**SLT/Concentration Inequalities** **2025**

# Exam 2: August 2025

*Instructor: Shashank Vatedka*

**Instructions:** *This is a closed-book exam. You are not permitted to refer to any material or discuss the problem with anyone. Malpractice will be severely punished. Please mention your ROLL Number and name clearly in the answer sheet.*

Justify all your statements clearly. You may use any result proved in class (but clearly state which results you are using), but everything else needs to be proved.

**Question 2.1.** Consider two independent random sources where the first one $X_1(1), X_1(2), X_1(3), \ldots$ whose components are iid with pdf $f_{X_1}$, and the second is $X_2(1), X_2(2), \ldots$ whose components are iid with pdf $f_{X_2}$. This is passed through a system which at each time instant $t$, outputs a sample $X(t)$ equal to the $t$'th symbol from the first source with probability $p$, or equal to the $t$'th symbol from the second source (with probability $1-p$).

What is the density of $X(t)$? (1pt)

Assume that $f_{X_1}$ and $f_{X_2}$ are subgaussian with subgaussian norms $K_1, K_2$ respectively and $0 < K_1 < K_2$. Is $X$ subgaussian? (1pt) If yes, prove whether it can satisfy $\|X\|_{\psi_2} \leqslant K_1$? (2pts) or $\|X\|_{\psi_2} \leqslant K_2$? (2pts)

Now suppose that we want to estimate the mean of the output $X$ from $n$ iid samples $X(1), \ldots, X(n)$. Design an estimator which for any $\epsilon, \delta > 0$ takes (where $c$ is a universal constant that does not depend on the distribution)

$$n = c\frac{K_2^2}{\epsilon^2}\log\frac{2}{\delta}$$

samples and outputs an estimate $\hat{\mu}$ satisfying $\Pr[|\hat{\mu} - \mu| \geqslant \epsilon] \leqslant \delta$. (5pts)

**Question 2.2.** Consider the probability mass function $p_X$ defined on the set of positive integers $\mathbb{Z}_+ = \{1, 2, \ldots\}$:

$$p_X(i) = \frac{6}{\pi^2 i^2}$$

Find the mean of $X$ (2pts)

Justify whether $X$ is subgaussian (1pt) or subexponential (1pt)

For what values of $\alpha > 2$ is the following pmf is subgaussian (1pt) or subexponential (1pt) and why:

$$p_X(i) = \frac{\beta}{i^\alpha}, \quad i \in \mathbb{Z}_+$$

where $\beta$ is a normalizing constant to ensure that $p_X$ is a pmf.

**Question 2.3** (6pts). Let $X_1$ and $X_2$ be independent subgaussian random variables with subgaussian norms $K_1$ and $K_2$ respectively. Is $X = \max\{X_1, X_2\}$ subgaussian? If yes, give an upper bound on the subgaussian norm as a function of $K_1, K_2$.

**Question 2.4.** In class, we derived bounds on the performance of empirical risk minimization when the hypothesis class was finite. Let us now consider an arbitrary hypothesis class $\mathcal{G}$ consisting of functions/classifiers $g : \mathcal{X} \to \{0, 1\}$ where $\mathcal{X}$ could be an arbitrary set, and $\mathcal{G}$ need not be finite.

(8 pts)

Consider a dataset of independent and identically distributed labeled samples

$$\mathcal{S} = \Big((X_1, Y_1), \ldots, (X_n, Y_n)\Big)$$

where each sample $(X_i, Y_i) \sim p_{XY}$ (the distribution $p_{XY}$ is unknown and arbitrary).

The loss function for binary classification is $\ell(y, \hat{y}) = 1_{\{y \neq \hat{y}\}}$, the empirical risk is defined as

$$R_{\mathcal{S}}(g) = \frac{1}{n} \sum_{i=1}^{n} \ell(Y_i, g(X_i)),$$

the average risk is defined as

$$R(g) = \Pr[Y_i \neq g(X_i)].$$

Prove the following uniform convergence property for the empirical risk: there is a universal constant $c > 0$ such that for every $\delta > 0$,

$$\Pr\left[ R(g) \leqslant R_{\mathcal{S}} + \sqrt{\frac{\log(2/\delta)}{cn}}, \text{ for every } g \in \mathcal{G} \right] \geqslant 1 - \delta.$$

Hint: Define $\phi(\mathcal{S}) = \sup_{g \in \mathcal{G}}(R(g) - R_{\mathcal{S}}(g))$, and express the above in terms of $\phi(\mathcal{S})$. Does $\phi(\mathcal{S})$ satisfy the bounded differences property? If yes, you can try to use McDiarmid's inequality.