LECTURE 14

- Last class, we introduced the class of <u>subgaussian distributions</u>

$X \sim$ Subgaussian if:

(1) $\exists K_1$: $\mathbb{P}\{|X| \geq t\} \leq 2\exp(-t^2/K_1^2)$ $\forall t \geq 0$

$\Updownarrow$

(2) $\exists K_2$: $\|X\|_{L^p} \leq K_2\sqrt{p}$ $\forall p \geq 1$

$\Updownarrow$

(3) $\exists K_3$: $\mathbb{E}\exp(X^2/K_3^2) \leq 2$

$\Updownarrow$ if $\mathbb{E}X = 0$

(4) $\exists K_4$: $\mathbb{E}\exp(\lambda X) \leq \exp(K_4^2\lambda^2)$ $\forall \lambda \in \mathbb{R}$.

The smallest K_3 is called <u>subgaussian norm</u> $\|X\|_{\psi_3}$;

all other K_i's are <u>equivalent</u> to it.

- <u>Ex</u> (a) <u>Normal</u>: $g \sim N(0,1)$

(b) <u>Bernoulli</u>: $X \sim Ber(p)$

(c) <u>Uniform</u>: $X \sim Unif[0,1]$

$\Big\}$ $\|X\|_{\psi_2} \leq C$.

(c) <u>$\forall$ Bounded</u>: $|X| \leq M$ a.s. $\Rightarrow \|X\|_{\psi_2} \leq CM$. (Check!)

- Fact: X_i subgauss $\Rightarrow$ $\sum X_i$ subgauss (normed space).

- $\|X\|_{\psi_2}^2$ is a "variance proxy" (see HW5).

$\mathrm{Var}(\sum X_i) = \sum \mathrm{Var}(X_i)$ for indep. X_i. Similarly:

Prop If X_i are independent, mean zero, subgaussian, then

$$\boxed{\Big\|\sum_{i=1}^N X_i\Big\|_{\psi_2}^2 \le C \sum_{i=1}^N \|X_i\|_{\psi_2}^2}$$

$$\mathbb{E}\exp\Big(\lambda \sum_i X_i\Big) = \underbrace{\mathbb{E}\prod_i}_{\text{indep}} \exp(\lambda X_i) = \prod_i \underbrace{\mathbb{E}\exp(\lambda X_i)}_{\le\ \exp(CK_i^2\lambda^2)} \le$$

where $K_i = \|X_i\|_{\psi_2}$ due to subgauss. property (4)

$$\le \exp\Big(C\lambda^2 \underbrace{\sum_i K_i^2}_{=K^2}\Big)$$

$\Rightarrow \sum_i X_i$ is subgaussian and $\|\sum_i X_i\|_{\psi_2} \le CK$, due to subgauss. property (4).

THM (Subgaussian Hoeffding ineq.) If X_i are independent, mean zero, subgaussian, then

$$P\Big\{\Big|\sum_{i=1}^N X_i\Big| \ge t\Big\} \le 2\exp\Big(-\frac{ct^2}{\sigma^2}\Big)$$

where $\sigma^2 = \sum_{i=1}^N \|X_i\|_{\psi_2}^2$ ← variance proxy of the sum

Apply Prop and then the subgauss. tail (property 1).

Ex (a) $X_i \sim$ Sym. Bernoulli.

$$\mathbb{P}\left\{\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} X_i\right| \geq t\right\} = \mathbb{P}\left\{\left|\sum_{i=1}^{N} X_i\right| \geq t\sqrt{N}\right\} \leq 2\exp\left(-\frac{(ct\sqrt{N})^2}{\sigma^2}\right) \circledleq$$

where $\sigma^2 = \sum_{i=1}^{N} \underbrace{\|X_i\|_{\psi_2}^2}_{\leq C} \leq CN$

$$\circledleq\ 2\exp\left(-\frac{ct^2\cancel{N}}{C\cancel{N}}\right) = 2\exp(-c_1 t^2)$$ Gauss. tail.

$\Downarrow$

K.I. in lec. 6

(b) $X_i \sim N(0,1)$: $\frac{1}{\sqrt{N}}\sum_1^N X_i \sim N(0,1) \Rightarrow$ same.

Subexponential distributions.

- $X \sim$ Subgaussian $\not\Rightarrow$ $X^2 \sim$ Subgaussian :

  e.g. $g \sim N(0,1)$. $\mathbb{P}\{g^2 > t\} = \mathbb{P}\{g > \sqrt{t}\} \sim \exp\left(-\frac{\sqrt{t}^2}{2}\right) = \boxed{\exp(-t/2)}$

  exponential tail

  heavier than $\exp(-ct^2)$

- Fix : define subexponential distr. similarly :

Prop $\forall$ r.v. X, TFAE :

(1) $\exists K_1$: $\mathbb{P}\{|X| \geq t\} \leq 2\exp(-t/K_1)$ $\forall t \geq 0$

$\Updownarrow$

(2) $\exists K_2$: $\|X\|_{L^p} \leq K_2 p$ $\forall p \geq 1$

$\Updownarrow$

(3) $\exists K_3$: $\mathbb{E}\exp(|X|/K_3) \leq 2$

$\Updownarrow$ if $\mathbb{E}X = 0$

(4) $\exists K_4$: $\mathbb{E}\exp(\lambda X) \leq \exp(K_4^2\lambda^2)$ $\forall |\lambda| \leq \frac{1}{K_4}$

Such X is called subexponential. The smallest K_3 is called the subexponential norm $\|X\|_{\psi_1}$

Orlicz norm for $\psi_1(x) = e^x - 1$

Proof: DIY

- Examples

(a) $\forall$ subgaussian, e.g. $g \sim N(0,1)$ (by property 2)

(b) $\forall$ subgaussian squared, e.g. g^2.

$\forall$ r.v. X: $\quad \|X^2\|_{\psi_1} = \|X\|_{\psi_2}^2$ trivially.

(c) Exponential $\quad X \sim \mathrm{Exp}(1) \quad X \geq 0$,

$$\mathbb{P}\{X \geq t\} = e^{-t} \qquad \forall t \geq 0$$

$$\text{pdf} = \frac{d}{dt}(\text{cdf}) = \frac{d}{dt}(1-e^{-t}) = e^{-t}, \quad t \geq 0$$

$$\Rightarrow \mathbb{E}\exp(\lambda X) = \int_0^\infty e^{\lambda x} \cdot e^{-x}\, dx < \infty \text{ only if } \underline{\lambda \leq 1}$$

$\Rightarrow$ restriction on λ in (4) is essential.

(d) Poisson: not subgaussian (HW5) but subexponential (tail t^{-t})

Q Does Hoeffding hold for subexponential X_i?

- NO! Even for $N=1$: $\quad \mathbb{P}\{|X| \geq t\} \leq \underbrace{\exp\left(-\frac{ct}{K}\right)}$ where $K = \|X\|_{\psi_1}$

too heavy! $\Rightarrow$ Hoeffding's $\exp(-ct^2)$

- Surprisingly, $N=1$ is "the only obstruction":

BERNSTEIN'S INEQUALITY If X_i are indep, mean 0, subexponential, then

$$\boxed{\mathbb{P}\left\{\left|\sum_{i=1}^N X_i\right| \geq t\right\} \leq 2\exp\left[-c \cdot \min\left(\frac{t^2}{\sigma^2}, \frac{t}{K}\right)\right]}$$

where $\sigma^2 = \sum_{i=1}^N \|X_i\|_{\psi_1}^2$, $\quad K = \max_i \|X_i\|_{\psi_1}$

Proof: the same MGF method as in Hoeffding, Chernoff: $\quad S = \sum X_i$

$$\mathbb{P}\{S \geq t\} = \mathbb{P}\{e^{\lambda S} \geq e^{\lambda t}\} \leq e^{-\lambda t}\mathbb{E}e^{\lambda S} = e^{-\lambda t}\boxed{\mathbb{E}e^{\lambda X_i}}$$

now use Property 4 p.3

Restriction on λ leads to