Ex 6: Hoeffding's inequality. $\mathbb{P}[S_n - n\mu \geq \delta] \leq e^{-\frac{2\delta^2}{n(b-a)^2}}$

Let $X_1, X_2, \cdots$ iid real valued r.v. w/ mean μ and moment: $G(\lambda) = \log(\mathbb{E}[e^{-\lambda(X_n-\mu)}])$. ~~Assume all X_i's are bounded: $X_i \in [a, b]$~~

Define $S_n = \sum_{i=1}^{n} X_i$.

Conc. Ineq.

$$\mathbb{P}[S_n - n\mu \geq \delta] \overset{\{\lambda>0\}}{=} \mathbb{P}[e^{\lambda(S_n - n\mu)} \geq e^{\lambda\delta}] =$$

$$= (\text{Markov Inequality}) \leq \mathbb{E}[e^{\lambda[S_n - n\mu]}] \cdot e^{-\lambda\delta} =$$

$$= e^{-\lambda\delta} \cdot \mathbb{E}[e^{\sum(X_i-\mu)\cdot\lambda}] = e^{-\lambda\delta} \cdot \prod_{i=1}^{n} \mathbb{E}[e^{(X_i-\mu)\cdot\lambda}] =$$

$$= e^{-\lambda\delta} \cdot \prod_{i=1}^{n} e^{G(\lambda)} = e^{-(\lambda\delta - nG(\lambda))}$$

For a fixed δ, we would like to pick $\lambda(\delta)$ which will result the tightest bound, i.e. we want $(\lambda\delta - nG(\lambda))$ as large as possible.

$$\Rightarrow \boxed{\mathbb{P}[S_n - n\mu \geq \delta] \leq e^{-\sup_{\lambda>0}(\lambda\delta - nG(\lambda))}} \quad (1)$$

$\boxed{\text{Hoeffding's Lemma}}$ $\mathbb{E}[e^{\lambda X}] \leq \exp\left(\frac{\lambda^2(b-a)^2}{8}\right)$ given that $\mathbb{E}[X] = 0$ and X is bounded $[a', b']$

proof: Since $e^{-\lambda x}$ is a convex function:

$$e^{\lambda x} \leq \frac{b'-x}{b'-a'} e^{\lambda a'} + \frac{x-a'}{b'-a'} e^{\lambda b'}$$

then,

$$\mathbb{E}[e^{\lambda x}] \leq e^{\lambda a'} \cdot \frac{b' - \mathbb{E}[X]}{b'-a'} + \frac{\mathbb{E}[X] - a'}{b'-a'} \cdot e^{\lambda b'} = \frac{b'}{b'-a'} e^{\lambda a'} - \frac{a'}{b'-a'} e^{\lambda b'} =$$

$$= \exp\left(\log\left(e^{\lambda a'} \cdot \left[\frac{b'}{b'-a'} - \frac{a'}{b'-a'} \cdot e^{\lambda(b'-a')}\right]\right)\right) =$$

$$\exp\left[\lambda a' + \log\left(1 + \frac{a'}{b'-a'} - \frac{a'}{b'-a'} e^{-\lambda(b'-a')}\right)\right] \boxed{=}$$

Change in variables: $h = \lambda(b'-a')$ $\quad p = -\frac{a'}{b'-a'} \Rightarrow h \cdot p = -\lambda \cdot a'$

$$\Rightarrow \boxed{=} \exp\Big[\underbrace{-hp + \log(1 + p\cdot e^{h} - p)}_{L(h)}\Big] = \exp\left[L(h)\right]$$

$$\mathbb{E}\left[e^{\lambda x}\right] \leq e^{L(h)}$$

From the Taylor's expansion, we know that there is an $\varepsilon \in (0,h)$ such that

$$L(h) = L(0) + h\cdot L'(0) + \frac{h^2}{2}\cdot L''(\varepsilon)$$

$L(0) = 0$

$L'(0) = 0$

$$L''(h) = e^{h}\frac{p}{1+pe^{h}-p} - e^{2h}\frac{p^2}{(1+pe^{h}-p)^2} =$$

$$= \underbrace{\frac{pe^{h}}{pe^{h}+1-p}}_{z}\Big[1 - \underbrace{\frac{pe^{h}}{pe^{h}+1-p}}_{z}\Big] = z\cdot(1-z) \leq \frac{1}{4} \quad \text{for } \forall h$$

$$\Rightarrow L(h) \leq \frac{h^2}{2}\cdot\frac{1}{4} = \frac{\lambda^2(b'-a')^2}{8}$$

Then $\mathbb{E}\left[e^{\lambda x}\right] \leq e^{L(h)} \leq e^{\frac{\lambda^2(b'-a')^2}{8}}$ ■

From (1) we have $\mathbb{P}\left[S_n - n\mu \geq \delta\right] \leq e^{-\sup_{\lambda>0}(\lambda\delta - n\cdot G(\lambda))}$

$$G(\lambda) = \log\left[\mathbb{E}\left[e^{-\lambda(x_n-\mu)}\right]\right] \leq \log e^{\frac{\lambda^2(b-a)^2}{8}} = \frac{\lambda^2(b-a)^2}{8}$$

$$\Rightarrow \mathbb{P}\left[S_n - n\mu \geq \delta\right] \leq e^{-\sup_{\lambda>0}\left(\lambda\delta - n\cdot\frac{\lambda^2(b-a)^2}{8}\right)} \quad (2)$$

The exponent of the right hand side is quadratic =>
we can find the maximum by diff. w/ respect to λ and setting
to 0

$$\Rightarrow \delta - \lambda \cdot \frac{n(b-a)^2}{4} = 0 \Rightarrow \lambda^* = \frac{4\delta}{n(b-a)^2}$$

Plugging back into (2) we get:

$$\mathbb{P}[S_n - n\mu \geq \delta] \leq e^{-\frac{2\delta^2}{n(b-a)^2}}$$

Ex. 6: UCB proof

Proof of Upper Confidence Bound algorithm

$$\boxed{\text{UCB:} \quad b_a(t) = \hat{\theta}_a(t) + \sqrt{\frac{2\log(t)}{N_a(t)}}}$$

where $\hat{\theta}(t)$ = empirical reward of a at time t

$N_a(t)$ = # of times a played up to time t

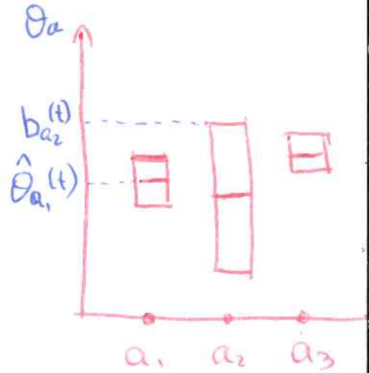
At each round t, select an arm w/ highest index $b_a(t)$

Theorem: Under UCB, the # of times $a \neq a^*$ satisfies $\boxed{\mathbb{E}[N_a(T)] \leq \frac{8\log T}{(\theta_{a^*} - \theta_a)^2} + \frac{\pi^2}{6}}$

Preliminaries: Hoeffding inequality

Bounded random variable $X_n \in [a, b]$

$$P\left(\sum_{i=1}^{n} X_i - n \cdot \mu > \delta\right) \leq e^{-\frac{2\delta^2}{n(b-a)^2}}$$



Proof:

$$N_a(t) = \underbrace{1}_{\text{select each of the K arms 1 time}} + \sum_{t=K+1}^{T} \mathbb{1}\{a(t) = a\} = 1 + \sum_{t=K+1}^{T} \Big[\mathbb{1}\{a(t)=a, N_a(t) \geq \ell\} + \underbrace{\mathbb{1}\{a(t)=a, N_a(t) < \ell\}}_{\leq \ell - 1}\Big]$$

$$\leq \ell + \sum_{t=K+1}^{T} \Big[\mathbb{1}\{a(t)=a, N_a(t) \geq \ell\}\Big] = \ell + \sum_{t=K+1}^{T} \mathbb{1}\{N_a(t) \geq \ell\} \cdot \mathbb{1}\{a(t)=a\}$$

$$= \ell + \sum_{t=K+1}^{T} \mathbb{1}\{N_a(t) \geq \ell\} \cdot \mathbb{1}\{b_a(t-1) \geq b_{a^*}(t-1), b_a(t-1) \geq b_{a'}(t-1)\}$$

$a' \in \{1, \dots K\} \setminus \{a, a^*\}$

$$\leq \ell + \sum_{t=K+1}^{T} \mathbb{1}\{N_a(t) \geq \ell\} \cdot \mathbb{1}\{b_a(t-1) \geq b_{a^*}(t-1)\} =$$

$$= \ell + \sum_{t=K+1}^{T} \mathbb{1}\{N_a(t) \geq \ell\} \cdot \mathbb{1}\left\{\hat{\theta}_a + \sqrt{\frac{2\log(t-1)}{N_a(t-1)}} \geq \hat{\theta}_{a^*} + \sqrt{\frac{2\log(t-1)}{N_{a^*}(t-1)}}\right\}$$

use this information here

$$\leq \ell + \sum_{t=K+1}^{T} \mathbb{1}\left\{\max_{\ell \leq s < t}\left(\hat{\theta}_a^{s} + \sqrt{\frac{2\log(t-1)}{s}}\right) \geq \min_{s' < t}\left(\hat{\theta}_{a^*}^{s'} + \sqrt{\frac{2\log(t-1)}{s'}}\right)\right\}$$

$$\leq \ell + \sum_{t=K+1}^{T} \sum_{s<t} \sum_{\ell \leq s' < t} \mathbb{1}\left\{\hat{\theta}_a^{s} + \sqrt{\frac{2\log(t-1)}{s}} \geq \hat{\theta}_{a^*}^{s'} + \sqrt{\frac{2\log(t-1)}{s'}}\right\}$$

Next, we try to analyze $\mathbb{1}\left\{\underbrace{\hat{\theta}_a^{s}}_{A} + \underbrace{\sqrt{\frac{2\log(t-1)}{s}}}_{B} \geq \underbrace{\hat{\theta}_{a^*}^{s'}}_{C} + \underbrace{\sqrt{\frac{2\log(t-1)}{s'}}}_{D}\right\}$

This inequality implies that at lease one of (*), (**), (***) is true.

$$A + B \geq C + D \iff \left[\begin{array}{l} A + B \geq \theta^* \\ C + D \leq \theta^* \end{array}\right. \longrightarrow \left[\begin{array}{ll} C + D \leq \theta^* & (*) \\ \left[\begin{array}{ll} A - B \geq \theta^* & (**) \\ 2B > \theta^* - \theta_a & (***) \end{array}\right. \end{array}\right.$$

Check...

Now we compute the probabilities of each of these events

$$(*): \mathbb{P}[C + D \leq \theta^*] = \mathbb{P}\left[\hat{\theta}_{a^*}^{s'} - \theta^* \leq -\sqrt{\frac{2\log(t-1)}{s'}}\right] \leq (\text{Hoefding's inequality})$$

$$\leq e^{-2\left[\sqrt{\frac{2\log(t-1)}{s'}}\, s'\right]^2} = \frac{1}{(t-1)^{4s'}} \leq \frac{1}{(t-1)^4}$$

$$(**): \mathbb{P}[A - B \geq \theta_a] = \mathbb{P}\left[\hat{\theta}_a^{s} - \theta_a \geq \sqrt{\frac{2\log(t-1)}{s}}\right] \leq \cdots = \frac{1}{(t-1)^4}$$

(***): Lets take $\ell = 8\log(T)/(\theta^* - \theta_a)^2$ and $s \geq \ell \Rightarrow$

$$2B = 2\sqrt{\frac{2\log(t-1)}{s}} \leq 2\sqrt{\frac{2\log(t-1)}{\ell}} = 2\sqrt{\frac{2\log(t-1)}{8\cdot\log T}}\cdot|\theta^* - \theta_a| \leq$$

$$\leq |\theta^* - \theta_a|$$

On the other hand we know that $2B > \theta^* - \theta_a$ from (***)

$\Rightarrow$ This event is false for $s > \frac{8\log T}{(\theta^* - \theta_a)^2}$

$$\Rightarrow N_a(t) \leq \ell + \sum_{t=k+1}^{T} \sum_{s<t} \sum_{\ell \leq s' \leq t} \mathbb{1}\{A + B \geq C + D\} \leq$$

$$\leq \frac{8\log T}{(\theta^* - \theta_a)^2} + \sum_{t=k+1}^{T} \sum_{s,s'=1}^{t} 2t^{-4} \leq \boxed{\frac{8\log T}{(\theta^* - \theta_a)^2} + \frac{\pi^2}{3}}$$