Seminar 3

High-dimensional Probability and Statistics

HSE University, spring 2023

Theorem 1 (Bernstein's inequality)**.** *Let $\xi_1, \dots, \xi_n$ be i.i.d. centered random variables supported on $[-B, B]$. Then, for any $t \geqslant 0$, it holds that*

$$\mathbb{P}\left(\sum_{i=1}^n \xi_i \geqslant t\right) \leqslant \exp\left\{-\frac{t^2/2}{n\sigma^2 + Bt/3}\right\},$$

where $\sigma^2 = \mathrm{Var}(\xi_1)$.

Problem 1. Deduce from Bernstein's inequality that, for any $\delta \in (0,1)$, we have

$$\frac{1}{n}\sum_{i=1}^n \xi_i \leqslant \sigma\sqrt{\frac{2\log(1/\delta)}{n}} + \frac{2B\log(1/\delta)}{3n}$$

with probability at least $1 - \delta$.

Problem 2. Let $\xi_1, \dots, \xi_n$ be i.i.d. Bernoulli random variables with parameter p. Fix any $\delta \in (0,1)$. Prove that

$$\mathbb{P}\left(\sum_{i=1}^n \xi_i \leqslant t\right) \leqslant e^{-\frac{3t}{16}}, \quad \forall\, t < np/2.$$

Definition 2 (Orlicz norm)**.** The Ψ-Orlicz norm of a random variable ξ is given by

$$\|\xi\|_\Psi = \inf\left\{t > 0 : \mathbb{E}\Psi(|\xi|/t) \leqslant 1\right\}.$$

Remark 3. Particular cases of the Orlicz norm include $\psi_p(x) = e^{x^p} - 1$.

Remark 4. For any $p \geqslant 1$, the $\|\cdot\|_{\psi_p}$-norm is indeed a norm. In particular, it satisfies the triangle inequality.

Problem 3. Let ξ be a centered random variable.

1. Show that if ξ is a sub-Gaussian random variable with variance proxy σ^2, then there exists a constant $C_1 > 0$ such that $\|\xi\|_{\psi_2} \leqslant C_1\sigma$.

2. Prove that $\|\xi\|_{\psi_2} < \infty$ implies that $\xi \in \mathrm{SG}(C_2\|\xi\|_{\psi_2}^2)$ for an absolute constant $C_2 > 0$.

Problem 4. Let ξ be a centered random variable.

1. Show that if ξ is a sub-exponential random variable with parameters $\sigma > 0$ and $b > 0$, then there exists a constant $C_1 > 0$ such that $\|\xi\|_{\psi_1} \leqslant C_1\sigma$.

2. Prove that $\|\xi\|_{\psi_1} < \infty$ yields $\|\xi\|_{L_k} \leqslant C_2 k\,\|\xi\|_{\psi_1}$ for an absolute constant $C_2 > 0$ and all $k \in \mathbb{N}$.

Theorem 5 (Bernstein's inequality for unbounded random variables)**.** *Let* $\xi_1, \dots, \xi_n$ *be centered i.i.d. random variables, such that* $\mathrm{Var}(\xi_1) = \sigma$ *and* $\|\xi_1\|_{\psi_1} < \infty$*. Then, for any* $t > 0$ *and any* $\rho > \|\xi_1\|_{\psi_1} \log(4n/t)$*, it holds that*

$$\mathbb{P}\left(\sum_{i=1}^n \xi_i > t\right) \leqslant \exp\left\{-\frac{t^2/8}{n\sigma^2 + \rho t/6}\right\} + 2\exp\left\{-\frac{\rho}{\big\|\max\limits_{1 \leqslant i \leqslant n} |\xi_i|\big\|_{\psi_1}}\right\}.$$

Problem 5. Let $\xi_1, \dots, \xi_n$ be i.i.d. random variables, such that $\|\xi_1\|_{\psi_1} < \infty$. Show that

$$\big\|\max_{1 \leqslant i \leqslant n} |\xi_i|\big\|_{\psi_1} \leqslant \|\xi_1\|_{\psi_1} \log_2 n.$$