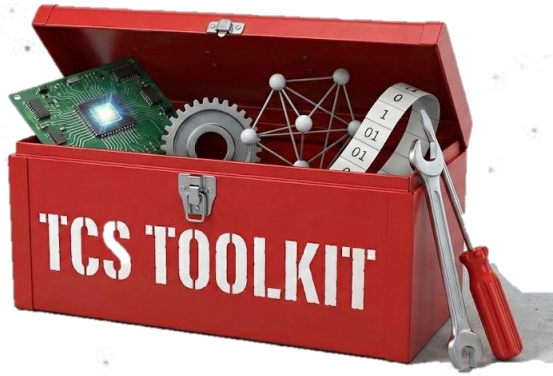
CS 58500 – Theoretical Computer Science Toolkit

Lecture 3 (01/27)

Concentration Inequality II



https://ruizhezhang.com/course_spring_2026.html

Today's Lecture

- Chernoff Bound (Additive Form)
- Chernoff Bound (Multiplicative Form)
- Application: Discrepancy Theory

Remark: in many TCS papers,

"Chernoff bound" ≈ exponential convergence ≈ Chernoff/Hoeffding/Bernstein inequality

Question

Suppose you have a coin with unknown biased probability $p \in [0, 1]$. How many flips to estimate p with approximation error $\pm\epsilon$?

- Flip n times and output $\hat{p} \coloneqq \#\,\mathrm{Heads}/n$
- Suppose we want to achieve the guarantee

$$\Pr[|\hat{p} - p| \leq \epsilon] \geq 1 - \delta$$

- How small can we choose n?

Question

Suppose you have a coin with unknown biased probability $p \in [0, 1]$. How many flips to estimate p with approximation error $\pm\epsilon$?

- Define *iid* random variables $X_1, X_2, \ldots, X_n$ such that

$$X_i := \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

- For $Z := X_1 + \cdots + X_n$, how large should n be such that

$$\Pr[|Z - pn| \geq \epsilon n] \leq \delta$$

- $\mathbb{E}[Z] = pn$
- $\mathrm{Var}[X_i] = p - p^2$ and $\mathrm{Var}[Z] = np(1-p)$

Question

Suppose you have a coin with unknown biased probability $p \in [0, 1]$. How many flips to estimate p with approximation error $\pm\epsilon$?

- $\mathbb{E}[Z] = pn$ and $\mathrm{Var}[Z] = np(1-p)$
- Chebyshev inequality:

$$\Pr[|Z - \mathbb{E}[Z]| \geq t] \leq \frac{\mathrm{Var}[Z]}{t^2}$$

$$\Pr[|Z - pn| \geq \epsilon n] \leq \frac{np(1-p)}{\epsilon^2 n^2} = \frac{p(1-p)}{\epsilon^2 n} \leq \delta$$

$$n \geq \frac{p(1-p)}{\epsilon^2}\frac{1}{\delta}$$

Chernoff Bound

- Let $X_i' \coloneqq X_i - \mathbb{E}[X_i]$ be the centered version of X_i and $Z' \coloneqq X_1' + \cdots + X_n' = Z - \mathbb{E}[Z]$
- Consider the moment generating function:

$$\mathbb{E}\left[e^{\theta Z'}\right] = \mathbb{E}\left[e^{\theta \sum_i X_i'}\right] = \mathbb{E}\left[\prod_i e^{\theta X_i'}\right] = \prod_i \mathbb{E}\left[e^{\theta X_i'}\right]$$

- By Markov inequality,

$$\Pr[Z' \geq \epsilon n] = \Pr\left[e^{\theta Z'} \geq e^{\theta \epsilon n}\right] \leq \frac{\mathbb{E}\left[e^{\theta Z'}\right]}{e^{\theta \epsilon n}} = \frac{\prod_i \mathbb{E}\left[e^{\theta X_i'}\right]}{e^{\theta \epsilon n}}$$

- This inequality holds for any $\theta > 0$. So, we can optimize over t to get the tightest bound:

$$\Pr[Z' \geq \epsilon n] \leq \inf_{\theta > 0} e^{-\theta \epsilon n} \prod_i \mathbb{E}\left[e^{\theta X_i'}\right]$$

Chernoff Bound

Lemma (MGF bound). If $a \le X \le b$ then for every $\theta \in \mathbb{R}$ we have

$$\mathbb{E}\left[e^{\theta(X-\mathbb{E}[X])}\right] \le e^{\theta^2(b-a)^2/8}$$

Proof.

- Wlog, we may assume $\mathbb{E}[X] = 0$ and $\theta \ge 0$
- Define the log-MGF:

$$\psi(\theta) \coloneqq \log \mathbb{E}\left[e^{\theta X}\right]$$

- $\psi(0) = \log \mathbb{E}[1] = 0$
- We compute the derivatives:

$$\psi'(\theta) = \frac{\mathbb{E}\left[Xe^{\theta X}\right]}{\mathbb{E}[e^{\theta X}]}, \qquad \psi''(\theta) = \frac{\mathbb{E}\left[X^2e^{\theta X}\right]}{\mathbb{E}[e^{\theta X}]} - \left(\frac{\mathbb{E}\left[Xe^{\theta X}\right]}{\mathbb{E}[e^{\theta X}]}\right)^2$$

Chernoff Bound

Lemma (MGF bound). If $a \le X \le b$ then for every $\theta \in \mathbb{R}$ we have

$$\mathbb{E}\left[e^{\theta(X-\mathbb{E}[X])}\right] \le e^{\theta^2(b-a)^2/8}$$

Proof.

- We compute the derivatives:

$$\psi'(\theta) = \frac{\mathbb{E}\left[Xe^{\theta X}\right]}{\mathbb{E}[e^{\theta X}]}, \qquad \psi''(\theta) = \frac{\mathbb{E}\left[X^2e^{\theta X}\right]}{\mathbb{E}[e^{\theta X}]} - \left(\frac{\mathbb{E}\left[Xe^{\theta X}\right]}{\mathbb{E}[e^{\theta X}]}\right)^2$$

← Looks like a variance

- $\psi'(0) = \mathbb{E}[X] = 0$ by assumption
- Let $\mu := \mathrm{Law}(X)$. Then we can define its exponential tilt μ_θ:

$$\frac{\mu_\theta(\mathrm{d}x)}{\mu(\mathrm{d}x)} := \frac{e^{\theta x}}{\mathbb{E}[e^{\theta X}]}$$

- You can easily check that μ_θ is indeed a probability distribution for any $\theta \ge 0$

Chernoff Bound

Lemma (MGF bound). If $a \le X \le b$ then for every $\theta \in \mathbb{R}$ we have

$$\mathbb{E}\left[e^{\theta(X-\mathbb{E}[X])}\right] \le e^{\theta^2(b-a)^2/8}$$

Proof.

- We compute the derivatives:

$$\psi'(\theta) = \mathbb{E}_{\mu_\theta}[X], \qquad \psi''(\theta) = \mathrm{Var}_{\mu_\theta}[X]$$

- Recall the range bound for variance:

$$\psi''(\theta) = \mathrm{Var}_{\mu_\theta}[X] \le (b-a)^2/4$$

- Now, we integrate ψ'' twice from 0 to θ and using the fact that $\psi(0) = \psi'(0) = 0$:

$$\int_0^\theta \psi''(\lambda)\mathrm{d}\lambda = \psi'(\theta), \qquad \int_0^\theta \psi'(\lambda)\mathrm{d}\lambda = \psi(\theta)$$

$$\Longrightarrow \quad \psi(\theta) = \int_0^\theta \int_0^s \psi''(\lambda)\mathrm{d}\lambda\,\mathrm{d}s \le \int_0^\theta \int_0^s \frac{1}{4}(b-a)^2\mathrm{d}\lambda\,\mathrm{d}s = \int_0^\theta \frac{s}{4}(b-a)^2\mathrm{d}s = \frac{\theta^2}{8}(b-a)^2$$

■

Chernoff Bound

Lemma (MGF bound). If $a \le X \le b$ then for every $\theta \in \mathbb{R}$ we have

$$\mathbb{E}\left[e^{\theta(X-\mathbb{E}[X])}\right] \le e^{\theta^2(b-a)^2/8}$$

- This lemma implies that

$$\Pr[Z' \ge \epsilon n] \le \inf_{\theta>0} e^{-\theta\epsilon n} \prod_i \mathbb{E}\left[e^{\theta X_i'}\right] \le \inf_{\theta>0} e^{-\theta\epsilon n}\left(e^{\theta^2(b-a)^2/8}\right)^n$$

$$= \exp\left(\inf_{\theta>0} -\theta\epsilon n + \theta^2(b-a)^2 n/8\right)$$

- The quadratic function is minimized at $\theta = \frac{4\epsilon}{(b-a)^2}$:

$$\Pr[Z - \mathbb{E}[Z] \ge \epsilon n] = \Pr[Z' \ge \epsilon n] \le e^{-\frac{2\epsilon^2}{(b-a)^2}n}$$

- What about the other direction (the lower tail)?

Chernoff Bound

Lemma (MGF bound). If $a \le X \le b$ then for every $\theta \in \mathbb{R}$ we have

$$\mathbb{E}\left[e^{\theta(X-\mathbb{E}[X])}\right] \le e^{\theta^2(b-a)^2/8}$$

- For the lower tail:

$$\begin{aligned}
\Pr[Z - \mathbb{E}[Z] \le -\epsilon n] &= \Pr[Z' \le -\epsilon n] \\
&= \Pr\left[e^{\theta Z'} \ge e^{-\theta\epsilon n}\right] \qquad (\theta \le 0) \\
&\le \inf_{\theta \le 0} e^{+\theta\epsilon n} \prod_i \mathbb{E}\left[e^{\theta X_i'}\right] \\
&= \exp\left(\inf_{\theta \le 0} \theta\epsilon n + \theta^2(b-a)^2 n/8\right) \\
&= \exp\left(-\frac{2\epsilon^2}{(b-a)^2} n\right)
\end{aligned}$$

Chernoff Bound (Additive Form)

Theorem (Hoeffding's inequality)**.**

Let $X_1, \dots, X_n$ be independent real random variables such that $a_i \le X_i \le b_i$ for all $i \in [n]$.
Let $Z \coloneqq \sum_{i=1}^{n} X_i$ and introduce the variance proxy

$$\nu \coloneqq \frac{1}{4}\sum_{i=1}^{n}(b_i - a_i)^2 .$$

Then

$$\Pr[|Z - \mathbb{E}[Z]| \ge t\sqrt{\nu}] \le 2e^{-t^2/2} \qquad \forall t \ge 0$$

- For our biased coin problem, $X_i \in \{0,1\}$ and thus, $\nu = n/4$

$$\Pr[|Z - \mathbb{E}[Z]| \ge \epsilon n] \le 2e^{-2\epsilon^2 n} \le \delta \quad \Longrightarrow \quad n \ge \frac{1}{2\epsilon^2}\log\left(\frac{2}{\delta}\right)$$

Chebyshev:

$$n \ge \frac{p(1-p)}{\epsilon^2}\frac{1}{\delta}$$

Improvement?

In the whole proof, there are only two inequalities:

1. By Markov inequality, we obtain the so-called Laplace transform:

$$\Pr[Z - \mathbb{E}[Z] \geq t] \leq \inf_{\theta \geq 0} e^{-t\theta} \prod_i \mathbb{E}\left[e^{\theta(X_i - \mathbb{E}[X_i])}\right],$$

$$\Pr[Z - \mathbb{E}[Z] \leq -t] \leq \inf_{\theta \leq 0} e^{t\theta} \prod_i \mathbb{E}\left[e^{\theta(X_i - \mathbb{E}[X_i])}\right]$$

Cramér's theorem:

This is tight for iid sum

2. MGF bound: If $a \leq X \leq b$ then for every $\theta \in \mathbb{R}$ we have

$$\mathbb{E}\left[e^{\theta(X - \mathbb{E}[X])}\right] \leq e^{\theta^2 (b-a)^2/8}$$

- In our case, $X_1, \ldots, X_n$ are iid Bernoulli random variables

$$\mathbb{E}\left[e^{\theta(X_i - p)}\right] = pe^{\theta(1-p)} + (1-p)e^{-\theta p} = e^{-\theta p}\left(1 + \left(e^{\theta} - 1\right)p\right)$$

Improvement?

$$\begin{aligned}\Pr[Z - \mathbb{E}[Z] \geq t] &\leq \inf_{\theta \geq 0} e^{-t\theta} \prod_i \mathbb{E}\left[e^{\theta(X_i - \mathbb{E}[X_i])}\right] \\ &= \inf_{\theta \geq 0} e^{-t\theta} e^{-np\theta} \left(1 + \left(e^{\theta} - 1\right)p\right)^n \\ &\leq \inf_{\theta \geq 0} e^{-t\theta} e^{-np\theta} e^{np\left(e^{\theta}-1\right)} \\ &= \exp\left(-np + \inf_{\theta \geq 0} -(t + np)\theta + npe^{\theta}\right)\end{aligned}$$

- The minimizer is at $\theta = \log\left(\frac{t+np}{np}\right)$:

$$\begin{aligned}\Pr[Z - \mathbb{E}[Z] \geq t] &\leq \exp\left(-np - (t + np)\log\left(\frac{t + np}{np}\right) + (t + np)\right) \\ &= \exp\left(-(t + np)\log\left(\frac{t + np}{np}\right) + t\right)\end{aligned}$$

Improvement?

$$\Pr[Z - \mathbb{E}[Z] \geq t] \leq \inf_{\theta \geq 0} e^{-t\theta} \prod_i \mathbb{E}\left[e^{\theta(X_i - \mathbb{E}[X_i])}\right]$$
$$= \inf_{\theta \geq 0} e^{-t\theta} e^{-np\theta} \left(1 + \left(e^{\theta} - 1\right)p\right)^n$$
$$\leq \inf_{\theta \geq 0} e^{-t\theta} e^{-np\theta} e^{np\left(e^{\theta} - 1\right)}$$
$$= \exp\left(-np + \inf_{\theta \geq 0} -(t + np)\theta + npe^{\theta}\right)$$

- The minimizer is at $\theta = \log\left(\frac{t+np}{np}\right)$:

$$\Pr[Z - \mathbb{E}[Z] \geq tnp] \leq \exp\left(-np - (tnp + np)\log\left(\frac{tnp + np}{np}\right) + (tnp + np)\right)$$
$$= \exp(-(1 + t)np\log(1 + t) + tnp)$$
$$= \left(\frac{e^t}{(1 + t)^{1+t}}\right)^{np}$$

Improvement?

$$\begin{aligned}\Pr[Z - \mathbb{E}[Z] \le -tnp] &\le \inf_{\theta \le 0} e^{tnp\theta} \prod_i \mathbb{E}\left[e^{\theta(X_i - \mathbb{E}[X_i])}\right] \\ &= \inf_{\theta \le 0} e^{tnp\theta} e^{-np\theta} \left(1 + \left(e^{\theta} - 1\right)p\right)^n \\ &\le \inf_{\theta \le 0} e^{tnp\theta} e^{-np\theta} e^{np\left(e^{\theta} - 1\right)} \\ &= \exp\left(-np + \inf_{\theta \le 0}(t-1)np\theta + npe^{\theta}\right)\end{aligned}$$

- The minimizer is at $\theta = \log(1-t)$:

$$\begin{aligned}\Pr[Z - \mathbb{E}[Z] \le -tnp] &\le \exp(-np - (1-t)np\log(1-t) + (1-t)np) \\ &= \exp(-(1-t)np\log(1-t) - tnp) \\ &= \left(\frac{e^{-t}}{(1-t)^{1-t}}\right)^{np}\end{aligned}$$

Chernoff Bound (Multiplicative Form)

Theorem (Chernoff bound).

Let $X_1, \dots, X_n$ be iid Bernoulli random variables with $\mathbb{E}[X_i] = p$. Let $Z \coloneqq \sum_{i=1}^{n} X_i$. Then

$$\Pr[Z \geq (1+t)np] \leq \left(\frac{e^t}{(1+t)^{1+t}}\right)^{np} \qquad \forall t \geq 0$$

$$\Pr[Z \leq (1-t)np] \leq \left(\frac{e^{-t}}{(1-t)^{1-t}}\right)^{np} \qquad \forall t \in [0,1]$$

Notice the asymmetric tails:

- For the right tail, when t is large, it behaves like t^{-t} (i.e., a Gamma distribution)
- For the left tail, when t is small, it behaves like $e^{-t^2/2}$ (i.e., a Gaussian distribution)

$$\log(\cdots) = -t - (1-t)\log(1-t) = -t - (1-t)\left(-t - t^2/2 + \mathcal{O}(t^3)\right) = -t^2/2 + \mathcal{O}(t^3)$$

Chernoff Bound (Multiplicative Form)

(Source: Joel A. Tropp)

- $Z \sim \mathrm{Bin}(n, p) \to \mathrm{Poi}(np)$ for large n
- The probability of a Poisson random variable $X \sim \mathrm{Poi}(np)$ at $k > np$ is

$$\frac{e^{-np}(np)^k}{k!} = e^{-\Theta(k \log k)} < e^{-\Omega(k^2)}$$

Chernoff Bound (Multiplicative Form)

Theorem (Chernoff bound, user-friendly version)**.**

Let $X_1, \dots, X_n$ be independent Bernoulli random variables with $\mathbb{E}[X_i] = p_i$. Let $Z \coloneqq \sum_{i=1}^{n} X_i$ and $\mu \coloneqq \mathbb{E}[Z]$. Then

$$\Pr[Z \geq (1+t)\mu] \leq e^{-\frac{t^2}{2+t}\mu} \qquad \forall t \geq 0$$

$$\Pr[Z \leq (1-t)\mu] \leq e^{-t^2\mu/2} \qquad \forall t \in [0,1]$$

Moreover,

$$\Pr[|Z-\mu| \geq t\mu] \leq 2e^{-t^2\mu/3} \qquad \forall t \in [0,1]$$

Chernoff Bound (Multiplicative Form)

Theorem (Chernoff bound, user-friendly version)**.**

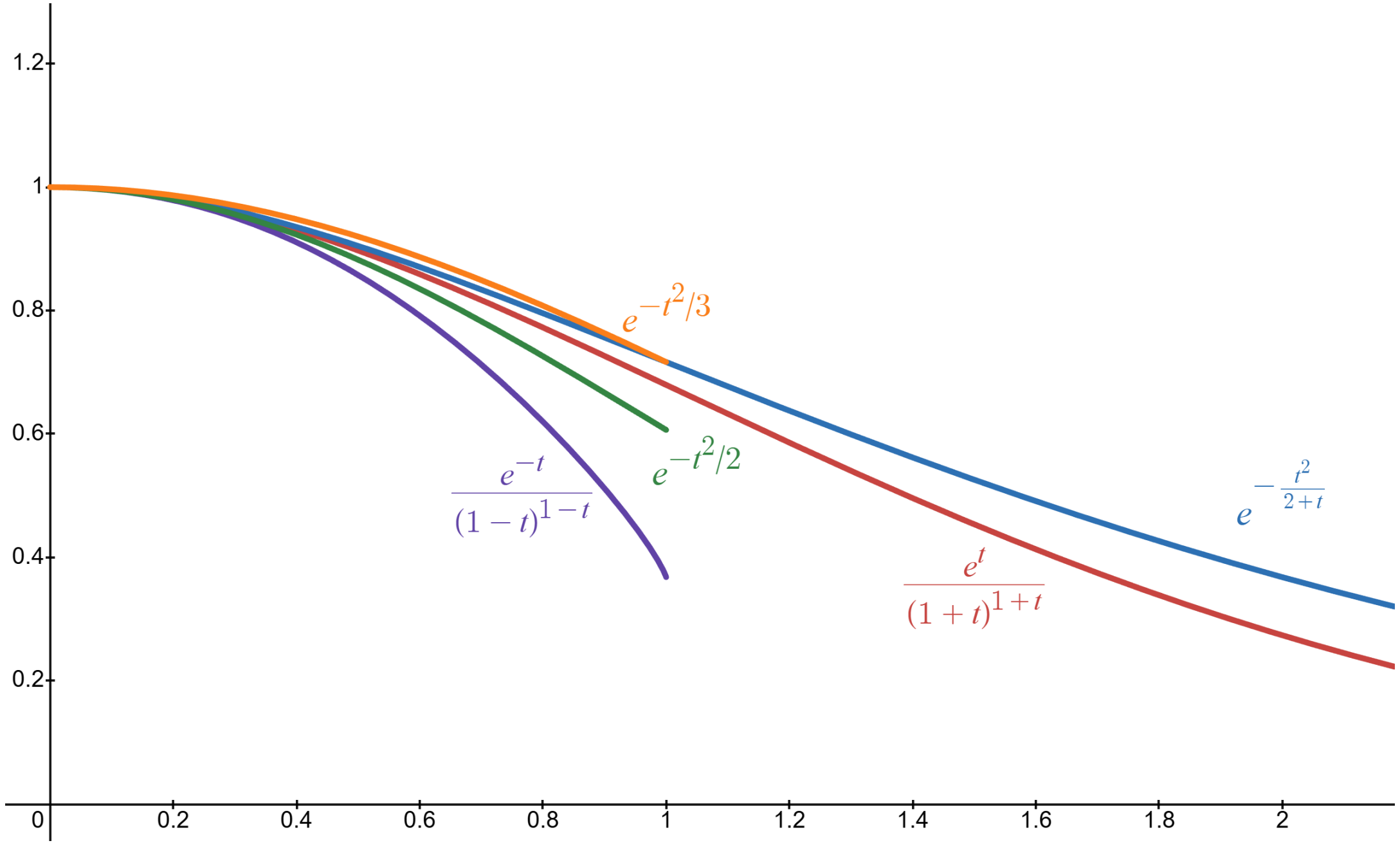
Let $X_1, \dots, X_n$ be independent Bernoulli random variables with $\mathbb{E}[X_i] = p_i$. Let $Z \coloneqq \sum_{i=1}^{n} X_i$ and $\mu \coloneqq \mathbb{E}[Z]$. Then

$$\Pr[Z \geq (1+t)\mu] \leq e^{-\frac{t^2}{2+t}\mu} \qquad \forall t \geq 0$$

Proof.

- $\log\left(\frac{e^t}{(1+t)^{1+t}}\right) = t - (1+t)\log(1+t)$
- $\log(1+t) \geq \frac{2t}{2+t}$
- $t - (1+t)\log(1+t) \leq t - (1+t)\frac{2t}{2+t} = \frac{-t^2}{2+t}$

■


1.2
1
0.8
0.6
0.4
0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
$e^{-t^2/3}$
$e^{-t^2/2}$
$\frac{e^{-t}}{(1-t)^{1-t}}$
$e^{-\frac{t^2}{2+t}}$
$\frac{e^{t}}{(1+t)^{1+t}}$

Chernoff Bound (Multiplicative Form)

Theorem (Chernoff bound, user-friendly version)**.**

Let $X_1, \dots, X_n$ be independent Bernoulli random variables with $\mathbb{E}[X_i] = p_i$. Let $Z \coloneqq \sum_{i=1}^{n} X_i$ and $\mu \coloneqq \mathbb{E}[Z]$. Then

$$\Pr[|Z - \mu| \geq t\mu] \leq 2e^{-t^2\mu/3} \qquad \forall t \in [0,1]$$

- For our biased coin question,

$$\Pr[|Z - np| \geq (\epsilon/p)np] \leq 2\exp\left(-\frac{\epsilon^2 n}{3p}\right) \leq \delta \quad \Longrightarrow \quad n \geq \frac{3p}{\epsilon^2}\log\left(\frac{2}{\delta}\right)$$

Chernoff Bound (Multiplicative Form)

Theorem (Chernoff for positive, bounded random variables)**.**

Let $X_1, \dots, X_n$ be independent random variables such that $0 \leq X_i \leq b$ for all $i \in [n]$. Let $Z \coloneqq \sum_{i=1}^{n} X_i$ and $\mu \coloneqq \mathbb{E}[Z]$. Then

$$\Pr[Z \geq (1+t)\mu] \leq \left(\frac{e^t}{(1+t)^{1+t}}\right)^{\mu/b} \qquad \forall t \geq 0$$

$$\Pr[Z \leq (1-t)\mu] \leq \left(\frac{e^{-t}}{(1-t)^{1-t}}\right)^{\mu/b} \qquad \forall t \in [0,1]$$

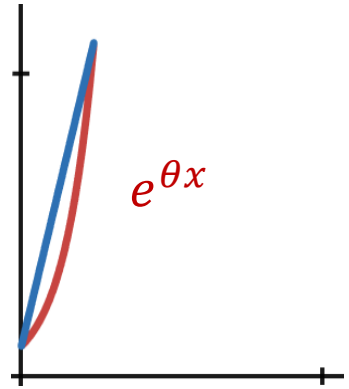
Proof of Chernoff for Positive Bounded RVs

- When X is a Bernoulli random variable,

$$\mathbb{E}\left[e^{\theta X}\right] = 1 + (e^{\theta} - 1)p$$

- If we only know $0 \leq X \leq b$, then by the convexity of $e^{\theta x}$ and Jensen inequality,

$$e^{\theta x} \leq 1 + \left(e^{\theta b} - 1\right)x/b \qquad \forall x \in [0, b]$$



- Thus,

$$\mathbb{E}\left[e^{\theta X}\right] \leq 1 + \frac{e^{\theta b} - 1}{b}\mathbb{E}[X]$$

- Following the same proof, you'll get the term involving

$$\prod_i \left(1 + \frac{e^{\theta b} - 1}{b}\mathbb{E}[X_i]\right) = \exp\left(\sum_i \log\left(1 + \frac{e^{\theta b} - 1}{b}\mathbb{E}[X_i]\right)\right) \leq \exp\left(\sum_i \frac{e^{\theta b} - 1}{b}\mathbb{E}[X_i]\right)$$

$$= \exp\left(\frac{e^{\theta b} - 1}{b}\mu\right)$$

Chernoff Bound (Additive Form)

Theorem (Chernoff-Hoeffding inequality)**.**

Let $X_1, \dots, X_n$ be iid Bernoulli random variables with $\mathbb{E}[X_i] = p$. Let $Z \coloneqq \sum_{i=1}^{n} X_i$. Then

$$\Pr[Z \geq (p+\epsilon)n] \leq e^{-nD(p+\epsilon\|p)} \qquad \forall \epsilon \in (0, 1-p)$$
$$\Pr[Z \leq (p-\epsilon)\mu] \leq e^{-nD(p-\epsilon\|p)} \qquad \forall \epsilon \in (0, p)$$

$D(p\|q)$ is the relative entropy defined as:

$$D(p\|q) = p \log\left(\frac{p}{q}\right) + (1-p)\log\left(\frac{1-p}{1-q}\right)$$

Proof idea:

- Apply the Laplace transform and use the closed-form of MGF (same as our proofs)
- No more approximation and directly optimize it

Application: Discrepancy Theory

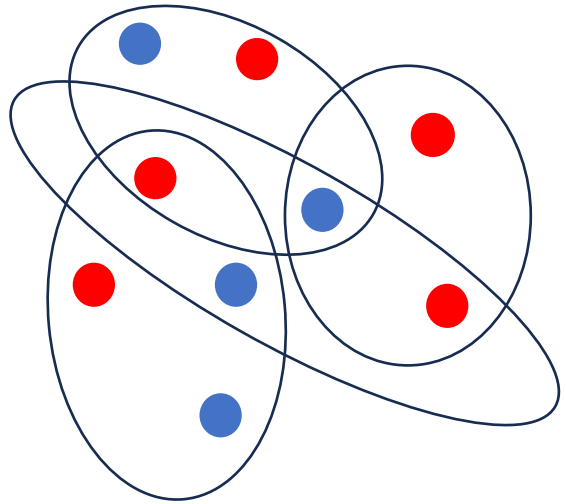
Given a hypergraph (or a set system), how to color each vertex in red or blue such that in each hyperedge (or set), the number of red vertices is roughly equal to the number of blue vertices?

- Universe $U = \{1,2,\dots,n\}$
- Set system $\mathcal{F} = \{S_1, S_2, \dots, S_m\}$
- For a coloring $\chi\colon [n] \to \{-1,1\}$, its discrepancy is defined as:

$$\mathrm{Disc}(\chi) \coloneqq \max_{i\in[m]} \left|\sum_{j\in S} \chi(j)\right|$$

- And the discrepancy of the set system is defined as

$$\mathrm{Disc}(\mathcal{F}) \coloneqq \min_{\chi} \mathrm{Disc}(\chi)$$



$\mathrm{Disc}(\mathcal{F}) = 1$

Application: Discrepancy Theory

Lemma. Let $\mathcal{F}$ be a collection of m subsets of $[n]$. Then there is a coloring $\chi\colon [n] \to \{-1,1\}$ such that $\mathrm{Disc}(\chi) = \mathcal{O}\left(\sqrt{n \log m}\right)$.

Proof.

- Let χ be a uniformly random coloring, i.e., we assign ± 1 to each element iid uniformly at random
- $\mathbb{E}[\chi(i)] = 0$ for any $i \in [n]$
- For each subset $S \in \mathcal{F}$, let $Z \coloneqq \sum_{i \in S} \chi(i)$. By Hoeffding's inequality,

$$\Pr[|Z| \geq t\sqrt{v}] \leq 2e^{-t^2/2}, \qquad v = \frac{1}{4}\sum_{i \in S}\left(1-(-1)\right)^2 = |S| \leq n$$

$$t \coloneqq 2\sqrt{\log m} \quad \Longrightarrow \quad \Pr\left[|Z| \geq 2\sqrt{n \log m}\right] \leq 2e^{-2\log m} = 2/m^2$$

- By union bound over all m subsets, with probability $\geq 1 - m \cdot 2/m^2 > 0$, $\mathrm{disc}(\chi) \leq 2\sqrt{n \log m}$

■

Application: Discrepancy Theory

Lemma. Let $\mathcal{F}$ be a collection of m subsets of $[n]$. Then there is a coloring $\chi: [n] \to \{-1,1\}$ such that $\mathrm{Disc}(\chi) = \mathcal{O}\left(\sqrt{n \log m}\right)$.

- This bound is not tight!

Theorem (Six standard deviations suffice, Spencer '85)**.**

Let $\mathcal{F}$ be a collection of n subsets of $[n]$. Then there is a coloring $\chi: [n] \to \{-1,1\}$ such that $\mathrm{Disc}(\chi) \leq 6\sqrt{n}$.

More generally, if $\mathcal{F}$ is a collection of $m \geq n$ subsets of $[n]$. Then there is a coloring $\chi: [n] \to \{-1,1\}$ such that $\mathrm{Disc}(\chi) \leq \mathcal{O}\left(\sqrt{n \log(2m/n)}\right)$.

Bernstein Inequality

Theorem (Bernstein)**.**

Let $X_1, \dots, X_n$ be independent random variables such that $|X_i - \mathbb{E}[X_i]| \leq b$ for all $i \in [n]$. Let $Z \coloneqq \sum_{i=1}^{n} X_i$. Then

$$\Pr[|Z - \mathbb{E}[Z]| \geq t] \leq 2 \exp\left(-\frac{t^2/2}{\mathrm{Var}[Z] + bt/3}\right) \qquad \forall t > 0$$

- For large t, the tail bound looks like $\exp(-2t/(3b))$, an exponential tail
- For small or medium t, the tail bound looks like $\exp(-t^2/(2\mathrm{Var}[Z]))$, a Gaussian tail with the true variance (instead of the variance proxy ν in Hoeffding's inequality)