The Annals of Mathematical Statistics
1968, Vol. 39, No. 5, 1719–1723

BOUNDS ON THE MOMENTS OF MARTINGALES

By S. W. Dharmadhikari,[1] V. Fabian[2] and K. Jogdeo[3]

Michigan State University and Courant Institute of Mathematical Sciences

1. Summary and related work. We prove the following

Theorem. *Let* $\{S_n, n \geq 1\}$ *be a martingale,* $S_0 = 0$, $X_n = S_n - S_{n-1}$, $\gamma_{\nu n} = E(|X_n|^\nu)$ *and* $\beta_{\nu n} = (1/n)\sum_{j=1}^n \gamma_{\nu j}$. *Then for all* $\nu \geq 2$ *and* $n = 1, 2, \cdots$

$$E(|S_n|^\nu) \leq C_\nu n^{\nu/2}\beta_{\nu n}, \tag{1.1}$$

where

$$C_\nu = [8(\nu - 1)\max(1, 2^{\nu-3})]^\nu. \tag{1.2}$$

As shown by Chung ([3], pp. 348–349) an inequality of Marcinkiewicz and Zygmund ([5], p. 87) implies that the theorem holds (possibly with a different value of C_ν) whenever the X's are independent. In the same way the above theorem is implied by the generalization of the Marcinkiewicz-Zygmund result given by Burkholder ([2], Theorem 9). However, our proof is elementary. Although our choice of C_ν is not the best possible, it is explicit. For the case of independent X's, von Bahr ([6], p. 817) has given a bound for $E(|S_n|^\nu)$ which may sometimes involve powers of $\beta_{\nu n}$ higher than 1. Finally Doob ([4], Chapter V, Section 7) has treated the case when the X's form a Markov chain.

After proving some lemmata in Section 2, we give the proof of the theorem in Section 3. The case of exchangeable random variables is dealt with in Section 4.

2. Two lemmata. We use the following two lemmata in the proof of the theorem.

Lemma 1. $0 < \nu \leq \mu \Rightarrow \beta_{\nu n}^{1/\nu} \leq \beta_{\mu n}^{1/\mu}$.

Proof. Observe that $\gamma_{\nu j}^{\mu/\nu} \leq \gamma_{\mu j}$. Therefore

$$\beta_{\nu n}^{\mu/\nu} = (n^{-1}\sum_{j=1}^n \gamma_{\nu j})^{\mu/\nu} \leq n^{-1}\sum_{j=1}^n \gamma_{\nu j}^{\mu/\nu} \leq n^{-1}\sum_{j=1}^n \gamma_{\mu j} = \beta_{\mu n}.$$

Lemma 2. *Let* $\{y_n, n \geq 1\}$ *be a sequence of non-negative numbers. Let* $z_n = (y_1 + \cdots + y_n)/n$. *Then for all* $x \geq 1$ *and for* $n = 1, 2, \cdots$

$$x^{-1}(n-1)^x z_{n-1} + (n-1)^{x-1} z_{n-1}^{(x-1)/x} y_n^{1/x} \leq x^{-1} n^x z_n. \tag{2.1}$$

Proof. If $x = 1$, the lemma holds for all n. So let $x > 1$. If $y_n = 0$ or $n = 1$ the lemma is easy to verify. So we may assume that $y_n > 0$ and $n > 1$. If, in (2.1), we replace y_i by cy_i, $c > 0$, we get an equivalent inequality. We may there-

Received 24 January 1968.

[1] On leave from the Indian Statistical Institute.

[2] On leave from the Czechoslovak Academy of Sciences.

[3] Research supported at the Courant Institute of Mathematical Sciences under a Ford Foundation grant for probability and statistics. Now at the Department of Mathematics, University of Illinois.

fore assume that $y_n = 1$. Writing $A = z_{n-1}$, we see that (2.1) is equivalent to

$$(2.2) \qquad (n-1)^x A + x(n-1)^{x-1}A^{(x-1)/x} \leqq n^{x-1}[(n-1)A + 1].$$

Write $\theta = (n-1)^{-1}$. Then (2.2) is the same as

$$(2.3) \qquad A + \theta x A^{(x-1)/x} \leqq A(1+\theta)^{x-1} + \theta(1+\theta)^{x-1}.$$

If $A = 0$, the inequality (2.3) is immediate. So let $A > 0$. Then (2.3) is equivalent to

$$(2.4) \qquad f(x) \leqq 0 \quad \text{for} \quad x \geqq 1,$$

where

$$(2.5) \qquad f(x) = (1 + \theta x A^{-1/x})(1+\theta)^{-(x-1)} - (1 + \theta A^{-1})$$

Clearly $f(1) = 0$. So to prove (2.4), it suffices to show that $f'(x) \leqq 0$ for $x \geqq 1$. We have

$$(2.6) \qquad \begin{aligned} f'(x) &= (1+\theta)^{-(x-1)}[\theta(A^{-1/x} + x^{-1}A^{-1/x}\log A) \\ &\qquad - (1 + x\theta A^{-1/x})\log(1+\theta)] \\ &= (1+\theta)^{-(x-1)}A^{-1/x}g(x, B), \end{aligned}$$

where $B = A^{1/x}$ and

$$(2.7) \qquad g(x, B) = \theta(1 + \log B) - (B + x\theta)\log(1+\theta).$$

We note that $f'(x) \leqq 0 \Leftrightarrow g(x, B) \leqq 0$. Now it is easy to verify that $g(x, B)$, as a function of B, reaches its maximum at $B = B_0$, where

$$B_0 \log(1+\theta) = \theta.$$

But

$$g(x, B_0) = \theta[\log B_0 - x\log(1+\theta)],$$

which is negative for $x \geqq 1$, because of the elementary inequality $[\theta/(1+\theta)] \leqq \log(1+\theta)$ for $\theta > 0$. This completes the proof of the lemma.

COROLLARY. *Use the notation of Lemma 2. Then, for all $x \geqq 1$ and for $n = 1, 2, \cdots$,*

$$(2.8) \qquad \textstyle\sum_{j=1}^{n} (j-1)^{x-1} z_{j-1}^{(x-1)/x} y_j^{1/x} \leqq n^x x^{-1} z_n.$$

PROOF. Let T_n denote the left side of (2.8). The corollary holds trivially for $n = 1$. Suppose it holds for $n = m$. Then

$$\begin{aligned} T_m &= T_{m-1} + (T_m - T_{m-1}) \\ &\leqq (m-1)^x x^{-1} z_{m-1} + (m-1)^{x-1} z_{m-1}^{(x-1)/x} y_m^{1/x} \\ &\leqq m^x x^{-1} z_m, \qquad \text{by Lemma 2.} \end{aligned}$$

This proves the corollary.

3. Proof of the theorem. Suppose $\{S_n\}$ is a given martingale. The inequality (1.1) clearly holds if $\nu = 2$ or if $\beta_{\nu n} = \infty$. Suppose $\mu > 2$ and $\beta_{\mu n} < \infty$. To prove the theorem it is enough to prove (1.1) for all $\nu \in [2, \mu]$. Suppose we are able to prove

$$(*) \qquad \text{(1.1) holds for all} \quad \nu \in [2, \mu).$$

Then, since $|S_n|^\nu$ is bounded by the integrable function $1 + |S_n|^\mu$ for all $\nu \in [2, \mu)$, we have

$$E\,|S_n|^\mu = \lim_{\nu \uparrow \mu} E|S_n|^\nu \leqq \lim_{\nu \uparrow \mu} C_\nu n^{\nu/2}\beta_{\nu n} = C_\mu n^{\mu/2}\beta_{\mu n}.$$

Thus the theorem will be proved as soon as $(*)$ is proved. This is accomplished by induction as follows. Assuming that (1.1) holds for all $\nu \in [2, \nu_0]$, with $\nu_0 \geqq 2$, we show that (1.1) holds for all $\nu \in (\nu_0, \nu_0 + \epsilon]$ where ϵ is a suitable positive number. This, and the fact that (1.1) holds for $\nu = 2$, will imply $(*)$.

By Taylor expansion, for all $\nu \in [2, \mu]$,

$$(3.1) \quad |S_n|^\nu = |S_{n-1}|^\nu + \nu \operatorname{sgn}(S_{n-1})|S_{n-1}|^{\nu-1} X_n + \tfrac{1}{2}\nu(\nu - 1)|S_{n-1} + \theta X_n|^{\nu-2} X_n^2,$$

where $0 < \theta < 1$. We note that

$$(3.2) \qquad |S_{n-1} + \theta X_n|^{\nu-2} \leqq \max(1, 2^{\nu-3})[|S_{n-1}|^{\nu-2} + |X_n|^{\nu-2}],$$

and that the expectation of the middle term on the right side of (3.1) is zero. Setting

$$(3.3) \qquad \delta_\nu = (\nu - 1)\max(1, 2^{\nu-3})$$

and

$$(3.4) \qquad \Delta_n(\nu) = E(|S_n|^\nu - |S_{n-1}|^\nu),$$

we get from (3.1) and (3.2)

$$(3.5) \qquad \Delta_n(\nu) \leqq 2^{-1}\nu\delta_\nu[E(|S_{n-1}|^{\nu-2}X_n^2) + \gamma_{\nu n}].$$

Assume that $\nu = \nu_1 < \nu_2 \leqq \mu$. The Hölder inequality gives

$$(3.6) \qquad E(|S_{n-1}|^{\nu_1-2}X_n^2) \leqq [E\,|S_{n-1}|^{\nu_0}]^{(\nu_2-2)/\nu_2}[E\,|X_n|^{\nu_2}]^{2/\nu_2},$$

where

$$(3.7) \qquad \nu_0 = [(\nu_1 - 2)\nu_2]/(\nu_2 - 2) < \nu_1.$$

Let $\nu_0 \in [2, \mu)$ be fixed. For any $\nu_1 \in (\nu_0, \nu_0 + \epsilon]$, we can choose ν_2 using (3.7). If ϵ is positive and sufficiently small, then again $\nu_0 < \nu_1 < \nu_2 \leqq \mu$ and (3.6) holds. Adjusting ϵ if necessary we can obtain ν_2 so close to ν_0 that, for all $\nu_1 \in (\nu_0, \nu_0 + \epsilon]$,

$$(3.8) \qquad \gamma_{\nu_2 n}^{1/\nu_2} \leqq 2\gamma_{\nu_1 n}^{1/\nu_1}$$

and

$$(3.9) \qquad \nu_2 \leqq 2\nu_1.$$

Suppose now that (1.1) holds for all $\nu \in [2, \nu_0]$. Let $\nu_1 \in (\nu_0, \nu_0 + \epsilon]$ with ϵ positive and determined as above. We shall show that (1.1) holds for $\nu = \nu_1$.

Lemma 1 shows that $\beta_{\nu_0 n} \leqq \beta_{\nu_1 n}^{\nu_0/\nu_1}$. Therefore

$$\beta_{\nu_0 n}^{(\nu_2-2)/\nu_2} \leqq (\beta_{\nu_1 n})^{\nu_0(\nu_2-2)/(\nu_1\nu_2)} = \beta_{\nu_1 n}^{(\nu_1-2)/\nu_1}. \tag{3.10}$$

Now (3.6) and the inductive hypothesis imply

$$\begin{aligned} &E(|S_{n-1}|^{\nu_1-2}X_n^2) \\ &\leqq [C_{\nu_0}\beta_{\nu_0,n-1}(n-1)^{\nu_0/2}]^{(\nu_2-2)/\nu_2}\cdot\gamma_{\nu_2 n}^{2/\nu_2} \\ &\leqq C_{\nu_0}^{(\nu_2-2)/\nu_2}\beta_{\nu_1,n-1}^{(\nu_1-2)/\nu_1}(n-1)^{(\nu_1-2)/2}\cdot\gamma_{\nu_2 n}^{2/\nu_2} \quad \text{[from (3.7), (3.10)]} \\ &\leqq C_{\nu_0}^{(\nu_2-2)/\nu_2}\beta_{\nu_1,n-1}^{(\nu_1-2)/\nu_1}(n-1)^{(\nu_1-2)/2}4\gamma_{\nu_1 n}^{2/\nu_1} \quad \text{[from (3.8)]}. \end{aligned} \tag{3.11}$$

It is seen from (1.2) and (3.3) that $C_\nu = [8\delta_\nu]^\nu$. Therefore C_ν is increasing in ν and $C_\nu > 1$. Hence

$$\begin{aligned} C_{\nu_0}^{(\nu_2-2)/\nu_2} &\leqq C_{\nu_1}^{(\nu_2-2)/\nu_2} = C_{\nu_1}C_{\nu_1}^{-2/\nu_2} \\ &\leqq C_{\nu_1}C_{\nu_1}^{-1/\nu_1} \quad \text{[Use (3.9)]} \\ &= C_{\nu_1}(8\delta_{\nu_1})^{-1}. \end{aligned} \tag{3.12}$$

Now the last lines of (3.11) and (3.12) yield

$$E(|S_{n-1}|^{\nu_1-2}X_n^2) \leqq C_{\nu_1}(2\delta_{\nu_1})^{-1}(n-1)^{(\nu_1-2)/2}\beta_{\nu_1,n-1}^{(\nu_1-2)/\nu_1}\cdot\gamma_{\nu_1,n}^{2/\nu_1}. \tag{3.13}$$

From (3.5), with $\nu = \nu_1$, and (3.13), we have

$$\Delta_n(\nu_1) \leqq 2^{-1}\nu_1\delta_{\nu_1}[C_{\nu_1}(2\delta_{\nu_1})^{-1}(n-1)^{(\nu_1-2)/2}\beta_{\nu_1,n-1}^{(\nu_1-2)/\nu_1}\gamma_{\nu_1,n}^{2/\nu_1} + \gamma_{\nu_1,n}]. \tag{3.14}$$

From the corollary of Lemma 2, we have

$$\textstyle\sum_{j=1}^n (j-1)^{(\nu_1-2)/2}\beta_{\nu_1,j-1}^{(\nu_1-2)/\nu_1}\gamma_{\nu_1,j}^{2/\nu_1} \leqq 2\nu_1^{-1}n^{\nu_1/2}\beta_{\nu_1,n}. \tag{3.15}$$

Now (3.14) and (3.15) yield

$$\begin{aligned} E(|S_n|^{\nu_1}) &= \textstyle\sum_{j=1}^n \Delta_j(\nu_1) \\ &\leqq 2^{-1}\nu_1\delta_{\nu_1}[C_{\nu_1}(2\delta_{\nu_1})^{-1}\cdot 2\nu_1^{-1}n^{\nu_1/2}\beta_{\nu_1,n} + n\beta_{\nu_1,n}] \\ &= 2^{-1}\nu_1\delta_{\nu_1}[(\nu_1\delta_{\nu_1})^{-1}C_{\nu_1}n^{\nu_1/2} + n]\beta_{\nu_1,n}. \end{aligned} \tag{3.16}$$

Multiplying the two inequalities $(\nu_1\delta_{\nu_1})^{-1}C_{\nu_1} > 1$ and $n^{\nu_1/2} \geqq n$, we see that the second term in the bracket on the right side of (3.16) is smaller than the first. Therefore (3.16) yields

$$\begin{aligned} E(|S_n|^{\nu_1}) &\leqq 2^{-1}\nu_1\delta_{\nu_1}2(\nu_1\delta_{\nu_1})^{-1}C_{\nu_1}n^{\nu_1/2}\beta_{\nu_1,n} \\ &= C_{\nu_1}n^{\nu_1/2}\beta_{\nu_1,n}. \end{aligned}$$

This completes the proof of the theorem.

4. The case of exchangeable random variables. Our theorem has the following

COROLLARY. *Let* $\{X_n, n \geqq 1\}$ *be an exchangeable process with* $E(X_1X_2) = 0$.

Let $S_n = \sum_{j=1}^{n} X_j$ *and* $\beta_\nu = E(|X_1|^\nu)$. *Then, for all* $\nu \geqq 2$ *and* $n = 1, 2, \cdots$

$$E(|S_n|^\nu) \leqq C_\nu n^{\nu/2}\beta_\nu,$$

with C_ν *given by* (1.2).

PROOF. We shall assume that $E(X_1^2) < \infty$ because, otherwise, the corollary is trivially true. In what follows equalities among random variables will mean equalities with probability one. According to the de Finetti theorem (see also Bühlmann [2], Theorem 2.2.1 and the remark following Corollary 2.4.2″), there is a σ-algebra $\mathfrak{F}$ such that, given $\mathfrak{F}$, the X_i's are independent and identically distributed, and, with $\mu = E(X_1 \mid \mathfrak{F})$, for all $n \geqq 1$,

$$E(X_1X_2 \mid \mathfrak{F}) = \mu^2 \tag{4.1}$$

$$E(X_{n+1} \mid S_n, \mathfrak{F}) = \mu. \tag{4.2}$$

From the assumption $E(X_1X_2) = 0$ from (4.1) we obtain $\mu = 0$. Now (4.2) shows that $E(X_{n+1} \mid S_n) = E(\mu \mid S_n) = 0$. Thus $\{S_n\}$ is a martingale and the theorem applies.

REMARK. The condition $E(X_1X_2) = 0$ of the corollary is necessary. Suppose that $E(X_1X_2) = \xi > 0$. [It cannot be negative because of (4.1).] Then $E(S_n^2) = n\beta_2 + n(n-1)\xi \sim n^2\xi$.

REFERENCES

[1] BÜHLMANN, HANS (1960). Austauschbare stochastische Variabeln und ihre Grenzwertsaetze. *Univ. of California Publ. Statist.* **3** No. 1, 1–36.

[2] BURKHOLDER, D. L. (1966). Martingale transforms. *Ann. Math. Statist.* **37** 1494–1504.

[3] CHUNG, K. L. (1951). The strong law of large numbers. *Proc. Second Berkeley Symp. Math. Statist. Prob.* 341–352. Univ. of California Press.

[4] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[5] MARCINKIEWICZ, J. and ZYGMUND, A. (1937). Sur les fonctions independantes. *Fund. Math.* **29** 60–90.

[6] VON BAHR, B. (1965). On the convergence of moments in the central limit theorem. *Ann. Math. Statist.* **36** 808–818.