} \tag{6.1}$$

In view of our assumption that $M > K$, it follows from (2.3) that

$$2\{\textstyle\sum_{i=1}^{n} E(X_i'/M)^2 + \sum_{i=1}^{n} E\Phi(X_i'')/\Phi(M)\} < 1.$$

Hence there exist nonnegative Q, Q' with $Q + Q' < 1$ such that

$$2 \textstyle\sum_{i=1}^{j} E(X_i')^2 = QM^2 \tag{6.2}$$

and

$$2 \textstyle\sum_{i=1}^{j} E\Phi(X_i'')^2 = Q'\Phi(M). \tag{6.3}$$

It follows from (6.3) that $\sum_{i=1}^{j} P(X_i'' \ne 0) \le \tfrac{1}{2}Q'$. Therefore by Lemma 1.3,

$$P(\cap_{i=1}^{j} \{X_i'' = 0\}) = \textstyle\prod_{i=1}^{j} P(X_i'' = 0) \ge 1 - \tfrac{1}{2}Q' \ge \tfrac{1}{2}(1 + Q). \tag{6.4}$$

In addition,

$$\begin{aligned} P(T_j < -\tfrac{1}{2}M) &= P(-T_j I(T_j < 0) > \tfrac{1}{2}M) \\ &\le -2ET_j I(T_j < 0)/M && \text{(by Markov)} \\ &= E|T_j|/M && \text{(since } ET_j = 0) \\ &\le (ET_j^2)^{1/2}/M \\ &\le (\textstyle\sum_{i=1}^{j} E(X_i')^2)^{1/2}/M \\ &= (\tfrac{1}{2}Q)^{1/2} && \text{(by (6.2))}. \end{aligned} \tag{6.5}$$

Inequalities (6.1), (6.4) and (6.5) now yield

$$\begin{aligned} E\Phi(a + S_j) &\ge 2^{-\alpha}\Phi(M)\{\tfrac{1}{2}(1 + Q) - (\tfrac{1}{2}Q)^{1/2}\} \\ &\ge 2^{-\alpha-2}\Phi(M) \end{aligned}$$

where we took the infimum over $Q \in [0, 1]$ in the last inequality. This proves the lemma for continuous X's.

For the general case, introduce rv's $U_1, U_2, \cdots, U_n$ uniform on $(-1, 1)$ such that $X_1, U_1, X_2, U_2, \cdots, X_n, U_n$ forms an independent sequence. Let

$$X_i(\epsilon) = X_i + \epsilon U_i, \qquad \vec{X}_i(\epsilon) = (X_1(\epsilon), \cdots, X_i(\epsilon)), \qquad 1 \le i \le n,$$

$$S_n(\epsilon) = \textstyle\sum_{i=1}^{n} X_i(\epsilon), \qquad M(\epsilon) = M_{\vec{X}_j(\epsilon)}(1, a), \qquad K(\epsilon) = K_{\vec{X}_n(\epsilon)}(2).$$

By what has already been proved,

$$\Phi(M(\epsilon)) \le \max\{\Phi(K(\epsilon)), 2^{\alpha+2}E\Phi(a + S_n(\epsilon))\}.$$

It is easily verified that $\lim_{\epsilon\to0}\Phi(M(\epsilon)) = \Phi(M)$ and $\lim_{\epsilon\to0}\Phi(K(\epsilon)) = \Phi(K)$. Furthermore, since $S_n(\epsilon) \to S_n$ a.s. as $\epsilon \to 0$, and $\Phi(a + S_n(\epsilon)) \le 2^{\alpha}\{\Phi(a + S_n) + \Phi(n)\}$ for $|\epsilon| \le 1$, dominated convergence yields $\lim_{\epsilon\to 0} E\Phi(a + S_n(\epsilon)) = E\Phi(a + S_n)$. The obvious limit argument then completes the proof of the lemma. □

As a result of the above lemma and Theorem 3.1 it is clear (and is proved in Theorem 7.1) that whenever $\Phi(K_{\vec{X}_n}(2))$ is too much smaller than $E\Phi(a + S_j)$, an accurate approximation of $E\Phi(a + S_j)$ is obtained by using $\Phi(M_{\vec{X}_j}(1, a))$. We asserted at the beginning of this section (and also in Section 4) that in such cases $M_{\vec{X}_j}(1, a)$ is useful because it in turn approximates the absolute value of the "center" of the distribution of $a + S_j$. We now quantify this assertion.

COROLLARY 6.1. *Let $\Phi \in \mathscr{F}_\alpha$ for some $\alpha > 0$. Let $\vec{X}_j = (X_1, X_2, \cdots, X_j)$ be a vector of independent* rv's *and $S_j = \sum_{i=1}^{j} X_i$ for $1 \le j \le n$. Suppose that for some real number a and $1 \le j \le n$, $M_{\vec{X}_j}(1, a) \ge K_{\vec{X}_n}(2) > 0$. Let $M = M_{\vec{X}_j}(1, a)$ and*

$$\tilde{M} = \begin{cases} M & \text{if } a + \sum_{i=1}^{j} EX_iI(|X_i| \le M) \ge 0 \\ -M & \text{otherwise.} \end{cases}$$

Then

$$P(a + S_j - \tilde{M} \le \tfrac{1}{2}M) \ge \tfrac{1}{4} \tag{6.6}$$

and

$$P(a + S_j - \tilde{M} \ge -\tfrac{1}{2}M) \ge \tfrac{1}{4}. \tag{6.7}$$

PROOF. Suppose first that each X_i has a continuous distribution. Then $\tilde{M} = a + \sum_{i=1}^{j} EX_iI(|X_i| \le M)$. The proof of Lemma 6.1 can be adapted easily to establish both (6.6) and (6.7). For the general case, we let $X_i(\epsilon)$, $M(\epsilon)$, $K(\epsilon)$ be defined as in the proof of the preceding lemma and define $\tilde{M}(\epsilon)$ as indicated above. Note that $M(\epsilon) \to M$ and $K(\epsilon) \to K$ as $\epsilon \to 0$. Since $M \ge K$, it follows that $\sum_{i=1}^{n} P(|X_i| \ge M) \le \tfrac{1}{2}$. Consequently, $\tilde{M}(\epsilon)$ assumes a definite sign for all small ϵ. It therefore converges to $\tilde{M}$ as $\epsilon \to 0$. Now a standard weak convergence argument completes the proof. □

7. The main theorem and some remarks.

THEOREM 7.1. *Let $\Phi \in \mathscr{F}_\alpha$ for some $\alpha > 0$. Let $X_1, X_2, \cdots$ be independent* rv's *and for each $j \ge 1$, write $S_j = \sum_{i=1}^{j} X_i$ and $\vec{Y}_j = (Y_1, Y_2, \cdots, Y_j)$, where $Y_j = X_j -$* med *X_j. For any $y > 0$, let $K_j(y) = K_{\vec{Y}_j}(y)$, with $K_j = K_j(1)$ for short. For any real a, let*

$$M_j(a) = \sup\{m: |a + \textstyle\sum_{i=1}^{j} (\text{med } X_i + EY_iI(|Y_i| \le m))| \ge m\}.$$

Let J be a finite nonempty subset of the positive integers and set

$$M_J(a) = \max_{j\in J} M_j(a), \qquad K_J = \max_{j\in J} K_j.$$

Then, if n is the largest element in J,

$$K_J = K_n$$

and

$$A(\alpha)\Phi(K_J \vee M_J(a)) \le E \max_{j\in J}\Phi(a + S_j) \le B(\alpha)\Phi(K_J \vee M_J(a)) \tag{7.1}$$

where

$$A(\alpha) = 2^{-\alpha-1}G(\alpha) \tag{7.2}$$

and $G(\alpha)$ is defined in (5.11) *and $B(\alpha)$ in Theorem* 3.1.

PROOF. The fact that $K_J = K_n$ and the right-hand side of (7.1) follow immediately from Theorem 3.1 after first centering the X's at their medians and adjusting the constant. It therefore only remains to prove the left-hand side. Two cases are considered.

Case 1. $M_J(a) > K_n(2)$. Let $k \in J$ be such that $M_k(a) = M_J(a)$. It is then immediate from Lemma 6.1 that

$$\begin{aligned} E \max_{j\in J}\Phi(a + S_j) &\geq E\Phi(a + S_k) \geq 2^{-\alpha-2}\Phi(M_J(a)) \\ &> 2^{-\alpha-1}G(\alpha)\Phi(K_n \vee M_J(a)), \end{aligned}$$

where the last inequality follows from the fact that $G(\alpha) < \frac{1}{2}$ and $K_n < K_n(2) < M_J(a)$.

Case 2. $M_J(a) \leq K_n(2)$. Let $(\bar{Y}_1, \bar{Y}_2, \cdots)$ be an independent copy of $(Y_1, Y_2, \cdots)$, and write

$$\bar{X}_i = \bar{Y}_i + \text{med } X_i, \qquad \bar{S}_j = \textstyle\sum_{i=1}^{j} \bar{X}_i, \qquad \vec{Y}_n^* = (Y_1, Y_2, \cdots, Y_n, \bar{Y}_1, \bar{Y}_2, \cdots, \bar{Y}_n).$$

It is easy to see that $K_n(2) = K_{\vec{Y}_n^*}(1)$. Applying Theorem 5.2 to obtain the second inequality below,

$$\begin{aligned} G(\alpha)\Phi(K_n \vee M_J(a)) &\leq G(\alpha)\Phi(K_n(2)) \\ &\leq E\Phi(2a + S_n + \bar{S}_n) \\ &\leq 2^{\alpha}E\{\Phi(a + S_n) + \Phi(a + \bar{S}_n)\} \\ &= 2^{\alpha+1}E\Phi(a + S_n) \\ &\leq 2^{\alpha+1}E \max_{j\in J}\Phi(a + S_j). \end{aligned}$$

COROLLARY 7.1. *Assuming the conditions and notation in Theorem* 7.1, *write* $M_n^*(a) = \max_{1\leq j\leq n} M_j(a)$. *Then*

$$A(\alpha)\Phi(K_n \vee M_n(a)) \leq E\Phi(a + S_n) \leq B(\alpha)\Phi(K_n \vee M_n(a))$$

and

$$A(\alpha)\Phi(K_n \vee M_n^*(a)) \leq E \max_{1\leq j\leq n}\Phi(a + S_j) \leq B(\alpha)\Phi(K_n \vee M_n^*(a)).$$

PROOF. Obvious from Theorem 7.1 by appropriate choice of J. □

REMARK 7.1. If in addition Φ is convex and nondecreasing on $[0, \infty)$ (whence $\Phi(x)/x$ is also nondecreasing on $[0, \infty)$), and the X's have zero means, then we do not have to center at medians. In fact, letting $K_n = K_{\vec{X}_n}(1)$, there exist constants $c_1(\alpha)$, $c_2(\alpha)$ such that

$$c_1(\alpha)\Phi(K_n) \leq E\Phi(S_n) \leq E \max_{1\leq j\leq n}\Phi(S_j) \leq c_2(\alpha)\Phi(K_n). \tag{7.3}$$

The middle inequality is obvious. From Theorem 3.1 and Remark 3.3, it is clear that the right-hand side holds with $c_2(\alpha) = B(\alpha)$. Next we require use of the fact that

$$E\Phi(\max_{1\leq j\leq n} |S_j|) \leq 3(2)^{\alpha}E\Phi(S_n). \tag{7.4}$$

This inequality (which we prove below) shows that to complete the proof of (7.3), it suffices to show that $\Phi(K_n) \leq cE\Phi(\max_{1\leq j\leq n} |S_j|)$ for some universal constant c. This fact can be proved directly by elementary arguments. However, for reasons of economy, we choose to prove it via the left-hand side of the following result of Burkholder (1973, Theorem 15.1):

There exist positive constants $c_3(\alpha)$, $c_4(\alpha)$ such that

$$c_3(\alpha)E\Phi(S(\vec{X}_n)) \leq E \max_{1\leq j\leq n}\Phi(S_j) \leq c_4(\alpha)E\Phi(S(\vec{X}_n)) \tag{7.5}$$

where $S(\vec{X}_n) = (\sum_{i=1}^{n} X_i^2)^{1/2}$.

Two additional facts are needed:

(7.6) $$E\Phi(S(\vec{X}'_n)) \geq c_5(\alpha)\Phi(K_n) \quad \text{if} \quad \sum_{i=1}^n E(X'_i)^2 \geq \tfrac{1}{2}K_n^2$$

(7.7) $$E\Phi(S(\vec{X}''_n)) \geq c_6(\alpha)\Phi(K_n) \quad \text{if} \quad \sum_{i=1}^n E(X'_i)^2 < \tfrac{1}{2}K_n^2$$

where $c_5(\alpha)$, $c_6(\alpha)$ are positive constants and $X'_i = X_i I(|X_i| \leq K_n)$, $X''_i = X_i - X'_i$, $\vec{X}'_n = (X'_1, \cdots, X'_n)$, etc. The proof of (7.6) uses the proof of Lemma 5.2 while that of (7.7) is similar to the first part of the proof of Theorem 5.1. Since trivially

(7.8) $$E\Phi(S(\vec{X}_n)) \geq \max\{E\Phi(S(\vec{X}'_n)), E\Phi(S(\vec{X}''_n))\},$$

the left-hand side of (7.3) follows from inequalities (7.4) through (7.8).

As may be surmised, there is a close connection between this paper and the individual and joint works of Burkholder, Davis and Gundy (see Burkholder (1973) for references) on square function and maximal inequalities and martingale decompositions. In particular, their results can be used to obtain ours when Φ is convex and the X_i's have mean zero.

PROOF OF (7.4). By Ottavianni's inequality,

$$\begin{aligned} P(\max_{1\leq j\leq n} S_j \geq y + 2ES_n^-) &\leq P(S_n \geq y)/\min_{1\leq j\leq n} P(S_n - S_j \geq -2ES_n^-) \\ &\leq P(S_n \geq y)\{1 - P(\max_{1\leq j\leq n}(S_n - S_j)^- > 2ES_n^-)\}^{-1} \\ &\leq 2P(S_n \geq y) \qquad \text{(by Doob's submartingale inequality).} \end{aligned}$$

A similar inequality may be obtained with $\{-S_j\}$ in place of $\{S_j\}$. Hence, observing that $ES_n = 0$ implies $E|S_n| = 2ES_n^-$, we get

$$\mu[y, \infty) \equiv P(\max_{1\leq j\leq n}|S_j| \geq y + E|S_n|) \leq 2P(|S_n| \geq y) \equiv \nu[y, \infty).$$

Using Lemma 1.4,

$$\begin{aligned} E\Phi(\max_{1\leq j\leq n}|S_j|) &\leq \Phi(2E|S_n|) + \int_{E|S_n|}^{\infty} \Phi(y + E|S_n|)\mu\,(dy) \\ &\leq 2^\alpha\Phi(E|S_n|) + \int_{E|S_n|}^{\infty} \Phi(2y)\mu\,(dy) \\ &\leq 2^\alpha\Phi(E|S_n|) + 2^\alpha\int_{E|S_n|}^{\infty} \Phi(y)\nu\,(dy) \\ &\leq 2^\alpha E\Phi(|S_n|) + 2^{\alpha+1}E\Phi(|S_n|) \\ &\qquad \text{(by convexity and definition of } \nu) \\ &= 3(2)^\alpha E\Phi(S_n). \end{aligned}$$

REMARK 7.2. The same bounds we derived for $E\max_{j\in J}\Phi(a + S_j)$ based on $\Phi(K_n \vee M_J(a))$ in Theorem 7.1 can be obtained using slightly different truncation points. In the notation of that theorem, let

$$t_{\vec{Y}_n}(y) = \sup\{t: y\sum_{i=1}^n E\Phi(Y_i)I(|Y_i| > t) \geq \Phi(t)\}$$

and

$$v_{\vec{Y}_n}(y) = \sup\{v: y\sum_{i=1}^n EY_i^2 I(|Y_i| \leq v) \geq v^2\}.$$

For simplicity, write $t_n = t_{\vec{Y}_n}(1)$ and $v_n = v_{\vec{Y}_n}(1)$. Then

THEOREM 7.2.

$$A(\alpha)\Phi(v_n \vee t_n \vee M_J(a)) \le E \max_{j\in J}\Phi(a + S_j) \le B(\alpha)\Phi(v_n \vee t_n \vee M_J(a)).$$

PROOF. The left-hand side follows from the observation that $v_n \vee t_n \le K_n$. A perusal of Section 3 shows that

$$E \max_{j\in J}\Phi(a + S_j) \le B(\alpha)\Phi(b)$$

for any $b \ge v_n \vee t_n \vee M_J(a)$. Hence the right-hand side holds also.

We think of $M_J(a)$ as approximating the absolute value of an appropriate quantile of $a + S_{j^*}$ for $j^* \in J$ with $M_{j^*}(a) = M_J(a)$. Similarly v_n approximates the range of the distribution of $a + S_n$. Finally, $\Phi(t_n)$ always approximates $E \max_{1\le j\le n}\Phi(Y_j)$. In this manner, $\Phi(t_n \vee v_n \vee M_J(a))$ gives weight to each of the three aspects of the distribution of $W = \max_{j\in J}\Phi(a + S_j)$ identified in Section 4 as governing the order of magnitude of EW. Although we have provided but an heuristic discussion, with some effort these notions can be made rigorous.

REMARK 7.3. In practice it may often be possible to obtain one-sided bounds of $Ef(S_n)$ by other methods. For example, suppose $f(x) = xg(x)$[6] and $X_1, X_2, \cdots, X_n$ are i.i.d. nonnegative rv's with mean μ. Then, assuming all expectations are well-defined and finite,

$$\begin{aligned} Ef(S_n) &= ES_ng(S_n) = \textstyle\sum_{i=1}^n EX_ig(S_n) = nEX_ng(S_n) \\ &= nEX_nE(g(S_{n-1} + X_n)\,|\,X_n) \\ &\le nEX_ng((n-1)\mu + X_n) \quad \text{if } g \text{ is concave} \\ &\ge nEX_ng((n-1)\mu + X_n) \quad \text{if } g \text{ is convex.} \end{aligned}$$

Such a bound satisfies one of our criteria for acceptability: It may be calculated directly from the one-dimensional X-distribution (by computer if necessary). One integration is required to compute μ and a second to construct the bound. In many cases such a bound is extremely close to the true value of $Ef(S_n)$. This seems to occur whenever the L_1-interaction between X_n and $g(S_{n-1} + X_n)$ brings out the dominance of the maximal X_i, placing X_n in that role.

NOTE ADDED IN PROOF. The auther has recently been informed by J. Kuelbs of related and pioneering work by J. Bretangnolle and D. Dacunha-Castelle (see Application de l'étude certaines formes linéaires aléatories au plongement d'espaces de Banach dans des espaces L^p, *Ann. Ecole Normale Supérieure*, **2**(1969), 437–480). Let $X_1, X_2, \ldots,$ be i.i.d. random variables, $S_n = X_1 + \ldots + X_n$, and $\vec{X}_n = (X_1, \ldots, X_n)$. Using the notation of this paper, they show that if X_1 is symmetric and has finite mean, there exist constants $0 < c < c^* < \infty$ (possibly depending on X_1) such that, $c \le E \mid S_n \mid /K_{\vec{X}_n}(1) \le c^*$ for all n. Thus their paper is perhaps the first to approximate $E \mid S_n \mid$ in terms of $K_{\vec{X}_n}(1)$. However, explicit results on the precision and generality of the approximation do not seem to exist in print prior to Klass (1980).

[6] An example of this kind cropped up while the author was working at the Jet Propulsion Laboratory in 1973. To determine the capacity of a communication channel, it was necessary to approximate a quantity which could be expressed in the form $E(S_n/n)\log(S_n/n)$ where the X's were modified Bessel functions of the first kind of $N(O, \sigma^2)$ r.v.'s and $n = \exp \sigma^2/2$. Using this method of approximation, numerical integration indicated a relative error not exceeding 0.05. For further details, consult Butman-Klass (1973).

Acknowledgment. I want to express my gratitude to Ian Abramson who provided valuable comments while this work was in progress, and to Wei Yin Loh whose editorial assistance, mathematical expertise and tireless effort helped produce this finished product.

REFERENCES

VON BAHR, B. and ESSEEN, C. G. (1965). Inequalities for the rth absolute moment of a sum of rv's $1 \le r \le 2$.

BRILLINGER, D. R. (1962). A note on the rate of convergence of the mean. *Biometrika* **49** 574–576.

BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probability* **1** 19–42.

BUTMAN and KLASS, M. J. (1973). Capacity of noncoherent channels. JPL Technical Report 32-1526. **18** 85–93.

CHUNG, K. L. (1974). *A Course in Probability Theory*. 2nd Ed. Academic, New York.

DAVIS, B. (1973). Moments of random walk having infinite variance and the existence of certain optimal stopping rules for S_n/n. *Illinois J. Math.* **17** 75–81.

KLASS, M. J. (1975). A survey of the S_n/n problem. *Proc. Summer Res. Inst. Statist. Inference for Stochastic Processes*, Indiana University **2** 179–201. Academic, New York.

KLASS, M. J. (1976). Toward a universal law of the iterated logarithm, Part I. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **36** 165–178.

KLASS, M. J. (1979). Precision bounds for the relative error in the approximation of $E|S_n|$ and extensions. *Ann. Probability* **8** 350–367.

KLASS, M. J. Approximating expectations of functions of normed sums of independent Banach space random elements. Unpublished manuscript.

ROSENTHAL, H. P. (1970). On the sequences of $L^p(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8**(3) 273–303.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720