A COMPARISON INEQUALITY FOR SUMS OF INDEPENDENT RANDOM VARIABLES*

STEPHEN J. MONTGOMERY-SMITH AND ALEXANDER R. PRUSS

Abstract. We give a comparison inequality that allows one to estimate the tail probabilities of sums of independent Banach space valued random variables in terms of those of independent *identically distributed* random variables. More precisely, let $X_1, \ldots, X_n$ be independent Banach-valued random variables. Let I be a random variable independent of $X_1, \ldots, X_n$ and uniformly distributed over $\{1, \ldots, n\}$. Put $\tilde{X}_1 = X_I$, and let $\tilde{X}_2, \ldots, \tilde{X}_n$ be independent identically distributed copies of $\tilde{X}_1$. Then, $P(\|X_1 + \cdots + X_n\| \geq \lambda) \leq cP(\|\tilde{X}_1 + \cdots + \tilde{X}_n\| \geq \lambda/c)$ for all $\lambda \geq 0$, where c is an absolute constant.

The independent Banach-valued random variables $X_1, \ldots, X_n$ are said to *regularly cover* (the distribution of) a random variable Y provided that

$$E[g(Y)] = \frac{1}{n} \sum_{k=1}^{n} E[g(X_k)],$$

for all Borel functions g for which either side is defined [8]. An easy way of constructing Y, given the independent Banach-valued random variables $X_1, \ldots, X_n$, is to let I be a random variable independent of $X_1, \ldots, X_n$, with values in $\{1, 2, \ldots, n\}$ and with each value having equal probability $1/n$, and then put $Y = X_I$. It is easy to see that then $X_1, \ldots, X_n$ regularly cover Y. This construction will be useful for our proofs.

If the variables are real valued, then the regular covering condition is easily seen to be equivalent to the condition that the distribution

Date: October 15, 1999.

1991 *Mathematics Subject Classification.* Primary 60G50, 60E15; Secondary 60F15.

Key words and phrases. comparison inequalities, sums of independent random variables, sums of independent identically distributed random variables, rates of convergence in the law of large numbers.

*The first author was supported in part by the NSF and the Research Board of the University of Missouri.

function F of Y is the arithmetic mean of the respective distribution functions $F_1, \dots, F_n$ of $X_1, \dots, X_n$.

A variable X' is said to be a *copy* of X if it has the same distribution as X. The main purpose of this paper is then to prove the following result.

Theorem 1. *There exists an absolute constant $c \in (0, \infty)$ such that if $X_1, \dots, X_n$ are independent Banach-valued random variables which regularly cover a random variable $\tilde{X}_1$, then:*

$$P(\|X_1 + \cdots + X_n\| \geq \lambda) \leq cP(\|\tilde{X}_1 + \cdots + \tilde{X}_n\| \geq \lambda/c), \tag{1}$$

for all $\lambda \geq 0$, where $\tilde{X}_2, \dots, \tilde{X}_n$ are independent copies of $\tilde{X}_1$.

Remark 1. In the case where the random variables are symmetric, this was shown in [9] (strictly speaking, it was only shown in the real-valued case, but the proof also works for the Banach-valued case).

Remark 2. The inequality converse to (1) is false, even in the special cases of symmetric real random variables. For, suppose that c is an absolute constant such that

$$P(|\tilde{X}_1 + \cdots + \tilde{X}_n| \geq \lambda) \leq cP(|X_1 + \cdots + X_n| \geq \lambda/c), \tag{2}$$

for all $\lambda \geq 0$, whenever the conditions of Theorem 1 hold with symmetric variables. Fix any $n > \max(1, c)$. Put $X_2 \equiv \cdots \equiv X_n \equiv 0$. Let X_1 be such that $P(X_1 = 1) = P(X_1 = -1) = \frac{1}{2}$. Put $\lambda = n$. Then the right hand side of (2) is zero, since $|X_1 + \cdots + X_n| \equiv 1$. But the left hand side of (2) is non-zero, since it is easy to see that $P(\tilde{X}_i = 1) = 1/2n$ for each i (as the $\tilde{X}_i$ are identically distributed, and as $\tilde{X}_1$ can be taken to be X_I where I is independent of everything else and uniformly distributed on $\{1, \dots, n\}$), so that $P(|\tilde{X}_1 + \cdots + \tilde{X}_n| \geq n) \geq (1/2n)^n > 0$.

Remark 3. The main consequence of Theorem 1 is that any upper bound on tail probabilities of sums of independent identically distributed random variables automatically gives a bound on tail probabilities of sums of non-identically distributed independent random variables. Our result explains why upper bounds for sums of independent identically distributed cases have so readily and consistently generalized to non-identically distributed cases. We cannot think of any new applications at this time.

Remark 4. For a very simple, though not new, application, we give another proof of one side of a result from [8] on randomly sampled Riemann sums. Let $f \in L^2[0,1]$. For $1 \le k \le n$, let x_{nk} be uniformly distributed over $[(k-1)/n, k/n]$, and assume $x_{n1}, \dots, x_{nn}$ are independent for each fixed n. Define the randomly sampled Riemann sum $R_n f = n^{-1} \sum_{k=1}^n f(x_{nk})$. Then the result says that $R_n f$ converges almost surely to the Lebesgue integral $A = \int_0^1 f$. (For a converse in the case where *all* the x_{nk} are independent, not just for fixed n, see [8].) For, by Borel-Cantelli it suffices to show that

$$\sum_{n=1}^{\infty} P(|R_n f - A| \ge \varepsilon) < \infty, \tag{3}$$

for all $\varepsilon > 0$. Let $X_1, X_2, \dots$ be independent identically distributed random variables with the same distribution as f. Note that $f(x_{n1}), \dots, f(x_{nn})$ regularly cover X_1, and $f(x_{n1}) - A, \dots, f(x_{nn}) - A$ regularly cover $X_1 - A$. Since $f \in L^2$, we have X_1 having a finite second moment, and moreover $E[X_1] = A$, so that by the Hsu-Robbins law of large numbers [6] (see also [3, 4]), we have

$$\sum_{n=1}^{\infty} P(|(X_1 - A) + \cdots + (X_n - A)|/n \ge \varepsilon) < \infty,$$

for all $\varepsilon > 0$. By Theorem 1 and the fact that $f(x_{n1}) - A, \dots, f(x_{n1}) - A$ regularly cover $X_1 - A$, we obtain (3).

To prove Theorem 1, we need some definitions and lemmata. If X is a random variable, then let $X^s = X - X'$ be the *symmetrization* of X, where X' is an independent copy of X. We shall always choose symmetrizations so that we have $(X_1 + \cdots + X_k)^s = X_1^s + \cdots + X_k^s$ whenever we need this identity.

Write $\|X\|_p = (E[\|X\|^p])^{1/p}$, where $\|\cdot\|$ is the norm on the Banach space in which our random variables take values.

The following lemma must surely be well known. However, we could find no reference, and so we include a short proof.

Lemma 1. *Let X be a Banach-valued random variable with $\|X\|_2 < \infty$. Then, $\|X\|_2 \le \|X^s\|_2 + \|E[X]\| \le 3\|X\|_2$*

Proof. Let X' be an independent copy of X so that $X^s = X - X'$. Let $\mathcal{A}$ be the sigma-algebra generated by X. Then $E[X^s \mid \mathcal{A}] = X - E[X'] =$

$X - E[X]$, and so

$$\|X\|_2 = \|E[X^s + E[X] \mid \mathcal{A}]\|_2 \le \|X^s + E[X]\|_2 \le \|X^s\|_2 + \|E[X]\|,$$

where the first inequality used the fact that conditional expectation is a contraction on the Banach-valued L^p spaces, $p \ge 1$ (see, e.g., [2, Chapter V, Theorem 4]). The rest of the Lemma follows from the triangle inequality. □

Lemma 2. *Let $X_1, \dots, X_n$ be independent random variables, and let $\tilde{X}_1, \dots, \tilde{X}_n$ be independent identically distributed random variables such that $X_1, \dots, X_n$ regularly cover $\tilde{X}_1$. Put $S_n = X_1 + \cdots + X_n$ and $\tilde{S}_n = \tilde{X}_1 + \cdots + \tilde{X}_n$. Then:*

$$\|S_n\|_2 \le 12\|\tilde{S}_n\|_2.$$

Proof. Let $I_1, \dots, I_n$ be independent random variables uniformly distributed on the set $\{1, \dots, n\}$. Let $\{X_{i,j}\}_{1\le i,j\le n}$ and $\{X'_{i,j}\}_{1\le i,j\le n}$ be independent arrays of independent random variables, with the arrays independent of the I_i, and such that $X_{i,j}$ and $X'_{i,j}$ both have the same distribution as X_j for all i and j. Without loss of generality we can put $\tilde{X}_i = X_{i,I_i}$. Set $\tilde{X}'_i = X'_{i,I_i}$. Let $\tilde{S}'_n = \tilde{X}'_1 + \cdots + \tilde{X}'_n$. Let $(X'_1, \dots, X'_n)$ be an independent copy of $(X_1, \dots, X_n)$, and put $S'_n = X'_1 + \cdots + X'_n$. Observe that $X_1 - X'_1, \dots, X_n - X'_n$ regularly cover $\tilde{X}_i - \tilde{X}'_i$ for all i, and that moreover the $X_i - X'_i$ are symmetric. Thus, by [9, Proposition 1] (which though stated for real valued random variables, holds for the Banach-valued case as well, and with the same proof) we have:

$$\|S_n - S'_n\|_2 \le 4\|\tilde{S}_n - \tilde{S}'_n\|_2. \tag{4}$$

Also, it is clear that $E[S_n] = E[\tilde{S}_n]$. Combining this with Lemma 1, we see that:

$$\|S_n\|_2 \le \|S_n - S'_n\|_2 + \|E[S_n]\| \le 4\|\tilde{S}_n - \tilde{S}'_n\|_2 + 4\|E[\tilde{S}_n]\| \le 12\|\tilde{S}_n\|_2,$$

as desired. □

The following Lemma is in effect a special case of a result of Hitczenko [5].

Lemma 3. *Let $X_1, \dots, X_n$ be independent identically distributed Banach-valued random variables with $\|X_i\| < L$ almost surely for all i. Let $S_k = X_1 + \cdots + X_k$. Then:*

$$\|S_n\|_1 \ge c_1\|S_n\|_2 - c_2 L,$$

where $c_1, c_2 \in (0,\infty)$ *are absolute constants.*

Proof. By the work of Hitczenko [5], we have that if $S^* = \max_k \|S_k\|$ and $X^* = \max_k \|X_k\|$, then for $q \geq p$:

$$\|S^*\|_q \leq c_3 \frac{q}{p}(\|S^*\|_p + \|X^*\|_q),$$

for a finite absolute constant c_3. By [7, Corollary 4] we have $\|S^*\|_p \leq c_4 \|S_n\|_p$ for an absolute constant c_4, as the X_i are identically distributed. The desired inequality easily follows from this if we let $q = 2$ and $p = 1$. □

Proof of Theorem 1. Let $I_1, \ldots, I_n$, $\{X_{i,j}\}_{1\leq i,j\leq n}$, $\{X'_{i,j}\}_{1\leq i,j\leq n}$, S'_n and $\tilde{S}'_n$ be as in the proof of Lemma 2. As in that proof, $X_{n,1} - X'_{n,1}, \ldots, X_{n,n} - X_{n,n}$ regularly cover $X_{n,I_k} - X'_{n,I_k}$ for any fixed $k \leq n$. Applying [9, Proposition 1] (which works for Banach-valued variables as already stated), we see that

$$\begin{aligned} P(\|S_n - S'_n\| \geq \lambda) &\leq 8P(\|\tilde{S}_n - \tilde{S}'_n\| \geq \lambda/2) \\ &\leq 8P(\|\tilde{S}_n\| \geq \lambda/4) + 8P(\|\tilde{S}'_n\| \geq \lambda/4) \\ &\leq 16P(\|\tilde{S}_n\| \geq \lambda/4), \end{aligned} \tag{5}$$

for all λ, where the second inequality followed from the inequality that $P(\|X^s\| \geq t) \leq P(\|X\| \geq t/2) + P(\|X'\| \geq t/2) = 2P(\|X\| \geq t/2)$, where X' is an independent copy of X such that $X^s = X - X'$. Note that $S^s_n = S_n - S'_n$.

Let M be a median of $\|S_n\|$. It is easy to see that

$$P(|\|S_n\| - M| \geq \lambda) \leq 2P(\|S^s_n\| \geq \lambda), \tag{6}$$

for all λ. (For, if $\|S_n\| - M \geq \lambda$, there is at least probability 1/2 that $\|S'_n\| \leq M$ in which case $\|S_n - S'_n\| \geq \|S_n\| - \|S'_n\| \geq \|S_n\| - M \geq \lambda$ and a similar calculation can be done if $\|S_n\| - M \leq -\lambda$.)

We now claim that in general in our present setting:

$$P(\|\tilde{S}_n\| \geq \varepsilon M) > \delta, \tag{7}$$

for absolute constants $\varepsilon, \delta \in (0,1)$ to be determined later. (They will be determined in accordance with (12), (18), (20), (25) and (26), below). To prove (7), suppose that on the contrary we have:

$$P(\|\tilde{S}_n\| \geq \varepsilon M) \leq \delta. \tag{8}$$

Since the $\tilde{X}_i$ are independent and identically distributed, by a maximal inequality for sums of independent and identically distributed random variables [7, Corollary 4] together with (8), we have:

$$P\big(\max_{1\le k\le n} \|\tilde{S}_k\| \ge c_1\varepsilon M\big) \le c_1 P(\|\tilde{S}_n\| \ge \varepsilon M) \le c_1\delta, \tag{9}$$

where $c_1 \in [1,\infty)$ is an absolute constant. By the elementary inequality

$$P(\max_{1\le k\le n} \|U_k\| \ge 2t) \le P\left(\max_{1\le k\le n} \left\|\sum_{i=1}^k U_i\right\| \ge t\right),$$

valid for all t if the U_i are independent (since if $\|U_k\| \ge 2t$ then $\|\sum_{i=1}^k U_i\| \ge t$ or $\|\sum_{i=1}^{k-1} U_i\| \ge t$), it follows from (9) that

$$P\big(\max_{1\le k\le n} \|\tilde{X}_k\| \ge 2c_1\varepsilon M\big) \le c_1\delta. \tag{10}$$

Let $L = 2c_1\varepsilon M$. Set $Y_k = X_k \cdot 1_{\{\|X_k\|<L\}}$. Put $\tilde{Y}_k = \tilde{X}_k \cdot 1_{\{\|\tilde{X}_k\|<L\}}$. Note that $Y_1,\dots,Y_n$ regularly cover $\tilde{Y}_k$ for each k. Let $T_n = Y_1+\cdots+Y_n$ and put $\tilde{T}_n = \tilde{Y}_1+\cdots+\tilde{Y}_n$. By (10), we have:

$$P\left(\bigcup_{k=1}^n \{\tilde{X}_k \ne \tilde{Y}_k\}\right) \le c_1\delta. \tag{11}$$

Let $p = P(\|\tilde{X}_k\| \ge L)$. Note that this does not depend on k since the $\tilde{X}_k$ are identically distributed. Henceforth we will assume that

$$\delta < 1/(2c_1). \tag{12}$$

Now, if $x \in [0,1]$ is such that $1-(1-x)^n \le 1/2$, then $x \le 2n^{-1}(1-(1-x)^n)$. Then, since the left hand side of (11) equals $1-(1-p)^n$, it follows from (11) that

$$p \le 2n^{-1}(1-(1-p)^n) \le 2n^{-1}c_1\delta.$$

Combining this with (12) and the condition that $X_1,\dots,X_n$ regularly cover $\tilde{X}_1$, we see that:

$$\begin{aligned} P\left(\bigcup_{k=1}^n \{X_k \ne Y_k\}\right) &\le \sum_{k=1}^n P(X_k \ne Y_k) \\ &= \sum_{k=1}^n P(\|X_k\| \ge L) \\ &= nP(\|\tilde{X}_1\| \ge L) = np \le 2c_1\delta. \end{aligned} \tag{13}$$

Now, by (5), (6) and (8), it follows that

$$P(|\|S_n\| - M| \geq 4\varepsilon M) \leq 32\delta.$$

Using (13), it then follows that:

$$P(|\|T_n\| - M| \geq 4\varepsilon M) \leq (32 + 2c_1)\delta. \tag{14}$$

Moreover, by (8) and (11):

$$P(\|\tilde{T}_n\| \geq \varepsilon M) \leq (1 + c_1)\delta. \tag{15}$$

Observe that $\|\tilde{Y}_i\| < L$ almost surely. Lemma 3 then shows that:

$$\|\tilde{T}_n\|_1 \geq c_2\|\tilde{T}_n\|_2 - c_3 L, \tag{16}$$

where $c_2, c_3 \in (0, \infty)$ are absolute constants.

Now, by (14) we have:

$$E[\|T_n\|^2] \geq [1 - (32 + 2c_1)\delta](1 - 4\varepsilon)^2 M^2. \tag{17}$$

Henceforth, we will assume that δ and ϵ are sufficiently small that

$$(1 - (32 + 2c_1)\delta)(1 - 4\varepsilon)^2 \geq \tfrac{1}{4}. \tag{18}$$

Using Lemma 2 we see that $E[\|T_n\|^2] \leq 144 E[\|\tilde{T}_n\|^2]$. Combining this with (17) and (18), we see that

$$\|\tilde{T}_n\|_2 \geq M/24. \tag{19}$$

Recall that $L = 2c_1\varepsilon M$, and choose $\epsilon \leq c_2/(96c_1c_3)$, so that $c_3 L \leq c_2 M/48$. This assumption is equivalent to:

$$\varepsilon \leq c_2/(96c_1c_3). \tag{20}$$

Thus by (19):

$$c_3 L \leq \tfrac{1}{2}c_2\|\tilde{T}_n\|_2. \tag{21}$$

Then, by (16),

$$\|\tilde{T}_n\|_1 \geq \tfrac{1}{2}c_2\|\tilde{T}_n\|_2. \tag{22}$$

The elementary inequality $P(|\Xi| \geq \lambda E[|\Xi|]) \geq (1-\lambda)^2 (E[|\Xi|])^2/E[|\Xi|^2]$ (see, e.g., [1, Exercise 3.3.11]) then implies that

$$P(\|\tilde{T}_n\| \geq \tfrac{1}{2}\|\tilde{T}_n\|_1) \geq (1 - \tfrac{1}{2})^2 \cdot \tfrac{1}{4}c_2^2. \tag{23}$$

Now, by (19) and (22) we have $\|\tilde{T}_n\|_1 \geq c_2 M/48$, so that (23) gives:

$$P(\|\tilde{T}_n\| \geq c_2 M/96) \geq c_2^2/16. \tag{24}$$

If we choose ε and δ such that

$$0 < \varepsilon \leq c_2/96 \tag{25}$$

and

$$0 < (1 + c_1)\delta < c_2^2/16 \tag{26}$$

and satisfying the other conditions required in the above argument (namely (12), (18) and (20)), we will obtain from (24) a contradiction to (15). Hence, if we take ε and δ to be absolute constants in $(0, 1)$ satisfying these assumptions, we obtain (7).

Now, combining (5) and (6), we see that:

$$P(\|S_n\| - M \geq \lambda) \leq 32P(\|\tilde{S}_n\| \geq \lambda/4), \tag{27}$$

for all λ. There are now two cases to be considered. Suppose first that $\lambda \leq 2M$. Then using (7):

$$P(\|S_n\| \geq \lambda) \leq 1 \leq \delta^{-1}P(\|\tilde{S}_n\| \geq \varepsilon M) \leq \delta^{-1}P(\|\tilde{S}_n\| \geq \varepsilon\lambda/2). \tag{28}$$

On the other hand, suppose that $\lambda > 2M$. In that case if $\|S_n\| \geq \lambda$ then $\|S_n\| - M > \lambda - \lambda/2 = \lambda/2$, so that

$$P(\|S_n\| \geq \lambda) \leq P(\|S_n\| - M \geq \lambda/2) \leq 32P(\|\tilde{S}_n\| \geq \lambda/4), \tag{29}$$

by (27). Inequality (1) follows from (28) for $\lambda \leq 2M$ and from (29) for $\lambda > 2M$, if we let $c = \max(32, 2/\varepsilon, \delta^{-1})$. □

References

1. K. L. Chung, "A Course in Probability Theory," 2nd ed., Academic Press, San Diego, 1974.
2. J. Diestel and J. J. Uhl, Jr., "Vector Measures," Math. Surveys, vol. 15, Amer. Math. Soc., Providence, RI, 1977.
3. P. Erdős, On a theorem of Hsu and Robbins, *Ann. Math. Statist.* **20** (1949), 286–291.
4. ______, Remark on my paper "On a theorem of Hsu and Robbins", *Ann. Math. Statist.* **21** (1950), 138.
5. P. Hitczenko, On a domination of sums of random variables by sums of conditionally independent ones, *Ann. Probab.* **22** (1994), 453–468.
6. P. L. Hsu and H. Robbins, Complete convergence and the law of large numbers, *Proc. Nat. Acad. Sci. U.S.A.* **33** (1947), 25–31.
7. S. J. Montgomery-Smith, Comparison of sums of independent identically distributed random vectors, *Probab. Math. Statist.* **14** (1993), 281–285.
8. A. R. Pruss, Randomly sampled Riemann sums and complete convergence in the law of large numbers for a case without identical distribution, *Proc. Amer. Math. Soc.* **124** (1996), 919–929.

9. ———, Comparisons between tail probabilities of sums of independent symmetric random variables, *Ann. Inst. Poincaré Probab. Statist.* **33** (1997), 651–671.

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

Department of Philosophy, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.

E-mail address: stephen@math.missouri.edu
E-mail address: pruss+@pitt.edu
URL: http://www.math.missouri.edu/~stephen
URL: http://www.pitt.edu/~pruss