Lecture 4, Tuesday afternoon

Notational note: $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ $\mathbb{E}_\varepsilon$
$G = (G_1, \dots, G_n)$ $\mathbb{E}_G$ independent

Chernoff; Crámer; Hoeffding; Bennett

$\underbrace{}_{\text{exponential methods}}$

* for coresets maybe get better bds?

$S = \sum_{i=1}^n a_i r_i$, $a_i \in \mathbb{R}$

$\forall t > 0$, $\underset{\substack{\text{Laplace transform} \\ \text{Moment Gen fcn}}}{\mathbb{E} e^{tS}} \le e^{t^2 (\sum a_i^2)/2}$

Chebyshev
$\Rightarrow$ $\forall u > 0$, $\mathbb{P}(S > u) \le e^{t^2 \sum a_i^2/2} e^{-tu}$
Markov

$\hookrightarrow \mathbb{P}(S > u) = \mathbb{P}(e^{tS} > e^{tu}) \le e^{-tu} \mathbb{E} e^{tS}$

min $\frac{t^2 \sum a_i^2}{2} - tu$ over $t > 0 \Rightarrow t_{opt} = \frac{u}{\sum a_i^2}$

So $\mathbb{P}(S > u) \le e^{-u^2/(2\sum a_i^2)}$ ← Gaussian tail bd.

Rademacher symmetric $\Rightarrow \mathbb{P}(|S| > u) \le 2 e^{-\frac{u^2}{2\sum a_i^2}}$

$\mathbb{E} e^{tS} = \mathbb{E}[e^{ta_1 r_1} \cdot e^{t a_2 r_2} \cdot \dots \cdot e^{t a_n r_n}]$

$= \cosh(a_1 t) \cdots \cosh(a_n t)$

$\forall s \in \mathbb{R}$, $0 \le \underset{= 1 + \frac{s^2}{2} + \frac{s^4}{4!} + \dots}{\cosh s} \overset{*}{\le} e^{s^2/2} = 1 + (s^2/2) + \frac{(s^2/2)^2}{2!} + \dots$

Term by term: $\frac{s^{2k}}{(2k)!} = \frac{(s^2/2)^k}{k!}$

$\Updownarrow$

$(2k)!! \le (2k)!$

$\rightarrow \cosh x \le e^{x^2/2}$
is very loose for large x

Azuma's ≠: (James Lee) $X_1, \dots, X_n$ bdd (a.s.) RV defined on the same probability space and s.t. $\forall i_1 < i_2 < \dots < i_k$ $\mathbb{E}[X_{i_1} \cdots X_{i_k}] = 0$

(ex: $(M_i, F_i)_{i=0}^n$ - martingale $\Rightarrow X_i = M_i - M_{i-1}$ satisfies this condition.

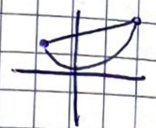
$$\mathbb{E}[X_{i_1}\cdots X_{i_k}] = \mathbb{E}\left[\mathbb{E}[X_{i_1}\cdots X_{i_{k-1}} X_{i_k} \mid \mathcal{F}_{i_{k-1}}]\right]$$

$$= \mathbb{E}\Big[X_{i_1}\cdots X_{i_{k-1}} \cdot \underbrace{\mathbb{E}[X_{i_k} \mid \mathcal{F}_{i_{k-1}}]}_{0}\Big] = 0$$

↳ This is a common ex. but not the only ex!

$\forall i$, $\|X_i\| \le d_i$. Then:

$$\mathbb{E}\, e^{t\sum_{i=1}^n X_i} \overset{?}{\le} e^{t^2 \sum_i d_i^2/2}$$



$s \mapsto e^s$ is convex $\Rightarrow$ $0 \le e^{ax} \underset{x=\pm 1}{\overset{=}{\le}} \frac{e^a + e^{-a}}{2} + x\frac{e^a - e^{-a}}{2}$

$$\mathbb{E}\, e^{t\sum_{i=1}^n X_i} = \mathbb{E}[e^{tX_1}\cdots e^{tX_n}] = \mathbb{E}[e^{\frac{tX_1}{d_1}d_1} \cdots e^{\frac{tX_n}{d_n}d_n}]$$

$$\le \mathbb{E}\Big[\prod_{i=1}^n \big(\cosh(td_i) + \frac{X_i}{d_i}\sinh(td_i)\big)\Big]$$

$$\overset{\text{by assumption}}{=} \prod_{i=1}^n \cosh(td_i) \le \prod_{i=1}^n e^{t^2 d_i^2/2}. \quad \square$$

$$S = \sum_{i=1}^n a_i r_i$$

We choose normalization s.t. $\sum_{i=1}^n a_i^2 = 1 \iff \|S\|_2 = 1$

We show: $\|S\|_p \lesssim \sqrt{p/e}$ as $p \to \infty$.

Using (stronger) CLT: $\frac{\|G\|_p}{\sqrt{p}} \le \limsup_{p\to\infty} \frac{C_{p,2}}{\sqrt{p}} \le 1/\sqrt{e}$

$\frac{\|G\|_p}{\sqrt{p}} \to \sqrt{1/e}$ as $p \to \infty$.

Assumption $\|S\|_2 = 1$

$\Rightarrow \forall t > 0$, $\mathbb{E}\, e^{tS} \le e^{t^2/2}$

$$\mathbb{E}\, e^{tS} = \frac{1}{2}\mathbb{E}\, e^{t|S|} + \frac{1}{2}\mathbb{E}\, e^{-t|S|} \quad \text{\{ symmetry}$$

$$\ge \frac{1}{2}\mathbb{E}\, e^{t|S|}$$

Take $t = \sqrt{p}$: Then

take p outside so that mass lies where bd is tight

$$2e^{p/2} \ge \mathbb{E}[e^{\sqrt{p}|S|}] = \mathbb{E}[(e^{|S|/\sqrt{p}})^p] = e^p\, \mathbb{E}[(e^{\frac{|S|}{\sqrt{p}} - 1})^p]$$

$$\ge e^p \cdot \mathbb{E}\left(\frac{|S|}{\sqrt{p}}\right)^p$$

$* \ \forall x \ge 0$: $e^{x-1} \ge x \ge 0$, $e^{p(x-1)} \ge x^p \ge 0$.

So: $2^{1/p} \to 1$

$$\frac{2^{1/p}}{\sqrt{e}}\sqrt{p} \ge \|S\|_p.$$

Or: $\forall u > 0$ $\mathbb{P}(|S| > u) \le 2e^{-u^2/2}$

$$\Rightarrow \mathbb{E}|S|^p = p\int_0^\infty u^{p-1}\,\mathbb{P}(|S| > u)\,du$$

$$\le 2p\int_0^\infty u^{p-1} e^{-u^2/2}\,du = 2p\int_0^\infty (2s)^{\frac{p-2}{2}} e^{-s}\,ds$$

$$= 2^{p/2}\, p\, \Gamma(p/2) \quad \Leftarrow \text{Stirling's formula}$$

$$\Rightarrow \|S\|_p \lesssim \sqrt{p/e}$$

There are ways of using exp. method to understand asymptotics of moments.

↳ In second, don't need Chebyshev!

$\forall t > 0, p > 0$: $\mathbb{P}(\|S\| > t) \le \frac{\mathbb{E}\|S\|^p}{t^p} = \left(\frac{\|S\|_p}{t}\right)^p$ Chebyshev/Markov

$\forall u > 0, p > 0$: $\mathbb{P}(\|S\| > u\|S\|_p) \le u^{-p}$

↳ If you know something about moments, can get. p.

The Paley-Zigmund $\neq$: $\alpha \in (0,1)$ ✓ s2lt prob space

$$\mathbb{E}|Z| = \mathbb{E}|Z|\,\mathbb{1}_{|Z| > \alpha\mathbb{E}|Z|} + \underbrace{\mathbb{E}|Z|\,\mathbb{1}_{|Z| \le \alpha\mathbb{E}|Z|}}$$

Schwarz: $\mathbb{E}|Z|\,\mathbb{1}_{|Z| > \alpha\mathbb{E}|Z|} \le \sqrt{\mathbb{E}|Z|^2}\sqrt{\mathbb{E}\mathbb{1}^2_{|Z| > \alpha\mathbb{E}|Z|}} = \sqrt{\mathbb{E}Z^2}\sqrt{\mathbb{P}(|Z| > \alpha\mathbb{E}|Z|)}$; $\quad \mathbb{E}|Z|\,\mathbb{1}_{|Z| \le \alpha\mathbb{E}|Z|} \le \alpha\mathbb{E}|Z|$

$$(1-\alpha^2)(\mathbb{E}|Z|)^2 \le \mathbb{E}Z^2 \cdot \mathbb{P}(|Z| > \alpha\mathbb{E}|Z|)$$

$$\Rightarrow \mathbb{P}(|Z| > \alpha\mathbb{E}|Z|) \ge (1-\alpha)^2 \frac{(\mathbb{E}|Z|)^2}{\mathbb{E}Z^2}$$

$$(*) \Rightarrow \mathbb{P}(|Z| > \tfrac{1}{2}\mathbb{E}|Z|) \ge \frac{(\mathbb{E}|Z|)^2}{4\mathbb{E}Z^2} \quad \text{or } Z \equiv 0 \text{ a.s.}$$

Taking $Z = \|S\|^p = \|\sum_{i=1}^{n} r_i v_i\|^p$, Then

$$\mathbb{P}\left(\|S\|^p > \tfrac{1}{2}\mathbb{E}\|S\|^p\right) \ge \frac{(\mathbb{E}\|S\|^p)^2}{4\mathbb{E}\|S\|^{2p}} = \frac{1}{4}\left(\frac{\|S\|_p}{\|S\|_{2p}}\right)^{2p}$$

$$\mathbb{P}\left(\|S\| > \tfrac{1}{2^{1/p}}\|S\|_p\right) \ge \tfrac{1}{4}3^{-p}$$

Hypercontractive bound: $\forall p > q > 1$,

$$\|S\|_p \le \sqrt{\frac{p-1}{q-1}}\,\|S\|_q$$

so $\dfrac{\|S\|_{2p}}{\|S\|_p} \le \dfrac{\sqrt{2p-1}}{\sqrt{p-1}} \le \sqrt{3}$

↳ For Rademacher sums: very close relation btwn tail decay and moments.
→ True any time you know moments don't grow too quickly!
Moments $\Longleftrightarrow$ tails

Paley-Zygmund (version 2.0) $\forall x \in \mathbb{R}$: $|x| = x_+ + x_-$, $x = x_+ - x_-$

Assume $\mathbb{E}[Z] = 0$
$= \mathbb{E}(Z_+ - Z_-) \Rightarrow \mathbb{E}Z_+ = \mathbb{E}Z_-$

$\Rightarrow \mathbb{E}|Z| = \mathbb{E}(Z_+ + Z_-) = 2\mathbb{E}Z_+$

$= 2\mathbb{E}Z\mathbb{1}_{Z>0} \le 2\sqrt{\mathbb{E}Z^2}\sqrt{\mathbb{E}\mathbb{1}^2_{Z>0}} = 2\sqrt{\mathbb{E}Z^2}\sqrt{P(Z>0)}$

So: $P(Z>0) \ge \frac{1}{4}\frac{(\mathbb{E}|Z|)^2}{\mathbb{E}Z^2} \ge \alpha(C) > 0$.

* Recall: $\mathbb{E}Z^4 \le C(\mathbb{E}Z^2)^2 \Rightarrow \mathbb{E}Z^2 \le \alpha(C)(\mathbb{E}|Z|)^2$

↳ Same argument $\Rightarrow P(Z<0) \ge \beta(C) > 0$.

Moral:
↳ If you control higher moments, you can bound prob that not 0.

A polynomial $V \in \mathbb{R}[x_1,\dots,x_n]$ is called a tetrahedral chaos if

$\forall i \quad \frac{\partial^2 V}{\partial x_i^2} \equiv 0$ [no powers higher than 1].

ex: $V(x_1,x_2) = x_1x_2 + 5x_1 - x_2 + 7$
non-ex: $V(x_1,x_2) = x_1^3x_2$

Rademacher: $P(r=\pm 1) = 1/2 \Rightarrow r^{2k} = 1$, $r^{2k+1} = r$
$\Rightarrow$ any such poly of Rademachers is tetrahedral chaos.

Bonami lemma: $Z_1,\dots,Z_n$ ind. s.t. $\mathbb{E}[Z_i] = \mathbb{E}[Z_i^3] = 0$, $\mathbb{E}[Z_i^2] = 1$, $\mathbb{E}[Z_i^4] \le 9$ C

$V \in \mathbb{R}[x_1,\dots,x_n]$, deg $V \le d$, V of (tetrahedral) chaos type. Then for $Z = V(Z_1,\dots,Z_n)$,

max(C, 9)

$$\mathbb{E}[Z^4] \le 9^d(\mathbb{E}[Z^2])^2$$

↳ Not optimal; $d=1 \Rightarrow$ can get 3, as we saw.

BUT: $(C_d)^{1/d} \xrightarrow[d\to\infty]{} 9$.

pf: deg $V \le d$; induct on n, # of variables

$V(x_1,\dots,x_n) = x_nQ(x_1,\dots,x_{n-1}) + P(x_1,\dots,x_{n-1})$
deg $Q \le d-1$ deg $P \le d$

↳ key: chaos $\Rightarrow x_n$ has power 0 or 1.

$X = P(Z_1,\dots,Z_{n-1})$, $Y = Q(Z_1,\dots,Z_{n-1})$

$\Rightarrow Z = Z_n \cdot Y + X$ (X,Y) ind of Z_n

Then $\mathbb{E}[Z^4] = \mathbb{E}[X^4 + 4X^3YZ_n + 6X^2Y^2Z_n^2 + 4XY^3Z_n^3 + Y^4Z_n^4]$

$= \mathbb{E}[X^4] + 6\mathbb{E}[X^2Y^2\overset{=1}{Z_n^2}] + \mathbb{E}[Y^4]\cdot\mathbb{E}[Z_n^4]$

$\le \mathbb{E}[X^4] + 6\underbrace{\mathbb{E}[X^2Y^2]}_{\le 6\sqrt{\mathbb{E}X^4}\sqrt{\mathbb{E}Y^4}} + 9\underbrace{\mathbb{E}[Y^4]}_{\text{ind.}\ \le\ 9^{d-1}(\mathbb{E}Y^2)^2}$

$$\le \mathbb{E}[X^4] + 6\sqrt{9^d}\,\mathbb{E}X^2\sqrt{9^{d-1}}\,\mathbb{E}Y^2 + 9^{d-1}(\mathbb{E}Y^2)^2$$

$$\le 9^d(\mathbb{E}X^2)^2 + 6\cdot 3^d\cdot 3^{d-1}\,\mathbb{E}X^2\cdot\mathbb{E}Y^2 + 9^d(\mathbb{E}Y^2)^2$$

$$= 9^d(\mathbb{E}X^2+\mathbb{E}Y^2)^2 = 9^d(\mathbb{E}Z^2)^2$$

Note: $\mathbb{E}[Z_nXY] = \mathbb{E}[\underbrace{Z_n}_{\text{ind}}\cdot P(z_1,\dots,z_{n-1})\cdot P(z_1,\dots,z_{n-1})] = 0$ ind

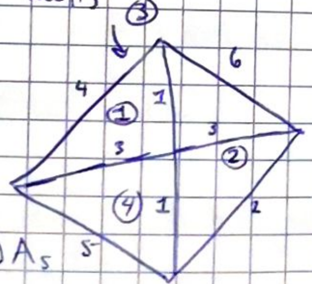
$$\mathbb{E}Z^2 = \mathbb{E}X^2 + \mathbb{E}[Z_n^2Y^2] + 2\,\mathbb{E}[Z_nXY] = \mathbb{E}X^2 + \mathbb{E}Y^2 \quad \square$$

Why can't you beat 9, asymptotically? Assume $3 \mid d$, more annoying o.w.

$$Z = \sum_{\substack{A\subseteq[n]\\|A|=d}}\prod_{i\in A} r_i$$ (let $n\to\infty$, normalize to keep Variance 1)

↳ In $\lim_{n\to\infty}$, behaves like $H_d(G)$, $G\sim N(0,1)$, H_d dth Hermite polynomial

$A_1,\dots,A_6\subseteq[n]$ pairwise disjoint, $|A_i| = d/3$ — by edges of faces



$S_1 = A_1\cup A_4\cup A_5$ $S_2 = A_1\cup A_2\cup A_6$

$S_3 = A_3\cup A_4\cup A_6$ $S_4 = A_2\cup A_3\cup A_5$

So: $A_1 = S_1\cap S_2$, so on (common edge of faces)

$$\mathbb{E}Z^2 = \mathbb{E}\Big[\sum_{\substack{S_1\subseteq[n]\\|S_1|=d}}\prod_{i\in S_1} r_i\cdot\sum_{\substack{S_2\subseteq[n]\\|S_2|=d}}\prod_{i\in S_2} r_i\Big]$$

$$= \sum_{\substack{S_1,S_2\subseteq[n]\\|S_1|=|S_2|=d}}\Big[\prod_{i=1}^n r_i^{\delta_i(S_1)+\delta_i(S_2)}\Big]$$

$$\sum\prod_{i=1}^n \mathbb{E}[r_i^{\delta_i(S_1)+\delta_i(S_2)}]$$

$$= \sum_{\substack{S\subseteq[n]\\|S|=d}} 1 = \binom{n}{d}$$

$$\mathbb{E}Z^4 = \mathbb{E}\Big[\Big(\sum_{\substack{S_1\subseteq[n]\\|S_1|=d}}\prod_{i\in S_1} r_i\Big)\cdots\Big(\sum_{\substack{S_4\subseteq[n]\\|S_4|=d}}\prod_{i\in S_4} r_i\Big)\Big]$$

$$= \sum_{\substack{S_1,\dots,S_4\subseteq[n]\\|S_i|=d\ \forall i\in[4]}}\prod_{i=1}^n \underbrace{\mathbb{E}[r_i^{\delta_i(S_1)+\dots+\delta_i(S_4)}]}_{\ge 0}$$

↳ can get rid of terms

$$\ge \sum_{\substack{A_1,\dots,A_6\subseteq[n]\\ \text{p.w. disjoint}\\ \forall i\ |A_i| = d/3}} \mathbb{E}[r_i^{\delta_i(A_1\cup A_4\cup A_5)+\delta_i(A_1\cup A_2\cup A_6)+\dots}]$$

(no double count b/c of intersection prop!)

$$= \sum \mathbb{E}[r_i^{\delta_i(A_1)+\delta_i(A_4)+\cdots}]$$

↳ each appears twice b/c sum over edges of △

$$= \underbrace{\binom{n}{d/3\ \cdots\ d/3}}_{6\text{ times}}$$

Look at ratio of these values, Stirling's formula, $n\to\infty$, d fixed $\Rightarrow 9^d$.

So even for Rademacher ~~...~~ RV, for large d, cannot beat constant 9!

Lecture 5, morning

$C_{p,2} = \|G\|_p$ $p\in(2,3)$

We have: $\mathbb{E}|S|^p \le \sum_{k=1}^n a_k^2 F_p(a_k^{-2})$

Technical Part: $F_p(s)\le F_p(\infty)$, $s\ge\sqrt{2}$ ← for $p\in(2,4)$

Recall

If all $|a_k|\le\frac{1}{\sqrt2}$, Then

$$\mathbb{E}|S|^p \le \sum_{k=1}^n a_k^2 F_p(\infty) = F_p(\infty) = \mathbb{E}|G|^p.$$

Case 2: $\exists k: |a_k| > \frac{1}{\sqrt2}$

$$\mathbb{E}|S|^4 = 3 - 2\sum_{k=1}^n a_k^4 \le 3-1 = 2$$

2 p 4

$$(\mathbb{E}|S|^p)^2 \le (\mathbb{E}|S|^2)^{4-p}(\mathbb{E}|S|^4)^{p-2}$$

$$\le 2^{p-2} \le (\mathbb{E}|G|^p)^2.$$