VU University Amsterdam
Faculty of Sciences

University of Warsaw
Faculty of Mathematics, Computer Science and Mechanics

Joint Master of Science Programme

Marcin Lis

Student no. 236089 (UW), 2005786 (VU)

Gaussian approximation of moments of sums of independent random variables

Master's thesis
in MATHEMATICS

Supervisor
Prof. dr hab. Rafał Latała
Institute of Mathematics
University of Warsaw

Under the co-supervision of
Prof. dr. R.W.J. Meester
Department of Mathematics
Vrije Universiteit Amsterdam

June 2010

Supervisor's statement

Hereby I confirm that the present thesis was prepared under my supervision and that it fulfils the requirements for the degree of Master of Mathematics.

Date Supervisor's signature

Author's statement

Hereby I declare that the present thesis was prepared by me and none of its contents was obtained by means that are against the law.

I also declare that the present thesis is a part of the Joint Master of Science Programme of the University of Warsaw and the Vrije Universiteit in Amsterdam. The thesis has never before been a subject of any procedure of obtaining an academic degree.

Moreover, I declare that the present version of the thesis is identical to the attached electronic version.

Date Author's signature

Abstract

We try to find good bounds for Gaussian approximation of moments of sums of independent random variables. We generalize some results of Latała from [7] in case when $2 \leq p \leq 4$ and present a combinatorial approach to this problem for even moments.

Keywords

Gaussian approximation, moments of random variables, sums of independent random variables

Thesis domain (Socrates-Erasmus subject area codes)

11.1 Mathematics

Subject classification

60 Probability theory and stochastic processes
60E Distribution theory
60E15 Inequalities; stochastic orderings
60F Limit theorems
60F05 Central limit and other weak theorems

Contents

Introduction

We will recall the standard formulation of the central limit theorem with the Lindeberg's condition. Let us begin with the definition of the useful notion of a scheme of series.

Definition 1. *A scheme of series* is an array of random variables of the form $(X_{n,i_n})_{n\geq 1}$, where $i_n \to \infty$ and $X_{n,1}, X_{n,2}, \ldots, X_{n,i_n}$ are independent for every n. We define $s_n^2 = \sum_{i=1}^{i_n} \mathrm{Var}(X_{n,i})$. We call a scheme of series *normalized* if $\mathbf{E}X_{n,i} = 0$ and $s_n^2 = 1$ for all n, i.

We also need to know how the Lindeberg's condition is defined.

Definition 2. We say that a scheme of series $(X_{n,i})$ satisfies *Lindeberg's condition of order* $p \geq 2$, if for every $r > 0$

$$L_n^p(r) = \frac{1}{s_n^p} \sum_{i=1}^{i_n} \mathbf{E}\big(|X_{n,i} - \mathbf{E}X_{n,i}|^p \mathbf{1}_{\{|X_{n,i} - \mathbf{E}X_{n,i}| > r s_n\}}\big) \overset{n\to\infty}{\longrightarrow} 0.$$

One can show that the Lindeberg's condition of order $p > 2$ holds if and only if $L_n^p = \frac{1}{s_n^p} \sum_{i=1}^{i_n} \mathbf{E}|X_{n,i} - \mathbf{E}X_{n,i}|^p \overset{n\to\infty}{\longrightarrow} 0$ for $p > 2$. We can now formulate the central limit theorem.

Theorem 1 (The Lindeberg-Levy Central Limit Theorem)**.** *Let $(X_{n,i})$ be a normalized scheme of series. If it satisfies the Lindeberg's condition of order 2, i.e. for every $r > 0$*

$$\lim_{n\to\infty} \sum_{i=1}^{i_n} \mathbf{E}X_{n,i}^2 \mathbf{1}_{\{|X_{n,i}| > r\}} = 0,$$

then

$$\sum_{i=1}^{i_n} X_{n,i} \overset{d}{\longrightarrow} \mathcal{N}(0,1).$$

If we suppose that there is no element of a row in $(X_{n,i})$ that dominates other elements then the assumptions in this theorem are the most general one can assume i.e. the following theorem holds.

Theorem 2 (Feller)**.** *If $(X_{n,i})$ is a normalized scheme of series and $\sum_{i=1}^{i_n} X_{n,i} \overset{d}{\longrightarrow} \mathcal{N}(0,1)$ and moreover*

$$\max_{i \leq i_n} \mathbf{E}X_{n,i}^2 \overset{n\to\infty}{\longrightarrow} 0,$$

then $(X_{n,i})$ satisfies the Lindeberg's condition of order 2.

Proofs of these two theorems can be found in [1]. Since convergence in distribution does not imply convergence of moments one can ask what are the additional assumptions to Theorem 1 which assure that $\big\| \sum_{i=1}^{i_n} X_{n,i} \big\|_p = \big(\mathbf{E}\big| \sum_{i=1}^{i_n} X_{n,i}\big|^p\big)^{1/p} \overset{n\to\infty}{\longrightarrow} \|g\|_p$, where $g \sim \mathcal{N}(0,1)$. The answer to this question was given by Bernstein in 1939 and can be expressed as the following theorem.

Theorem 3 (Bernstein)**.** *Let $(X_{n,i})$ be a normalized scheme of series such that $X_{n,i}$ are asymptotically negligible, i.e.* $\max_{i\leq i_n} \mathbf{P}(|X_{n,i}| \geq r) \stackrel{n\to\infty}{\longrightarrow} 0$ *for every $r > 0$ then for $p > 2$*

$$\Big\| \sum_{k=1}^{i_n} X_{n,i} \Big\|_p \stackrel{n\to\infty}{\longrightarrow} \gamma_p \text{ if and only if } L_n^p \stackrel{n\to\infty}{\longrightarrow} 0,$$

where $\gamma_p = \|g\|_p$ for $g \sim \mathcal{N}(0,1)$.

Remark. In the whole thesis we will denote by γ_p the p-th absolute moment of a normalized Gaussian variable which is equal to $2^{1/2}(\Gamma(\frac{p+1}{2})/\sqrt{\pi})^{1/p}$.

A proof of Theorem 3 based on characteristic functions is given in [2] and [3]. This theorem shows that with reasonable assumptions considering variables X_i moments of $\sum X_i$ can be approximated by corresponding Gaussian moments. One can ask how good such approximation can be. In [7] Latała gave precise estimations when X_i have logarithmically concave tails (i.e. the function $t \mapsto \ln(|X_i| \geq t)$ is concave from $[0,\infty)$ to $[-\infty, 0]$). This wide class of random distributions contains, among others, normal distributions $\mathcal{N}(0,\sigma^2)$, scaled Rademacher distributions (i.e. symmetric distributions concentrated in $\{-a, a\}$) and symmetric exponential distributions which have the density function of the form $\frac{\lambda}{2}\exp(-\lambda|x|)$ $(\lambda > 0)$. The main concept presented in [7] is to try to bound moments by removing summands with biggest variance, i.e. if $X_1, X_2, \ldots, X_n$ are independent symmetric random variables with variance 1 and $|a_1| \geq |a_2| \geq \ldots \geq |a_n|$ then we search for such a constant C_p depending on p and somehow on the distributions of X_i, for which we can prove that

$$\Big\| \sum_{i=C_p}^{n} a_i X_i \Big\|_p \leq \Big\| \sum_{i=1}^{n} a_i \varepsilon_i \Big\|_p.$$

From this inequality we almost immediately get the following approximation

$$\Big| \Big\| \sum_{i=1}^{n} a_i X_i \Big\|_p - \gamma_p \|a\|_2 \Big| \leq 2C_p \|a\|_\infty,$$

whose proof will be seen in this thesis at least twice. This proof is very simple but it uses a result from [4] which can be stated as the following theorem.

Theorem 4. *Let $X_1, X_2, \ldots, X_n$ be independent symmetric random variables with variance 1 and let $a_1, a_2, \ldots, a_n$ be a sequence of real numbers. Let $\varphi(t)$ be an Orlicz function (i.e. $\varphi : \mathbf{R} \to \mathbf{R}$ is an even, convex function which satisfies $\varphi(0) = 0$) such that φ is C^2 and $\varphi''(t)$ is convex. Then*

$$\mathbf{E}\varphi\Big(\sum_{i=1}^{n} a_i \varepsilon_i\Big) \leq \mathbf{E}\varphi\Big(\sum_{i=1}^{n} a_i X_i\Big),$$

where ε_i are independent Rademacher variables (i.e. $\mathbf{P}(\varepsilon_i = \pm 1) = 1/2$).

Proof. Let $\Phi_b(t) = \mathbf{E}\varphi(b + \sqrt{t}\varepsilon) = \frac{1}{2}(\varphi(b - \sqrt{t}) + \varphi(b + \sqrt{t}))$, $b \in \mathbf{R}$. We will prove that $\Phi_b(t)$ is convex on $[0, \infty)$. It is enough to prove that $\Phi_b'(t)$ is increasing on $(0,\infty)$. We have $\Phi_b'(t) = \frac{1}{4\sqrt{t}}(\varphi'(b + \sqrt{t}) - \varphi'(b - \sqrt{t}))$ and one only needs to show that $\Psi_b(t) = \frac{1}{t}(\varphi'(b+t) - \varphi'(b-t))$ is increasing on $(0, \infty)$. Thus one must show that $\Psi_b'(t) \geq 0$ and this is equivalent to showing that $t(\varphi''(b+t) + \varphi''(b-t)) \geq \varphi'(b+t) - \varphi'(b-t)$. We have

$$\varphi'(b+t) - \varphi'(b-t) = \int_0^t (\varphi''(b+s) + \varphi''(b-s))ds \leq t(\varphi''(b+t) + \varphi''(b-t)),$$

where in the inequality we used the bound $\varphi''(b+s)+\varphi''(b-s)\le\varphi''(b+t)+\varphi''(b-t)$ for $0\le s\le t$, which follows by the convexity of φ''. Let X be a symmetric random variable with variance 1. We have $\sigma X\sim\sqrt{\sigma^2X^2}\varepsilon$, where ε is a Rademacher variable independent of X. Now by Jensen's inequality we get for any b and σ

$$\mathbf{E}\varphi(b+\sigma X)=\mathbf{E}\varphi(b+\sqrt{\sigma^2X^2}\varepsilon)=\mathbf{E}\Phi_b(\sigma^2X^2)\ge\Phi_b(\mathbf{E}\sigma^2X^2)=\Phi_b(\sigma^2)=\mathbf{E}\varphi(b+\sigma\varepsilon).$$

Thus for every sequence $X_1,X_2,\ldots,X_n$ of independent, symmetric random variables satisfying $\mathbf{E}X_i^2=1$ we get

$$\begin{aligned}\mathbf{E}\varphi\Big(\sum_{i=1}^{n-1}\sigma_iX_i+\sigma_nX_n\Big)&\ge\mathbf{E}\varphi\Big(\sum_{i=1}^{n-1}\sigma_iX_i+\sigma_n\varepsilon_n\Big)=\mathbf{E}\varphi\Big(\sum_{i=1}^{n-2}\sigma_iX_i+\sigma_n\varepsilon_n+\sigma_{n-1}X_{n-1}\Big)\\&\ge\mathbf{E}\varphi\Big(\sum_{i=1}^{n-2}\sigma_iX_i+\sigma_n\varepsilon_n+\sigma_{n-1}\varepsilon_{n-1}\Big)\ge\ldots\ge\mathbf{E}\varphi\Big(\sum_{i=1}^{n}\sigma_i\varepsilon_i\Big).\end{aligned}$$

□

This theorem applies for $\varphi(t)=|t|^p$ when $p\ge3$ and guarantees good estimation of moments of sums of symmetric random variables from below. The problem left is to give equally good bounds from above. In this thesis we will generalize results of Latała in case when $2\le p\le4$ and in case when $p=2k$ for $k\in\mathbb{N}$ to all symmetric random variables and will give some estimates for nonsymmetric variables also in the case when $p=2k$.

Chapter 1

Gaussian approximation for moments of order $2 \le p \le 4$

We will try to imitate and develop proofs and methods used in [7]. We begin with a simple lemma about characteristic functions of random variables.

Lemma 1.
Let φ be a characteristic function of a symmetric distribution with variance 1 *and fourth moment less or equal M. Then for any x,*

$$1 - \frac{1}{2}x^2 \le \varphi(x) \le 1 - \frac{1}{2}x^2 + \frac{M}{4!}x^4.$$

Proof. Let X be a random variable with the characteristic function φ. Since X is symmetric we have $\varphi(t) = \mathbb{E}\cos(tX)$. To prove the lower estimate it is enough to use the well known inequality $\cos(x) \ge 1 - \frac{1}{2}x^2$. The upper estimate follows by a less popular bound which we will prove here, namely $\cos(x) \le 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4 = A(x)$. All functions occurring in this inequality are even and thus we can assume that $x \ge 0$. By Taylor's theorem we get $\cos(x) = A(x) + \frac{1}{5!}\cos^{(5)}(\xi)x^5 = A(x) - \frac{1}{5!}\sin(\xi)x^5$ for some $0 \le \xi \le x$. If $x \in [0, \pi]$ then $\sin(\xi) \ge 0$ and the desired estimate holds. If $x \in [\pi, \frac{3}{2}\pi]$ then $\cos(x) \le 0$. We can calculate the biggest root of $A(x)$ which is $x_0 = \sqrt{6 + 2\sqrt{3}}$. Furthermore $A(x) \overset{x\to\infty}{\longrightarrow} \infty$ and thus $A(x) \ge 0$ for $x \ge x_0$. One can easily check that $x_0 \le \pi$ and thus our inequality holds on $[\pi, \frac{3}{2}\pi]$. If $x \in [\frac{3}{2}\pi, \infty)$ then $\frac{x^4}{4!} \ge \frac{x^2}{2}$ and thus $A(x) \ge 1 \ge \cos(x)$. □

Next lemma is a generalization of Lemma 1 from [7].

Lemma 2.
Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be independent symmetric random variables with variance 1*. Let $|a_1| \ge |a_2| \ge \ldots \ge |a_n|$, $M_k = \mathbf{E}Y_k^4/6 < \infty$ and*

$$m = \min\left\{ i : \sum_{k=1}^{i} a_k^2 \ge \max_{k>i} a_k^2 M_k \right\}. \tag{1.1}$$

By $\varphi_1, \varphi_2, \ldots, \varphi_n$ and $\psi_1, \psi_2, \ldots, \psi_n$ we denote the characteristic functions of $X_1, X_2, \ldots, X_n$ and $Y_1, Y_1, \ldots, Y_n$. Then for any t,

$$\prod_{k=1}^{n} \varphi_k(a_k t) + \frac{1}{2}\sum_{k=1}^{m} a_k^2 t^2 \ge \prod_{k=m+1}^{n} \psi_k(a_k t). \tag{1.2}$$

Proof. We will consider 4 cases.

Case I $|a_1 t| \le \sqrt{2}$ and $|a_k t| \le \sqrt{2}/\sqrt{M_k}$ for $k \ge m+1$.

Let $x_k = a_k^2 t^2/2$. By Lemma 1 we can easily get that $|\psi_k(x)| \le 1 - \frac{x^2}{2} + \frac{M_k x^4}{4} \le 1$ for $0 \le x \le \min\{\sqrt{2}, \sqrt{2}/\sqrt{M_k}\}$. Since $\varphi_k(a_k t) \ge 1 - a_k^2 t^2/2 \ge 0$, to establish (1.2) it is enough to show that

$$\prod_{k=1}^{n}(1-x_k) + \sum_{k=1}^{m} x_k \ge \prod_{k=m+1}^{n}(1 - x_k + M_k x_k^2), \tag{1.3}$$

for $1 \ge x_1 \ge x_2 \ge \ldots \ge x_n \ge 0$ and $x_k \le 1/M_k$ for $k \ge m+1$. We will show this inequality by induction on n. The base case is when $n = m+1$, we have then to show the following inequality

$$\prod_{k=1}^{m+1}(1-x_k) \ge 1 - \sum_{k=1}^{m+1} x_k + M_{m+1} x_{m+1}^2. \tag{1.4}$$

From (1.1) we have that $M_{m+1} x_{m+1}^2 \le x_{m+1} \sum_{k=1}^{m} x_k$ and thus

$$\begin{aligned}\prod_{k=1}^{m+1}(1-x_k) &\ge (1-x_{m+1})(1 - \sum_{k=1}^{m} x_k) = 1 - \sum_{k=1}^{m+1} x_k + x_{m+1}\sum_{k=1}^{m} x_k \\ &\ge 1 - \sum_{k=1}^{m+1} x_k + M_{m+1} x_{m+1}^2.\end{aligned}$$

Assume (1.3) holds for $n \le l-1$, we will show that (1.3) holds for $n = l$.

$$\begin{aligned}\prod_{k=1}^{l}(1-x_k) + \sum_{k=1}^{m} x_k &= (1-x_l)\Big[\prod_{k=1}^{l-1}(1-x_k) + \sum_{k=1}^{m} x_k\Big] + x_l \sum_{k=1}^{m} x_k \\ &\ge (1-x_l)\prod_{k=m+1}^{l-1}(1 - x_k + M_k x_k^2) + x_l \sum_{k=1}^{m} x_k \\ &\ge (1-x_l)\prod_{k=m+1}^{l-1}(1 - x_k + M_k x_k^2) + M_l x_l^2 \prod_{k=m}^{l-1}(1 - x_k + M_k x_k^2) \\ &= \prod_{k=m+1}^{l}(1 - x_k + M_m x_k^2),\end{aligned}$$

where in the last inequality we used that $\prod_{k=m}^{l-1}(1 - x_k + M_k x_k^2) \le 1$ and that $M_l x_l^2 \le x_l \sum_{k=1}^{m} x_k$.

Case II $|a_1 t| \le \sqrt{2}$ and there is such $k_0 \ge m+1$ such that $|a_{k_0} t| \ge \sqrt{2}/\sqrt{M_{k_0}}$.

We have that $x_{k_0} = a_{k_0}^2 t^2/2 \ge 1/M_{k_0}$ and thus

$$\prod_{k=1}^{n}\varphi_k(a_k t) + \frac{1}{2}\sum_{k=1}^{m} a_k^2 t^2 \ge \frac{1}{2}\sum_{k=1}^{m} a_k^2 t^2 \ge M_{k_0} a_{k_0}^2 t^2/2 \ge 1 \ge \prod_{k=m+1}^{n}\psi_k(a_k t).$$

Case III $|a_1 t| \ge \sqrt{2}$ and $\prod_{k=1}^{n}\varphi_k(a_k t) < 0$.

There is such k_0 that $\varphi_{k_0}(a_{k_0} t) < 0$. Using Lemma 1 we get that $|\varphi_{k_0}(a_{k_0} t)| \le a_{k_0}^2 t^2/2 - 1 \le a_1^2 t^2/2 - 1$ and thus

$$\prod_{k=1}^{n}\varphi_k(a_k t) + \frac{1}{2}\sum_{k=1}^{m} a_k^2 t^2 \ge \prod_{k=1}^{n}\varphi_k(a_k t) + \frac{1}{2}a_1^2 t^2 \ge \frac{1}{2}a_1^2 t^2 - |\varphi_{k_0}(a_{k_0} t)| \ge 1 \ge \prod_{k=m+1}^{n}\psi_k(a_k t).$$

Case IV $|a_1 t| \ge \sqrt{2}$ and $\prod_{k=1}^{n}\varphi_k(a_k t) \ge 0$. Then (1.2) is obvious. □

Remark 1. The number m from the preceding lemma can be estimated from above as follows

$$m \leq \lceil \max_{1\leq k\leq n} \mathbf{E}Y_k^4/6\rceil .$$

Proof. Is is enough to use the fact that $|a_1| \geq |a_2| \geq \ldots \geq |a_n|$. □

Remark 2. If $Y_1, Y_2, \ldots, Y_n$ have logarithmically concave tails we have $m \leq 1$.

Proof. Let $\mathcal{E}$ be a variable with a symmetrical exponential distribution with variance 1. Then we have

$$m \leq \lceil \max_{1\leq k\leq n} \mathbf{E}Y_k^4/6\rceil \leq \lceil \mathbf{E}\mathcal{E}^4/6\rceil = 1,$$

where in the second inequality we used Proposition 1 from [7]. □

Now we will prove a lemma which corresponds to Lemma 2 from [7] and our proof will be very similar to that in [7].

Lemma 3. *Let $|a_1| \geq |a_2| \geq \ldots \geq |a_n|$ and $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n$ and m be as in Lemma 2. Then*

$$\mathbf{E}\Big|\sum_{k=1}^{n} a_k X_k\Big|^p \geq \mathbf{E}\Big|\sum_{k=m+1}^{n} a_k Y_k\Big|^p \qquad \text{for } 2 \leq p \leq 4.$$

Proof. Let $S_1 = \sum_{k=1}^n a_k X_k$ and $S_2 = \sum_{k=m+1}^n a_k Y_k$. We may of course assume that $2 < p < 4$. By Lemma 4.2 of [5] we have for any random variable X with finite fourth moment,

$$\mathbf{E}|X|^p = C_p \int_0^\infty \Big(\varphi_X(t) - 1 + \frac{1}{2}t^2\mathbf{E}|X|^2\Big)t^{-p-1},$$

where φ_X is the characteristic function of X and $C_p = -\frac{2}{\pi}\sin(\frac{p\pi}{2})\Gamma(p+1) > 0$. By Lemma 2 we have

$$\varphi_{S_1} - \varphi_{S_2} = \prod_{k=1}^{n} \varphi_k(a_k t) - \prod_{k=m+1}^{n} \psi_k(a_k t) \geq -\frac{1}{2}\sum_{k=1}^{m} a_k^2 t^2$$

and thus

$$\mathbf{E}|S_1|^p - \mathbf{E}|S_2|^p = C_p \int_0^\infty \Big(\varphi_{S_1} - \varphi_{S_2} + \frac{1}{2}\sum_{k=1}^{m} a_k^2 t^2\Big) \geq 0.$$

□

The next lemma will be useful in most of the proofs in this thesis.

Lemma 4. *Let $X_1, X_2, \ldots, X_n$ be symmetric independent random variables. If $p > 1$ and $\mathbb{E}|X_k|^p < \infty$ for $1 \leq k \leq n$ then*

$$\Big\|\sum_{k=1}^{n} a_k X_k\Big\|_p \leq \|a\|_\infty \Big\|\sum_{k=1}^{n} X_k\Big\|_p .$$

Proof. It is enough to prove this inequality for $n = 2$ and to do this it is sufficient to show that $\mathbb{E}|X + b_1|^p \le \mathbb{E}|X + b_2|^p$ for $|b_2| \ge |b_1|$ and X symmetric. Since $X \sim \varepsilon|X|$, where ε is a Rademacher variable independent of X, we have

$$\mathbb{E}|X + b|^p = \mathbb{E}\big|\varepsilon|X| + b\big|^p = \frac{\mathbb{E}\big||X| + b\big|^p + \mathbb{E}\big||X| - b\big|^p}{2}$$

and it is enough to prove that for $a \ge 0$ the function $f(x) = |a+x|^p + |a-x|^p$ is non-decreasing on $[0, \infty)$. We have $f'(x) = p(|a + x|^{p-1} - |a - x|^{p-1}) \ge 0$ which ends the proof. □

Next proposition and corollary give some estimations for moments of sums of independent random variables.

Proposition 1. *Let $X_1, X_2, \dots, X_n$ be independent symmetric random variables with variance 1. Let $|a_1| \ge |a_2| \ge \dots \ge |a_n|$, $M_k = \mathbf{E}X_k^4/6 < \infty$ and*

$$m = \min\Big\{i : \sum_{k=1}^{i} a_k^2 \ge \max_{k>i} a_k^2 M_k\Big\}.$$

Then

$$\gamma_p\Big(\sum_{k=2}^{n} a_k^2\Big)^{1/2} \le \Big\|\sum_{k=1}^{n} a_k X_k\Big\|_p \le \gamma_p\|a\|_2 + 2m\|a\|_\infty \quad \textit{for } 2 \le p \le 4. \tag{1.5}$$

Proof. Let $S = \sum_{k=1}^{n} a_k X_k$. The lower bound follows by Lemma 3 for $Y_k \sim \mathcal{N}(0, 1)$ and by Remark 1. To show the upper bound we use the triangle inequality, Lemma 4 and we apply Lemma 3 twice

$$\begin{aligned}\|S\|_p - \gamma_p\|a\|_2 &\le \|S\|_p - \Big\|\sum_{k=m+1}^{n} a_k X_k\Big\|_p \le \Big\|\sum_{k=1}^{m} a_k X_k\Big\|_p \le \|a\|_\infty\Big\|\sum_{k=1}^{m} X_k\Big\|_p \\ &\le \|a\|_\infty\Big\|\sum_{k=1}^{2m} \varepsilon_k\Big\|_p \le 2m\|a\|_\infty.\end{aligned}$$

□

Corollary 1. *Let $|a_1| \ge |a_2| \ge \dots \ge |a_n|$, $X_1, X_2, \dots, X_n$ and m be as in Proposition 1. Then*

$$\Big|\Big\|\sum_{k=1}^{n} a_k X_k\Big\|_p - \gamma_p\|a\|_2\Big| \le 2m\|a\|_\infty \quad \textit{for } 2 \le p \le 4.$$

Proof. The upper bound follows by Proposition 1. To prove the lower bound we use Lemma 3 and the lower bound from Proposition 1

$$\Big\|\sum_{k=1}^{n} a_k X_k\Big\|_p \ge \gamma_p\Big(\sum_{k=2}^{n} a_k^2\Big)^{1/2} \ge \gamma_p(\|a\|_2 - \|a\|_\infty) \ge \gamma_p\|a\|_2 - 2\|a\|_\infty.$$

□

These results allow us to improve some statements from [7]. Corollary 3 from [7] says that if X_k are independent, symmetric with logarithmically concave tails, $\mathbb{E}X_k^2 = 1$ and $p \geq 3$ then

$$\left| \left\| \sum_{k=1}^{n} a_k X_k \right\|_p - \gamma_p \|a\|_2 \right| \leq p \|a\|_\infty .$$

Corollary 2. *The estimation above is true for $p \geq 2$.*

Proof. It is enough to use Corollary 1 and Remark 2. □

The same applies to Theorem 2 from [7] which states that if X_k are like above, $|a_1| \geq |a_2| \geq \ldots \geq |a_n|$ and $p \geq 3$ then

$$\max\left\{ \gamma_p \Big(\sum_{k \geq \lceil p/2 \rceil} a_k^2 \Big)^{1/2}, \left\| \sum_{k<p} a_k X_k \right\|_p \right\} \leq \left\| \sum_{k=1}^{n} a_k X_k \right\|_p$$
$$\leq \gamma_p \Big(\sum_{k \geq \lceil p/2 \rceil} a_k^2 \Big)^{1/2} + \left\| \sum_{k<p} a_k X_k \right\|_p .$$

Corollary 3. *The inequalities above hold for $p \geq 2$.*

Proof. For $p = 2$ it is obvious. Let $2 < p \leq 3$. We have $\lceil p/2 \rceil = 2$ and the lower bound is a consequence of Proposition 1. The upper estimate can be obtained as follows

$$\left\| \sum_{k=1}^{n} a_k X_k \right\|_p \leq \left\| \sum_{k \leq 2} a_k X_k \right\|_p + \left\| \sum_{k>2} a_k X_k \right\|_p \leq \left\| \sum_{k<p} a_k X_k \right\|_p + \gamma_p \Big(\sum_{k \geq \lceil p/2 \rceil} a_k^2 \Big)^{1/2},$$

where the last inequality follows by Lemma 3 for $X_k \sim \mathcal{N}(0,1)$ and Remark 2. □

Chapter 2

A combinatorial approach to approximation of moments of order greater than 4

We will try to use combinatorial methods to give estimations for moments of sums of independent random variables where the degree is a natural even number and then to extend these results to all positive degrees greater than 4.

Notation. We will use the multi-index notation to simplify formulae appearing in our statements. An n-dimensional multi-index α is an n-tuple of non-negative integers $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}_0^n$. We define standard operations on n-dimensional multi-indices α, β

* $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$,
* $s\alpha = (s\alpha_1, s\alpha_2, \ldots, s\alpha_n)$ for $s \in \mathbb{N}_0$,
* $\alpha + \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$,
* $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!$,
* $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ for $x \in \mathbb{R}^n$.

We also introduce some other helpful functions

* $p_k(\alpha) = (\alpha_1, \alpha_2, \ldots, \alpha_k)$,
* $s(\alpha) = \{i : \alpha_i \neq 0\}$,
* $\mathrm{sing}(\alpha) = \{i : \alpha_i = 1\}$,
* $\mathrm{doub}(\alpha) = \{i : \alpha_i = 2\}$.

We call the set $\mathrm{sing}(\alpha)$ the set of single indicies of α. We say that an index i is single in a multi-index α if $i \in \mathrm{sing}(\alpha)$. For simpler notation we also write $\alpha = 0$ if $\alpha_i = 0$ for all i.

Such notation is very useful if we want to write the power of a sum in terms of powers of summands. This is called the multinomial theorem and can be written as follows

$$\left(\sum_{i=1}^{n} x_n\right)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha.$$

From now on all appearing multi-indicies will be n-dimensional and by $a = (a_1, a_2, \ldots, a_n)$ we denote an n-tuple of real numbers, such that $|a_1| \geq |a_2| \geq \ldots \geq |a_n|$. If A is a set we write $|A|$ to denote the cardinality of A. We begin with the case of symmetric random variables.

2.1. Moments of natural even order

2.1.1. Symmetric random variables

In this section we will use only one property of symmetric random variables, namely if X is symmetric and $m \in \mathbb{N}$ then $\mathbf{E}X^{2m+1} = 0$.

Lemma 5. *Let $X_1, X_2, \ldots, X_n$ be independent symmetric random variables with variance 1 and let $C \geq 1$ be such that*

$$\mathbb{E}X_i^{2l} \leq C^{2l}\frac{(2l)!}{2^l} \text{ for } l, k \in \mathbb{N},\ 2 \leq l \leq k,\ 1 \leq i \leq n,$$

then

$$\mathbb{E}\Big(\sum_{i=\lceil C^4(k-1)\rceil+1}^{n} a_iX_i\Big)^{2k} \leq \frac{(2k)!}{2^k}\sum_{\substack{|\alpha|=k \\ |s(\alpha)|=k}} a^{2\alpha}. \tag{2.1}$$

Proof. Let $D = \lceil C^4(k-1)\rceil + 1$. We have

$$\begin{aligned}
\mathbb{E}\Big(\sum_{i=D}^{n} a_iX_i\Big)^{2k} &= \sum_{\substack{|\alpha|=k \\ p_{D-1}(\alpha)=0}} \frac{(2k)!}{(2\alpha)!}a^{2\alpha}\mathbb{E}X_D^{2\alpha_D}\mathbb{E}X_{D+1}^{2\alpha_{D+1}}\cdots\mathbb{E}X_n^{2\alpha_n} \\
&\leq \sum_{\substack{|\alpha|=k \\ p_{D-1}(\alpha)=0}} \frac{(2k)!}{(2\alpha)!}a^{2\alpha}\frac{(2\alpha_D)!(2\alpha_{D+1})!\cdots(2\alpha_n)!}{2^{\alpha_D}2^{\alpha_{D+1}}\cdots 2^{\alpha_n}}C^{2(k-|\mathrm{sing}(\alpha)|)} \\
&= \frac{(2k)!}{2^k}\sum_{\substack{|\alpha|=k \\ p_{D-1}(\alpha)=0}} a^{2\alpha}C^{2(k-|\mathrm{sing}(\alpha)|)} \leq \frac{(2k)!}{2^k}\sum_{\substack{|\alpha|=k \\ p_{D-1}(\alpha)=0}} a^{2\alpha}C^{4(k-|s(\alpha)|)}.
\end{aligned}$$

The last inequality follows from a simple estimate $|\mathrm{sing}(\alpha)| \geq |s(\alpha)| - (|\alpha| - |s(\alpha)|) = 2|s(\alpha)| - k$. Thus it is enough to show

$$\sum_{\substack{|\alpha|=k \\ p_{D-1}(\alpha)=0}} a^{2\alpha}C^{4(k-|s(\alpha)|)} \leq \sum_{\substack{|\alpha|=k \\ |s(\alpha)|=k}} a^{2\alpha},$$

where we sum over n-dimensional indicies on both sides. If I is a nonempty subset of $\{D, D+1, \ldots, n\}$ and $|I| = i \leq k$ then there are exactly $\binom{k-1}{i-1} = \binom{k-1}{k-i}$ multi-indicies α on left-hand side which satisfy $s(\alpha) = I$ (because $\binom{k-1}{i-1}$ is the number of different ways one can put k indistinguishable balls in i distinguishable urns, without leaving any urn empty). If we look on the right-hand side and count these multi-indicies α for which $s(\alpha)\cap\{D, D+1, \ldots, n\} = I$, we get $\binom{D-1}{k-i}$, because we can add to these i fixed indicies any $k-i$ indicies from the set $\{1, 2, \ldots, D-1\}$. Of course any term corresponding to the selected multi-indicies on the right-hand side is bigger than any term from the chosen terms on the left-hand side (recall that $|a_1| \geq |a_2| \geq \ldots \geq |a_n|$). Thus it is enough to show that $\binom{D-1}{k-i} \geq C^{4(k-i)}\binom{k-1}{k-i}$, but we have

$$\begin{aligned}
\binom{D-1}{k-i}\Big/\binom{k-1}{k-i} &= \frac{(D-1)(D-2)\cdots(D-k+i)}{(k-1)(k-2)\cdots(k-(k-i))} \\
&= \frac{\lceil C^4(k-1)\rceil}{k-1}\,\frac{\lceil C^4(k-1)\rceil - 1}{k-2}\cdots\frac{\lceil C^4(k-1)\rceil - (k-i-1)}{i} \\
&\geq C^{4(k-i)}.
\end{aligned}$$

Since we can repeat this procedure for every $I \subseteq \{D, D+1, \dots, n\}$, (2.1) holds. $\square$

As an immediate consequence we get the following corollary.

Corollary 4. *Let $X_1, X_2, \dots, X_n$ and C be as in Lemma 5 then*

$$\mathbb{E}\Big(\sum_{i=\lceil C^4(k-1)\rceil+1}^{n} a_i X_i\Big)^{2k} \le \mathbb{E}\Big(\sum_{i=1}^{n} a_i \varepsilon_i\Big)^{2k}.$$

Proof. We have

$$\mathbb{E}\Big(\sum_{i=1}^{n} a_i \varepsilon_i\Big)^{2k} = \sum_{|\alpha|=k} \frac{(2k)!}{(2\alpha)!} a^{2\alpha} \mathbb{E}\varepsilon_1^{2\alpha_1} \mathbb{E}\varepsilon_2^{2\alpha_2} \cdots \mathbb{E}\varepsilon_n^{2\alpha_n} = \sum_{|\alpha|=k} \frac{(2k)!}{(2\alpha)!} a^{2\alpha} \ge \frac{(2k)!}{2^k} \sum_{\substack{|\alpha|=k \\ |s(\alpha)|=k}} a^{2\alpha}$$

$$\ge \mathbb{E}\Big(\sum_{i=\lceil C^4(k-1)\rceil+1}^{n} a_i X_i\Big)^{2k}.$$

where the last inequality follows by Lemma 5. $\square$

If $\mathcal{E}$ is a random variable with symmetric exponential distribution with variance 1 then $\mathbb{E}\mathcal{E}^{2l} = (2l)!/2^l$. Therefore if $X_1, X_2, \dots, X_n$ have logarithmically concave tails then $C = 1$ by Proposition 1 from [7] and we get a different proof of Theorem 1 from [7] for $p = 2k$. From the corollary above we can conclude the following results.

Corollary 5. *Let $X_1, X_2, \dots, X_n$ and C be as in Lemma 5 then*

$$\gamma_{2k}\Big(\sum_{i=k}^{n} a_i^2\Big)^{1/2} \le \Big\|\sum_{i=1}^{n} a_i X_i\Big\|_{2k} \le \gamma_{2k}\|a\|_2 + 2\lceil C^4(k-1)\rceil \|a\|_\infty.$$

Proof. The lower bound follows by Corollary 4 for $X_i \sim \mathcal{N}(0,1)$ and by the fact that $\|\sum_{i=1}^n a_i \varepsilon_i\|_{2k} \le \|\sum_{i=1}^n a_i X_i\|_{2k}$. Let $S = \sum_{i=1}^n a_i X_i$ and $D = \lceil C^4(k-1)\rceil$. We have

$$\|S\|_{2k} - \gamma_{2k}\|a\|_2 \le \|S\|_{2k} - \Big\|\sum_{i=D+1}^{n} a_i X_i\Big\|_{2k} \le \Big\|\sum_{i=1}^{D} a_i X_i\Big\|_{2k} \le \|a\|_\infty \Big\|\sum_{i=1}^{D} X_i\Big\|_{2k}$$

$$\le \|a\|_\infty \Big\|\sum_{i=1}^{2D} \varepsilon_i\Big\|_{2k} \le 2D\|a\|_\infty,$$

where we used the fact that $\gamma_{2k}\|a\|_2 \ge \|\sum_{i=1}^n a_i \varepsilon_i\|_{2k}$, the triangle inequality, Lemma 4 and Corollary 4 twice. $\square$

Corollary 6. *Let $X_1, X_2, \dots, X_n$ and C be as in Lemma 5 then*

$$\Big|\Big\|\sum_{i=1}^{n} a_i X_i\Big\|_{2k} - \gamma_{2k}\|a\|_2\Big| \le 2\lceil C^4(k-1)\rceil \|a\|_\infty.$$

Proof. The upper bound follows by Corollary 5. To prove the lower bound we use the lower estimate from Corollary 5, the triangle inequality, Lemma 4 and Corollary 4

$$\Big\|\sum_{i=1}^n a_iX_i\Big\|_{2k} \geq \gamma_{2k}\Big(\sum_{i=k}^n a_i^2\Big)^{1/2} \geq \gamma_{2k}\|a\|_2 - \gamma_{2k}\Big(\sum_{i=1}^{k-1} a_i^2\Big)^{1/2}$$
$$\geq \gamma_{2k}\|a\|_2 - \Big\|\sum_{i=1}^{2(k-1)} \varepsilon_i\Big\|_{2k}\|a\|_\infty \geq \gamma_{2k}\|a\|_2 - 2(k-1)\|a\|_\infty.$$

□

2.1.2. Random variables with mean 0

We begin with a lemma which is very similar to Lemma 5 but since our variables are no longer symmetric we get a weaker conclusion.

Lemma 6. *Let $X_1, X_2, \ldots, X_n$ be independent random variables with mean 0 and variance 1 and let $C \geq 1$ be such that*

$$|\mathbb{E}X_i^m| \leq C^m\frac{m!}{2^{m/2}} \text{ for } m,k \in \mathbb{N},\ 3 \leq m \leq 2k,\ 1 \leq i \leq n,$$

then

$$\mathbb{E}\Big(\sum_{i=\lceil C^6\frac{k(k-1)}{2}\rceil+1}^n a_iX_i\Big)^{2k} \leq \frac{(2k)!}{2^k}\sum_{\substack{|\alpha|=k\\|s(\alpha)|=k}} a^{2\alpha}. \tag{2.2}$$

Proof. Let $D = \lceil C^6\frac{k(k-1)}{2}\rceil + 1$. We have

$$\mathbb{E}\Big(\sum_{i=D}^n a_iX_i\Big)^{2k} = \sum_{\substack{|\alpha|=2k,\ \mathrm{sing}(\alpha)=\emptyset\\ p_{D-1}(\alpha)=0}} \frac{(2k)!}{\alpha!}a^\alpha\mathbb{E}X_D^{\alpha_D}\mathbb{E}X_{D+1}^{\alpha_{D+1}}\cdots\mathbb{E}X_n^{\alpha_n}$$
$$\leq \sum_{\substack{|\alpha|=2k,\ \mathrm{sing}(\alpha)=\emptyset\\ p_{D-1}(\alpha)=0}} \frac{(2k)!}{\alpha!}a^\alpha\frac{\alpha_D!\alpha_{D+1}!\cdots\alpha_n!}{2^{\alpha_D/2}2^{\alpha_{D+1}/2}\cdots2^{\alpha_n/2}}C^{2(k-|\mathrm{doub}(\alpha)|)}$$
$$= \frac{(2k)!}{2^k}\sum_{\substack{|\alpha|=2k,\ \mathrm{sing}(\alpha)=\emptyset\\ p_{D-1}(\alpha)=0}} a^\alpha C^{2(k-|\mathrm{doub}(\alpha)|)} \leq \frac{(2k)!}{2^k}\sum_{\substack{|\alpha|=2k,\ \mathrm{sing}(\alpha)=\emptyset\\ p_{D-1}(\alpha)=0}} a^\alpha C^{6(k-|s(\alpha)|)}.$$

The last inequality follows from the estimate $|\mathrm{doub}(\alpha)| \geq |s(\alpha)| - (|\alpha| - 2|s(\alpha)|) = 3|s(\alpha)| - 2k$ which holds for such α that $\mathrm{sing}(\alpha) = \emptyset$. Thus it is enough to show

$$\sum_{\substack{|\alpha|=2k,\ \mathrm{sing}(\alpha)=\emptyset\\ p_{D-1}(\alpha)=0}} a^\alpha C^{6(k-|s(\alpha)|)} \leq \sum_{\substack{|\alpha|=k\\|s(\alpha)|=k}} a^{2\alpha},$$

where we sum over n-dimensional indicies on both sides. Again if I is a nonempty subset of $\{D, D+1, \ldots, n\}$ and $|I| = i \leq k$ then there are exactly $\binom{2k-i-1}{i-1} = \binom{2k-i-1}{2(k-i)}$ multi-indicies α on left-hand side which satisfy $s(\alpha) = I$ (because $\binom{2k-i-1}{i-1}$ is the number of different ways one can put $2k$ indistinguishable balls in i distinguishable urns in such a way that in each urn there are at least two balls). If we look on the right-hand side and count these multi-indicies α for which $s(\alpha) \cap \{D, D+1, \dots, n\} = I$, we get $\binom{D-1}{k-i}$. We notice again that any term corresponding to the selected multi-indicies on the right-hand side is bigger than any term from the chosen terms on the left-hand side. Thus it is enough to show that $\binom{D-1}{k-i} \geq C^{6(k-i)}\binom{2k-i-1}{2(k-i)}$, but we have

$$\binom{D-1}{k-i} \Big/ \binom{2k-i-1}{2(k-i)} = \frac{(2(k-i))!}{(k-i)!} \frac{(D-1)(D-2)\cdots(D-(k-i))}{(2k-i-1)\cdots(i+1)i}$$
$$= \frac{(2(k-i))!}{(k-i)!} \frac{\lceil C^6k(k-1)/2\rceil - (k-i-1)}{i(2k-i-1)} \frac{\lceil C^6k(k-1)/2\rceil - (k-i-2)}{(i+1)(2k-i-2)} \cdots \frac{\lceil C^6k(k-1)/2\rceil}{(k-1)k}$$
$$\geq 2^{k-i}C^{6(k-i)} \frac{k(k-1)/2 - (k-i-1)}{i(2k-i-1)} \frac{k(k-1)/2 - (k-i-2)}{(i+1)(2k-i-2)} \cdots \frac{k(k-1)/2}{(k-1)k}$$
$$= C^{6(k-i)} \frac{k(k-1) - 2(k-i-1)}{i(2k-i-1)} \frac{k(k-1) - 2(k-i-2)}{(i+1)(2k-i-2)} \cdots \frac{k(k-1)}{(k-1)k},$$

and therefore it suffices to show that $k(k-1) - 2(k-j-1) \geq j(2k-j-1)$ for $1 \leq j \leq k-1$. By substituting j for $k-j$ we get a simpler inequality $j^2 - 3j + 2 \geq 0$, which is true for $j \in \mathbb{N}_+$. Since we can repeat this procedure for every $I \subseteq \{D, D+1, \dots, n\}$, (2.2) holds. □

We get as a consequence the following corollaries.

Corollary 7. *Let $X_1, X_2, \dots, X_n$ and C be as in Lemma 6 then*

$$\mathbb{E}\Big(\sum_{i=\lceil C^6 \frac{k(k-1)}{2}\rceil+1}^{n} a_iX_i\Big)^{2k} \leq \mathbb{E}\Big(\sum_{i=1}^{n} a_i\varepsilon_i\Big)^{2k}.$$

Corollary 8. *Let $X_1, X_2, \dots, X_n$ and C be as in Lemma 6 then*

$$\Big\|\sum_{i=1}^{n} a_iX_i\Big\|_{2k} \leq \gamma_{2k}\|a\|_2 + 2\Big\lceil C^6\frac{k(k-1)}{2}\Big\rceil\|a\|_\infty.$$

Proof. Let $S = \sum_{i=1}^n a_iX_i$ and $D = \lceil C^6 \frac{k(k-1)}{2}\rceil$. We have

$$\|S\|_{2k} - \gamma_{2k}\|a\|_2 \leq \|S\|_{2k} - \Big\|\sum_{i=D+1}^{n} a_iX_i\Big\|_{2k} \leq \Big\|\sum_{i=1}^{D} a_iX_i\Big\|_{2k}$$
$$\leq \Big\|\sum_{i=1}^{2D} a_i^\circ\varepsilon_i\Big\|_{2k} \leq \|a\|_\infty\Big\|\sum_{i=1}^{2D}\varepsilon_i\Big\|_{2k} \leq 2D\|a\|_\infty,$$

where $a_i^\circ = a_1$ for $i \leq D$ and $a_i^\circ = a_{i-D}$ for $D < i \leq 2D$. We used here the triangle inequality, Corollary 7 twice, Lemma 4 and the fact that $\gamma_{2k}\|a\|_2 \geq \|\sum_{i=1}^n a_i\varepsilon_i\|_{2k}$. □

We will now try to generalize results for natural even orders to all real orders bigger than 4.

2.2. An attempt to generalize previous results to arbitrary orders

Lemma 7. *Let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ be independent Rademacher variables, then for $k \geq 1$*

$$\frac{2k+1}{2k-1}\left(\mathbb{E}\Big(\sum_{i=1}^n a_i\varepsilon_i\Big)^{2k}\right)^2 \geq \frac{(2k+2)!}{2^{k+1}}\Bigg(\sum_{\substack{|\alpha|=k+1\\|s(\alpha)|=k+1}} a^{2\alpha}\Bigg)\mathbb{E}\Big(\sum_{i=1}^n a_i\varepsilon_i\Big)^{2k-2}. \tag{2.3}$$

Proof. We have

$$\begin{aligned}\left(\mathbb{E}\Big(\sum_{i=1}^n a_i\varepsilon_i\Big)^{2k}\right)^2 &= \left(\sum_{|\alpha|=k}\frac{(2k)!}{(2\alpha)!}a^{2\alpha}\cdot\mathbb{E}\varepsilon_1^{2\alpha_1}\mathbb{E}\varepsilon_2^{2\alpha_2}\cdots\mathbb{E}\varepsilon_n^{2\alpha_n}\right)^2\\ &= \left(\sum_{|\alpha|=k}\frac{(2k)!}{(2\alpha)!}a^{2\alpha}\right)^2 = (2k!)^2\sum_{|\alpha|=|\beta|=k}\frac{a^{2(\alpha+\beta)}}{(2\alpha)!(2\beta)!}\end{aligned}$$

and

$$\frac{(2k+2)!}{2^{k+1}}\Bigg(\sum_{\substack{|\alpha|=k+1\\|s(\alpha)|=k+1}} a^{2\alpha}\Bigg)\mathbb{E}\Big(\sum_{i=1}^n a_i\varepsilon_i\Big)^{2k-2} = (2k+2)!(2k-2)!\sum_{\substack{|\alpha|=k+1,|\beta|=k-1\\|s(\alpha)|=k+1}}\frac{a^{2(\alpha+\beta)}}{2^{k+1}(2\beta)!}.$$

We define two subsets of $\mathbb{N}_0^n \times \mathbb{N}_0^n$ as follows

$$\begin{aligned}L &= \{(\alpha,\beta) : |\alpha| = |\beta| = k\},\\ R &= \{(\gamma,\delta) : |\gamma| = k-1,\ |\delta| = k+1,\ |s(\delta)| = k+1\}.\end{aligned}$$

Since $\frac{2k+1}{2k-1}(2k)!^2 = \frac{k}{k+1}(2k+2)!(2k-2)!$ to prove (2.3) we need to show

$$\frac{k}{k+1}\sum_{(\alpha,\beta)\in L}\frac{a^{2(\alpha+\beta)}}{(2\alpha)!(2\beta)!} \geq \sum_{(\gamma,\delta)\in R}\frac{a^{2(\gamma+\delta)}}{2^{k+1}(2\gamma)!}. \tag{2.4}$$

To prove (2.4) we will divide R into disjoint subsets. For each such subset we will find a corresponding subset of L. We will make sure that these subsets are also disjoint and that the sums over corresponding subsets satisfy the desired inequality.
For $(\gamma,\delta) \in R$ we define

$$\begin{aligned}R(\gamma,\delta) = \{(\gamma',\delta') :\ &\gamma'+\delta' = \gamma+\delta,\\ &\gamma'_i = \gamma_i \text{ and } \delta'_i = \delta_i \text{ for } i \notin \operatorname{sing}(\gamma+\delta),\\ &|\gamma'| = k-1,\ |\delta'| = k+1,\\ &|s(\delta)| = k+1\} \subseteq R\end{aligned}$$

and

$$\begin{aligned}L(\gamma,\delta) = \{(\alpha,\beta) :\ &\alpha+\beta = \gamma+\delta,\\ &\alpha_i = \gamma_i \text{ and } \beta_i = \delta_i \text{ for } i \notin \operatorname{sing}(\gamma+\delta),\\ &|\gamma'| = |\delta'| = k\} \subseteq L.\end{aligned}$$

One can easily see that both families $\mathcal{R} = \{R(\gamma,\delta) : (\gamma,\delta) \in R\}$ and $\mathcal{L} = \{L(\gamma,\delta) : (\gamma,\delta) \in R\}$ are pairwise disjoint. Moreover $\mathcal{R}$ is a partition of R. It is also not difficult to notice that $R(\gamma,\delta) = R(\gamma',\delta')$ if and only if $L(\gamma,\delta) = L(\gamma',\delta')$. By the definition of $R(\gamma,\delta)$ and $L(\gamma,\delta)$ it follows also that the function $(\alpha,\beta) \mapsto \frac{a^{2(\alpha+\beta)}}{(2\alpha)!(2\beta)!}$ is constant on $R(\gamma,\delta)$ and $L(\gamma,\delta)$ and takes the same value on both sets. Thus to prove (2.6) it is enough to show

$$\frac{k}{k+1}|L(\gamma,\delta)| \geq |R(\gamma,\delta)|, \text{ for every } (\gamma,\delta) \in R. \tag{2.5}$$

Let us fix $(\gamma,\delta) \in R$. We have

$$s(\gamma) \setminus (\mathrm{sing}(\gamma) \cap \mathrm{sing}(\gamma+\delta)) \supseteq s(\delta) \setminus (\mathrm{sing}(\delta) \cap \mathrm{sing}(\gamma+\delta)),$$

since all indicies in δ are single and these indicies from δ which are not single in $\gamma+\delta$ must appear in γ. Thus we have

$$\begin{aligned} k-1-|\mathrm{sing}(\gamma)\cap\mathrm{sing}(\gamma+\delta)| &\geq |s(\gamma) \setminus (\mathrm{sing}(\gamma)\cap\mathrm{sing}(\gamma+\delta))| \\ &\geq |s(\delta) \setminus (\mathrm{sing}(\delta)\cap\mathrm{sing}(\gamma+\delta))| \\ &= k+1-|\mathrm{sing}(\delta)\cap\mathrm{sing}(\gamma+\delta)|, \end{aligned}$$

and therefore

$$|\mathrm{sing}(\delta)\cap\mathrm{sing}(\gamma+\delta)| \geq |\mathrm{sing}(\gamma)\cap\mathrm{sing}(\gamma+\delta)| + 2. \tag{2.6}$$

To simplify notation we define $g = |\mathrm{sing}(\gamma)\cap\mathrm{sing}(\gamma+\delta)|$ and $d = |\mathrm{sing}(\delta)\cap\mathrm{sing}(\gamma+\delta)|$. The idea behind the definition of the set $R(\gamma,\delta)$ is that it contains elements resulting from (γ,δ) by replacing the original sets $\mathrm{sing}(\gamma)\cap\mathrm{sing}(\gamma+\delta)$ and $\mathrm{sing}(\delta)\cap\mathrm{sing}(\gamma+\delta)$ by new subsets of $\mathrm{sing}(\gamma+\delta)$ while preserving the cardinality of these sets. Thus $|R(\gamma,\delta)| = \binom{d+g}{d}$. We can interpret $L(\gamma,\delta)$ similarly but this time we change cardinality of the sets of single indicies in a proper way and we get $|L(\gamma,\delta)| = \binom{d+g}{d-1}$. From (2.6) we get $|L(\gamma,\delta)| = \binom{d+g}{d-1} = \frac{d}{g+1}\binom{d+g}{d} \geq \frac{g+2}{g+1}\binom{d+g}{d} \geq \frac{k+1}{k}\binom{d+g}{d} = \frac{k+1}{k}|R(\gamma,\delta)|$, because $g+2 \leq d \leq k+1$, which ends the proof. □

Remark 3. One can see that the constant $\frac{2k+1}{2k-1}$ in the above lemma is optimal. If $a_1 = a_2 = \ldots = a_n = 1/\sqrt{n}$ then by Theorem 3 $\left\|\frac{\sum_{i=1}^n \varepsilon_i}{\sqrt{n}}\right\|_{2k}^{2k} \overset{n\to\infty}{\longrightarrow} \gamma_{2k}^{2k} = (2k-1)!!$. Thus we have that the left-hand side of (2.3) converges to $(2k+1)!!(2k-3)!!$ and we see that

$$\frac{(2k+2)!}{2^{k+1}} \sum_{|\alpha|=|s(\alpha)|=k+1} a^{2\alpha} = \frac{(2k+2)!}{2^{k+1}n^{k+1}}\binom{n}{k+1} \overset{n\to\infty}{\longrightarrow} (2k+1)!!,$$

and thus the right-hand side of (2.3) also goes to $(2k+1)!!(2k-3)!!$ as n approaches infinity.

Now we can prove an inequality about moments of arbitrary order which is weaker than previous results for even moments because of its non-unital constant.

Proposition 2. *Let $X_1, X_2, \ldots, X_n$ be independent random variables with variance 1 and mean 0 and let $C \geq 1$ be such that*

$$|\mathbb{E}X_i^m| \leq C^m \frac{m!}{2^{m/2}} \text{ for } m,k \in \mathbb{N},\ 3 \leq m \leq 2k,\ 1 \leq i \leq n,$$

then for $2 \leq p \leq 2k$

$$\frac{2\lfloor p/2\rfloor + 1}{2\lfloor p/2\rfloor - 1}\mathbb{E}\Big|\sum_{i=1}^n a_i\varepsilon_i\Big|^p \geq \mathbb{E}\Big|\sum_{i=C_p}^n a_iX_i\Big|^p, \tag{2.7}$$

where $C_p = \lceil C^4\lfloor p/2\rfloor\rceil + 1$ when X_i are symmetric and $C_p = \lceil C^{6\frac{\lfloor p/2\rfloor(\lfloor p/2\rfloor+1)}{2}}\rceil + 1$ otherwise.

Proof. We will use the fact that for every random variable X the function $p \mapsto \log \mathbb{E}|X|^p$ is convex. We define the following two functions

$$f(p) = \log \left[\mathbb{E}\Big| \sum_{i=1}^{n} a_i \varepsilon_i \Big|^p \right],$$

$$g(p) = \log \left[\frac{2\lfloor p/2 \rfloor - 1}{2\lfloor p/2 \rfloor + 1} \mathbb{E}\Big| \sum_{i=C_p}^{n} a_i X_i \Big|^p \right].$$

$f(p)$ is convex on $(0, \infty)$ and $g(p)$ is convex on $(2l, 2l+2)$ for every $1 \le l \le k$, since $\lfloor p/2 \rfloor$ is constant on such intervals. Let us fix l. By convexity of $g(p)$ on $(2l, 2l+2)$ we can identify this function with its continuous extension to the closed interval $[2l, 2l+2]$. We will show that (2.7) holds on $[2l, 2l+2]$. Let $p \in [2l, 2l+2]$. Since $f(p)$ is convex we have

$$f(p) \ge f(2l) + \frac{p - 2l}{2} \big(f(2l) - f(2l-2) \big) =: \underline{f}(p)$$

and since $g(p)$ is convex on $[2l, 2l+2]$ we get

$$g(p) \le g(2l) + \frac{p - 2l}{2} \big(g(2l+2) - g(2l) \big) =: \overline{g}(p)$$

Therefore it is enough to show that $\underline{f}(p) \ge \overline{g}(p)$. By Corollaries 4 and 7 we have $\underline{f}(2l) = f(2l) \ge g(2l) = \overline{g}(2l)$ and because of linearity of both $\underline{f}(p)$ and $\overline{g}(p)$ we only need to show that $\underline{f}(2l+2) = 2 \log \mathbb{E}| \sum_{i=1}^{n} a_i \varepsilon_i|^{2l} - \log \mathbb{E}| \sum_{i=1}^{n} a_i \varepsilon_i|^{2l-2} \ge \log \frac{2l-1}{2l+1} \mathbb{E}| \sum_{i=C_{2l}}^{n} a_i X_i|^{2l+2} = \overline{g}(2l+2)$. That follows by Corollary 5, Corollary 7 and Lemma 7. $\square$

Bibliography

[1] Billingsley, P. (1986). *Probability and measure, 2nd ed.*, John Wiley & Sons.

[2] Brown, B. M. (1969). *Moments of a stopping rule related to the central limit theorem*, Ann. Math. Statist. 40, 1236–1249.

[3] Brown, B. M. (1970). *Characteristic functions, moments and the central limit theorem*, Ann. Math. Statist. 41, 658–664.

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

[5] Haagerup, U. (1982). *The best constants in the Khintchine inequality*, Studia Math. 70, 231–283.

[6] Hall, P. (1978). *On the rate of convergence of moments in the central limit theorem*, J. Austral. Math. Soc. 25 (Series A), 250–256.

[7] Latała, R. (2009). *Gaussian approximation of moments of sums of independent symmetric random variables with logarithmically concave tails*, IMS Collections 5 , High Dimensional Probability V: The Luminy Volume, 37–42.

[8] von Bahr, B. (1965). *On the convergence of moments in the central limit theorem*, Ann. Math. Statist. 36, 808–818.