PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 108, Number 1, January 1990

BOUNDS ON THE EXPECTATION OF FUNCTIONS OF MARTINGALES AND SUMS OF POSITIVE RV'S IN TERMS OF NORMS OF SUMS OF INDEPENDENT RANDOM VARIABLES

VICTOR H. DE LA PEÑA

(Communicated by William D. Sudderth)

ABSTRACT. Let (x_i) be a sequence of random variables. Let (w_i) be a sequence of independent random variables such that for each i, w_i has the same distribution as x_i. If $S_n = x_1 + x_2 + \cdots + x_n$ is a martingale and Ψ is a convex increasing function such that $\Psi(\sqrt{x})$ is concave on $[0, \infty)$ and $\Psi(0) = 0$ then,

$$E\Psi\left(\max_{j\le n}\left|\sum_{i=1}^{j} x_i\right|\right) \le CE\Psi\left(\left|\sum_{i=1}^{n} w_i\right|\right)$$

for a universal constant C, $(0 < C < \infty)$ independent of Ψ, n, and (x_i).

The same inequality holds if (x_i) is a sequence of nonnegative random variables and Ψ is now any nondecreasing concave function on $[0, \infty)$ with $\Psi(0) = 0$. Interestingly, if $\Psi(\sqrt{x})$ is convex and Ψ grows at most polynomially fast, the above inequality reverses.

By comparing martingales to sums of independent random variables, this paper presents a one-sided approximation to the order of magnitude of expectations of functions of martingales. This approximation is best possible among all approximations depending only on the one-dimensional distribution of the martingale differences.

1. INTRODUCTION

Several authors have worked on the problem of obtaining the order of magnitude of expectations of functions of discrete time martingales. Among others, Brillinger [1], von Bahr and Essen [8], Dharmadhikari and Sreehari [3] have obtained upper bounds depending only on the one-dimensional distributions of the martingale differences. Klass [7] obtains exact bounds in the case of sums

Received by the editors August 2, 1988 and, in revised form, March 20, 1989. The main results in this paper where first presented at the Conference in honor of Paul Levy in Paris the summer of 1987.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 60G42, 60G50, 60E15.

Key words and phrases. Martingales, sums of independent random variables.

This research was partially supported by grants from CONACyT-MEXICO and NSF.

The results of this paper are partially contained in the author's Ph.D thesis (1988) at University of California at Berkeley, under Dr. M. J. Klass.

of independent random variables. By comparing martingales to sums of independent random variables, this paper presents a one-sided approximation to the order of magnitude of expectations of functions of martingales. The method used provides the best possible approximation among all methods depending only on the one-dimensional distributions of the martingale differences. Similar results are presented dealing with sums of nonnegative random variables. The tools used in the proofs, mainly developed by Klass [7], are shown to be quite powerful when dealing with expectations in general.

Let $S_n = x_1 + x_2 + \cdots + x_n$ be a real-valued martingale, $R_n = w_1 + w_2 + \cdots + w_n$ be a sum of independent random variables, such that w_i has the same distribution as x_i, and $\Pi = [\Phi\colon \Phi(0) = 0, \Phi(x) = \Phi(-x),\ \Phi(x)$ is convex and increasing on $[0,\infty)]$. Let $\Pi_\alpha = [\Phi$ in $\Pi\colon \Phi(cx) \le c^\alpha \Phi(x)$ for all $x > 0$, and all $c \ge 2]$. For Φ in Π, and $\Phi(\sqrt{x})$ concave on $[0,\infty)$, we show that

$$E\Phi\left(\max_{j\le n}|S_j|\right) \le C_1 E\Phi(R_n), \qquad (\text{ Theorem 1 })$$

where $0 > C_1 < \infty$ is independent of Φ, n, X_n.

In Theorem 2, assuming that Φ is in Π_α with $\Phi(\sqrt{x})$ a convex function on $[0,\infty)$, we prove that,

$$E\Phi\left(\max_{j\le n}|S_j|\right) \ge C_{2,\alpha} E\Phi\left(\max_{j\le n}|R_j|\right),$$

where $0 < C_{2,\alpha} < \infty$ is a constant depending only on α. In Propositions 1 and 2, we obtain similar results for sums of positive random variables. In Example 1 it is shown that the inequalities for martingales cannot always be reversed once Φ is fixed.

2. Preliminaries

Let Y_n be a vector of random variables. $Y_n = (y_1, y_2, \ldots, y_n)$. For Φ in Π_α let $K_\Phi(Y_n)$ be equal to $\sum_{i=1}^n E\Phi(y_i)$ if this sum is 0 or ∞, otherwise let $K_\Phi(Y_n)$ be the unique positive real number such that

$$(2.1) \qquad \sum_{i=1}^n E(y_i)^2 1(|y_i| \le K) + (K^2/\Phi(K)) \sum_{i=1}^n E\Phi(y_i) 1(|y_i| > K) = K^2.$$

Then Klass [7, equation 7.3], shows that if the y_i are independent mean zero r.v.'s, then there exists $0 < A_\alpha, B_\alpha < \infty$, constants depending on α only, such that

$$(2.2) \qquad A_\alpha \Phi(K_\Phi(Y_n)) \le E\Phi(y_1 + y_2 + \cdots + y_n) \le B_\alpha E\Phi(K_\Phi(Y_n)))$$

3. Proof

We will first consider the case S_n is mean zero. Let $X_n = (x_1, x_2, \ldots, x_n)$, $W_n = (w_1, w_2, \ldots, w_n)$. From now on, a generic constant C will be used, this constant may change from application to application but depends on α only.

Denote by $K_\Phi(X_n)$ the K-functions associated with X_n and W_n. Then obviously $K_\Phi(X_n) = K_\Phi(W_n)$ since X_n and W_n have the same marginal distributions. (See equation (2.1).)

Define $x_i^{|} = x_i 1(|x_i| \le K_\Phi(X_n)), x_i^{||} = x_i - x_i^{|}$

Following Klass [1976, Theorem 2.1] let λ_n be such that

$$\lambda_n K_\Phi(X_n)^2 = \left(\sum_{i=1}^n E(x_i)^2 1(|x_i| \le K_\Phi(X_n))\right)$$

$$(1-\lambda_n)\Phi(K_\Phi(X_n)) = \sum_{i=1}^n E\Phi(x_i)1(|x_i| > K_\Phi(X_n)).$$

We present only the proof of Theorem 2 as the proof of Theorem 1 is essentially the same. Since $\Phi(\sqrt{x})$ is a convex function for $x > 0$, then by the Burkholder–Davis–Gundy [2] inequality,

$$\begin{aligned}E\Phi\left(\max_{j\le n}|S_j|\right) &\ge CE\Phi\left(\left(\sum_{i=1}^n x_i^2\right)^{1/2}\right)\\ &\ge C\left(E\Phi\left(\left(\sum_{i=1}^n \left(x_i^{|}\right)^2\right)^{1/2}\right) + E\Phi\left(\left(\sum_{i=1}^n \left(x_i^{||}\right)^2\right)^{1/2}\right)\right)\\ &\ge C\left(\Phi\left(\left(\sum_{i=1}^n E\left(x_i^{|}\right)^2\right)^{1/2}\right) + \sum_{i=1}^n E\Phi\left(x_i^{||}\right)\right) =\end{aligned}$$

by convexity of $\Phi(\sqrt{x})$ and Jensen's inequality

$$C(\Phi(\lambda_n^{1/2}K_\Phi(X_n)) + (1-\lambda_n)\Phi(K_\Phi(X_n))) \ge (\min \tfrac{3}{4}, \tfrac{1}{2^\alpha})C)\Phi(K_\Phi(X_n)) =$$

by minimizing over λ_n and using Φ as in Π_α

$$C\Phi(K_\Phi(W_n)) \ge CE\Phi(|w_1 + w_2 + \cdots + w_n|) \ge CE\Phi(\max_{j\le n}|R_j|)$$

by R.H.S. of (2.2) and Klass [7, equation 7.4].

Note. The minimization can be easily carried by considering the cases $0 \le \lambda_n < \frac{1}{4}$ and $\frac{1}{4} \le \lambda_n < 1$. The proof of the upper bound is exactly the same, first assuming S_n is mean zero, and replacing $\ge$'s for $\le$'s. In line 4 we upper bound by $2C\Phi(K_\Phi(X_n))$ and proceed by using the L.H.S of (2.2). Moreover, since $\Phi(\sqrt{x})$ is now concave, $\Phi(cx) \le c^3\Phi(x)$ for all $c > 2$ and all $x > 0$ hence Φ is automatically on Π_3. The case when S_n is not mean zero can be handled easily by centering at the mean of S_n, and then using the convexity of Φ along with Jensen's inequality.

That the method is optimal for obtaining bounds on the expectation of functions of martingales follows once it is noticed that the K function introduced by Klass [7] provides the exact order of magnitude for the case of sums of independent random variables and that the one-dimensional distributions of the martingale differences are not affected in the process of obtaining the bounds.

The following results are close to equivalent to the ones shown above and deal with the problem of estimating expectations of functions of sums of non-negative random variables.

Proposition 1. *Let* Ψ *be a nondecreasing concave function on* $[0,\infty)$ *such that* $\Psi(0)=0$ *and* $\Psi(x)>0$ *if* $x>0$.

Let $x_1, x_2, \dots$ *be any sequence of nonnegative random variables and let* $w_1, w_2, \dots$ *be independent random variables such that* w_i *has the same distribution as* x_i. *There is a constant* $C>0$, *not depending on anything, such that*

$$E\Psi\left(\sum x_i\right) \le CE\Psi\left(\sum w_i\right).$$

Proposition 2. *Let* Γ *be in* Π_α, x_i *and* w_i *be nonnegative and as above. Then there is a constant* $C_\alpha > 0$ *depending only on* α *such that*

$$E\Gamma\left(\sum x_i\right) \ge C_\alpha E\Gamma\left(\sum w_i\right).$$

To see that these propositions are essentially equivalent to the theorems, we note that if Δ_i, $i\ge 0$, is any sequence of nonnegative random variables such that $\Delta_i^{1/2}$ is integrable then there is a mean 0 martingale with difference sequence $d_1, d_2, \dots$ such that $d_i^2, i\ge 1$, has the same distribution as Δ_i, $i \ge 1$, and use the Burkholder–Davis–Gundy [2] inequalities, setting Ψ and Γ above equal to $\Phi(\sqrt{x})$. To obtain the proof of Proposition 2 in this way, notice that if $\Gamma(x)$ is in Π_α then $\Gamma(x^2)$ is in $\Pi_{2\alpha}$.

We now proceed to prove Proposition 1 from first principles, as the proof suggested earlier restricts the class of functions one can work with. I am indebted to B. Davis for this suggestion.

Proof. Let $A=\inf\{a>0\colon \sum P(x_i>a)\le 1/2\}$. Assume $A<\infty$, the other case being easy. Let $(r \min s)$ denote the minimum between r and s. Put $\varepsilon_i = E(x_i \min A)$, $\beta_i = E\Psi(x_i)1(x_i>A)$, $\varepsilon=\sum\varepsilon_i$, $\beta=\sum\beta_i$. Then,

$$\text{(i)}\quad E\Psi\left(\sum x_i\right)\le \Psi(\varepsilon)+\beta$$

$$\text{(ii)}\quad E\Psi\left(\sum w_i\right)\ge C(\Psi(\varepsilon)+\beta).$$

The proof of (i) is immediate from Jensen's inequality and the facts that $\Psi(x+y)\le\Psi(x)+\Psi(y)$ and $x_i\le(x_i \min A)+x_i 1(x_i>A)$ as follows:

$$E\Psi\left(\sum x_i\right)\le E\left(\Psi\left(\sum(x_i \min A)\right)+\sum\Psi(x_i)1(x_i>a)\right)$$
$$\Psi\left(E\sum(x_i \min A)\right)+\beta=\Psi(\varepsilon)+\beta.$$

To prove (ii), note that if $h_i=(w_i \min A)$, and $p_i=P(w_i>A)$, then we have

$$P\left(\sum h_i>\varepsilon/10\right)>1/10.$$

To see this note that the h_i are independent, satisfy $0 \le h_i \le A$, and $\varepsilon \ge A/2$. Put $\tau = \inf[k\colon \sum_{i=1}^{k} h_i > \varepsilon/10]$, and $\tau = \infty$ if no such k exists. If $P(\tau = \infty) \ge \frac{9}{10}$, then

$$E\left(\sum_{i=1}^{\tau} h_i\right) = \sum_{i=1}^{\infty} Eh_i I(\tau \ge i) = \sum_{i=1}^{\infty} Eh_i EI(\tau \ge i) \quad \text{(by independence)}$$

$$\ge P(\tau = \infty) \sum_{i=1}^{\infty} Eh_i \ge \left(\frac{9}{10}\right)\varepsilon, \qquad \text{while}$$

$$\begin{aligned}\sum_{i=1}^{\tau} h_i &\le (\varepsilon/10)1(\tau = \infty) + ((\varepsilon/10) + A)1(\tau < \infty)\\ &\le (\varepsilon/10)(1 - 1(\tau < \infty) + (21\varepsilon/10)1(\tau < \infty)\\ &= (\varepsilon/10) + (2\varepsilon)1(\tau < \infty).\end{aligned}$$

Taking expectations one comes to a contradiction. Thus $E\Psi(\sum h_i) \ge c\Psi(\varepsilon)$.

To complete the proof of (ii), we show $E\Psi(\sum w_i) \ge \beta/2$. We obtain this result by using the following lemma.

Lemma 1. *For all sequences $\{f_n\}$ of nonnegative independent random variables, all increasing functions $\Phi\colon R_+ \to R_+$ and all positive γ 's, one has*

$$E\Phi\left(\sup_n f_n\right) \ge \sum_n E\Phi(f_n 1(f_n > \gamma))P\left(\sup_i f_i \le \gamma\right).$$

Proof. Let $\tau' = \{\inf_n\colon f_n > \gamma\}$, then

$$\begin{aligned}E\Phi\left(\sup_n f_n\right) &\ge E\Phi\left(\sup_{n\le\tau'} f_n\right) = \sum_n E\Phi\left(\max_{i\le n} f_i\right) 1(\tau' = n)\\ &= \sum_n E\Phi(f_n 1(f_n > \gamma))1\left(\max_{i\le n-1} f_i \le \gamma\right)\\ &\ge \sum_n E\Phi(f_n 1(f_n > \gamma))P\left(\sup_i f_i \le \gamma\right).\end{aligned}$$

The above computations follow from the independence of $\{f_i\}$. Applying Lemma 1 to the sequence $f_n = w_n 1(w_n > A)$, and letting $\Phi = \Psi$ and $\gamma = A$ we obtain

$$\begin{aligned}E\Psi\left(\sum_i w_i\right) &\ge E\Psi\left(\sum_i w_i 1(w_i > A)\right) \ge E\Psi\left(\sup_n w_n 1(w_n > A)\right)\\ &\ge \sum_n E\Psi(w_n 1(w_n > A)P\left(\sup_i w_i 1(w_i > A) \le A\right).\end{aligned}$$

The last probability is equal to $1 - P(\sup_i w_i 1(w_i > A) > A) \ge 1 - \sum_i P(w_i > A) \ge \frac{1}{2}$. This completes the proof of (ii). (i) and (ii) put together complete the proof of Proposition 1.

Before we continue some notation is needed: Let U_n, V_n be two sequences of r.v.'s, denote by $EU_n \approx EV_n$ if the ratio of adjacent quantities is bounded away from zero and infinity by positive finite constants independent of n and of the distributions of the r.v.'s.

4. Example

The following example due to Professor M. J. Klass, shows that the reverse inequalities to the ones in Theorems 1 and 2 cannot hold in general.

Let (x_i) be an i.i.d. sequence of $N(0,1)$ random variables. Let $(\varepsilon_i, i = 0, \dots, n)$ be an i.i.d. sequence of random variables such that for fixed p, $0 < p < 1$, $\varepsilon_0 = 1$ with probability p and $\varepsilon_0 = 0$ otherwise. We also assume (ε_i) to be independent of (x_i).

Then $S_n = \varepsilon_0 x_1 + \varepsilon_0 x_2 + \cdots + \varepsilon_0 x_n$ is a martingale with respect to $\sigma(\varepsilon_0 x_1, \dots, \varepsilon_0 x_n)$, with martingale differences $d_i = \varepsilon_0 x_i$. Furthermore, $R_n = \varepsilon_1 x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_n x_n$ is a sum of i.i.d. random variables and $\varepsilon_i x_i$ has the same distribution as $d_i = \varepsilon_0 x_i$.

Pick n, p such that np is an integer. It is easy to check that the K-function for $E\Phi(\sum_{i=1}^n \varepsilon_i x_i)$ is the same as the one for $E\Phi(\sum_{i=1}^{np} x_i)$. Therefore, by using (2.3) one gets

$$E\Phi\left(\max_{j\le n}\left|\sum_{i=1}^{j} \varepsilon_i x_i\right|\right) \approx E\Phi\left(\sum_{i=1}^{n} \varepsilon_i x_i\right) \approx E\Phi\left(\sum_{i=1}^{np} x_i\right).$$

Recalling that the x_i's are i.i.d. $N(0,1)$, one sees that

$$E\Phi\left(\sum_{i=1}^{np} x_i\right) = E\Phi(x_1\sqrt{np}) \approx \Phi(\sqrt{np}).$$

We also have

$$E\Phi\left(\max_{j\le n}\left|\sum_{i=1}^{j} \varepsilon_0 x_i\right|\right) \approx E\Phi\left(sum_{i=1}^{n} \varepsilon_0 x_i\right) = pE\Phi\left(\sum_{i=1}^{n} x_i\right)$$
$$= pE\Phi(x_1\sqrt{n}) \approx p\Phi(\sqrt{n}).$$

Now assume $\Phi(x)/x^2 \to \infty$ or 0. Then, letting $p = 1/n$, we have $\Phi(1)$ versus $\Phi(\sqrt{n})/n$ and these are of different orders as $n \to \infty$.

For results closely related to the ones presented here see Hitczenko [4], and Johnson and Schechtman [5].

5. Acknowledgments

Many thanks to Professor M. J. Klass for several important discussions while preparing this paper. I also wish to thank the referee and Burgess Davis for their comments and suggestions.

Bibliography

1. D. R. Brillinger, *A note on the rate of convergence of a mean*, Biometrika **49** (3,4) (1962), 574.
2. D. L. Burkholder, B. J. Davis, and R. F. Gundy, *Inequalities for convex functions of operators on martingales*, Proc. Sixth Berkeley Sympos. Math. Stat. Prob. **2**, 1972, pp. 223–240.
3. S. W. Dharmadhikari and M. Sreehari, *On convergence in r-mean of normalized partial sums*, Ann. Probability **3** (6), (1975) 1023–1024.
4. P. Hitczenko, *Comparison of mements for tangent sequences of random variables*, Probab. Theory Related Fields **78** (1988), 223–230.
5. W. B. Johnson and G. Schechtman, *Sums at independent random variables in rearrangement invariant function spaces*, Ann. Probability **17** (2) (1989), 789–807.
6. M. J. Klass, *Toward a universal law of the iterated logarithm, Part* I, Z. Wahrscheinlichkeiths-theorie und verw. Gebiete **36** (1976), 165–178.
7. ——, *A method of approximating expectations of functions of sums of independent random variables*, Ann. Probability **9** (3) (1981), 413–428.
8. B. von Bahr and C. G. Esseen, *Inequalities for the rth absolute moment of a sum of random variables,* $1 \leq r \leq 2$, Ann. Math. Statist., **36** (1965), 299–303.

Department of Statistics, University of California at Berkeley, Berkeley, California 94720

Current address: Department of Statistics, 618 Mathematics Building, Columbia University, New York, New York 10027