Seminar 4

High-dimensional Probability and Statistics

HSE University, spring 2023

Problem 1. Let $\xi_1, \dots, \xi_n$ be i.i.d. random variables, such that $\|\xi_1\|_{\psi_1} < \infty$. Show that

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

Problem 2. Let $\xi_1, \dots, \xi_n$ be centered sub-Gaussian random variables with variance proxy σ^2. Prove that

$$\mathbb{E} \max_{1 \leqslant i \leqslant n} \xi_i \leqslant \sigma \sqrt{2 \log n}.$$

Hint. Show that $\mathbb{E} \max\limits_{1 \leqslant i \leqslant n} \xi_i \leqslant (1/\lambda) \log \mathbb{E} \exp \left\{ \lambda \max\limits_{1 \leqslant i \leqslant n} \xi_i \right\} \leqslant (1/\lambda) \log \mathbb{E} \sum\limits_{1 \leqslant i \leqslant n} \exp\{\lambda \xi_i\}$.

Theorem 1 (Azuma-Hoeffding inequality)**.** *Let $\{S_t : 0 \leqslant t \leqslant T\}$ be a martingale with respect to filtration $\{\mathcal{F}_t : 0 \leqslant t \leqslant T\}$ and assume that*

$$\mathbb{E}\left(e^{\lambda(S_t - S_{t-1})} | \mathcal{F}_{t-1}\right) \leqslant e^{\lambda^2 \sigma_t^2 / 2}, \quad a.\ s.\ \forall\, \lambda \in \mathbb{R}.$$

Then for any $t > 0$, it holds that

$$\mathbb{P}\left(S_T - S_0 \geqslant t\right) \leqslant e^{-\frac{t^2}{2 \sum_{t=1}^{T} \sigma_t^2}}.$$

Theorem 2 (McDiarmid inequality)**.** *Let $\xi_1, \dots, \xi_n$ be independent random variables. Assume that a function $\varphi : \mathbb{R}^n \to \mathbb{R}$ satisfies a bounded difference inequality, i.e. there exist $c_1, \dots, c_n > 0$ such that*

$$|\varphi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) - \varphi(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_n)| \leqslant c_i$$

holds for all $i \in \{1, \dots, n\}$ and for all $x_1, \dots, x_n, x_1', \dots, x_n' \in \mathbb{R}$. Then for all $t \geqslant 0$, we have

$$\mathbb{P}\left(\varphi(\xi_1, \dots, \xi_n) - \mathbb{E}\varphi(\xi_1, \dots, \xi_n) \geqslant t\right) \leqslant e^{-\frac{t^2}{2 \sum_{i=1}^{n} c_i^2}}.$$

Problem 3. Let $(X_1, Y_1), \dots, (X_n, Y_n) \in \mathcal{X} \times \mathcal{Y}$ be i.i.d. pairs of random elements. Let $\mathcal{G}$ be a separable space of functions g, taking their values in $\mathcal{Y}$ and let $\ell : \mathcal{Y} \times \mathcal{Y} \to [0, B]$ be a loss function. For any $g \in \mathcal{G}$, define

$$R_n(g) = \frac{1}{n} \sum_{i=1}^{n} \ell(Y_i, g(X_i)) \quad \text{and} \quad R(f) = \mathbb{E} R_n(g)$$

Show that

$$\mathbb{P}\left(\sup_{g \in \mathcal{G}} (R(g) - R_n(g)) - \mathbb{E} \sup_{g \in \mathcal{G}} (R(g) - R_n(g)) \geqslant t \right) \leqslant e^{-nt^2/(2B^2)}.$$

Problem 4. Let $X = (X_1, \dots, X_n)^\top$ be a random vector with independent Rademacher components, that is, $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$ for all $i \in \{1, \dots, n\}$. Prove that

$$\mathbb{P}\left(\|X\| - \mathbb{E}\|X\| \geqslant t\right) \leqslant e^{-t^2/(8n)} \quad \text{and} \quad \mathbb{P}\left(\|X\| \geqslant \sqrt{n} + t\right) \leqslant e^{-t^2/(8n)}.$$