Week 7 — Lecture 35

$$\Pr\left\{ \hat{\mu}_i - \sqrt{\frac{\alpha \log t}{T_i(t-1)}} > \mu_i \right\}$$

$$= \Pr\left\{ \hat{\mu}_i - \mu_i > \sqrt{\frac{\alpha \log t}{T_i(t-1)}} \right\} \quad \text{—— (1)}$$

We know:

For n samples are all distributions being 1-subgaussia
↳ for estimate $\hat{\mu}_i$

$$\Pr\left\{ \hat{\mu}_i - \mu_i > \varepsilon \right\} \leq \exp\left\{ -n\varepsilon^2/2 \right\} \quad \text{—— (2)}$$

But cannot be applied completely on (1) as
$T_i(t-1)$ is a random quantity

∴ Need to do something about $T_i(t-1)$

$$1 \leq T_i(t-1) \leq t-1$$

$$\Pr\left\{ \hat{\mu}_i - \mu_i > \sqrt{\frac{\alpha \log t}{T_i(t-1)}} \right\}$$

$$\leq \Pr\left\{ \exists\, 1 \leq s \leq t-1,\ \hat{\mu}_{is} - \mu_i > \sqrt{\frac{\alpha \log t}{s}} \right\}$$

[s is deterministic thus can apply (2)]

$$\leq \sum_{s=1}^{t-1} \Pr\left\{ \hat{\mu}_{is} - \mu_i > \sqrt{\frac{\alpha \log t}{s}} \right\}$$

$$\leq \sum_{s=1}^{t-1} \exp\left\{ -\frac{s\alpha \log t}{2s} \right\}$$

$$= \sum_{s=1}^{t-1} t^{-\alpha/2}$$

$$\leq t^{-\alpha/2} \cdot t = t^{1-\alpha/2}$$

If $\alpha = 6$ then

$$\leq 1/t^2$$

So now we have:

$$\sum_{t=1}^{n} \Pr\left\{ I_t = i, \text{①} \text{ holds} \right\}$$ (① from previous lecture)

$$\leq \sum_{t=1}^{n} \Pr\left\{ \text{①} \text{ holds} \right\}$$

$$\leq \sum_{t=1}^{n} 1/t^2$$

$$\leq \sum_{t=1}^{\infty} 1/t^2 = \pi^2/6$$

Similarly:

$$\sum_{t=1}^{n} \Pr\{ I_t = i, \text{ ② holds} \} \quad (\text{② from previous lecture})$$

$$\leq \pi^2/6$$

Thus finally:

$$\mathbb{E}[T_i(n)] \leq \left\lceil \frac{4\alpha \log n}{\Delta_i^2} \right\rceil + \pi^2/6 + \pi^2/6$$

$$\leq \frac{4\alpha \log n}{\Delta_i^2} + \pi^2/3 + 1$$

Thus Regret for UCB is:

$$R(UCB, n) = \sum_{i=1}^{K} \mathbb{E}[T_i(n)] \Delta_i$$

$$= \sum_{i \neq i^*} \frac{4\alpha \log n}{\Delta_i} + \sum_{i \neq i^*} \Delta_i (\pi^2/3 + 1)$$

All this was for $\alpha = 6$

$$\therefore R(UCB, n) \leq 24 \sum_{i \neq i^*} \frac{\log n}{\Delta_i} + \sum \Delta_i (\pi^2/3 + 1)$$

$$\leq 24(K-1) \frac{\log n}{\Delta} + \sum_{i \neq i^*} (\pi^2/3 + 1) \Delta_i \quad \text{③}$$

$$R(UCB, n) = O(K \log n / \Delta)$$

Thus we get sub-linear regret

This bound is fixed for an instant, that is when the instance is fixed, the gaps are thus defined.

These bounds are thus called:

<u>Problem-dependent Bounds</u>

<u>Problem independent bounds</u>

$$R(\pi, n) = \sum_{i=1}^{K} \mathbb{E}[T_i(n)] \Delta_i \quad , \text{ we know.}$$

$$= \sum_{i=1}^{K} \sqrt{\mathbb{E}[T_i(n)]} \sqrt{\mathbb{E}[T_i(n)]} \; \Delta_i$$

Applying Cauchy-Schwartz inequality

by treating: $\sqrt{\mathbb{E}[T_i(n)]} = a_i$ where $a = (a_1, a_2 \ldots a_K)$

$\sqrt{\mathbb{E}[T_i(n)]} \Delta_i = b_i$ where $b = (b_1, b_2 \ldots b_K)$

$\therefore$

$$\Rightarrow R(\pi, n) \leq \sqrt{\sum \mathbb{E}[T_i(n)] \sum \mathbb{E}[T_i(n)] \Delta_i^2}$$

$$= \sqrt{n \sum_{i=1}^{K} \mathbb{E}[T_i(n)]\, \Delta_i^2}$$

$$\leq \sqrt{n \sum_{i \neq i^*} \left(\frac{4 \alpha \log n}{\Delta_i^2} + \pi^2/3 + 1 \right) \Delta_i^2}$$

$$= \sqrt{n \sum_{i \neq i^*} (4 \alpha \log n) + n \sum_{i \neq i^*} (\pi^2/3 + 1)\, \Delta_i^2}$$

$$= \sqrt{n(K-1)\, 4 \alpha \log n + n \sum_{i \neq i^*} (\pi^2/3 + 1)\, \Delta_i^2}$$

Assume: Support of distributions $[0, 1]$

Then all $\mu_i \in [0, 1]$

$\Delta_i \in [0, 1]$

$\therefore$ Can upper-bound the above as:

$$\leq \sqrt{n(K-1)\, 4 \alpha \log n + n(K-1)(\pi^2/3 + 1)}$$

$$R(\pi, n) \leq \sqrt{n(K-1)\, 4 \alpha \log n + n(K-1)(\pi^2/3 + 1)}$$

Problem independent bounds

for special case of support bounded by $[0, 1]$

Here the order is:

$$\underline{R(\pi, n) = O\left(\sqrt{n(K-1)}\right)}$$

Both Problem dependent and independent bounds are sublinear in n, but being $\sim O(\log n)$, problem dependent bounds decay faster with n than problem independent bound.