PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 14, Fasc. 2 (1993), pp. 281–285

COMPARISON OF SUMS OF INDEPENDENT IDENTICALLY DISTRIBUTED RANDOM VECTORS

BY

S. J. MONTGOMERY-SMITH (COLUMBIA, MISSOURI)

Abstract. Let S_k be the k-th partial sum of Banach space valued independent identically distributed random variables. In this paper, we compare the tail distribution of $\|S_k\|$ with that of $\|S_j\|$, and deduce some tail distribution maximal inequalities.

The main result of this paper* was inspired by the inequality from [1] that says that

$$\Pr(\|X_1\| > t) \leqslant 5 \Pr(\|X_1+X_2\| > t/2)$$

whenever X_1 and X_2 are independent identically distributed. Similar results for L_p $(p \geqslant 1)$ such as $\|X_1\|_p \leqslant \|X_1+X_2\|_p$ are straightforward, at least if X_2 has zero expectation. This inequality is also obvious if either X_1 is symmetric or X_1 is real valued positive. However, for arbitrary random variables, this result is somewhat surprising to the author. Note that the identically distributed assumption cannot be dropped, as one could take $X_1 = 1$ and $X_2 = -1$.

In this paper, we prove a generalization to sums of arbitrarily many independent identically distributed random variables. Note that all results in this paper are true for Banach space valued random variables. The author would like to thank Victor de la Peña for helpful conversations.

THEOREM 1. *There exist universal constants $c_1 = 3$ and $c_2 = 10$ such that if $X_1, X_2, \ldots$ are independent identically distributed random variables, and if we set $S_k = \sum_{i=1}^{k} X_i$, then for $1 \leqslant j \leqslant k$*

$$\Pr(\|S_j\| > t) \leqslant c_1 \Pr(\|S_k\| \, t/c_2).$$

This result cannot be asymptotically improved. Consider, for example, the case where $X_1 = 1$ with very small probability, and $X_1 = 0$ otherwise. This shows that for the inequality to be true for all X_1 the constant c_2 must be larger than some universal constant for all j and k. Also, it is easy to see that c_1 must

* Research supported in part by N.S.F. Grant DMS 9201357.

be larger than some universal constant because it is easy to select X_1 and t so that $\Pr(\|S_j\| > t)$ is close to 1.

However, Latała [3] has been able to obtain the same theorem with $c_1 = 4$ and $c_2 = 5$, or $c_1 = 2$ and $c_2 = 7$. In the case $j = 1$, he has shown that

$$\Pr(\|X_1\| > t) \leqslant 2\Pr\big(\|S_k\| > kt/(2k-1)\big),$$

and these constants cannot be improved.

In oder to show this result, we will use the following definition. We will say that x is a *t-concentration point* for a random variable X if $\Pr(\|X-x\| \leqslant t) > 2/3$.

LEMMA 2. *If x is a t-concentration point for X, and y is a t-concentration point for Y, and z is a t-concentration point for $X+Y$, then*

$$\|x+y-z\| \leqslant 3t.$$

Proof. We have

$$\begin{aligned}\Pr(\|x+y-z\| > 3t)& \\ &\leqslant \Pr\big(\|X-x+Y-y-(X+Y-z)\| > 3t\big)\\ &\leqslant \Pr(\|X-x\| > t)+\Pr(\|Y-y\| > t)+\Pr(\|X+Y-z\| > t) < 1.\end{aligned}$$

Hence $\Pr(\|x+y-z\| \leqslant 3t) > 0$. Since x, y and z are fixed vectors, the result follows. ∎

COROLLARY 3. *If $X_1, X_2, \ldots$ are independent identically distributed random variables, and if the partial sums $S_j = \sum_{i=1}^{j} X_i$ have t-concentration points s_j for $1 \leqslant j \leqslant k$, then*

$$\|ks_j - js_k\| \leqslant 3(k+j)t.$$

Proof. We prove the result by induction. It is obvious if $j = k$. Otherwise,

$$\begin{aligned}\|js_k - ks_j\| &\leqslant \|js_{k-j} - (k-j)s_j\| + \|js_k - js_{k-j} - js_j\|\\ &\leqslant 3(k-j+j)t + 3jt = 3(k+j)t.\end{aligned}$$

(The observant reader will notice that we are, in fact, following the steps of Euclidean algorithm. The same proof could show that $\|ks_j - js_k\| \leqslant 3(j+k-2h)t$, where h is the greatest common divisor of j and k.) ∎

Proof of Theorem 1. We consider three cases. First suppose that

$$\Pr(\|S_{k-j}\| > 9t/10) \leqslant 1/3.$$

Note that $S_k - S_j$ is independent of S_j, and identically distributed to S_{k-j}. Then

$$\begin{aligned}\Pr(\|S_j\| > t) &\leqslant (3/2)\Pr(\|S_j\| > t \text{ and } \|S_k - S_j\| \leqslant 9t/10)\\ &\leqslant (3/2)\Pr(\|S_k\| > t/10).\end{aligned}$$

For the second case, suppose that there is an i $(1 \leqslant i \leqslant k)$ such that S_i does not have any $(t/10)$-concentration point. Then

$$\Pr(\|S_i+X_{i+1}+ \ldots +X_k\| > t/10 \,|\, \sigma(X_{i+1}, \ldots, X_k)) \geqslant 1/3,$$

and hence $\Pr(\|S_k\| > t/10) \geqslant 1/3 \geqslant (1/3)\Pr(\|S_i\| > t)$.

Finally, we are left with the third case where $\Pr(\|S_{k-j}\| > 9t/10) > 1/3$, and S_i has a $(t/10)$-concentration point s_i for all $1 \leqslant i \leqslant k$. Clearly, $\|s_{k-j}\| \geqslant 8t/10$. Also, by Corollary 3,

$$\|s_k\| \geqslant \frac{k}{k-j}\|s_{k-j}\| - \frac{3(2k-j)t}{10(k-j)} \geqslant \frac{8kt}{10(k-j)} - \frac{6kt}{10(k-j)} \geqslant \frac{2t}{10}.$$

Therefore,

$$\Pr(\|S_k\| \geqslant t/10) \geqslant \Pr(\|S_k - s_k\| \leqslant t/10) \geqslant 2/3 \geqslant (2/3)\Pr(\|S_j\| > t),$$

and we are done. ∎

One might be emboldened to conjecture the following. Suppose that $X_1, X_2, \ldots$ are independent identically distributed random variables, and that $\alpha_i > 0$. Let

$$S_k = \sum_{i=1}^{k} \alpha_i X_i.$$

Then one might conjecture that there is a universal constant such that for $1 \leqslant j \leqslant k$

$$\Pr(\|S_j\| > t) \leqslant c\Pr(\|S_k\| > t/c).$$

It turns out that this is not the case. Let $Y_1, Y_2, \ldots$ be real valued independent identically distributed random variables such that

$$\Pr(Y_i = N-1) = 1/N, \qquad \Pr(Y_i = -1) = (N-1)/N.$$

Then, by the central limit theorem, there exists $M \geqslant N^3$ such that

$$\Pr\left(\left|\frac{1}{M^{2/3}}\sum_{i=1}^{M} Y_i\right| > \frac{1}{N}\right) \leqslant \frac{1}{N}.$$

Now let $X_i = Y_i + 1/M^{1/3}$, and let

$$S_M = \frac{1}{M^{2/3}}\sum_{i=1}^{M} X_i.$$

Then $\Pr(|S_M| > 1/2) \geqslant 1-1/N$, whereas $\Pr(|S_M + X_{M+1}| > 3/N) \leqslant 2/N$.

We can obtain several corollaries to Theorem 1.

Corollary 4. *There is a universal constant c such that if $X_1, X_2, \ldots$ are independent identically distributed random variables, and if we set $S_k = \sum_{i=1}^{k} X_i$, then*

$$\Pr\left(\sup_{1 \leqslant j \leqslant k} \|S_j\| > t\right) \leqslant c\Pr(\|S_k\| > t/c).$$

Latała [3] has been able to obtain this result with $c_1 = 4$ and $c_2 = 6$, or with $c_1 = 2$ and $c_2 = 8$.

Proof. This follows from Proposition 1.1.1 of [2] that states that if $X_1, X_2, \ldots$ are independent (not necessarily identically distributed), and if $S_k = \sum_{i=1}^k X_i$, then

$$\Pr(\sup_{1 \leqslant j \leqslant k} \|S_j\| > t) \leqslant 3 \sup_{1 \leqslant j \leqslant k} \Pr(\|S_j\| > t/3).$$

It is also possible to prove this result directly using the techniques of the proof of Theorem 1. The third case only requires that $\Pr(\|S_{k-j}\| > 9t/10) > 1/3$ for one of $j = 1, 2, \ldots, k$. Hence, for the first case we may assume that $\Pr(\|S_k - S_j\| > 9t/10) \leqslant 1/3$ for all $1 \leqslant j \leqslant k$. Let A_j be the event $\{\|S_i\| \leqslant t$ for all $i < j$ and $\|S_j\| > t\}$. Then

$$\Pr(A_j) \leqslant (3/2)\Pr(A_j \text{ and } \|S_k - S_j\| \leqslant 9t/10) \leqslant (3/2)\Pr(A_j \text{ and } \|S_j\| > t/10).$$

Summing over j, the result follows. ∎

COROLLARY 5. *There is a universal constant c such that if $X_1, X_2, \ldots$ are independent identically distributed random variables, and if $|\alpha_i| \leqslant 1$, then*

$$\Pr(\|\sum_{i=1}^k \alpha_i X_i\| > t) \leqslant c \Pr(\|\sum_{i=1}^k X_i\| > t/c).$$

Proof. The technique used in this proof is well known (see [2], Proposition 1.2.2), but is included for completeness.

By taking real and imaginary parts of α_i, we may suppose that the α_i are real. Without loss of generality, $1 \geqslant \alpha_1 \geqslant \ldots \geqslant \alpha_k \geqslant -1$. Then we may write $\alpha_j = -1 + \sum_{i=j}^k \sigma_i$, where $\sigma_i \geqslant 0$. Thus $\sum_{i=1}^k |\sigma_i| \leqslant 2$, and hence

$$\|\sum_{j=1}^k \alpha_j X_i\| = \|\sum_{j=1}^k (-1 + \sum_{i=j}^k \sigma_i) X_i\| = \|-(\sum_{i=1}^k X_i) + (\sum_{j=1}^k \sigma_j \sum_{i=1}^j X_i)\|$$
$$\leqslant \|\sum_{i=1}^k X_i\| + (\sum_{j=1}^k |\sigma_j|) \sup_{1 \leqslant j \leqslant k} \|\sum_{i=1}^j X_i\|.$$

Applying Corollary 4, we obtain the result.

COROLLARY 6. *There are universal constants c_1 and c_2 such that if $X_1, X_2, \ldots$ are independent identically distributed random variables, and if we set $S_k = \sum_{i=1}^k X_i$, then for $1 \leqslant k \leqslant j$*

$$\Pr(\|S_j\| > t) \leqslant (c_1 j/k) \Pr(\|S_k\| > kt/c_2 j).$$

Proof. Let m be the least integer such that $mk \geqslant j$. By Theorem 1, it follows that

$$\Pr(\|S_j\| > t) \leqslant c \Pr(\|S_{mk}\| > t/c).$$

The relation $\Pr(\|S_{mk}\| > t) \leqslant m \Pr(\|S_k\| > t/m)$ is straightforward. ∎

The example where X_1 is constant shows that c_2 cannot be made smaller than some universal constant. The example where $X_1 = 1$ with very small probability and $X_1 = 0$ otherwise shows the same is true for c_1.

REFERENCES

[1] V. H. de la Peña and S. J. Montgomery-Smith, *Bounds on the tail probability of U-statistics and quadratic forms*, preprint.

[2] S. Kwapień and W. A. Woyczyński, *Random Series and Stochastic Integrals: Simple and Multiple*, Birkhäuser, New York 1992.

[3] R. Latała, *A maximal inequality for sums of independent, identically distributed random vectors*, preprint, Warsaw University, Warszawa 1993.

Department of Mathematics
University of Missouri
Columbia, MO 65211, U.S.A.

Received on 20.12.1993