CS 6220: Machine Learning Theory

Lecture 12 — Concentration Inequalities

November 13, 2026

1 Motivation

We often need to bound how far a random variable deviates from its expectation, e.g. to argue that empirical risk converges to true risk as sample size grows.

2 Basic Inequalities

Definition 1. A random variable X is *sub-Gaussian* with parameter σ if $\mathbb{E}[e^{\lambda(X-\mathbb{E}X)}] \le e^{\lambda^2\sigma^2/2}$ for all $\lambda \in \mathbb{R}$.

Theorem 2 (Hoeffding's Inequality). If $X_1, \dots, X_n$ are independent, bounded in $[a, b]$, then for $\bar{X} = \frac{1}{n}\sum_i X_i$,

$$P\left(|\bar{X} - \mathbb{E}\bar{X}| \ge t\right) \le 2\exp\left(-\frac{2nt^2}{(b-a)^2}\right).$$

Proof. Apply Markov's inequality to $e^{\lambda(\bar{X}-\mathbb{E}\bar{X})}$, use independence to factor the expectation, bound each factor via Hoeffding's lemma, then optimize over λ. Symmetrize for the two-sided bound. ■

2.1 Example

Suppose we estimate the accuracy of a classifier from $n = 1000$ i.i.d. test examples with $X_i \in \{0, 1\}$. Hoeffding's inequality gives, for $t = 0.02$:

- Bound width shrinks as $1/\sqrt{n}$.
- Confidence improves exponentially in nt^2.
- The bound is distribution-free — no assumption beyond boundedness.

3 Summary

- Sub-Gaussian variables satisfy exponential concentration around their mean.
- Hoeffding's inequality requires only boundedness, not variance information.
- These bounds underlie generalization guarantees in learning theory.