Measuring the magnitude of sums of independent random variables

Paweł Hitczenko*
Department of Mathematics and Computer Science
Drexel University
Philadelphia, PA 19104
phitczen@mcs.drexel.edu
http://www.mcs.drexel.edu/~phitczen

Stephen Montgomery-Smith†
Department of Mathematics
University of Missouri-Columbia
Columbia, MO 65211
stephen@math.missouri.edu
http://www.math.missouri.edu/~stephen

Abstract

This paper considers how to measure the magnitude of the sum of independent random variables in several ways. We give a formula for the tail distribution for sequences that satisfy the so called Lévy property. We then give a connection between the tail distribution and the pth moment, and between the pth moment and the rearrangement invariant norms.

Keywords: sum of independent random variables, tail distributions, decreasing rearrangement, pth moment, rearrangement invariant space, disjoint sum, maximal function, Hoffmann-Jørgensen/Klass-Nowicki Inequality, Lévy Property.

A.M.S. Classification (1991): Primary 60G50, 60E15, 46E30; Secondary 46B09.

*The first named author was partially supported by NSF grant DMS 9401345.

†The second named author was partially supported by NSF grants DMS 9424396 and DMS 9870026, and by the University of Missouri Research Board.

1 Introduction

This paper is about the following type of problem: given independent (not necessarily identically distributed) random variables $X_1, X_2, \ldots, X_N$, find the 'size' of $|S|$, where

$$S = \sum_{n=1}^{N} X_n.$$

We will examine several ways to measure this size. The first will be through tail distributions, that is, $\Pr(|S| > t)$. Finding an exact solution to this problem would be a dream of probabilists, so we have to temper our desires in some manner. In fact, this problem goes back to the foundations of probability in the following form: if the sequence (X_n) consists of random variables that are mean zero, identically distributed and have finite variance, find the asymptotic value of $\Pr(|S| > \sqrt{N}t)$ as $N \to \infty$. This is answered, of course, by the Central Limit Theorem, which tells us that the answer is the Gaussian distribution. There has been a tremendous amount of work on generalizing this. We refer the reader to almost any advanced work on probability.

Our approach is different. Instead of seeking asymptotic solutions, we will look for approximate solutions. That is, we seek a function $f(t)$, computed from (X_n), such that there is a positive constant c with

$$c^{-1} f(ct) \leq \Pr(|S| > t) \leq c\, f(c^{-1}t).$$

The second measurement of the size of $|S|$ will be through the pth moments, $\|S\|_p = (\mathbb{E}|S|^p)^{1/p}$. Again, we shall be searching for approximate solutions, that is, finding a quantity A such that there is a positive constant c so that

$$c^{-1} A \leq \|S\|_p \leq c\, A.$$

While this may seem like quite a different problem, in fact, as we will show, there is a precise connection between the two, in that obtaining an approximate formula for $\|S\|_p$ with constants that are uniform as $p \to \infty$ is equivalent to obtaining an approximate formula for the tail distribution.

The third way that we shall look at is to find the size of $|S|$ in a rearrangement invariant space. This line of research was began by Carothers and Dilworth (1988) who obtained results for Lorentz spaces, and was completed by Johnson and Schechtman (1989). Our results will give a comparison of the size of $|S|$ in the rearrangement invariant space with $\|S\|_p$, obtaining a greater control on the sizes of the constants involved than the previous works.

Many of the results of this paper will be true for all sums of independent random variables, even those that are vector valued, with the following proviso. Instead of considering the sum $S = \sum_n X_n$, we will consider the maximal function $U = \sup_n |\sum_{k=1}^{n} X_k|$. We will define a property for sequences called the *Lévy property*, which will imply that U is comparable to S. Sequences with this Lévy property will include positive random variables, symmetric random variables, and identically distributed random variables. The result of this paper that gives the

tail distribution for S is only valid for real valued sequences of random variables that satisfy the Lévy property. However the results connecting the L_p and the rearrangement invariant norms to the tail distributions of U are valid for all sequences of vector valued independent random variables. (Since this paper was submitted, Mark Rudelson pointed out to us that some of the inequalities can be extended from U to S by a simple symmetrization argument. We give details at the end of each relevant section.)

Let us first give the historical context for these results, considering first the problem of approximate formulae for the tail distribution. Perhaps the earliest works are the Paley-Zygmund inequality (see for example Kahane (1968, Theorem 3, Chapter 2)), and Kolmogorov's reverse maximal inequality (see for example Shiryaev (1980, Chapter 4, section 2.)) Both give (under an extra assumption) a lower bound on the probability that a sum of independent, mean zero random variables exceeds a fraction of its standard deviation and both may be regarded as a sort of converse to the Chebyshev's inequality. Next, in 1929, Kolmogorov, proved a two-sided exponential inequality for sums of independent, mean-zero, uniformly bounded, random variables (see for example Stout (1974, Theorem 5.2.2) or Ledoux and Talagrand (1991, Lemma 8.1)). All of these results require some restriction on the nature of the sequence (X_n), and on the size of the level t.

Hahn and Klass (1997) obtained very good bounds on one sided tail probabilities for sums of independent, identically distributed, real valued random variables. Their result had no restrictions on the nature of the random variable, or on the size of the level t. In effect, their result worked by removing the very large parts of the random variables, and then using an exponential estimate on the rest. We will take a similar approach in this paper.

Let us next look at the pth moments. Khintchine (1923) gave an inequality for Rademacher (Bernoulli) sums. This very important formula has found extensive applications in analysis and probability. Khintchine's result was extended to any sequence of positive or mean zero random variables by the celebrated result of Rosenthal (1970). The order of the best constants as $p \to \infty$ was obtained by Johnson, Schechtman and Zinn (1983), and Pinelis (1994) refined this still further. Now even more precise results are known, and we refer the reader to Figiel, Hitczenko, Johnson, Schechtman and Zinn (1997) (see also Ibragimov and Sharakhmetov (1997)). However, the problem with all these results is that the constants were not uniformly bounded as $p \to \infty$.

Khintchine's inequality was generalized independently by Montgomery and Odlyzko (1988) and Montgomery-Smith (1990). They were able to give approximate bounds on the tail probability for Rademacher sums, with no restriction on the level t. Hitczenko (1993) obtained an approximate formula for the L_p norm of Rademacher sums, where the constants were uniformly bounded as $p \to \infty$. (A more precise version of this last result was obtained in Hitczenko-Kwapień (1994) and it was used to give a simple proof of the lower bound in Kolmogorov's exponential inequality.)

Continuing in the direction of Montgomery and Odlyzko, Montgomery-Smith and Hitczenko, Gluskin and Kwapień (1995) extended tail and moment estimates from Rademacher sums to weighted sums of random variables with logarithmically concave tails (that is, $P(|X| \geq t) = \exp(-\phi(t))$, where $\phi : [0, \infty) \to [0, \infty)$ is convex). After that, Hitczenko, Montgomery-Smith, and Oleszkiewicz (1997) treated the case of logarithmically convex tails (that is, the ϕ above is concave rather than convex). It should be emphasized that in the last paper, the

result of Hahn and Klass (1997) played a critical role.

The breakthrough came with the paper of Latała (1997), who solved the problem of finding upper and lower bounds for general sums of positive or symmetric random variables, with uniform constants as $p \to \infty$. His method made beautiful use of special properties of the function $t \mapsto t^p$. In a short note, Hitczenko and Montgomery-Smith (1999) showed how to use Latała's result to derive upper and lower bounds on tail probabilities. Latała's result is the primary motivation for this paper.

The main tool we will use is the Hoffmann-Jørgensen Inequality. In fact, we will use a stronger form of this inequality, due to Klass and Nowicki (1998). The principle in many of our proofs is the following idea. Given a sequence of random variables (X_n), we choose an appropriate level $s > 0$. Each random variable X_n is split into the sum $X_n^{(\leq s)} + X_n^{(>s)}$, where $X_n^{(\leq s)} = X_n I_{|X_n| \leq s}$, and $X_n^{(>s)} = X_n I_{|X_n| > s}$. It turns out that the quantity $(X_n^{(>s)})$ can either be disregarded, or it can be considered as a sequence of disjoint random variables. (By "disjoint" we mean that the random variables are disjointly supported as functions on the underlying probability space.) As for the quantity $\sum_n X_n^{(\leq s)}$, it will turn out that the level s allows one to apply the Hoffmann-Jørgensen/Klass-Nowicki Inequality so that it may be compared with quantities that we understand rather better.

Let us give an outline of this paper. In Section 2, we will give definitions. This will include the notion of decreasing rearrangement, that is, the inverse to the distribution function. Many results of this paper will be written in terms of the decreasing rearrangement. Section 3 is devoted to the Klass-Nowicki Inequality. Since our result is slightly stronger than that currently in the literature, we will include a full proof. In Section 4, we will introduce and discuss the Lévy property. This will include a "reduced comparison principle" for sequences with this property. Section 5 contains the formula for the tail distribution of sums of real valued random variables. Then in Section 6, we demonstrate the connection between L_p-norms of such sums and their tail distributions. In Section 7 we will discuss sums of independent random variables in rearrangement invariant spaces.

2 Notation and definitions

Throughout this paper, a random variable will be a measurable function from a probability space to some Banach space (often the real line). The norm in the implicit Banach space will always be denoted by $|\cdot|$.

Suppose that $f : [0, \infty) \to [0, \infty]$ is a non-increasing function. Define the *left continuous inverse* to be

$$f^{-1}(x-) = \sup\{y : f(y) \geq x\},$$

and the *right continuous inverse* to be

$$f^{-1}(x+) = \sup\{y : f(y) > x\}.$$

In describing the tail distribution of a random variable X, instead of considering the function $t \mapsto \Pr(|X| > t)$, we will consider its right continuous inverse, which we will denote by $X^*(t)$. In fact, this quantity appears very much in the literature, and is more commonly referred to as the *decreasing rearrangement* (or more correctly the *non-increasing*

rearrangement) of $|X|$. Notice that if one considers X^* to be a random variable on the probability space $[0,1]$ (with Lebesgue measure), then X^* has exactly the same law as $|X|$. We might also consider the left continuous inverse $t \mapsto X^*(t-)$. Notice that $X^*(t) \le x \le X^*(t-)$ if and only if $\Pr(|X| > x) \le t \le \Pr(|X| \ge x)$.

If A and B are two quantities (that may depend upon certain parameters), we will write $A \approx B$ to mean that there exist positive constants c_1 and c_2 such that $c_1^{-1}A \le B \le c_2 A$. We will call c_1 and c_2 the constants of approximation. If $f(t)$ and $g(t)$ are two (usually non-increasing) functions on $[0,\infty)$, we will write $f(t) \underset{t}{\approx} g(t)$ if there exist positive constants c_1, c_2, c_3 and c_4 such that $c_1^{-1} f(c_2 t) \le g(t) \le c_3 f(c_4^{-1} t)$ for all $t \ge 0$. Again, we will call c_1, c_2, c_3 and c_4 the constants of approximation.

Suppose that X and Y are random variables. Then the statement $\Pr(|X| > t) \underset{t}{\approx} \Pr(|Y| > t)$ is the same as the statement $X^*(t) \underset{t}{\approx} Y^*(t)$. Since $X^*(t) = 0$ for $t \ge 1$ the latter statement is equivalent to the existence of positive constants c_1, c_2, c_3, c_4 and c_5 such that $c_1^{-1} X^*(c_2 t) \le Y^*(t) \le c_3 X^*(c_4^{-1} t)$ for $t \le c_5^{-1}$.

To avoid bothersome convergence problems, we will always suppose that our sequence of independent random variables (X_n) is of finite length. Given a sequence of independent random variables (X_n), when no confusion will arise, we will use the following notations. If A is a finite subset of $\mathbb{N}$, we will let $S_A = \sum_{n \in A} X_n$, and $M_A = \sup_{n \in A} |X_n|$. If k is a positive integer, then $S_k = S_{\{1,\ldots,k\}}$ and $M_k = M_{\{1,\ldots,k\}}$. We will define the maximal function $U_k = \sup_{1 \le n \le k} |S_n|$. Furthermore, $S = S_N$, $M = M_N$, and $U = U_N$, where N is the length of the sequence (X_n).

If s is a real number, we will write $X_n^{(>s)} = X_n I_{|X_n| > s}$ and $X_n^{(\le s)} = X_n I_{|X_n| \le s} = X_n - X_n^{(>s)}$. For $A \subset \mathbb{N}$, we will write $S_A^{(\le s)} = \sum_{n \in A} X_n^{(\le s)}$. Similarly we define $S_A^{(>s)}$, $S_k^{(\le s)}$, etc.

Another quantity that we shall care about is the decreasing rearrangement of the disjoint sum of random variables. This notion was used by Johnson, Maurey, Schechtman and Tzafriri (1979), Carothers and Dilworth (1988), and Johnson and Schechtman (1989), all in the context of sums of independent random variables. The disjoint sum of the sequence (X_n) is the measurable function on the measure space $\Omega \times \mathbb{N}$ that takes (ω, n) to $X_n(\omega)$. We shall denote the decreasing rearrangement of the disjoint sum by $\tilde{\ell} : [0,\infty) \to [0,\infty]$, that is, $\tilde{\ell}(t)$ is the least number such that

$$\sum_n \Pr(|X_n| > \tilde{\ell}(t)) \le t.$$

Define $\ell(t)$ to be $\tilde{\ell}(t)$ if $0 \le t \le 1$, and 0 otherwise. Since $\ell(t)$ is only non-zero when $0 \le t \le 1$, we will think of ℓ as being a random variable on the probability space $[0,1]$ with Lebesgue measure. The quantity ℓ is effectively M in disguise. This next result (and its proof) essentially appears in Giné and Zinn (1983).

Proposition 2.1 *If $0 < t < 1$, then*

$$\ell(2t) \le \ell(t/(1-t)) \le M^*(t) \le \ell(t).$$

Proof: The first inequality follows easily once one notices that both sides of this inequality are zero if $t > 1/2$.

To get the second inequality, note that, by an easy argument, if $\alpha_1, \alpha_2, \cdots \geq 0$ with $\sum_n \alpha_n \leq 1$, then

$$1 - \sum_n \alpha_n \leq \prod_n (1 - \alpha_n) \leq 1 - \frac{\sum_n \alpha_n}{1 + \sum_n \alpha_n}.$$

So, if $\Pr(\ell > x) = \sum_n \Pr(|X_n| > x) \leq 1$, then

$$\Pr(M > x) = 1 - \prod_n (1 - \Pr(|X_n| > x)),$$

and hence

$$\frac{\Pr(\ell > x)}{1 + \Pr(\ell > x)} \leq \Pr(M > x) \leq \Pr(\ell > x).$$

Taking inverses, the result follows. □

3 The Klass-Nowicki Inequality

This section is devoted to the following result — the Klass-Nowicki Inequality.

Theorem 3.1 *Let (X_n) be a sequence of Banach valued independent random variables. Then for all positive integers K we have*

$$\Pr(U > 2Kt + (K-1)s) \leq \frac{1}{K!} \left(\frac{\Pr(U > t)}{1 - \Pr(U > t)} \right)^K + \Pr(M > s),$$

whenever $\Pr(U > t) < 1$.

The original inequality of this form was for Rademacher (or Bernoulli) sums and $K = 2$, and was due to Kahane (1968). This was extended by Hoffmann-Jørgensen (1974) to general sums, at least for positive or symmetric random variables, for the case $K = 2$. Indeed, if one wants Theorem 3.1 for $K > 2$, but without the $K!$ factor, this may be obtained by iterating the Hoffmann-Jørgensen Inequality, as was done by Johnson and Schechtman (1989, Lemmas 6 and 7). (Both Kahane and Hoffmann-Jørgensen obtained slightly different constants than those we have presented. Also, in neither case did a factor like $(1 - \Pr(U > t))$ appear in their formulae.)

Klass and Nowicki (1998) were able to obtain Theorem 3.1, at least in the case when the random variables are positive or symmetric. (However their constants are better than ours.) Removing the positive or symmetric condition is really not so hard, but because it does not appear in the literature in this manner, we will give a complete proof of Theorem 3.1.

We also note that this inequality has some comparison with a result that appears in Ledoux and Talagrand (1991, Theorem 6.17.)

Proof: Let N be the length of the sequence (X_n). During this proof, let us write $(m, n]$ for the set of integers greater than m and not greater than n.

We start with the observation

$$\Pr(U > 2Kt + (K-1)s) \leq \Pr(U > 2Kt + (K-1)s \text{ and } M \leq s) + \Pr(M > s).$$

Now, if we have that both $U > 2Kt + (K-1)s$ and $M \leq s$, then we ensure the existence of an increasing sequence of non-negative integers $m_0, \ldots, m_K$, bounded by N, and defined as follows. Set $m_0 = 0$. If we have picked m_{l-1}, let m_l be the smallest positive integer greater than m_{l-1} such that $|S_{(m_{l-1},m_l]}| > 2t$. For $l = 1$, it is clear that such an integer exists. Let us explain why the integer $m_l \leq N$ exists if $2 \leq l \leq K$.

For $1 \leq l' \leq l-1$, and $m_{l'-1} < k \leq m_{l'} - 1$, we have that $|S_{(m_{l'-1},k]}| \leq 2t$ and $|X_{m_{l'}}| \leq s$. Hence for $1 \leq l' \leq l-1$

$$|S_{m_{l'}}| = \left| \sum_{j=1}^{l'} (S_{(m_{j-1},m_j-1]} + X_{m_j}) \right| \leq 2l't + l's,$$

and for $m_{l'-1} < k \leq m_{l'} - 1$

$$|S_k| = \left| \left(\sum_{j=1}^{l'-1} (S_{(m_{j-1},m_j-1]} + X_{m_j}) \right) + S_{(m_{l'-1},k]} \right| \leq 2l't + (l'-1)s.$$

But we know that there exists a number m such that $|S_m| > 2Kt + (K-1)s$. Hence, we must have that $m > m_{l-1}$, and that $|S_{(m_{l-1},m]}| > 2Kt + (K-1)s - 2(l-1)t - (l-1)s \geq 2t$.

Therefore

$$\Pr(U > 2Kt + (K-1)s \text{ and } M \leq s) \leq \sum_{1 \leq m_1 < \cdots < m_K \leq N} p_{0,m_1} p_{m_1,m_2} \cdots p_{m_{K-1},m_K},$$

where

$$p_{m,n} = \Pr(|S_{(m,k]}| \leq 2t \text{ for } m \leq k < n, \text{ and } |S_{(m,n]}| > 2t).$$

Now let us show the following inequality:

$$\sum_{k=m+1}^{n} p_{m,k} \leq \frac{1}{1 - \Pr(U > t)} \sum_{k=m+1}^{n} \tilde{p}_k,$$

where

$$\tilde{p}_n = \Pr(|S_k| \leq t \text{ for } 1 \leq k < n, \text{ and } |S_n| > t).$$

Using independence, we have that

$$\begin{aligned}
\sum_{k=m+1}^{n} p_{m,k} &= \Pr(\sup_{m<k\leq n} |S_k - S_m| > 2t) \\
&= \Pr(\sup_{m<k\leq n} |S_k - S_m| > 2t \big| U_m \leq t) \\
&\leq \Pr(\sup_{m<k\leq n} |S_k| > t \big| U_m \leq t) \\
&= \frac{\Pr(\sup_{m<k\leq n} |S_k| > t \text{ and } \sup_{1\leq k\leq m} |S_k| \leq t)}{\Pr(U_m \leq t)} \\
&\leq \frac{1}{1 - \Pr(U > t)} \sum_{k=m+1}^{n} \tilde{p}_k,
\end{aligned}$$

as required.

Now we rearrange the sum as follows:

$$\begin{aligned}
&\sum_{1\le m_1<\cdots<m_K\le N} p_{0,m_1}p_{m_1,m_2}\cdots p_{m_{K-1},m_K} \\
=\ &\sum_{1\le m_1<\cdots<m_{K-1}\le N} p_{0,m_1}p_{m_1,m_2}\cdots p_{m_{K-2},m_{K-1}} \sum_{m_K=m_{K-1}+1}^{N} p_{m_{K-1},m_K} \\
\le\ &\frac{1}{1-\Pr(U>t)} \sum_{1\le m_1<\cdots<m_{K-1}\le N} p_{0,m_1}p_{m_1,m_2}\cdots p_{m_{K-2},m_{K-1}} \sum_{m_K=m_{K-1}+1}^{N} \tilde{p}_{m_K}.
\end{aligned}$$

Now we rearrange this last quantity to get

$$\begin{aligned}
&\sum_{1\le m_1<\cdots<m_K\le N} p_{0,m_1}p_{m_1,m_2}\cdots p_{m_{K-1},m_K} \\
\le\ &\frac{1}{1-\Pr(U>t)} \sum_{1\le m_1<\cdots<m_{K-2}<m_K\le N} p_{0,m_1}p_{m_1,m_2}\cdots p_{m_{K-3},m_{K-2}}\tilde{p}_{m_K} \times \\
&\times \sum_{m_{K-1}=m_{K-2}+1}^{m_K} p_{m_{K-2},m_{K-1}} \\
\le\ &\frac{1}{(1-\Pr(U>t))^2} \sum_{1\le m_1<\cdots<m_{K-2}<m_K\le N} p_{0,m_1}p_{m_1,m_2}\cdots p_{m_{K-3},m_{K-2}}\tilde{p}_{m_K} \times \\
&\times \sum_{m_{K-1}=m_{K-2}+1}^{m_K} \tilde{p}_{m_{K-1}}.
\end{aligned}$$

Repeating this argument $(K-2)$ more times, we eventually see that

$$\begin{aligned}
&\sum_{1\le m_<\cdots<m_K\le N} p_{0,m_1}p_{m_1,m_2}\cdots p_{m_{K-1},m_K} \\
\le\ &\frac{1}{(1-\Pr(U>t))^K} \sum_{1\le m_1<\cdots<m_K\le N} \tilde{p}_{m_1}\tilde{p}_{m_2}\cdots\tilde{p}_{m_K}.
\end{aligned}$$

Now, since K distinct numbers may be rearranged in $K!$ different ways, we have that

$$\begin{aligned}
&\sum_{1\le m_1<\cdots<m_K\le N} \tilde{p}_{m_1}\tilde{p}_{m_2}\cdots\tilde{p}_{m_K} \\
=\ &\frac{1}{K!} \sum_{\substack{1\le m_1,m_2,\ldots,m_K\le N \\ m_1,m_2,\ldots,m_K \text{ distinct}}} \tilde{p}_{m_1}\tilde{p}_{m_2}\cdots\tilde{p}_{m_K} \\
\le\ &\frac{1}{K!} \sum_{1\le m_1,m_2,\ldots,m_K\le N} \tilde{p}_{m_1}\tilde{p}_{m_2}\cdots\tilde{p}_{m_K} \\
=\ &\frac{1}{K!}\left(\sum_{k=1}^{N}\tilde{p}_k\right)^K.
\end{aligned}$$

Since

$$\sum_{k=1}^{N} \tilde{p}_k = \Pr(U > t),$$

we obtain the result. □

Let us now understand what this result means in terms of the decreasing rearrangement.

Corollary 3.2 *There exists a universal positive constant c_1 such that for any sequence of Banach valued independent random variables (X_n), and for $0 < t \leq s \leq 1/2$ we have*

$$U^*(t) \leq c_1 \frac{\log(1/t)}{\max\{\log(1/s), \log\log(4/t)\}} \big(U^*(s) + M^*(t/2)\big).$$

Proof: Notice that if $f, g : [0, \infty) \to [0, \infty]$ are non-increasing functions, then $(\max\{f, g\})^{-1} = \max\{f^{-1}, g^{-1}\}$, and if $f \leq g$, then $f^{-1} \leq g^{-1}$, where here f^{-1} denotes either the left or right continuous inverse of f. Since $A + B \leq 2\max\{A, B\}$ for any two positive numbers A and B, from Theorem 3.1, and setting $s = t$, we have that if $\Pr(U > t) \leq 1/2$, then for all positive integers K

$$\Pr(U > (3K-1)t) \leq 2\max\left\{\frac{1}{K!}\big(2\Pr(U > t)\big)^K, \Pr(M > t)\right\}.$$

Taking inverses, we see that if $(K!t/2)^{1/K} \leq 1/2$, then

$$\frac{1}{3K-1} U^*(t) \leq \max\left\{U^*\left(\frac{1}{2}\left(\frac{K!t}{2}\right)^{1/K}\right), M^*\left(\frac{t}{2}\right)\right\}.$$

Now, using the fact that $\max\{A, B\} \leq A + B$ for any positive numbers A and B, and by choosing K to be the smallest integer such that $s \leq (K!t/2)^{1/K}$, and by some elementary but tedious algebra, the result follows. □

Since this paper was submitted, Mark Rudelson pointed out to us a couple of ways that Theorem 3.1 can be improved. First, we may obtain a result closer to that of Ledoux and Talagrand (1991, Theorem 6.17. Let $(|X_n|^*)$ be the order statistics of $(|X_n|)$, that is, the values of $(|X_n|)$ rearranged in decreasing order. Then exactly the same proofs gives the following strengthening: for all positive integers K

$$\Pr(U > 2Kt + (K-1)s) \leq \frac{1}{K!}\left(\frac{\Pr(U > t)}{1 - \Pr(U > t)}\right)^K + \Pr\left(\sum_{n=1}^{K} |X_n|^* > (K-1)s\right),$$

whenever $\Pr(U > t) < 1$.

Secondly, a similar result is also true if we replace U by $|S|$. This is certainly the case if the sequence (X_n) consists of symmetric random variables, since they satisfy the Lévy property. Now let $(\bar{X}_n)$ be an independent copy of (X_n), and let $\tilde{X}_n = X_n - \bar{X}_n$. Let $\bar{S}$ and

$\tilde{S}$ respectively denote the sums formed from these two sequences of random variables. Thus we have the result for $\left|\tilde{S}\right|$, since it is a sum of symmetric random variables. But

$$\Pr(\left|\tilde{S}\right| > ct) \geq \Pr(|S| > (c+1)t \text{ and } \left|\bar{S}\right| \leq t) = \Pr(|S| > (c+1)t)\Pr(|S| \leq t).$$

Arguing in this way, we quickly see that there are constants c_1, c_2, and c_3 such that

$$\Pr(S > c_1 K(s+t)) \leq \frac{\Pr(S > t)^K}{K!(1 - c_2 \Pr(S > c_3 t))^{K+1}} + \Pr(M > s),$$

whenever $\Pr(S > c_3 t) < c_2^{-1}$.

Thus a version of Corollary 3.2 is also true when U is replaced by $|S|$.

4 The Lévy Property

Let (X_n) be a sequence of independent random variables. We will say that (X_n) satisfies the *Lévy property* with constants c_1 and c_2 if whenever $A \subseteq B \subseteq \mathbb{N}$, with A and B finite, then for $t > 0$

$$\Pr(|S_A| > c_1 t) \leq c_2 \Pr(|S_B| > t).$$

The casual reader should beware that this property has nothing to do with Lévy processes.

The sequence (X_n) has the *strong Lévy property* with constants c_1 and c_2 if for all $s > 0$ the sequence $(X_n^{(\leq s)})$ has the Lévy property with constants c_1 and c_2.

Here are examples of sequences with the strong Lévy property. (It may be easily seen that in all these cases it is sufficient to show that they have the Lévy property.)

(i) Positive sequences, with constants 1 and 1.

(ii) Sequences of symmetric random variables with constants 1 and 2. This "reflection property" plays a major role in results attributed to Lévy, hence the name of the property.

(iii) Sequences of identically distributed random variables. This was shown independently by Montgomery-Smith (1993) with constants 10 and 3, and by Latała (1993) with constants 5 and 4, or 7 and 2.

We see that sequences with the Lévy property satisfy a maximal inequality.

Proposition 4.1 *Let (X_n) be a sequence of independent random satisfying the Lévy property with constants c_1 and c_2. Then for all $t > 0$*

$$\Pr(U > 3c_1 t) \leq 3c_2 \Pr(|S| > t).$$

Thus $M^*(t) \leq 6c_1 S^*(t/3c_2)$.

Proof: The first statement is an immediate corollary of the following result known as Lévy-Ottaviani inequality:

$$\Pr(U_N > 3t) \leq 3 \sup_{1 \leq k \leq N} \Pr(|S_k| > t).$$

(Billingsley (1995, Theorem 22.5, p. 288) attributes this result to Etemadi (1985) who proved it with constants 4 in both places, but the same proof gives constants 3; see, for example, Billingsley. However the first named author learned this result from Kwapień in 1980.)

The second statement follows from the first, since $M \leq 2U$. □

We end with a lemma that lists some elementary properties. Part (i) of the lemma might be thought of as a kind of reduced comparison principle.

Lemma 4.2 *Let (X_n) be a sequence of random variables satisfying the strong Lévy property.*

(i) There exist positive constants c_1 and c_2, depending only upon the Lévy constants of (X_n), such that if $s \leq 1/2$ and $0 \leq t \leq 1$, then

$$(S^{(\leq M^*(s))})^*(t) \leq c_1 S^*(c_2^{-1} t).$$

(ii) There exist positive constants c_1 and c_2, depending only upon the strong Lévy constants of (X_n), such that if $r \leq s \leq 1/2$, and if $0 \leq t \leq 1$, then $(S^{(\leq M^*(s))})^*(t) \leq c_1 (S^{(\leq M^*(r))})^*(c_2^{-1} t)$.

(iii) If $0 \leq s \leq t \leq 1$, then $S^*(t) \leq (S^{(\leq M^*(s))})^*(t-s)$, and $(S^{(\leq M^*(s))})^*(t) \leq S^*(t-s)$. In particular, $S^*(t) \leq (S^{(\leq M^*(t/2))})^*(t/2)$, and $(S^{(\leq M^*(t/2))})^*(t) \leq S^*(t/2)$.

(iv) For $\alpha, \beta > 0$, we have that

$$(S^{(\leq M^*(t))})^*(t) \underset{t}{\approx} (S^{(\leq M^*(\alpha t))})^*(\beta t)$$

where the constants of approximation depend only upon α, β and the strong Lévy constants of (X_n).

(v) We have that

$$S^*(t) \underset{t}{\approx} (S^{(\leq M^*(t))})^*(t) \underset{t}{\approx} (S^{(\leq \ell(t))})^*(t),$$

where the constants of approximation depend only upon the strong Lévy constants of (X_n).

Proof: Let us start with part (i). For each set $A \subseteq \mathbb{N}$, define the event

$$E_A = \{|X_n| \leq M^*(s) \text{ if and only if } n \in A\}.$$

Note that the whole probability space is the disjoint union of these events. Also

$$\{|S^{(\leq M^*(s))}| > x\} \cap E_A = \{|S_A| > x\} \cap E_A.$$

Furthermore, by independence, we see that

$$\begin{aligned}&\Pr(|S_A| > x \text{ and } E_A)\\&= \Pr(|S_A| > x \text{ and } |X_n| \le M^*(s) \text{ for } n \in A)\Pr(|X_n| > M^*(s) \text{ for } n \notin A).\end{aligned}$$

Hence

$$\begin{aligned}&\Pr(|S^{(\le M^*(s))}| > x)\\&= \sum_{A\subseteq\mathbb{N}} \Pr(|S_A| > x \text{ and } |X_n| \le M^*(s) \text{ for } n \in A)\Pr(|X_n| > M^*(s) \text{ for } n \notin A)\\&\le 2\sum_{A\subseteq\mathbb{N}} \Pr(|S_A| > x)\Pr(|X_n| \le M^*(s) \text{ for } n \in A)\Pr(|X_n| > M^*(s) \text{ for } n \notin A)\\&\le c_2 \Pr(|S| > c_1^{-1}x),\end{aligned}$$

where in the first inequality we have used the fact that

$$\Pr(|X_n| \le M^*(s) \text{ for } n \in A) \ge \Pr(M \le M^*(s)) \ge 1 - s \ge 1/2.$$

Part (ii) follows by applying part (i) to $S^{(\le M^*(r))}$.

Part (iii) follows from the observation that

$$\Pr(S \ne S^{(\le M^*(s))}) \le \Pr(M > M^*(s)) \le s.$$

Hence, if $\Pr(S > \alpha) \ge t$, then $\Pr(S^{(\le M^*(s))} > \alpha) \ge t - s$, and conversely, if $\Pr(S^{(\le M^*(s))} > \alpha) > t$ then $\Pr(S > \alpha) \ge t - s$.

To show part (iv), we may suppose without loss of generality that $\alpha = 1$ and $\beta > 1$. Clearly $S^{(\le M^*(t))}(t) \ge S^{(\le M^*(t))}(\beta t)$, so we need only show an opposite inequality. From part (ii), there are positive constants c_1 and c_2, depending only upon the strong Lévy constants of (X_n), such that for $0 \le t \le 1/2$

$$S^{(\le M^*(t))}(t) \le c_1 S^{(\le M^*(c_2^{-1}\beta^{-1})t)}(c_2^{-1}t) \le c_1 S^{(\le M^*(c_3^{-1}t)}(c_3^{-1}\beta t),$$

where $c_3 = c_2\beta$.

Part (v) follows easily by combining part (iii), part (iv), and Proposition 2.1. □

5 Tail distributions

In this section, we will state and prove the formula for the tail distribution of the sum of independent, real valued, random variables that satisfy the Lévy Property.

If one restricts the formula to the case of sums of independent, identically distributed random variables, one obtains a formula very similar to the main result of Hahn and Klass (1997). The main differences are that their inequality involves one sided inequalities, and also that their inequality is more precise.

This formula also has a strong resemblance to the result of Latała. As we shall show in Section 6, computing the L_p norm of U is effectively equivalent to computing $U^*(e^{-p})$. Then

if one notices that $(1+x)^p$ is very close to e^{xp} for small positive x, one can see that this result and the result of Latała are very closely related. Presumably one could derive Latała's result by combining Theorem 5.1 with Theorem 6.1. However the technical difficulties are quite tricky, and since Latała's proof is elegant, we will not carry out this program here.

Theorem 5.1 *Let (X_n) be a sequence of real valued independent random variables satisfying the strong Lévy property. Define the functions $F_1(t)$ and $F_2(t)$ to be 0 if $t > 1$, and if $0 \le t \le 1$,*

$$\begin{aligned} F_1(t) &= \inf\left\{\lambda > 0 : \prod_n \mathbb{E}(t^{X_n^{(\le \ell(t))}/\lambda}) \le t^{-1} \text{ and } \prod_n \mathbb{E}(t^{-X_n^{(\le \ell(t))}/\lambda}) \le t^{-1}\right\}, \\ F_2(t) &= \inf\left\{\lambda > 0 : \prod_n \mathbb{E}(t^{X_n^{(\le M^*(t))}/\lambda}) \le t^{-1} \text{ and } \prod_n \mathbb{E}(t^{-X_n^{(\le M^*(t))}/\lambda}) \le t^{-1}\right\}. \end{aligned}$$

Then

$$S^*(t) \underset{t}{\approx} F_1(t) \underset{t}{\approx} F_2(t),$$

where the constants of approximation depend only upon the strong Lévy constants of (X_n).

Let us start with gaining some understanding of *Orlicz spaces*. There is a huge literature on Orlicz spaces, see for example Lindenstrauss and Tzafriri (1977). Suppose that $\Phi : [0,\infty) \to [0,\infty]$ is an increasing function (usually convex with $\Phi(0) = 0$). Then the *Orlicz norm* of a random variable X is defined according to the formula

$$\|X\|_\Phi = \inf\{\lambda > 0 : \mathbb{E}\Phi(|X|/\lambda) \le 1\}.$$

We will be concerned with the special functions

$$\Phi_t(x) = \frac{t^{-x} - 1}{t^{-1} - 1}.$$

The following is a special case of results that appear in Montgomery-Smith (1992).

Lemma 5.2 *For any random variable X, and for $t \le 1/4$, we have that*

$$\|X\|_{\Phi_t} \approx \sup_{0 \le x \le 1} \frac{\log(t)}{\log(xt)} X^*(x),$$

with constants of approximation bounded by 2.

Proof: Suppose first that $\|X\|_{\Phi_t} \le 1$. Then $\mathbb{E}\Phi_t(X) \le 1$, which implies that

$$xt^{-X^*(x)} \le \int_0^1 t^{-X^*(y)}\, dy \le \mathbb{E}(t^{-|X|}) \le t^{-1},$$

that is, $X^*(x) \le \log(xt)/\log(t)$.

Conversely, suppose that $X^*(x) \le \log(xt)/\log(t)$ for $0 \le x \le 1$. Then

$$E\Phi_t(X/2) \le \int_0^1 \Phi_t\left(\frac{\log(xt)}{2\log(t)}\right) dx = \frac{2t^{-1/2} - 1}{t^{-1} - 1} \le 1.$$

□

Proof of Theorem 5.1: Let us start with the proof that $S^*(t) \underset{t}{\approx} F_1(t)$. Since the random variables $X_n^{(\leq \ell(t))}$ are independent, we have that

$$F_1(t) = \inf\left\{\lambda > 0 : \mathbb{E}(t^{S^{(\leq \ell(t))}/\lambda}) \leq t^{-1} \text{ and } \mathbb{E}(t^{-S^{(\leq \ell(t))}/\lambda}) \leq t^{-1}\right\}.$$

Now we notice that for any random variable Y, and $0 \leq t \leq 1$, we have that

$$\tfrac{1}{2}\mathbb{E}(t^{-|Y|}) \leq \max\{\mathbb{E}(t^Y), \mathbb{E}(t^{-Y})\} \leq \mathbb{E}(t^{-|Y|}).$$

Hence

$$F_1(t) \leq \inf\left\{\lambda > 0 : \mathbb{E}(t^{-|S^{(\leq \ell(t))}/\lambda|}) \leq t^{-1}\right\} = \|S^{(\leq \ell(t))}\|_{\Phi_t},$$

and

$$F_1(t) \geq \inf\left\{\lambda > 0 : \mathbb{E}(t^{-|S^{(\leq \ell(t))}/\lambda|}) \leq 2t^{-1}\right\} = \|S^{(\leq \ell(t))}\|_{\Psi_t},$$

where $\Psi_t(x) = \dfrac{t^{-x} - 1}{2t^{-1} - 1}$. However, we quickly see that for $x \geq 0$ that if $t \leq 1/2$ then $\Psi_t(x) \geq \frac{1}{3}\Phi_t(x) \geq \Phi_t(x/3)$, since Φ_t is a convex function. Hence

$$F_1(t) \approx \|S^{(\leq \ell(t))}\|_{\Phi_t}$$

with constants of approximation bounded by 3.

Next, we apply Lemma 5.2, and we see that

$$F_1(t) \approx \sup_{0 \leq x \leq 1} \frac{\log(t)}{\log(xt)} (S^{(\leq \ell(t))})^*(x).$$

Taking $x = t$, we see that the right hand side is bounded below by $\frac{1}{2}(S^{(\leq \ell(t))})^*(t)$. Also, if $t \leq x \leq 1$, then

$$\frac{\log(t)}{\log(xt)} (S^{(\leq \ell(t))})^*(x) \leq (S^{(\leq \ell(t))})^*(t).$$

Further, by Corollary 3.2 combined with Proposition 4.1, there exist constants c_1 and c_2, depending only on the Lévy constants of (X_n), such that if $0 \leq x \leq t \leq c_1^{-1}$, then

$$\begin{aligned} \frac{\log(t)}{\log(xt)} (S^{(\leq \ell(t))})^*(x) &\leq c_2 \frac{\log(x)}{\log(xt)} \big((S^{(\leq \ell(t))})^*(t) + (M^{(\ell(t))})^*(x/2)\big) \\ &\leq c_2 \big((S^{(\leq \ell(t))})^*(t) + \ell(t)\big). \end{aligned}$$

Now, applying Proposition 2.1, Proposition 4.1, and Lemma 4.2 part (v), we finally obtain the desired result.

To show that $S^*(t) \underset{t}{\approx} F_2(t)$ is an almost identical proof. □

6 L_p norms

The main result of this section establishes the relationship between the L_p norm of sums of random variables and their tail distributions.

Theorem 6.1 *Given $p_0 > 0$, if $p \geq p_0$, and (X_n) is a sequence of Banach valued independent random variables, then*

$$\|U\|_p \approx U^*(e^{-p}/4) + \|\ell\|_p \approx (U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/4) + \|\ell\|_p,$$

where the constants of approximation depend only upon p_0.

We should note that we are not able to get universal control over the constants as $p_0 \to 0$, as is shown by simple examples once one understands that $\|Y\|_p$ converges to the geometric mean of $|Y|$ as $p \to 0$.

Combining this with Corollary 3.2, we immediately obtain the following result that compares $\|S\|_q$ to $\|S\|_p$. This result extends results of Talagrand, (see Ledoux and Talagrand (1991, Theorem 6.20), Kwapień and Woyczyński (1992, Proposition 1.4.2 and comments following it; see also Hitczenko (1994, Proposition 4.1)) and Johnson, Schechtman and Zinn (1983). If this result is specialized to symmetric or positive real valued random variables, then by considering the cases $p = 2$ or $p = 1$, it implies the inequality of Rosenthal (1970), including the result of Johnson, Schechtman and Zinn (1983) that gives correct order of the constants as $p \to \infty$. Note that $\|\ell\|_q \leq 2^{1/q}\|M\|_q$ by Proposition 2.1.

Theorem 6.2 *Let (X_n) be a sequence of Banach valued independent random variables and let $p_0 > 0$. Then there exist positive constants c_1, c_2 and c_3, depending only upon p_0, such that for $q \geq p \geq p_0$ we have*

$$\begin{aligned} \|U\|_q &\leq c_1 \frac{q}{\max\{p, \log(e+q)\}} \big(\|U\|_p + M^*(c_2^{-1}e^{-q})\big) + c_1\|M\|_q \\ &\leq c_3 \frac{q}{\max\{p, \log(e+q)\}} \big(\|U\|_p + \|M\|_q\big). \end{aligned}$$

Let us proceed with the proofs. First we need a lemma that allows us to deal with the "large" parts of U, so that they might be effectively considered as a sum of disjoint random variables.

Lemma 6.3 *Let (X_n) be a sequence of Banach valued independent random variables, and let $0 < r < 1$. Then we may express $U^{(>\ell(r))} = \sum_{k=1}^{\infty} V_k$, where the random variables V_k are disjoint, and $V_k^*(t) \leq k\ell\big(t(k-1)!/r^{k-1}\big)$.*

Proof: In proving this result, we may suppose without loss of generality that $X_n = X_n^{(>\ell(r))}$, that is, we may suppose that $\sum_n \Pr(X_n \neq 0) \leq r$.

If A is a finite subset of $\mathbb{N}$, define the event

$$E_A = \{X_n \neq 0 \text{ if and only if } n \in A\}.$$

For each positive integer k, let $E_k = \bigcup_{\substack{A\subseteq\mathbb{N}\\|A|=k}} E_A$. Set $V_k = UI_{E_k}$. Notice that if $|A| = k$, then

$$\begin{aligned}\Pr(UI_{E_A} > x) &\leq \sum_{n\in A} \Pr(|X_n| > x/k \text{ and } E_A)\\ &= \sum_{n\in A} \Pr(|X_n| > x/k) \prod_{m\in A\setminus\{n\}} \Pr(X_m \neq 0).\end{aligned}$$

Hence,

$$\begin{aligned}\Pr(V_k > x) &= \sum_{\substack{A\subseteq\mathbb{N}\\|A|=k}} \Pr(UI_{E_A} > x)\\ &\leq \sum_{i_1<\cdots<i_k} \sum_{j=1}^{k} \Pr(|X_{i_j}| > x/k) \prod_{\substack{l=1\\l\neq j}}^{k} \Pr(X_{i_l} \neq 0)\\ &\leq \frac{1}{k!} \sum_{i_1} \cdots \sum_{i_k} \sum_{j=1}^{k} \Pr(|X_{i_j}| > x/k) \prod_{\substack{l=1\\l\neq j}}^{k} \Pr(X_{i_l} \neq 0)\\ &= \frac{k}{k!} \sum_{i_1} \cdots \sum_{i_k} \Pr(|X_{i_1}| > x/k) \prod_{l=2}^{k} \Pr(X_{i_l} \neq 0)\\ &= \frac{k}{k!} \left(\sum_n \Pr(|X_n| > x/k)\right) \left(\sum_n \Pr(X_n \neq 0)\right)^{k-1}\\ &\leq \frac{r^{k-1}}{(k-1)!} \Pr(\ell > x/k).\end{aligned}$$

□

Corollary 6.4 *Let (X_n) be a sequence of Banach valued independent random variables, let $0 < r < 1$, and let $0 < p < \infty$. Then*

$$\|U^{(>\ell(r))}\|_p \leq 2e^{2^p r/p} \|\ell\|_p.$$

Proof: Apply Lemma 6.3 to obtain the V_k. Using the fact that $k \leq 2^k$, we obtain that

$$\|V_k\|_p^p \leq k^p \frac{r^{k-1}}{(k-1)!} \|\ell\|_p^p \leq 2^p \frac{(2^p r)^{k-1}}{(k-1)!} \|\ell\|_p^p.$$

Thus

$$\|U^{(>\ell(r))}\|_p^p = \sum_{k=1}^{\infty} \|V_k\|_p^p \leq 2^p \|\ell\|_p^p \sum_{k=1}^{\infty} \frac{(2^p r)^{k-1}}{(k-1)!} = 2^p e^{2^p r} \|\ell\|_p^p.$$

□

Proof of Theorem 6.1: Applying Proposition 2.1, we see that

$$\|U\|_p \geq \frac{1}{2}\|M\|_p \geq 2^{-1-1/p}\|\ell\|_p.$$

Also, we have that

$$\|U\|_p^p = \int_0^1 (U^*(t))^p\,dt \geq 8^{-1}e^{-p}(U^*(e^{-p}/8))^p,$$

that is, $\|U\|_p \geq 8^{-1/p}e^{-1}U^*(e^{-p}/8) \geq 8^{-1/p}e^{-1}U^*(e^{-p}/4)$. Hence we have shown that there exists a constant c_1, depending only upon p_0, such that

$$\|U\|_p \geq c_1^{-1}(U^*(e^{-p}/4) + \|\ell\|_p).$$

Furthermore, by Proposition 2.1,

$$\Pr(U \neq U^{(\leq \ell(e^{-p}/8))}) \leq \Pr(M > \ell(e^{-p}/8)) \leq \Pr(M > M^*(e^{-p}/8)) \leq e^{-p}/8.$$

Hence $U^*(e^{-p}/8) \geq (U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/4)$, and so we have shown that there is a constant $c_2 > 0$, depending only upon p_0, such that

$$\|U\|_p \geq c_2^{-1}((U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/4) + \|\ell\|_p).$$

Now let us derive the converse inequalities. Corollary 3.2 tells us that for $t \leq e^{-p}/2$ that

$$(U^{(\leq \ell(e^{-p}/8))})^*(t) \leq c_2\frac{\log(1/t)}{p+\log(2)}\big((U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) + \ell(e^{-p}/8)\big).$$

Thus

$$\begin{aligned}
\|U^{(\leq \ell(e^{-p}/8))}\|_p^p &\leq \int_0^{e^{-p}/2} ((U^{(\leq \ell(e^{-p}/8))})^*(t))^p\,dt + (1-e^{-p}/2)(U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) \\
&\leq \frac{1}{(p+\log(2))^p}\int_0^1 (\log(1/t))^p\,dt\,\big((U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) + \ell(e^{-p}/8)\big)^p \\
&\qquad +(U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) \\
&\leq c_3^p\big((U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) + \ell(e^{-p}/8)\big)^p,
\end{aligned}$$

where $c_3 > 0$ depends only upon p_0. Furthermore,

$$\ell(e^{-p}/8)^p \leq 8e^p\int_0^1 \ell(t)^p\,dt = 8e^p\|\ell\|_p^p.$$

Hence, applying Corollary 6.4, and the (quasi-)triangle inequality for L_p, we deduce that there exists a constant c_4, depending only upon p_0, such that

$$\|U\|_p \leq c_4((U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) + \|\ell\|_p).$$

Finally the result follows by noticing that

$$(U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) \leq (U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/4),$$

and also, by an argument similar to one presented above, that

$$(U^{(\leq \ell(e^{-p}/8))})^*(e^{-p}/2) \leq U^*(3e^{-p}/8) \leq U^*(e^{-p}/4).$$

□

Finally we remark that from the results mentioned at the end of Section 3 we can obtain one sided versions of Theorem 6.1 with $|S|$ in place of U, for example, given $p \geq p_0$,

$$\|S\|_p \leq cS^*(e^{-p}/c) + c\|\ell\|_p.$$

where the constants depend only upon p_0.

Obviously if the sequence of random variables satisfy the Lévy property, then we can obtain the two sided inequality, but otherwise the other side of the inequality need not hold, as is shown by the example $X_1 = 1$, $X_2 = -1$, $X_n = 0$ $(n \geq 2)$.

7 Rearrangement invariant spaces

Rearrangement invariant spaces are studied in much of the literature, see for example Lindenstrauss and Tzafriri (1977). However, we will work with a definition that is a little less restrictive. A *rearrangement invariant space* on the random variables is a quasi-normed Banach space $\mathcal{L}$ of random variables such that $1 \in \mathcal{L}$, and if $X^* \leq Y^*$ and $Y \in \mathcal{L}$, then $X \in \mathcal{L}$ and $\|X\|_{\mathcal{L}} \leq \|Y\|_{\mathcal{L}}$. Obviously the spaces L_p for $0 < p \leq \infty$ are rearrangement invariant spaces.

Given a rearrangement invariant space $\mathcal{L}$, we define the *quasi-constant* of $\mathcal{L}$ to be the least constant $K > 0$ such that $\|X + Y\|_{\mathcal{L}} \leq K(\|X\|_{\mathcal{L}} + \|Y\|_{\mathcal{L}})$ for all $X, Y \in \mathcal{L}$. Notice that if $X^*(2t) \leq Y^*(t)$, and $Y \in \mathcal{L}$, then X may be written as the sum of two disjoint random variables Y_1 and Y_2 with $Y_1^*(t), Y_2^*(t) \leq Y^*(t)$, and hence $\|X\|_{\mathcal{L}} \leq 2K\|Y\|_{\mathcal{L}}$.

Given two rearrangement invariant spaces $\mathcal{L}$ and $\mathcal{M}$, we will say that $\mathcal{L}$ embeds into $\mathcal{M}$ if there is a positive constant c such that if $X \in \mathcal{L}$, then $X \in \mathcal{M}$ and $\|X\|_{\mathcal{M}} \leq c\|X\|_{\mathcal{L}}$. We will call the least such c the *embedding constant* of $\mathcal{L}$ into $\mathcal{M}$.

Theorem 7.1 *Let $p_0 > 0$, and let $\mathcal{L}$ be a rearrangement invariant space such that $\mathcal{L}$ embeds into L_p, and L_q embeds into $\mathcal{L}$, where $q \geq p \geq p_0$. Then there is a positive constant c, depending only upon the quasi-constant of $\mathcal{L}$, the embedding constants, p_0 and q/p, such that for any sequence of Banach valued independent random variables (X_n)*

$$c^{-1}(\|U\|_p + \|\ell\|_{\mathcal{L}}) \leq \|U\|_{\mathcal{L}} \leq c(\|U\|_p + \|\ell\|_{\mathcal{L}}).$$

Proof: Let us first obtain the left hand side inequality. It follows by hypothesis that $\|U\|_{\mathcal{L}} \geq c_1^{-1}\|U\|_p$, where c_1 is the embedding constant of $\mathcal{L}$ into L_p. Furthermore, $U \geq \frac{1}{2}M$, and by Proposition 2.1, $\ell(t) \leq M^*(2t)$. Hence $\|U\|_{\mathcal{L}} \geq (4K)^{-1}\|\ell\|_{\mathcal{L}}$, where K is the quasi-constant of $\mathcal{L}$.

Now let us obtain the right hand inequality. By Corollary 3.2, we have that there is a universal positive c_2 for $0 \leq t \leq 1$

$$U^*(t)I_{0 \leq t \leq 2^{-2q/p}} \leq c_2 \frac{2q}{p}(U^*(t^{p/2q}) + M^*(t/2)).$$

Now $U^*(t)I_{0 \leq t \leq 2^{-2q/p}} \geq U^*(2^{2q/p}t)$, and hence

$$\|U\|_{\mathcal{L}} \leq (2K)^{\lceil 2q/p \rceil} c_2 K \frac{2q}{p}(\|t \mapsto U^*(t^{p/2q})\|_{\mathcal{L}} + \|M\|_{\mathcal{L}}).$$

To finish the proof, suppose that $\|U\|_p = \lambda$. Then it is easily seen that $U^*(t) \leq \lambda t^{-1/p}$. Thus, if c_3 is the embedding constant of L_q into $\mathcal{L}$, then

$$\begin{aligned}
\|t \mapsto U^*(t^{p/2q})\|_{\mathcal{L}} &\leq c_3 \|t \mapsto U^*(t^{p/2q})\|_q \\
&= c_3 \left(\int_0^1 (U^*(t^{p/2q}))^q \, dt \right)^{1/q} \\
&\leq c_3 \lambda \left(\int_0^1 t^{-1/2} \, dt \right)^{1/q} \\
&= 2^{1/q} c_3 \lambda.
\end{aligned}$$

□

References

[1] Billingsley, P. (1995) *Probability and Measure*, 3rd ed., Wiley & Sons, New York.

[2] Carothers, N.L., Dilworth, S.J. (1988) Inequalities for sums of independent random variables, *Proc. Amer. Math. Soc.* **104**, 221–226.

[3] Etemadi, N. (1985) Tail probabilities for sums of independent Banach space valued random variables, *Sankhyā Ser. A* **47**, 209–214.

[4] Figiel, T., Hitczenko, P., Johnson, W.B., Schechtman, G., Zinn, J. (1997) Extremal properties of Rademacher functions with applications to Khintchine and Rosenthal inequalities, *Trans. Amer. Math. Soc.*, **349**, 997–1027.

[5] Giné, E.; Zinn, J. (1983) Central limit theorems and weak laws of large numbers in certain Banach spaces. *Z. Wahrsch. Verw. Gebiete,* **62***, 323–354.*

[6] Gluskin, E.D., Kwapień, S. (1995) Tail and moment estimates for sums of independent random variables with logarithmically concave tails, *Studia Math.* **114**, 303–309.

[7] Hahn, M.G., Klass, M.J. (1997) Approximation of partial sums of arbitrary i.i.d. random variables and the precision of the usual exponential bound, *Ann. Probab.* **25**, 1457–1470.

[8] Hitczenko, P. (1993) Domination inequality for martingale transforms of a Rademacher sequence, *Israel J. Math.* **84**, 161–178.

[9] Hitczenko, P. (1994) On a domination of sums of random variables by sums of conditionally independent ones, *Ann. Probab.* **22**, 453–468.

[10] Hitczenko, P., Kwapień, S. (1994) On the Rademacher series, *Probability in Banach Spaces, Nine, Sandbjerg, Denmark*, J. Hoffmann–Jørgensen, J. Kuelbs, M.B. Marcus, Eds. Birkhäuser, Boston, 31–36.

[11] Hitczenko, P., Montgomery-Smith, S.J. (1999) A note on sums of independent random variables, *Advances in Stochastic Inequalities*, Ed.: T. Hill and C. Houdré, Contemporary Mathematics 234, A.M.S., Providence R.I., 69–73.

[12] Hitczenko, P., Montgomery-Smith, S.J., Oleszkiewicz, K. (1997) Moment inequalities for sums of certain independent symmetric random variables, *Studia Math.* **123**, 15–42.

[13] Hoffmann-Jørgensen, J. (1974) Sums of independent Banach space valued random variables. *Studia Math.* **52**, 159–186.

[14] Ibragimov, R., Sharakhmetov, Sh. (1997) On an exact constant for the Rosenthal inequality. *Teor. Veroyatnost. i Primen.* **42**, 341 – 350 (translation in *Theory Probab. Appl.* **42** (1997), 294 – 302 (1998)).

[15] Johnson, W.B., Maurey, B., Schechtman, G., Tzafriri, L. (1979) Symmetric structures in Banach spaces, *Mem. Amer. Math. Soc.* **217**.

[16] Johnson, W.B., Schechtman, G. (1989) Sums of independent random variables in rearrangement invariant function spaces, *Ann. Probab.* **17**, 789–808.

[17] Johnson, W.B., Schechtman, G., Zinn, J. (1983) Best constants in moment inequalities for linear combinations of independent and exchangeable random variables, *Ann. Probab.* **13**, 234–253.

[18] Kahane, J.-P. (1968) *Some Random Series of Functions*, 2nd ed., Cambridge University Press. Cambridge.

[19] Khintchine, A., (1923) Über dyadische Brüche, *Math. Z.* **18**, 109–116.

[20] Klass, M. J., Nowicki, K (1998) An improvement of Hoffmann-Jørgensen's inequality, with applications, preprint.

[21] Kwapień, S., Woyczyński, W.A. (1992) *Random Series and Stochastic Integrals. Single and Multiple*, Birkhäuser, Boston.

[22] Latała, R. (1993) A maximal inequality for sums of independent identically distributed random variables, *Warsaw University Preprint* .

[23] Latała, R. (1997) Estimation of moments of sums of independent random variables, *Ann. Probab.*, **25**, 1502–1513.

[24] Ledoux, M., Talagrand, M. (1991) *Probability in Banach Spaces*, Springer, Berlin, Heidelberg.

[25] Lindenstrauss, J., Tzafriri, L. (1977) *Classical Banach Spaces, II. Function Spaces*, Springer, Berlin, Heidelberg.

[26] Montgomery, H.L., Odlyzko, A.M. (1988) Large deviations of sums of independent random variables, *Acta Arithmetica* **49**, 427–434.

[27] Montgomery-Smith, S.J. (1990) The distribution of Rademacher sums. *Proc. Amer. Math. Soc.* **109**, 517–522.

[28] Montgomery-Smith, S.J. (1992) Comparison of Orlicz-Lorentz spaces. *Studia Math.* **103**, 161–189.

[29] Montgomery-Smith, S.J. (1993) Comparison of sums of independent identically distributed random variables, *Probab. Math. Statist.* **14**, 281–285.

[30] Pinelis, I. (1994), Optimum bounds for the distributions of martingales in Banach spaces, *Ann. Probab.* **22**, 1679–1706.

[31] Rosenthal, H.P. (1970) On the subspaces of L_p $(p > 2)$ spanned by sequences of independent symmetric random variables, *Israel J. Math.* **8**, 273–303.

[32] Shiryaev, A. N. (1980) *Probability*, Springer, Berlin, Heidelberg.

[33] Stout, W.F. (1974) *Almost Sure Convergence*, Academic Press, New York.