# Lecture 06: Hoeffding's Bound

- One of the "easy to evaluate" forms of the Chernoff bound states the following. Let $(\mathbb{X}_1, \dots, \mathbb{X}_n)$ be independent random variables such that $0 \leqslant \mathbb{X}_i \leqslant 1$, for each $1 \leqslant i \leqslant n$. Let $\mathbb{S}_{n,p} = \mathbb{X}_1 + \cdots + \mathbb{X}_n$, where $np = \mathbb{E}\left[\mathbb{S}_{n,p}\right]$. The bound we proved was

$$\mathbb{P}\left[\mathbb{S}_{n,p} \geqslant n(p+\varepsilon)\right] \leqslant \exp(-2\varepsilon^2 n)$$

This bound can be equivalently be written as follows.

$$\mathbb{P}\left[\mathbb{S}_{n,p} \geqslant \mathbb{E}\left[\mathbb{S}_{n,p}\right] + E\right] \leqslant \exp(-2E^2/n),$$

where we substituted $E = n\varepsilon$.

# Introduction II

- Let us change the random variables $\mathbb{X}_i$ as follows. Define $\mathbb{X}_i^* = \mathbb{X}_i - \mathbb{E}[\mathbb{X}_i]$. Note that we have $\mathbb{E}\left[\mathbb{X}_i^*\right] = 0$, for each $1 \leqslant i \leqslant n$. Define $\mathbb{S}_n^* = \mathbb{X}_1^* + \cdots + \mathbb{X}_n^*$. We have $\mathbb{E}\left[\mathbb{S}_{n,p}^*\right] = 0$. Note that $\mathbb{P}[\mathbb{S}_n^* \geqslant E]$ is identical to $\mathbb{P}\left[\mathbb{S}_{n,p} \geqslant \mathbb{E}\left[\mathbb{S}_{n,p}\right] + E\right]$. So, we can claim that

$$\mathbb{P}\left[\mathbb{S}_n^* \geqslant E\right] \leqslant \exp(-2E^2/n)$$

- Hoeffding's bound shall generalize this bound. In the next slide we shall state the formal statement that we shall prove.

# Hoeffding's Bound

We formally state the Hoeffding's bound.

**Theorem**

*Let $\mathbb{X}_1, \ldots, \mathbb{X}_n$ are independent random variables such that $\mathbb{E}[\mathbb{X}_i] = 0$ and $a_i \leqslant \mathbb{X}_i \leqslant b_i$, for each $1 \leqslant i \leqslant n$. We define $\mathbb{S}_n = \mathbb{X}_1 + \cdots + \mathbb{X}_n$. Then, the following bound holds*

$$\mathbb{P}[\mathbb{S}_n \geqslant E] \leqslant \exp\left(-\frac{2E^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

**Comment.** When $b_i - a_i = 1$ for all $1 \leqslant i \leqslant n$, we get the "first form" of "easy to evaluate" Chernoff bound.
We emphasize that the Chernoff bound is much stronger than the Hoeffding's bound. It is only that the first form of the easy to evaluate Chernoff bound is a special case of the Hoeffding's bound.

# Proof of the Hoeffding's Bound I

- We are interested in upper-bounding

$$\mathbb{P}\left[\mathbb{S}_n \geqslant E\right]$$

- For any $h > 0$, we have

$$\mathbb{P}\left[\mathbb{S}_n \geqslant E\right] = \mathbb{P}\left[\exp(h\mathbb{S}_n) \geqslant \exp(hE)\right]$$

- By Markov inequality, we have

$$\mathbb{P}\left[\exp(h\mathbb{S}_n) \geqslant \exp(hE)\right] \leqslant \frac{\mathbb{E}\left[\exp(h\mathbb{S}_n)\right]}{\exp(hE)} = \frac{\mathbb{E}\left[\prod_{i=1}^{n} \exp(h\mathbb{X}_i)\right]}{\exp(hE)}$$

- By independence of $\mathbb{X}_1, \ldots, \mathbb{X}_n$, we have

$$\frac{\mathbb{E}\left[\prod_{i=1}^{n} \exp(h\mathbb{X}_i)\right]}{\exp(hE)} = \frac{\prod_{i=1}^{n} \mathbb{E}\left[\exp(h\mathbb{X}_i)\right]}{\exp(hE)}$$

- Now, our objective is to upper-bound $\mathbb{E}\left[\exp(h\mathbb{X}_i)\right]$, where $a_i \leqslant \mathbb{X}_i \leqslant b_i$ and $\mathbb{E}\left[\mathbb{X}_i\right] = 0$. We have done this type of upper-bound earlier. We showed that the maximum is achieved by a distribution $\mathbb{X}_i^*$ that puts the entire probability mass either at $a_i$ or $b_i$.

  Suppose the probability mass of $\mathbb{X}_i^*$ at $a_i$ is $p$. Then, the probability mass of $\mathbb{X}_i^*$ at $b_i$ is $(1-p)$. The expected value of $\mathbb{X}_i^*$ is 0. So, we have $pa_i + (1-p)b_i = 0$. That is, we have $p = b_i/(b_i - a_i)$. Therefore, we have $(1-p) = (-a_i)/(b_i - a_i)$.

  For this distribution $\mathbb{X}_i^*$, the expected value of $\exp(h\mathbb{X}_i^*)$ is

$$\frac{b_i}{b_i - a_i}\exp(ha_i) + \frac{-a_i}{b_i - a_i}\exp(hb_i)$$

So, we can conclude that

$$\mathbb{E}\left[\exp(h\mathbb{X}_i)\right] \leqslant \frac{b_i}{b_i - a_i}\exp(ha_i) - \frac{a_i}{b_i - a_i}\exp(hb_i)$$

- Substituting this upper-bound, we obtain

$$\frac{\prod_{i=1}^{n}\mathbb{E}\left[\exp(h\mathbb{X}_i)\right]}{\exp(hE)} \leqslant \frac{\prod_{i=1}^{n}\frac{b_i}{b_i-a_i}\exp(ha_i) - \frac{a_i}{b_i-a_i}\exp(hb_i)}{\exp(hE)}$$

- At this juncture, we need a crucial lemma, namely, Hoeffding's lemma.

**Lemma (Hoeffding's Lemma)**

*For $a \leqslant 0 \leqslant b$, we have*

$$\frac{b}{b-a}\exp(ha) - \frac{a}{b-a}\exp(hb) \leqslant \exp(h^2(b-a)^2/8)$$

We shall prove this result later in this lecture using Lagrange form of the Taylor's remainder theorem. Currently, let us use this result without proof and go forward.

Using Hoeffding's lemma, we have

$$\frac{\prod_{i=1}^n \frac{b_i}{b_i-a_i}\exp(ha_i) - \frac{a_i}{b_i-a_i}\exp(hb_i)}{\exp(hE)} \leqslant \frac{\prod_{i=1}^n \exp(h^2(b_i-a_i)^2/8)}{\exp(hE)}$$

$$= \frac{\exp\left(h^2\sum_{i=1}^n (b_i-a_i)^2/8\right)}{\exp(hE)}$$

- At this point, we have proven that, for any $h > 0$, we have

$$\mathbb{P}\left[\mathbb{S}_n \geqslant E\right] \leqslant \frac{\exp(h^2\alpha)}{\exp(hE)},$$

where $\alpha = \sum_{i=1}^n (b_i - a_i)^2/8$ Our objective, is to pick $h = h^*$ that minimized the RHS. That is, equivalently, minimize $h^2\alpha - hE$.

Clearly, at $h^* = E/2\alpha$ the quantity is minimized.

- For $h = h^*$, we have

$$\exp(h^2\alpha - hE) = \exp(E^2/4\alpha - E^2/2\alpha) = \exp(-E^2/4\alpha)$$

Substituting the value of $\alpha$, we get

$$\exp(-E^2/4\alpha) = \exp\left(-\frac{2E^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

This completes the proof of Hoeffding's bound.

# Proof of Hoeffding's Lemma I

- For $a \leqslant 0 \leqslant b$ and $h > 0$, we need to prove the following

$$\frac{b}{b-a}\exp(ha) - \frac{a}{b-a}\exp(hb) \leqslant \exp(h^2(b-a)^2/8)$$

- Let us perform some variable substitutions. Let $x = h(b-a)$ and $p = b/(b-a)$. Therefore, we can rewrite the old expressions using the new variables as follows

$$\begin{aligned} hb &= px \\ ha &= -(1-p)x \\ -\frac{a}{b-a} &= 1-p \end{aligned}$$

# Proof of Hoeffding's Lemma II

- So, we need to prove

$$p\exp(-(1-p)x) + (1-p)\exp(px) \leqslant \exp(x^2/8)$$

This statement is equivalent to proving

$$\exp(px)\left(p\exp(-x) + (1-p)\right) \leqslant \exp(x^2/8)$$

Taking log both sides, this statement is equivalent to proving

$$f(x) := px + \log(1-p+p\exp(-x)) \leqslant x^2/8$$

- Now, let us compute the derivatives of $f(x)$

$$f(x) = px + \log(1-p+p\exp(-x))$$

$$f'(x) = p - \frac{p\exp(-x)}{1-p+p\exp(-x)} = p - \frac{p}{(1-p)\exp(x)+p}$$

$$f''(x) = \frac{p(1-p)\exp(x)}{((1-p)\exp(x)+p)^2}$$

# Proof of Hoeffding's Lemma III

- We apply the Lagrange form of the Taylor's remainder theorem. For every $x$, there exists $\theta \in [0,1]$ such that

$$f(x) = f(0) + f'(0)x + f''(\theta x)\frac{x^2}{2}$$

Note that $f(0) = 0$ and $f'(0) = 0$. Let us upper-bound $f''(\theta x)$ as follows

$$\begin{aligned} f''(\theta x) &= \frac{(1-p)\exp(\theta x)\cdot p}{\left((1-p)\exp(\theta x)+p\right)^2} \\ &\leqslant \frac{1}{\left((1-p)\exp(\theta x)+p\right)^2}\left(\frac{(1-p)\exp(\theta x)+p}{2}\right)^2 \quad \text{By AM-GM} \\ &= \frac{1}{4} \end{aligned}$$

- So, we conclude that

$$f(x) \leqslant 0 + 0 \cdot x + \frac{1}{4} \cdot \frac{x^2}{2} = x^2/8$$

This completes the proof of Hoeffding's Lemma.