Week 9 – Lecture 41

An environment:

$$\nu = (\mu_1, 0, 0, 0 \ldots \ldots 0)$$

where $\mu_1 = [0, 1/2]$

$$i = \arg\min_{j>1} \mathbb{E}[N_j(T)]$$

where i has been played less than $T/(K-1)$ rounds

$$\therefore \mathbb{E}[N_i(T)] \leqslant T/(K-1)$$

2nd env:

$$\nu' = (\mu_1, 0, 0, \ldots \underset{\uparrow i}{2\mu_1}, \ldots 0)$$

$$\boxed{\text{We know: } R_T(\pi, \nu) = \sum_{i=1}^{K} \Delta_i \, \mathbb{E}_\nu[N_i(T)]}$$

1) $R_T(\pi, \nu) \geqslant P_\nu(N_1(T) \leqslant T/2) \dfrac{T\mu_1}{2}$

optimal arm has been played less than T/2 times [1 is optimal arm]

Thus sub-optimal arm has been played T/2 times.

And on those rounds would have incurred regret of $T\mu_1/2$.

2) $R_T(\pi, \nu') \geq P_{\nu'}(N_1(T) > T/2)\, \frac{T\mu_1}{2}$

$\therefore$

$R_T(\pi, \nu) + R_T(\pi, \nu')$

$$\geq \frac{T\mu_1}{2}\left(P_\nu(\underbrace{N_1(T) \leq T/2}_{\text{Event } A}) + \left(P_{\nu'}(\underbrace{N_1(T) > T/2}_{\text{Event } A^c})\right)\right)$$

[Event A & A^c w.r.t. two diff. probability measures]

$\therefore$ Using the divergence result:

$$\geq \frac{T\mu_1}{2}\, \frac{1}{2} \exp\left(-D(P_\nu \| P_{\nu'})\right)$$

[Underlying sample space for both environments

$\Omega = \mathbb{R}^T \times [K]^T$]

Invoking the lemma proven in the earlier lecture:

$$D(P_\nu \| P_{\nu'}) = \sum_{k=1}^{K} \mathbb{E}_\nu[N_k(T)]\, D(P_k \| P_k')$$

Since ν & ν' are same except at position i,

$\therefore$ apart from at position i, everything would be zero, $\therefore$ we have:

$$D(P_\nu \| P_{\nu'}) = \mathbb{E}_\nu[N_i(T)]\, D(P_i \| P_i')$$

The environments are gaussian $\therefore$ we can get expression for divergence : $\hookrightarrow$ [Assume $\sigma^2 = 1$]

$$D(P_\nu \| P_{\nu'}) = \mathbb{E}_\nu[N_i(T)] \frac{(0 - 2\mu_1)^2}{2\sigma^2}$$

$$= \mathbb{E}_\nu[N_i(T)] \frac{2\mu_1^2}{\sigma^2}$$

$$\leq \left(\frac{T}{K-1}\right) 2\mu_1^2 \quad [\text{Assume } \sigma^2 = 1]$$

Plugging this into the regret summation:

$$R_T(\pi, \nu) + R_T(\pi, \nu') \geq \frac{T\mu_1}{4} \exp\left(\frac{-2T\mu_1^2}{K-1}\right)$$

Let $\mu_1 = \sqrt{\frac{K-1}{4T}}$ [Already assumed that $T > K-1$]

$\leq \sqrt{1/4} = \frac{1}{2}$ [Thus satisfied $\mu_1 = [0, 1/2]$ & $2\mu_1 = [0, 1]$]

$\therefore$

$R_T(\pi, \nu) + R_T(\pi, \nu')$

______ .

$$\geq \frac{1}{2 \times 4} \sqrt{T(K-1)} \frac{1}{\sqrt{e}}$$

For minmax lower bound the regret to be incurred would be max of ½ of the summation of regrets

$\therefore$

$$2 \max\left\{R_T(\pi, \nu), R_T(\pi, \nu')\right\}$$
$$\geq R_T(\pi, \nu) + R_T(\pi, \nu')$$

$$\Rightarrow \max\left\{R_T(\pi, \nu), R_T(\pi, \nu')\right\}$$
$$\geq \left(R_T(\pi, \nu) + R_T(\pi, \nu')\right)/2$$

$$\therefore R_T^*(\mathcal{E}^K) \geq \frac{1}{16\sqrt{e}} \sqrt{T(K-1)}$$

Hence Proved