PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 10, Fasc. 2 (1989), p. 247-255

REARRANGEMENTS OF SEQUENCES OF RANDOM VARIABLES AND EXPONENTIAL INEQUALITIES

BY

BERNARD HEINKEL (STRASBOURG)

Abstract. Exponential bounds are studied for $P(\|X_1+\ldots+X_n\| > t)$, where $(X_1, \ldots, X_n)$ denotes a sequence of independent random variables with values in a real separable Banach space $(B, \| \|)$. In our results the usual boundedness assumptions on $\|X_1\|, \ldots, \|X_n\|$, are replaced by hypotheses on the weak l_p-norm of the sequence $(\|X_1\|, \ldots, \|X_n\|)$.

The following problem arises often in probability and statistics: "$X_1, \ldots, X_n$ being independent, centered and square integrable real-valued random variables (r.v.), having for sum S_n, how to bound for every positive x, the quantity $P(|S_n| > x)$?" A classical bound is of course given by Tchebichev's inequality

(0.1) $$P(|S_n| > x) \leq \operatorname{var}(S_n)/x^2,$$

but, if it is easy, that bound is usually not very sharp. Many authors have studied bounds in which the square function involved in (0.1) is replaced by a function f which is more efficient, at least for big values of x:

(0.2) $$P(|S_n| > x) \leq f(x).$$

The central-limit theorem and the strong law of large numbers have suggested to consider situations in which $f(x)$ is of the type $\exp(-ax^2)$ (Hoeffding's inequality [5], Bernstein's inequality [2]). The functions $\exp(-ax)$ and $\exp(-ax \operatorname{Log}(x+1))$ also appear in many efficient inequalities (Yurinskii's inequality [12] for big values of x, Bennett's inequality [1]). For such special functions f, inequalities like (0.2) are called *exponential inequalities*. These sharp inequalities of course don't apply to arbitrary r.v. X_k. Roughly speaking the r.v. X_k have to be relatively small either almost surely (if there exists a positive constant M bounding almost surely all the $|X_k|$) or in mean. For making more precise the kind of restrictions which have to be made on the r.v. X_k, we will state in a table some famous exponential inequalities. As above $(X_1, \ldots, X_n)$ will be a sequence of independent, centered and square integrable real-valued r.v., having for sum S_n.

We use the following notation: $s_n = \sqrt{E(S_n^2)}$; $\forall t > 0$, $P(|S_n| > t) = g(t)$.

The first column of the table gives the name of the inequality, the second one precises the restrictions made on the r.v. X_k, the third one the domain in which it applies, and the last one the bound for $g(t)$.

Bernstein [2]	$\lvert X_k\rvert \le M$ a.s.	$\left]0, \frac{s_n^2}{M}\right]$	$2\exp\left(\frac{-t^2}{2\left(s_n + \frac{tM}{2s_n}\right)^2}\right)$
Kolmogorov [6]	$\lvert X_k\rvert \le M s_n$ a.s.	$\left]0, \frac{s_n}{M}\right]$	$2\exp\left(-\frac{t^2}{2s_n^2}\left(1 - \frac{Mt}{2s_n}\right)\right)$
Prohorov [11]	$\lvert X_k\rvert \le M s_n$ a.s.	$]0, +\infty[$	$2\exp\left(-\frac{t}{2Ms_n}\operatorname{arc\,sin\,h}\left(\frac{tM}{2s_n}\right)\right)$
Hoeffding [5]	$a_k \le X_k \le b_k$ a.s.	$]0, +\infty[$	$2\exp\left(\frac{-2t^2}{\sum_{k=1}^n (b_k - a_k)^2}\right)$
Bennett [1]	$\forall r = 2, 3, \ldots$ $E\lvert X_k\rvert^r \le M^{r-2}E(X_k^2)$	$]0, +\infty[$	$2\exp\left(-\frac{t}{M}\left[\left(1 + \frac{s_n^2}{Mt}\right)\operatorname{Log}\left(1 + \frac{Mt}{s_n^2}\right) - 1\right]\right)$
Fuk-Nagaev [3]		$]0, +\infty[$	$\sum_{k=1}^n P(\lvert X_k\rvert > y_k) +$ $2\exp\left(\frac{t}{y} - \left(\frac{t-\mu}{y} + \frac{A}{y^2}\right)\operatorname{Log}\left(\frac{ty}{A} + 1\right)\right)$

Remarks. 1. In the above inequalities M denotes a positive constant.

2. In Fuk-Nagaev's result the following notation has been used: $(y_1, \ldots, y_n)$ denotes a sequence of positive numbers, y is such that $y \ge \sup(y_1, \ldots, y_n)$ and

$$A = \sum_{k=1}^n E(X_k^2 I_{(|X_k| \le y_k)}), \quad \mu = \sum_{k=1}^n E(X_k I_{(|X_k| \le y_k)}).$$

The second order integrability assumption made on the r.v. X_k are not needed in Fuk-Nagaev's result; but their inequality is sharp only if $\sum_{k=1}^n P(|X_k| > y_k)$ is small, which requires that the distribution functions of the r.v. X_k have a good tail behaviour.

3. Exponential inequalities have also been extended to vector valued r.v.; we will state a result of that kind, that we will need later:

PROPOSITION 0.1 (YURINSKII [12]). *Let $(B, \|\cdot\|)$ be a real separable Banach space, equipped with its Borel σ-field $\mathscr{B}$.* Let $X_1, \ldots, X_n$ *be independent $(B, \|\cdot\|)$-valued r.v., which are strongly square integrable and such that*

$$\exists M > 0: \forall k = 1, 2, \ldots, n, \forall m \geq 2, E\|X_k\|^m \leq \left(\frac{m!}{2}\right) M^{m-2} E\|X_k\|^2.$$

Then

$$\forall t > 0, P(\|S_n\| > t) \leq \exp\left(\frac{-\frac{t^2}{8} + \frac{t}{4} E\|S_n\|}{\sum_{k=1}^{n} E\|X_k\|^2 + \frac{Mt}{2}}\right),$$

where of course $S_n = X_1 + \ldots + X_n$.

The restrictions made on the r.v. X_k in the above exponential inequalities are of various kinds:

in Bernstein's or Hoeffding's result the r.v. are supposed to be small a.s.;

in Bennett's they are supposed to be small in mean;

in Bernstein's, Hoeffding's or Bennett's result the restrictions are made individually on every r.v.;

in Kolmogorov's, Prohorov's or Fuk-Nagaev's inequalities the restrictions involve the whole sequence of r.v., ...

The aim of the present work is to obtain exponential inequalities under a different kind of hypothesis on the r.v. X_k — which will be supposed $(B, \| \|)$ valued. The idea is to use information on the whole sequence $(\|X_1\|, \ldots, \|X_n\|)$, information under which each of the $\|X_k\|$ is relatively small. More precisely that information will be that the sequence $(\|X_1\|, \ldots, \|X_n\|)$ has a small weak l_p norm. Before to state the results we need to recall some facts on weak l_p spaces; this will be done in Section 1. In Section 2 we study exponential inequalities for sums of r.v. taking their values in a Banach space $(B, \| \|)$. In the appendix an efficient and simple exponential inequality in type r spaces will be derived from the method of proof used in Section 2.

1. Weak l_p spaces. Let $0 < p < +\infty$ be given and denote by $l_{p,\infty}$ the space of all sequences (a_n) of real numbers such that

$$\sup_{t>0} \left(t^p \operatorname{card}(n: |a_n| > t)\right) < +\infty.$$

That space $l_{p,\infty}$ is called the *weak l_p space.* Furthermore, let's define

$$\|(a_n)\|_{p,\infty} = \left(\sup_{t>0} \left(t^p \operatorname{card}(n: |a_n| > t)\right)\right)^{1/p}; \tag{1.1}$$

if $p > 1$, the functional $\| \|_{p,\infty}$ is equivalent to a norm on $l_{p,\infty}$ and $l_{p,\infty}$ equipped with that norm is a Banach space. In the sequel we will call the quantity $\|(a_n)\|_{p,\infty}$ the *weak l_p norm* of the sequence (a_n) and this for any value of p. It is obvious that a sequence (a_n) belonging to $l_{p,\infty}$ is also in c_0, so the non-increasing rearrangement (a_n^*) of $(|a_n|)$ can be defined without any problem; it is easy to check that

$$\|(a_n)\|_{p,\infty} = \sup(n^{1/p} a_n^*). \tag{1.2}$$

Before to state the result of Pisier, Rodin and Semyonov, which is the prototype for the exponential inequalities that we will prove in the next section, we need some more notations.

Let $q > 2$ be given and denote by ψ_q the following function:

$$\forall x \in \mathbf{R},\ \psi_q(x) = \exp|x|^q - 1.$$

For any probability space $(\Omega, \mathscr{F}, P)$ one denotes by $L^{\psi_q}(dP)$ the Orlicz space associated to ψ_q and P:

$$L^{\psi_q}(dP) = \{f:(\Omega, \mathscr{F}) \to (\mathbf{R}, B(\mathbf{R})): \exists c > 0: E\psi_q(|f|/c) < +\infty\};$$

that space will be equipped with the Luxemburg norm:

$$\|f\|_{\psi_q} = \inf(c > 0: E\psi_q(|f|/c) \leq 1).$$

The announced exponential inequality is as follows:

PROPOSITION 1.1 ([9]). *Let (α_n) belonging to $l_{p,\infty}$, $1 < p < 2$, be given. Consider (ε_n) a sequence of independent Rademacher r.v. (that is taking only the values $+1$ and -1, each with probability $1/2$) and define:*

$$S = \sum_{n \geqslant 1} \alpha_n \varepsilon_n.$$

Then $S \in L^{\psi_q}(dP)$, where $1/p + 1/q = 1$, and one has

$$(1/k_p)\|(\alpha_n)\|_{p,\infty} \leq \|S\|_{\psi_q} \leq k_p \|(\alpha_n)\|_{p,\infty}, \tag{1.3}$$

where k_p is a constant depending only on p.

Remark. Proposition 1.1 is not stated as an exponential inequality, but it contains such an inequality implicitly: it is easy to see that by definition of $\|\ \|_{\psi_q}$ one can derive the following bound from (1.3):

$$\forall t > 0,\ P(|S| > t) \leq \exp(1 - t^q/(k_p \|(\alpha_n)\|_{p,\infty})^q). \tag{1.4}$$

Exponential inequalities similar to (1.4), for scalar valued r.v. which are more general than weighted Rademacher ones, have been studied in [4]. The main result obtained in that paper is the following one:

PROPOSITION 1.2 ([4], Theorem 1.2). *Let $X_1, \ldots, X_n$ be independent, real-valued, symmetrically distributed and square integrable r.v. Denote by σ^2 the variance of their sum S_n. Then:*

$$\forall t > 0,\ P(|S_n| > t) \leq \inf_{c > 0} \left\{ P(\|(X_k)\|_{2,\infty} > c) + 2\exp\left(\frac{6\sigma t - t^2}{72\sigma^2 + 36c^2}\right) \right\}.$$

In fact, Propositions 1.1 and 1.2 above are only special cases of general results which apply in the more abstract setting of Banach space valued r.v. In the next section we will prove these general results, that we will call from now "weak l_p exponential inequalities".

2. Weak l_p exponential inequalities in Banach spaces. Let's first introduce some notations.

In the whole sequel $(B, \|\,\|)$ will be a real separable Banach space, equipped with its Borel σ-field $\mathscr{B}$. We will consider sequences $X = (X_1, \ldots, X_n)$ of independent r.v., defined on a probability space $(\Omega, \mathscr{F}, P)$ and taking their values in $(B, \mathscr{B})$; the sum of the terms of X will be denoted by $S(X)$. We will denote by $\Delta(X)$ the sum of the strong second moments of the X_k:

$$\Delta(X) = \sum E\|X_k\|^2. \tag{2.1}$$

Finally we define, for every $0 < p < +\infty$,

$$\varphi(X, p) = \| \|X_k\| \|_{p,\infty}. \tag{2.2}$$

In the symmetric case the tail of the distribution of $\|S(X)\|$ can be bounded in the following way:

THEOREM 2.1. *Let $p \in]1, +\infty[$ be given and let q be its conjugate. Consider $X = (X_1, \ldots, X_n)$ a sequence of B-valued r.v. which are independent, symmetrically distributed and strongly square integrable. Then*

$$\forall t > 0,\ \forall c > 0, \quad P(\|S(X)\|) > t \le P(\varphi(X, p) > c) + \tag{2.3}$$

$$+2\exp\left\{\frac{-\dfrac{t^2}{8(q+1)^2} + \dfrac{t}{4(q+1)} E\|S(X)\|}{\Delta(X) + \dfrac{c^q(q+1)^{q-2}}{2} t^{\frac{p-2}{p-1}}}\right\}.$$

Proof. Let $(\varepsilon_k)_{k \le n}$ be a sequence of independent Rademacher r.v. defined on $(\Omega', \mathscr{F}', P')$.

Let $c > 0$ be fixed and define:

$$A = (\omega \in \Omega\colon \varphi(X(\omega), p) \le c).$$

By symmetry the following inequality holds obviously for every $t > 0$:

$$\begin{aligned} P(\|S(X)\| > t) &= P \otimes P'(\|\sum \varepsilon_k X_k\| > t) \\ &\le P(A^c) + P \otimes P'(\|\sum \varepsilon_k X_k\| I_A > t). \end{aligned} \tag{2.4}$$

Let now $t = (q+1)cu$.

If $u \ge n^{1/q}$, it is clear that the second term in the righthand side part of the preceding inequality vanishes; so we suppose from now on that $u < n^{1/q}$.

Fix an $\omega \in A$ and define:

$$a(\omega) = P'(\|\sum \varepsilon_k X_k(\omega)\| > (q+1)cu).$$

There exists obviously a measurable function θ:

$$\theta\colon (\Omega \times \{1, 2, \ldots, n\}, \mathscr{F} \otimes \mathscr{P}(\{1, 2, \ldots, n\})) \to (\{1, 2, \ldots, n\}, \mathscr{P}(\{1, 2, \ldots, n\}))$$

such that, for every w, $(\|X_{\theta(w,k)}(w)\|)_{k \le n}$ is a non-increasing rearrangement of the sequence $(\|X_j(w)\|)_{j \le n}$. So

(2.5) $$\forall w \in A,\ \forall k = 1, \dots, n,\ \|X_{\theta(w,k)}(w)\| \le ck^{-1/p};$$

therefore

(2.6) $$a(\omega) \le P'\big(\|\sum_{k=[u^q]+1}^{n} \varepsilon_{\theta(\omega,k)} X_{\theta(\omega,k)}(\omega)\| > cu\big),$$

where [ ] stands for the integer part of a real number.

By using (2.5) and Lévy's inequality, it follows from (2.6) that

$$a(\omega) \le 2P'\big(\|\sum_{k=1}^{n} \varepsilon_k X_k(\omega) I_{(\|X_k\| \leqslant cu^{-1/(p-1)})}(\omega)\| > cu\big).$$

Integrating now the function a over the set A with respect to P, one obtains

(2.7) $$P\big(\|S(X)\| > (q+1)cu\big) \le P(A^c) + 2P\big(\|\sum_{k=1}^{n} X_k I_{(\|X_k\| \leqslant cu^{-1/(p-1)})}\| > cu\big).$$

Using finally Proposition 0.1 for bounding the last term in the righthand side of this inequality one easily obtains the claimed result by choosing $u = t/c(q+1)$ in (2.7).

Remarks. 1. Let's notice that the integrability assumptions made on the r.v. X_k are not needed in the proof of Theorem 2.1, because Yurinskii's inequality is applied to the truncated r.v. $\|X_k\| I_{(\|X_k\| \leqslant cu^{-1/(p-1)})}$. So it is possible to state Theorem 2.1 in a more sophisticated way in a spirit close to the one of Fuk-Nagaev's inequalities – without any integrability assumption on the r.v. X_k, by using truncated r.v. in the quantities $E\|S(X)\|$ and $\Delta(X)$ involved in (2.3). We have not stated Theorem 2.1 in that way because it gives a bound for $P(\|S(X)\| > t)$ which is too complicated; anyway (2.3) is efficient only if $P(\varphi(X, p) > c)$ is small, and this requires some integrability of the r.v. $\|X_k\|$.

2. Inequality (2.3) is efficient only if $P(\varphi(X, p) > c)$ is small. Does it exist an efficient bound for this probability? This question is answered positively by the following general result of Marcus and Pisier:

PROPOSITION 2.2 ([7] Theorem 3.3). *Let (Z_n) be a sequence of independent, positive r.v. Then, for any $0 < p < +\infty$ and all $c > 0$,*

$$c^p P\big(\|(Z_n)\|_{p,\infty} > c\big) \le 2e \sup_{t>0}\big(t^p \sum_{n=1}^{\infty} P(Z_n > t)\big).$$

In fact, in [7] this proposition is stated and proved with a constant 262 instead of $2e$; this constant has been reduced to $2e$ in a proof of Zinn which can be found in [10].

3. Let's now give an example of a situation in which Theorem 2.1 is more efficient than the classical inequalities that we recalled in the introduction.

Consider $X_1, \dots, X_n$ real-valued r.v. which are independent and such that

$$\forall k = 1, \ldots, n,\ P\left(X_k = \frac{1}{\sqrt{k}}\right) = P\left(X_k = -\frac{1}{\sqrt{k}}\right) = \frac{1}{2k},$$

$$P(X_k' = 0) = 1 - \frac{1}{k}.$$

Among the inequalities listed in our table, the only one which gives a bound in $\exp(-t^2)$ for $P(|S_n| > t)$ when t is large, is Hoeffding's result. More precisely,

$$P(|S_n| > t) \le 2\exp\left(\frac{-t^2}{2(1+\operatorname{Log} n)}\right).$$

Applying to this case Theorem 2.1, for $p = 2$, one gets

$$P(|S_n| > t) \le (2e)\exp\left(\frac{-t^2}{72+24\pi^2}\right);$$

for large values of n, this inequality is of course much better than the previous one.

4. It is easy to see that Theorem 2.1, applied in the scalar setting for $p = 2$, reduces to Proposition 1.2. A short computation gives also (1.4) as a corollary of Theorem 1.2. So the following question is natural: Does it exist — at least in some particular Banach spaces — an inequality analogous to (1.4)? In the next section we will see that a positive answer can be given to this question in type r spaces.

3. Appendix. The type r case. We will show that in type r spaces the same method of proof as for Theorem 2.1 gives an exponential inequality which is very similar to (1.4).

Recall that a Banach space $(B, \|\ \|)$ is said to be *of type* r, $1 < r \le 2$, if there exists a constant $K > 0$ such that for any finite sequence $(x_k)_{k \leqslant n}$ of elements of B one has

$$E\|\sum \varepsilon_k x_k\|^r \le K \sum \|x_k\|^r,$$

where $(\varepsilon_k)_{k \leqslant n}$ is a sequence of independent Rademacher r.v.

For instance, the space $L^r[0, 1]$, $1 < r \le 2$, is of type r.

In such spaces one has the following exponential inequality, which is more handy than that of Theorem 2.1:

THEOREM 3.1. *Let $1 < p < r \le 2$ be given and denote by q the conjugate of p: $1/p + 1/q = 1$. Suppose that $(B, \|\cdot\|)$ is a real separable Banach space of type r. Then there exist two positive constants $L(p)$ and $M(p, r, B)$ such that, for every sequence $X = (X_1, \ldots, X_n)$ of $(B, \|\cdot\|)$-valued r.v. which are independent and symmetrically distributed,*

$\forall t > 0$, $\forall c > 0$,

$$P(\|S(X)\| > t) \leq P(\varphi(X, p) > c) + M(p, r, B) \exp\left(-L(p)\left(\frac{t}{c}\right)^q\right).$$

Proof starts in the same way as for Theorem 2.1, till inequality (2.6):

$$a(\omega) \leq P'\left(\left\|\sum_{k=[u^q]+1}^{n} \varepsilon_{\theta(\omega,k)} X_{\theta(\omega,k)}(\omega)\right\| > cu\right). \tag{2.6}$$

According to (2.5) and the fact that B is of type r, one has

$$\mu(\omega) = \int \left\| \sum_{k=[u^q]+1}^{n} \varepsilon_{\theta(\omega,k)} X_{\theta(\omega,k)}(\omega) \right\| dP'$$

$$\leq K^{1/r} c \left(\sum_{k=[u^q]+1}^{n} \frac{1}{k^{r/p}}\right)^{1/r} \leq K^{1/r} c \left(\sum_{k=1}^{\infty} \frac{1}{k^{r/p}}\right)^{1/r} = c\alpha(p, r, K).$$

From this one easily deduces that

$$\sup\left(\frac{cu}{4}\mu(\omega) \times \frac{2}{c^2 u^{1-1/(p-1)}};\ 0 < u \leq 4\alpha(p, r, K)\right) = \beta(p, r, K).$$

For bounding the righthand side of (2.6), we apply now Proposition 0.1 with $M = cu^{-1/(p-1)}$:

$$a(\omega) \leq \exp\left(\beta(p, r, K) - \frac{c^2 u^2}{16 \sum_{k=[u^q]+1}^{n} \|X_{\theta(\omega,k)}(\omega)\|^2 + 8c^2 u^{(p-2)/(p-1)}}\right).$$

From (2.5) it easily follows that, for $u \geq 1$,

$$16 \sum_{k=[u^q]+1}^{n} \|X_{\theta(\omega,k)}(\omega)\|^2 + 8c^2 u^{(p-2)/(p-1)} \leq \gamma(p) c^2 u^{(p-2)/(p-1)};$$

so there exists a $\beta' = \beta'(p, r, K)$ such that, for every $u > 0$,

$$a(\omega) \leq \exp\left(\beta' - \frac{u^q}{\gamma(p)}\right).$$

The proof of Theorem 3.1 is then concluded in the same way as that of Theorem 2.1, by integrating a over A with respect to P and by choosing $u = t/c(q+1)$.

Several questions raise from the above weak l_p approach of the exponential inequalities. We will mention some of them as a conclusion to this paper:

In the proofs of Theorems 2.1 and 3.1, Yurinskii's bound is used as an ingredient; what kind of weak l_p exponential inequalities is it possible to obtain by replacing that ingredient by Bennett's or Nagaev's [8] results?

There exist exponential inequalities similar to Bernstein's result for martingales with almost sure small increments. Is it possible to improve these inequalities in the weak l_p setting?

Is it possible to study the asymptotic behaviour of trimmed sums of independent r.v. by using the same approach as for proving Theorem 2.1 or Theorem 3.1?

Acknowledgement. I am indebted to Michel Ledoux for several useful comments on the first draft of this paper.

REFERENCES

[1] G. Bennett, *Probability inequalities for the sum of independent random variables*, J. Amer. Stat. Assoc. 57 (1962), p. 33–45.

[2] S. N. Bernstein, *Sur une modification de l'inégalité de Tchebichev*, Ann. Sci. Inst. Sav. Ukraine, Sect. Math. I (1924).

[3] D. X. Fuk and S. V. Nagaev, *Probability inequalities for sums of independent random variables*, Th. Prob. Appl. 16 (1971), p. 643–660.

[4] B. Heinkel, *Some exponential inequalities with applications to the central-limit theorem in* $C(0, 1)$, Probability in Banach spaces 6, Sandbjerg, 1986, Progress in Probability, Birkhäuser (to appear).

[5] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Ann. Stat. Assoc. 58 (1963), p. 13–29.

[6] A. Kolmogorov, *Über das Gesetz des iterierten Logarithmus*, Math. Ann. 101 (1929), p. 126–135.

[7] M. B. Marcus and G. Pisier, *Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes*, Acta Math. 152 (1984), p. 245–301.

[8] S. V. Nagaev, *Large deviations of sums of independent random variables*, Ann. Prob. 7 (1979), p. 745–789.

[9] G. Pisier, *De nouvelles caractérisations des ensembles de Sidon*, Mathematical Analysis and Applications, Adv. in Math. Suppl. Stud. 7B (1981), p. 686–725.

[10] — *Probabilistic methods in the geometry of Banach spaces*, Probability and Analysis, Lecture Notes in Math. 1206, Springer, Berlin 1986, p. 167–241.

[11] Y. V. Prohorov, *Some remarks on the strong law of large numbers*, Theor. Prob. Appl. 4 (1959), p. 204–208.

[12] V. V. Yurinskii, *Exponential bounds for large deviations*, ibidem 19 (1974), p. 154–155.

Département de Mathématique
67084 Strasbourg Cedex
France

Received on 12. 10. 1987;
revised version on 21. 12. 1988