LECTURE 13

- Last class, we noted that $g \sim N(0,1)$ satisfies:

  (1) $\mathbb{P}\{|g| \geq t\} \leq 2\exp(-t^2/2) \quad \forall t \geq 0$

  (2) $\|g\|_{L^p} = (\mathbb{E}|g|^p)^{1/p} \leq \sqrt{p} \quad \forall p \geq 1$

  (3) $\mathbb{E}\, e^{g^2/4} \leq 2$

  (4) $\mathbb{E}\, e^{\lambda g} \leq e^{\lambda^2/4} \quad \forall \lambda \in \mathbb{R}$

- Today, we show that $\forall$ distribution, (1)–(4) are EQUIVALENT:

"the following are equivalent"

Subgaussian Lemma $\forall$ r.v. X, TFAE:

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

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

(3) (ψ_2): $\exists K_3$: $\quad \mathbb{E}\exp(X^2/K_3^2) \leq 2$

Moreover, if $\mathbb{E}X = 0$ then (1)–(3) are equivalent to:

(4) (MGF): $\exists K_5$: $\quad \mathbb{E}\exp(\lambda X) \leq \exp(K_4^2 \lambda^2) \quad \forall \lambda \in \mathbb{R}$.

A r.v. that satisfies one (and thus all) of properties (1)–(4) is called subgaussian.

HW: $\mathbb{E}X = 0$ is needed in (4)

$\mathrm{Var}(X) \lesssim \|X - \mathbb{E}X\|_{\psi_2}$

$\mathrm{Pois} \in L_{\psi_2}$

The proof of (1) ⇒ (2) is based on

Lem (Integral identity) ∀ nonnegative r.v. X,

$$\mathbb{E}X = \int_0^\infty \mathbb{P}\{X > t\}\,dt.$$

$\ulcorner$ $\forall x \in \mathbb{R}$, $\quad x = \int_0^x dt = \int_0^\infty \underbrace{\mathbb{1}_{\{t<x\}}}_{= \begin{cases} 1 & \text{if } t<x \\ 0 & \text{if } t \geq x \end{cases}} dt$

let $x = X$, take expectation ⇒

$$\mathbb{E}X = \int_0^\infty \underbrace{\mathbb{E}\,\mathbb{1}_{\{t<X\}}}_{= \mathbb{P}\{t<X\}} dt \qquad (\text{Fubini}) \qquad \lrcorner$$

Proof of Subgaussian Lemma

(1) ⇒ (2): WLOG $K_1 = 1$

$$\mathbb{E}|X|^p = \int_0^\infty \mathbb{P}\{|X|^p > \underset{\text{change var. } s^p}{t}\}\,dt \qquad (\text{Integral Lemma}) \qquad dt = ps^{p-1}ds$$

$$= \int_0^\infty \underbrace{\mathbb{P}\{|X| > s\}}_{\leq 2e^{-s^2}} p s^{p-1} ds \qquad \text{by Property (1).}$$

$$= p\left(\frac{p}{2} - 1\right)! \qquad (\text{by parts} - \text{DIY})$$

$$\leq p \cdot p^{p/2 - 1} = p^{p/2}$$

$$\Rightarrow \|X\|_{L^p} \leq \sqrt{p}.$$

$\underline{(2) \Rightarrow (3)}$: WLOG $K_2 = 1$

$$\mathbb{E}\exp(X^2/100) = \mathbb{E}\left[\sum_{p=0}^{\infty} \frac{(X^2/100)^p}{p!}\right] \quad \text{(Taylor series)}$$

$$= \sum_{p=0}^{\infty} \frac{\underbrace{\mathbb{E}X^{2p}}_{\le (\sqrt{2p})^{2p}}}{100^p\, p!} \le \sum_{p=0}^{\infty} \frac{(2p)^p}{100^p \underbrace{p!}_{\ge (p/e)^p}} \quad \text{(Stirling)}$$

$$\le \sum_{p=0}^{\infty} \Big(\underbrace{\frac{2e}{100}}_{\le 1/10}\Big)^p \le 2$$

$(3) \Rightarrow (4)$ WLOG $K_3 = 1$. Use $e^x \le x + e^{x^2}$ $\forall x \in \mathbb{R}$ $\Rightarrow$

$$\mathbb{E}e^{\lambda X} \le \underbrace{\mathbb{E}[\lambda X]}_{=0 \text{ by assumption}} + \mathbb{E}e^{\lambda^2 X^2}$$

- If $|\lambda| \le 1$, $\mathbb{E}e^{\lambda^2 X^2} \le \left(\mathbb{E}e^{X^2}\right)^{\lambda^2}$

$\left(\Leftrightarrow \left(\mathbb{E}e^{\lambda^2 X^2}\right)^{1/\lambda^2} \le \mathbb{E}e^{X^2} \right.$

$\left. \Leftrightarrow \|e^{X^2}\|_{L^{\lambda^2}} \le \|e^{X^2}\|_{L^1} \right.$ true since $\lambda^2 \le 1$ $\Big)$

$$\Rightarrow \mathbb{E}e^{\lambda X} \le \Big(\underbrace{\mathbb{E}e^{X^2}}_{\le 2 \text{ by property (3)}}\Big)^{\lambda^2} \le \exp(\lambda^2 \ln 2)$$

- If $|\lambda| > 1$ — DIY. (or see the Book).

$(4) \Rightarrow (1)$ WLOG $K_4 = 1$.

$$\mathbb{P}\{X \ge t\} = \mathbb{P}\{e^{\lambda X} \ge e^{\lambda t}\} \quad \text{(multiply by } \lambda \ge 0 \text{ and exponentiate)}$$

$$\le e^{-\lambda t}\, \underbrace{\mathbb{E}e^{\lambda X}}_{\le e^{\lambda^2} \text{ (property 4)}} \quad \text{(Markov)}$$

$$= \exp(\lambda^2 - \lambda t).$$

Minimize in $\lambda \Rightarrow \lambda = t/2 \Rightarrow$

$$= \exp(-t^2/4).$$

Repeat for $-X \Rightarrow$ QED

QUANTITATIVE ASPECTS

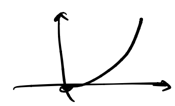
- Subgaussian is a qualitative notion (yes/no). But:
- K_1, K_2, K_3, K_4 in Subgaussian Lemma
  are all equivalent up to absolute const. factors:

$$K_i \le C_{ij} K_j \qquad \forall i,j = 1,2,3,4 \qquad \text{(CHECK!)}$$

$\Rightarrow$ they all capture the $\approx$ same quantity.

- We call this quantity the subgaussian norm
- Why norm? Property (3): "$\exists K_3 \quad \mathbb{E}\exp(X^2/K_3^2) \le 2$"

$\Leftrightarrow \exists K_3 : \mathbb{E}\,\psi_2(X/K_3) \le 1$

where $\psi_2(x) = e^{x^2} - 1$ is an Orlicz function 

Hence the smallest K_3 equals $\boxed{\|X\|_{\psi_2}}$ $\leftarrow$ subgaussian norm.

& all other K_i's are equivalent $\Rightarrow$

Subg. Lemma restated:

(1) (tails): $\mathbb{P}\{|X| \ge t\} \le 2\exp(-ct^2/\|X\|_{\psi_2}^2) \quad \forall t \ge 0$

(2) (moments): $\|X\|_{L^p} \le C\|X\|_{\psi_2}\sqrt{p} \quad \forall p \ge 1$

(3) (ψ_2): $\mathbb{E}\exp(cX^2/\|X\|_{\psi_2}^2) \le 2$

Moreover, if $\mathbb{E}X = 0$ then

(4) (MGF): $\mathbb{E}\exp(\lambda X) \le \exp(C\|X\|_{\psi_2}^2 \lambda^2) \quad \forall \lambda \in \mathbb{R}.$