Hoeffding's inequality

Let $X_1, X_2, \ldots, X_n$ be independent random variables such that $X_i \in [a_i, b_i]$ with probability one. Then, for all $\varepsilon > 0$,

$$① \quad \Pr\left[\sum_{i=1}^{n} X_i \geq E\left[\sum_{i=1}^{n} X_i\right] + \varepsilon\right] \leq e^{-2\varepsilon^2 / \sum_{i=1}^{n} (b_i - a_i)^2}$$

$$② \quad \Pr\left[\sum_{i=1}^{n} X_i \leq E\left[\sum_{i=1}^{n} X_i\right] - \varepsilon\right] \leq e^{-2\varepsilon^2 / \sum_{i=1}^{n} (b_i - a_i)^2}$$

$$③ \quad \Pr\left[\left|\sum_{i=1}^{n} X_i - E\left[\sum_{i=1}^{n} X_i\right]\right| \geq \varepsilon\right] \leq 2e^{-2\varepsilon^2 / \sum_{i=1}^{n} (b_i - a_i)^2}$$

---

① $\Rightarrow$ ② : Apply to ① to $-X_1, -X_2, \ldots, -X_n$

---

(1) AND (2) $\Rightarrow$ (3) : Union bound

It suffices to prove (1)

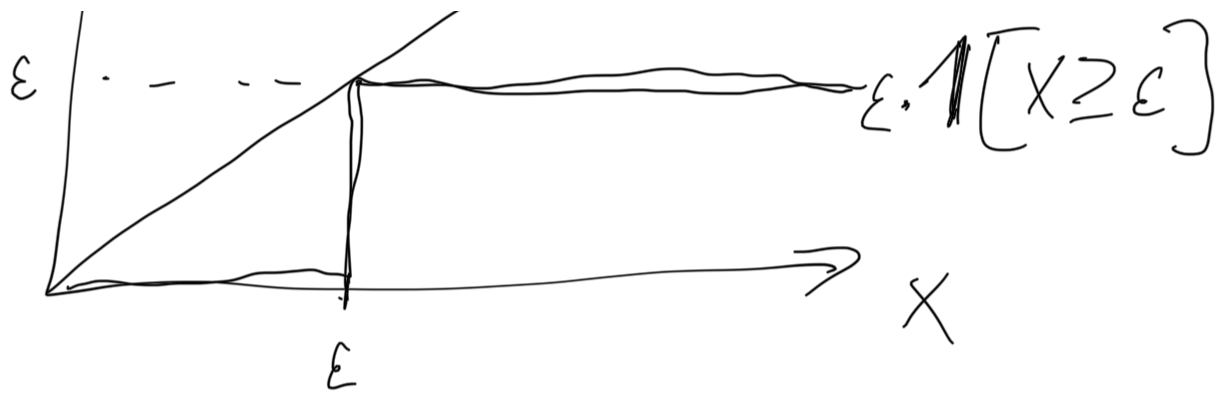
Detour #1 :

Markov's inequality

Let X be a non-negative random variable. Then for any $\varepsilon > 0$,

$$\Pr[X \geq \varepsilon] \leq \frac{E[X]}{\varepsilon}.$$

Proof of Markov's inequality:



$$\varepsilon \cdot \mathbb{1}[X \geq \varepsilon] \leq X$$

$$\varepsilon \cdot \underbrace{E\left[\mathbb{1}[X \geq \varepsilon]\right]}_{\Pr[X \geq \varepsilon]} \leq E[X]$$

$$\varepsilon \cdot \Pr[X \geq \varepsilon] \leq E[X] \quad // \tfrac{1}{\varepsilon}$$

$$\Pr[X \geq \varepsilon] \leq \frac{E[X]}{\varepsilon} \qquad \blacksquare$$

Proof Hoeffding's inequality (1):

Use Chernoff's bounding technique. The idea is to apply Markov's inequality to moment generating function of $\sum_{i=1}^{n} X_i$.

$$\Pr\left[\sum_i X_i \geq E\sum_i X_i + \varepsilon\right]$$ // $e^{s\cdot}$

$$= \Pr\left[e^{s\sum_i X_i} \geq e^{sE\sum_i X_i + s\varepsilon}\right]$$

$$\leq \frac{E\left[e^{s\sum_i X_i}\right]}{e^{sE\left[\sum_i X_i\right] + s\varepsilon}}$$

$$= \frac{\prod_{i=1}^{n} E[e^{sX_i}]}{\prod_{i=1}^{n} e^{sE[X_i]}} \cdot e^{-s\varepsilon}$$

$e^{sX_1}, e^{sX_2}, \dots$ are independent

$$= e^{-s\varepsilon} \prod_{i=1}^{n} E\left[e^{s(X_i - E[X_i])}\right]$$

Let $z_i = X_i - E[X_i]$ $z_i \in [a_i - E[X_i], b_i - E[X_i]]$

Detour #2

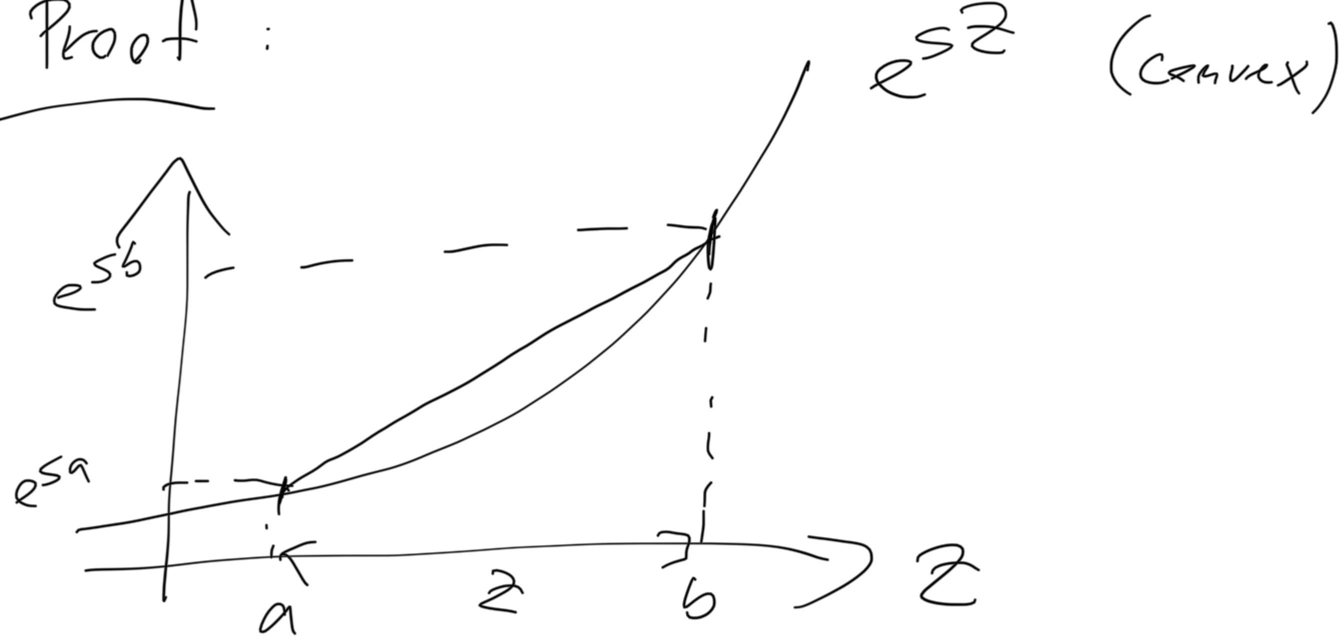
Hoeffding's lemma

Let z be random variable such that $E[z] = 0$ and $z \in [a, b]$ w.p. one.

Then for all $s \geq 0$,

$$E[e^{sz}] \leq e^{s^2(b-a)^2/8}$$

<u>Proof</u> :



$$e^{sz} \leq e^{sa} + (z-a) \cdot \frac{(e^{sb}-e^{sa})}{b-a}$$

$$E[e^{sz}] \leq e^{sa} + \frac{a(e^{sa}-e^{sb})}{b-a}$$

$$= \frac{b}{\ } \cdot e^{sa} - \frac{a}{\ } e^{sb}$$

$b-a$ $\qquad$ $b-a$

$\boxed{p = -\frac{a}{b-a}}$ $\qquad$ $\boxed{a = -p(b-a) = p(a-b)}$

$\boxed{p \geq 0}$ $\qquad = (1-p)e^{sa} + pe^{sb}$

$$= e^{sa}\left(1-p+pe^{s(b-a)}\right)$$

$$= e^{-ps(b-a)}\left(1-p+pe^{s(b-a)}\right)$$

$\boxed{u = s(b-a)}$

$$= e^{-pu}\left(1-p+pe^{u}\right)$$

$$\boxed{\phi(u) = -pu + \log(1-p+pe^{u})}$$

$$= e^{\phi(u)}$$

Need to prove $\quad \phi(u) \leq s^2(b-a)^2/8$.

$$\phi(u) = \phi(0) + u \cdot \phi'(0) + \frac{u^2}{2} \cdot \phi''(v)$$

for some $v \in [0,u]$

$$\phi(0) = 0$$

$$\phi'(u) = -p + \frac{pe^u}{1-p+pe^u}$$

$$\phi'(0) = 0$$

$$\phi''(u) = \frac{pe^u(1-p+pe^u) - (pe^u)^2}{(1-p+pe^u)^2}$$

$$= \frac{p(1-p)e^u}{(1-p+pe^u)^2} = \frac{AB}{(A+B)^2} \leq \frac{1}{4}$$

$$\sqrt{AB} \leq \frac{A+B}{2} \Rightarrow \frac{AB}{(A+B)^2} \leq \frac{1}{4}$$

$$\phi(u) = \frac{u^2}{2} \cdot \phi''(v) \leq \frac{u^2}{8} = \frac{s^2(b-a)^2}{8}$$

$$\Pr[\cdots] \le e^{-s\varepsilon} \prod_{i=1} E[e^{s t_i}]$$

$$\le e^{-s\varepsilon} \prod_{i=1}^{n} e^{s^2(b_i-a_i)^2/8}$$

We can choose s that minimizes the last expression

$$f(s) = -s\varepsilon + \frac{s^2}{8}\sum_{i=1}^{n}(b_i-a_i)^2$$

$$\Pr[\cdots] \le e^{f(s)}$$

Minimizer of $f(s)$ is at $f'(s^*)=0$.

$$f'(s) = -\varepsilon + \frac{s}{4}\sum_{i=1}^{n}(b_i-a_i)^2$$

Minimizer is at

$$s^* = \frac{4\varepsilon}{\sum_{i=1}^{n}(b_i-a_i)^2}$$

$$f(s^*) = -\frac{4\varepsilon^2}{\sum_{i=1}^{n}(b_i-a_i)^2} + \frac{16\varepsilon^2}{8\left(\sum(b_i-a_i)^2\right)^2}\cdot\sum_{i=1}^{n}(b_i-a_i)^2$$

$4\varepsilon^2$ $\qquad$ $2\varepsilon^2$

$$= -\frac{}{\sum_{i=1}^{n}(b_i - a_i)^2} + \frac{}{\sum_{i=1}^{n}(b_i - a_i)^2}$$

$$= -\frac{2\varepsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}$$