The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 1074–1077

ON BOUNDS ON THE CENTRAL MOMENTS OF EVEN ORDER OF A SUM OF INDEPENDENT RANDOM VARIABLES

BY BENGT ROSÉN

Uppsala University

1. The theorem. We shall prove the following theorem.

THEOREM. *Let* $X_1, X_2, \cdots, X_n$ *be independent random variables with mean* 0. *Let* p *be a natural number and* $\lambda_\nu(p)$ *and* $\rho_\nu(p)$ *real numbers such that*

$$EX_\nu^{2k} \leqq \lambda_\nu^{2k}(p)\rho_\nu(p), \qquad k = 1, 2, \cdots, p, \qquad \nu = 1, 2, \cdots, n. \tag{1}$$

Then

$$E(\sum_{\nu=1}^n X_\nu)^{2p} \leqq C(p) \max\left((\sum_{\nu=1}^n \lambda_\nu^2(p)\rho_\nu(p))^p, \quad \sum_{\nu=1}^n \lambda_\nu^{2p}(p)\rho_\nu(p)\right) \tag{2}$$

where $C(p)$ *is a number which only depends on* p.

Before we enter the proof of the theorem, we shall discuss its content somewhat. We list two particular cases, which are included in the theorem.

PARTICULAR CASE 1. *Let* $X_1, X_2, \cdots, X_n$ *be independent random variables with mean* 0. *Then we have for* $p = 1, 2, \cdots$

$$E\left|\sum_{\nu=1}^n X_\nu\right|^{2p} \leqq C(p)(\sum_{\nu=1}^n [E|X_\nu|^{2p}]^{1/p})^p. \tag{3}$$

REMARK 1. This is a special case of an inequality due to P. Whittle [4]. Whittle proved that (3) holds for $p \geqq 1$ (also for non-integral p). Whittle also gives a numerical value for $C(p)$.

REMARK 2. By applying Hölder's inequality to the bound in (3), the following inequality is obtained. For $p = 1, 2, \cdots$, we have

$$E\left|\sum_{\nu=1}^n X_\nu\right|^{2p} \leqq C(p)n^{p-1} \sum_{\nu=1}^n E|X_\nu|^{2p}. \tag{4}$$

This is a special case of a well-known inequality due to Marcinkievitz and Zygmund and Chung, who proved (4) for $p \geqq 1$, see [1] page 348. Whittle's numerical estimate of $C(p)$ works of course in (4) too. Other estimates of $C(p)$ can be found in the paper [2] by Dharmadhikari and Jogdeo.

PARTICULAR CASE 2. *Let* $X_1, X_2, \cdots, X_n$ *be independent random variables with mean* 0. *Let* p *be a natural number. Put*

$$\rho_\nu(p) = \max(EX_\nu^2, EX_\nu^{2p}) \qquad \nu = 1, 2, \cdots, n.$$

Then we have for $p = 1, 2, \cdots$

$$E(\sum_{\nu=1}^n X_\nu)^{2p} \leqq C(p) \max\left((\sum_{\nu=1}^n \rho_\nu(p))^p, \quad \sum_{\nu=1}^n \rho_\nu(p)\right). \tag{5}$$

Received June 9, 1969; revised October 30, 1969.

2. Derivation of the particular cases from the theorem. By using the well-known fact that $(E|X|^r)^{1/r}$ is nondecreasing as r increases, we see that condition (1) is fulfilled for $\lambda_v(p) = (EX_v^{2p})^{1/2p}$ and $\rho_v(p) = 1$, $v = 1, 2, \cdots, n$. By inserting these choices of λ_v and ρ_v into (2) and by using the inequality $\sum \alpha_v^{2p} \leqq (\sum \alpha_v^2)^p$, (3) follows. The second particular case follows from the fact that condition (1) is met for $\lambda_v(p) = 1$ and $\rho_v(p) = \max(EX_v^2, EX_v^4, \cdots, EX_v^{2p}) = \max(EX_v^2, EX_v^{2p})$. The last equality follows from the fact that $E|X|^r$ is convex as function of r.

Neither of the bounds (3) and (5) is generally better than the other, which is illustrated by the following two examples.

EXAMPLE 1. Let Y be Po(λ) (i.e., Poisson distributed with parameter λ). Then, for $r > 1$

$$E|Y-EY|^r = \lambda e^{-\lambda}(\lambda^{r-1} + |1-\lambda|^r + \tfrac{1}{2}\lambda|2-\lambda|^r + \cdots) \sim \lambda \quad \text{as} \quad \lambda \to 0. \tag{6}$$

Consider a double sequence $\{X'_{nv}, v = 1, 2, \cdots, n, n = 1, 2, \cdots\}$ where $X'_{n1}, \cdots, X'_{nn}$ are independent Po(α_n/n) random variables, $n = 1, 2, \cdots$, and $\alpha_n \to 0$ as $n \to \infty$. Put $X_{nv} = X'_{nv} - EX'_{nv}$. As $X'_{n1} + \cdots + X'_{nn}$ is Po(α_n), we get from (6)

$$E(\textstyle\sum_{v=1}^n X_{nv})^{2p} \sim \alpha_n \quad \text{as} \quad n \to \infty, \quad p = 1, 2, \cdots. \tag{7}$$

We calculate the values of the bounds in (3) and (5). By paying regard to (6) we get

$$(\textstyle\sum_{v=1}^n [EX_{nv}^{2p}]^{1/p})^p \sim (\sum_{v=1}^n (\alpha_n/n)^{1/p})^p = \alpha_n n^{p-1} \quad \text{as} \quad n \to \infty, \quad p = 1, 2, \cdots. \tag{8}$$

$$\max((\textstyle\sum_{v=1}^n \rho_{nv}(p))^p, \quad \sum_{v=1}^n \rho_{nv}(p)) \sim \alpha_n \quad \text{as} \quad n \to \infty, \quad p = 1, 2, \cdots. \tag{9}$$

From (8) and (9) we see that for $p = 2, 3, \cdots$ (5) yields a better bound than (3) when n becomes large.

EXAMPLE 2. Let $X_1, X_2, \cdots,$ be independent random variables where X_n is normally distributed with mean 0 and variance n, $n = 1, 2, \cdots$. Then we have, omitting some straightforward calculations

$$E(\textstyle\sum_{v=1}^n X_v)^{2p} \sim C_1(p) n^{2p} \quad \text{as} \quad n \to \infty, \quad p = 1, 2, \cdots. \tag{10}$$

$$(\textstyle\sum_{v=1}^n [EX_v^{2p}]^{1/p})^p \sim C_2(p) n^{2p} \quad \text{as} \quad n \to \infty, \quad p = 1, 2, \cdots. \tag{11}$$

$$\max((\textstyle\sum_{v=1}^n \rho_v(p))^p, \ \sum_{v=1}^n \rho_v(p)) \sim C_3(p) n^{p(p+1)} \quad \text{as} \quad n \to \infty, \quad p = 1, 2, \cdots. \tag{12}$$

From (11) and (12) we see that (3) is superior to (5) in this case, for $p = 2, 3, \cdots$.

We now turn to the proof of the theorem. First we prove a lemma.

LEMMA. *Let $X_1, X_2, \cdots, X_n$ be independent random variables with mean* 0. *Then we have for $p = 1, 2, \cdots$, k_s and u_s being integers,*

$$E(\textstyle\sum_{v=1}^n X_v)^{2p} \leqq C(p) \sum_{k_s \geqq 1, u_s \geqq 0, k_1 u_1 + k_2 u_2 + \cdots + k_p u_p = p} \prod_{s=1}^p (\sum_{v=1}^n EX_v^{2k_s})^{u_s} \tag{13}$$

where $C(p)$ is a number, which only depends on p.

PROOF. We first assume that $X_1, X_2, \cdots, X_n$ all have distributions which are symmetric around 0. Put

$$A(m, k) = E(\sum_{\nu=1}^{m} X_\nu)^{2k}, \qquad m = 1, 2, \cdots, n; \quad k = 1, 2, \cdots$$

and let $A(0, k) = 0$, $k = 1, 2, \cdots$ and $A(m, 0) = 1$, $m = 1, 2, \cdots, n$. According to independence and symmetry we have

$$(14) \qquad A(m, p) = E(\sum_{\nu=1}^{m-1} X_\nu + X_m)^{2p} = \sum_{s=0}^{2p} \binom{2p}{s} E(\sum_{\nu=1}^{m-1} X_\nu)^{2p-s} EX_m^{\ s}$$
$$= A(m-1, p) + \sum_{k=1}^{p} \binom{2p}{2k} A(m-1, p-k) EX_m^{\ 2k}.$$

As all terms in the last sum in (14) are nonnegative, we get

$$(15) \qquad A(m-1, k) \leqq A(m, k), \qquad m = 1, 2, \cdots, n; \quad k = 1, 2, \cdots.$$

From (14) and (15) we get

$$(16) \qquad A(m, p) - A(m-1, p) \leqq D(p) \sum_{k=1}^{p} A(n, p-k) EX_m^{\ 2k}$$

with $D(p) = \max(\binom{2p}{2}, \binom{2p}{4}, \cdots, \binom{2p}{2p})$. By summing over m from 1 to n in (16) we obtain

$$(17) \qquad A(n, p) \leqq D(p) \sum_{k=1}^{p} A(n, p-k) \sum_{\nu=1}^{n} EX_\nu^{\ 2k}.$$

Now (13) follows by iteration of (17), starting with $A(n, 1) = EX_1^{\ 2} + \cdots + EX_n^{\ 2}$.

Thus, the lemma is proved in the case when all X-variables have symmetric distributions. To prove the general case we shall use the following inequality which is well known. Let X and Y be independent random variables, Y having mean 0. Then

$$(18) \qquad E|X|^r \leqq E|X-Y|^r, \qquad r \geqq 1.$$

For the sake of completeness we indicate a proof. Let $r \geqq 1$. The curve $z = |x|^r$ is convex (in x), and thus it does not fall below any of its tangents. This yields $|x-y|^r \geqq |x|^r - y|x|^{r-1} \cdot \text{sign } x$, which gives $|X-Y|^r \geqq |X|^r - Y|X|^{r-1} \cdot \text{sign } X$. Now, take expectation in the last inequality and (18) is obtained.

Let $X_1', X_2', \cdots, X_n'$ be random variables such that X_ν and X_ν' have the same distribution and such that $X_1, X_1', X_2, X_2', \cdots, X_n, X_n'$ are independent. We have

$$(19) \qquad E(X_\nu - X_\nu')^{2k} \leqq 2^{2k} EX_\nu^{\ 2k}, \qquad \nu = 1, 2, \cdots, n; \quad k = 1, 2, \cdots.$$

According to (18) we have

$$(20) \qquad E(\sum_{\nu=1}^{n} X_\nu)^{2p} \leqq E(\sum_{\nu=1}^{n} (X_\nu - X_\nu'))^{2p}.$$

As $X_\nu - X_\nu'$ has a symmetric distribution, we can—from what is already proved—apply (13) to the right-hand side in (20). The proof is now easily completed by paying regard to (19).

3. Proof of the theorem. According to (1) and the convexity of $\log(\sum_{v=1}^{n}\lambda_v{}^r\rho_v)$ as a function of r (see e.g. [3] 2.9, 2.10 and 3.6) we get

$$(21)\quad \begin{aligned}\textstyle\sum_{v=1}^{n} EX_v{}^{2k} &\leqq \textstyle\sum_{v=1}^{n}\lambda_v{}^{2k}(p)\rho_v(p)\\ &\leqq \left(\textstyle\sum_{v=1}^{n}\lambda_v{}^{2p}(p)\rho_v(p)\right)^{(k-1)/(p-1)}\left(\textstyle\sum_{v=1}^{n}\lambda_v{}^{2}(p)\rho_v(p)\right)^{(p-k)/(p-1)},\\ &\qquad k=1,2,\cdots,p.\end{aligned}$$

It follows from (21) that the product in (13) is dominated by

$$(22)\qquad \left(\sum_{v=1}^{n}\lambda_v{}^{2p}(p)\rho_v(p)\right)^{\frac{p-\Sigma u_s}{p-1}}\left(\sum_{v=1}^{n}\lambda_v{}^{2}(p)\rho_v(p)\right)^{p\cdot\frac{\Sigma u_s-1}{p-1}}.$$

As $(p-\sum u_s)/(p-1)$ and $(\sum u_s-1)/(p-1)$ both are nonnegative and their sum is 1, the term in (22) is dominated by

$$(23)\qquad \max\left(\left(\textstyle\sum_{v=1}^{n}\lambda_v{}^{2}(p)\rho_v(p)\right)^p,\quad \textstyle\sum_{v=1}^{n}\lambda_v{}^{2p}(p)\rho_v(p)\right).$$

The theorem now follows from the lemma, as each term in (13) is dominated by (23) and the summation includes only finitely many terms.

REFERENCES

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

[2] Dharmadhikari, S. W. and Jogdeo, K. (1969). Bounds on moments of certain random variables. *Ann. Math. Statist.* **40** 1506–1509.

[3] Hardy, G. H., Littlewood, J. E. and Polya, G. (1959). *Inequalities.* Cambridge Univ. Press.

[4] Whittle, P. (1960). Bounds for the moments of linear and quadratic forms in independent random variables. *Theor. Probability Appl.* **5** 302–305.