DATA 37200: Learning, Decisions, and Limits (Winter 2026)

Lecture 2: Concentration Inequalities

Instructor: Frederic Koehler

Hoeffding's inequality (bounded independent sums)

Theorem (Hoeffding)

Let $X_1, \ldots, X_n$ be independent random variables with $X_i \in [a_i, b_i]$ almost surely. Let $S_n = \sum_{i=1}^n X_i$ and $\mu = \mathbb{E} S_n = \sum_{i=1}^n \mathbb{E} X_i$. Then for every $t > 0$,

$$\Pr\Big(S_n - \mu \geq t\Big) \;\leq\; \exp\Big(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\Big)$$

and

$$\Pr\Big(S_n - \mu \leq -t\Big) \;\leq\; \exp\Big(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\Big).$$

So

$$\Pr\Big(|S_n - \mu| \geq t\Big) \;\leq\; 2\exp\Big(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\Big).$$

Special case: if $X_i \in [0, 1]$ then $\Pr(S_n - \mu \geq t) \leq \exp(-2t^2/n)$.

Proof (Chernoff bound + Hoeffding's lemma)

Step 1: Chernoff (exponential Markov). For any $\lambda > 0$,

$$\Pr(S_n - \mu \geq t) = \Pr\left(e^{\lambda(S_n-\mu)} \geq e^{\lambda t}\right)$$

$$\Pr\left(e^{\lambda(S_n-\mu)} \geq e^{\lambda t}\right) \leq e^{-\lambda t}\,\mathbb{E}e^{\lambda(S_n-\mu)} = e^{-\lambda t}\prod_{i=1}^{n}\mathbb{E}e^{\lambda(X_i-\mathbb{E}X_i)}.$$

Step 2: Hoeffding's lemma. (We will prove it afterward.) If $Y \in [a, b]$ a.s. and $\mathbb{E}Y = 0$, then for all $\lambda \in \mathbb{R}$,

$$\mathbb{E}e^{\lambda Y} \;\leq\; \exp\left(\frac{\lambda^2(b-a)^2}{8}\right).$$

Apply to $Y_i = X_i - \mathbb{E}X_i \in [a_i - \mathbb{E}X_i,\ b_i - \mathbb{E}X_i]$ to get

$$\mathbb{E}e^{\lambda(X_i-\mathbb{E}X_i)} \leq \exp\left(\frac{\lambda^2(b_i-a_i)^2}{8}\right).$$

Step 3: Combine and optimize λ.

$$\Pr(S_n - \mu \geq t) \;\leq\; \exp\Big(-\lambda t + \frac{\lambda^2}{8} \sum_{i=1}^{n} (b_i - a_i)^2 \Big).$$

Minimizing the exponent gives $\lambda^\star = \dfrac{4t}{\sum_{i=1}^{n}(b_i - a_i)^2}$, yielding

$$\Pr(S_n - \mu \geq t) \;\leq\; \exp\Big(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2} \Big).$$

The lower-tail bound follows by applying the same argument to $-(S_n - \mu)$.

Hoeffding's lemma

Lemma (Hoeffding)

If $Y \in [a, b]$ almost surely and $\mathbb{E}Y = 0$, then for all $\lambda \in \mathbb{R}$,

$$\mathbb{E}e^{\lambda Y} \;\le\; \exp\Big(\frac{\lambda^2(b-a)^2}{8}\Big).$$

First step of proof: Take log of both sides, equivalent to show

$$\log \mathbb{E}e^{\lambda Y} \le \lambda^2(b-a)^2/8$$

In what follows we will analyze the degree-2 Taylor expansion of the lhs in λ.

Hoeffding's lemma: setup (cumulant generating function)

Let

$$K(\lambda) \;:=\; \log \mathbb{E}\, e^{\lambda Y} \qquad \text{(the cumulant generating function of } Y\text{)}.$$

We want to show

$$K(\lambda) \;\le\; \frac{\lambda^2(b-a)^2}{8} \quad \text{for all } \lambda \in \mathbb{R}.$$

Two easy facts:

$$K(0) = \log 1 = 0, \qquad K'(0) = \frac{\mathbb{E}[Ye^0]}{\mathbb{E}[e^0]} = \mathbb{E}Y = 0.$$

So if we can bound $K''(\lambda)$ uniformly, we can integrate twice ("2nd-order Taylor with remainder").

Key identity: $K''(\lambda)$ is a variance under a tilted measure

Define the tilted expectation

$$\mathbb{E}_\lambda[f(Y)] \;:=\; \frac{\mathbb{E}\big[f(Y)e^{\lambda Y}\big]}{\mathbb{E}[e^{\lambda Y}]}.$$

Differentiate:

$$K'(\lambda) = \frac{\mathbb{E}[Ye^{\lambda Y}]}{\mathbb{E}[e^{\lambda Y}]} = \mathbb{E}_\lambda[Y].$$

Differentiate again (quotient rule):

$$K''(\lambda) = \mathbb{E}_\lambda[Y^2] - \big(\mathbb{E}_\lambda[Y]\big)^2 = \mathrm{Var}_\lambda(Y).$$

Thus, Hoeffding's lemma reduces to bounding a variance.

Uniform bound: variance of a bounded variable

If $Z \in [a, b]$ almost surely, then for any probability measure,

$$\mathrm{Var}(Z) \;\le\; \frac{(b-a)^2}{4}.$$

Proof (one line): let $m = (a+b)/2$. Then $|Z - m| \le (b-a)/2$, hence

$$(Z-m)^2 \le \frac{(b-a)^2}{4}.$$

Taking expectations gives $\mathbb{E}(Z-m)^2 \le (b-a)^2/4$, and since $\mathrm{Var}(Z) \le \mathbb{E}(Z-m)^2$, the claim follows.
Apply this with $Z = Y$ under the tilted measure $\mathbb{P}_\lambda$:

$$K''(\lambda) = \mathrm{Var}_\lambda(Y) \;\le\; \frac{(b-a)^2}{4} \quad \text{for all } \lambda.$$

Integrate twice (a rigorous "2nd-order Taylor bound")

Using $K(0) = K'(0) = 0$,

$$K(\lambda) = \int_0^\lambda K'(s)\, ds = \int_0^\lambda \int_0^s K''(t)\, dt\, ds = \int_0^\lambda (\lambda - t)\, K''(t)\, dt.$$

Now plug in the uniform bound $K''(t) \le (b-a)^2/4$:

$$K(\lambda) \le \int_0^\lambda (\lambda - t)\, \frac{(b-a)^2}{4}\, dt = \frac{(b-a)^2}{4} \cdot \frac{\lambda^2}{2} = \frac{\lambda^2 (b-a)^2}{8}.$$

Exponentiating both sides yields

$$\mathbb{E} e^{\lambda Y} \;\le\; \exp\Big(\frac{\lambda^2 (b-a)^2}{8}\Big).$$

Analysis of NAIVE-EE (aka ETC, Explore-Then-Commit)

Recall: we have T rounds of interaction, k arms, and when we pull arm i we receive reward $r_i \sim \mathcal{D}_i$. For simplicity, we assume $r_i \in [0, 1]$ always.

Suppose we pull each arm m times in the exploration phase. Then, by Hoeffding's inequality, the sample mean

$$\hat{\mu}_i = \frac{1}{m} \sum_{j=1}^{m} r_i^{(j)}$$

satisfies with probability at least $1 - \delta/K$ that

$$|\hat{\mu}_i - \mu_i| \leq \sqrt{2 \log(2K/\delta)/m}.$$

By the union bound, this holds for all $i \in [K]$ with probability at least $1 - \delta$.

Completing analysis

Under the high probability event that for all $i \in [K]$

$$|\hat{\mu}_i - \mu_i| \leq \sqrt{2\log(2K/\delta)/m}.$$

we can now bound the regret. Let $\mu^* = \max \mu_i$ be the true optimal reward. Let $\tilde{\mu} = \max \hat{\mu}_i$ be the best sample reward. Then in each of the $T - Km$ exploitation rounds, our regret in expectation is at most

$$\epsilon := 2\sqrt{2\log(2K/\delta)/m}.$$

(Why?) In the first Km rounds we can lose at most Km reward. So

$$(\text{total regret}) \leq Km + (T - Km)\epsilon.$$

Take $m = T^{2/3}$, then we see the total regret is $O(KT^{2/3})$.

Looking forward

NAIVE-EE (also called ETC, Explore-Then-Commit) is easy to analyze but it's not very optimal. $T^{2/3}$ scaling of regret.

We want to understand how *combining* learning and decision-making can lead to improvements. Act more like an intelligent person ? This will lead to an improvement to regret scaling as $T^{1/2}\log^{1/2}(T)$.

(Think: how many times do you touch a stove before learning?)

To analyze better strategies, we need to spend some time learning/reviewing martingales (iterated betting games).

Motivation: “fair game” over time

We observe random information over time:

$$X_0,\ X_1,\ X_2,\ \ldots$$

Think: wealth of a gambler after t rounds, or a running estimate after seeing t samples.
Informal idea (fair game):

> *Given everything you know up to time t, your expected wealth at time $t+1$ is exactly your current wealth.*

That idea is a martingale.

Filtration = “information revealed so far”

Let $\mathcal{F}_t$ denote the information you have after observing the process up to time t.
Concrete picture: $\mathcal{F}_t$ is “everything you can compute from $X_0, \ldots, X_t$”.
Formally (but you can treat this as notation):

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$$

This increasing family $(\mathcal{F}_t)$ is called a filtration.

Definition: martingale

A sequence $(X_t)_{t\geq 0}$ is a martingale w.r.t. $(\mathcal{F}_t)$ if:

1. X_t is determined by time-t information (i.e. X_t is $\mathcal{F}_t$-measurable),
2. $\mathbb{E}|X_t| < \infty$,
3. Fairness: for all t,

$$\mathbb{E}[X_{t+1} \mid \mathcal{F}_t] \; = \; X_t.$$

Equivalent form (martingale differences): define $D_{t+1} := X_{t+1} - X_t$. Then

$$\mathbb{E}[D_{t+1} \mid \mathcal{F}_t] = 0.$$

Examples

1) Symmetric random walk. Let $\xi_1, \xi_2, \ldots$ be i.i.d. with $\mathbb{P}(\xi_i = +1) = \mathbb{P}(\xi_i = -1) = 1/2$, and set $S_t = \sum_{i=1}^{t} \xi_i$. With $\mathcal{F}_t = \sigma(\xi_1, \ldots, \xi_t)$,

$$\mathbb{E}[S_{t+1} \mid \mathcal{F}_t] = S_t + \mathbb{E}[\xi_{t+1} \mid \mathcal{F}_t] = S_t.$$

So (S_t) is a martingale.

2) "Conditional expectation process." For any integrable random variable Z, define

$$X_t := \mathbb{E}[Z \mid \mathcal{F}_t].$$

Then (X_t) is a martingale ("tower property"): $\mathbb{E}[X_{t+1} \mid \mathcal{F}_t] = X_t$.

Azuma–Hoeffding inequality (bounded differences)

Let $(X_t)_{t=0}^n$ be a martingale w.r.t. $(\mathcal{F}_t)$.
Assume bounded increments: for constants $c_1, \ldots, c_n$,

$$|X_t - X_{t-1}| \le c_t \qquad \text{almost surely for each } t = 1, \ldots, n.$$

Then for all $u > 0$,

$$\mathbb{P}(X_n - X_0 \ge u) \;\le\; \exp\left(-\frac{u^2}{2\sum_{t=1}^n c_t^2}\right),$$

and similarly

$$\mathbb{P}(|X_n - X_0| \ge u) \;\le\; 2\exp\left(-\frac{u^2}{2\sum_{t=1}^n c_t^2}\right).$$

How to read Azuma–Hoeffding (intuition + special case)

Intuition: $X_n - X_0$ is a sum of martingale differences,

$$X_n - X_0 = \sum_{t=1}^{n} (X_t - X_{t-1}),$$

and each term is mean-zero given the past and bounded by c_t.

Special case: if all steps are bounded by the same c (i.e. $c_t = c$), then

$$\mathbb{P}(|X_n - X_0| \geq u) \leq 2 \exp\left(-\frac{u^2}{2nc^2}\right).$$

So typical fluctuations are on the order of $c\sqrt{n}$.

Proof of Azuma-Hoeffding

Almost exactly the same as Hoeffding, if we are careful to use definitions correctly.

Azuma–Hoeffding: Step 1 (Chernoff + tower property)

Let $(X_t)_{t=0}^n$ be a martingale w.r.t. $(\mathcal{F}_t)$ and set

$$D_t := X_t - X_{t-1} \qquad (t = 1, \dots, n), \quad \text{so} \quad X_n - X_0 = \sum_{t=1}^{n} D_t,$$

with $\mathbb{E}[D_t \mid \mathcal{F}_{t-1}] = 0$. As before we start with Chernoff bound. For any $\lambda > 0$,

$$\Pr(X_n - X_0 \ge t) = \Pr\left(e^{\lambda(X_n - X_0)} \ge e^{\lambda t}\right) \le e^{-\lambda t}\, \mathbb{E} e^{\lambda(X_n - X_0)}.$$

Key difference vs. i.i.d.: we *cannot* factor the MGF as a product. Instead we peel one step at a time using conditional expectation:

$$\mathbb{E} e^{\lambda(X_n - X_0)} = \mathbb{E}\left[\exp\left(\lambda \sum_{s=1}^{n} D_s\right)\right] = \mathbb{E}\left[\exp\left(\lambda \sum_{s=1}^{n-1} D_s\right) \mathbb{E}\left[e^{\lambda D_n} \mid \mathcal{F}_{n-1}\right]\right].$$

Now apply Hoeffding's lemma *conditionally* to D_n given $\mathcal{F}_{n-1}$ (using $\mathbb{E}[D_n \mid \mathcal{F}_{n-1}] = 0$ and $|D_n| \le c_n$), to bound $\mathbb{E}[e^{\lambda D_n} \mid \mathcal{F}_{n-1}]$. Then iterate for $n-1, n-2, \dots, 1$.

Remainder of proof (same as before)

We obtain

$$\mathbb{E}e^{\lambda(X_n - X_0)} \leq \exp\left(\frac{\lambda^2}{8}\sum_{s=1}^{n} c_s^2\right)$$

and so

$$\Pr(X_n - X_0 \geq t) \;\leq\; \exp\Big(-\lambda t + \frac{\lambda^2}{8}\sum_{s=1}^{n} c_s^2\Big).$$

Minimizing the exponent gives $\lambda^\star = \dfrac{4t}{\sum_{s=1}^{n} c_s^2}$, yielding

$$\Pr(X_n - X_0 \geq t) \;\leq\; \exp\Big(-\frac{2t^2}{\sum_{s=1}^{n} c_s^2}\Big).$$

The lower-tail bound follows by applying the same argument to $-X_n$.

Next time

We will finish the description and analysis of the smarter UCB (Upper Confidence Bound) algorithm which has been very influential.