Assignment 7

Q.1 : The ε- Greedy Algorithm

$\forall\, t = 1, 2 \dots$

a) with probability $(1-\varepsilon)$ choose arm with highest estimate, $A_t = \arg\max_i \hat{\mu}_i(t-1)$

b) with probability ε, select arm at random with uniform probability, i.e. $1/K$ where K = no. of arms

We knew for Stochastic bandits :

$$\bar{R}(\pi, n) = \sum_{i=1}^{K} \mathbb{E}\left[T_i(n)\right] \Delta_i \qquad \text{--- (1)}$$

where $\Delta_i = \max_a \mu_a - \mu_i$

and $T_i(n)$ = no. of arm pulls

$$= \sum_{t=1}^{n} 1\{I_t = i\}$$

Now for ε-greedy

$$T_i(n) = \frac{\varepsilon}{K} + (1-\varepsilon) \Pr\{\hat{\mu}_i \geq \arg\max_j \hat{\mu}_j(n-1)_{\{j \neq i}$$

$$\mathbb{E}\left[T_i(n)\right] = \mathbb{E}\left[\frac{\varepsilon}{K} + (1-\varepsilon)\Pr\left\{\hat{\mu}_i \geq \underset{j}{\arg\max}\,\hat{\mu}_j(n-1)_{\{j \neq i\}}\right.\right.$$

$$= \frac{\varepsilon n}{K} + \mathbb{E}\left[(1-\varepsilon)\Pr\left\{\hat{\mu}_i \geq \underset{j}{\arg\max}\,\hat{\mu}_j(n-1)_{j\neq i}\right\}\right].$$

Substituting this in eq (1)

$$\bar{R}(\varepsilon\text{-greedy}, n) = \sum_{i=1}^{K}\left[\underbrace{\frac{\varepsilon n}{K}\Delta_i}_{\text{Term (i)}} + \underbrace{\mathbb{E}\left[(1-\varepsilon)\Pr\left\{\hat{\mu}_i \geq \underset{j}{\arg\max}\,\hat{\mu}_j(n-1)\right\}\right]\Delta_i}_{\text{Term (ii)}}\right]$$

Now if for any i $\hat{\mu}_i \geq \underset{j}{\arg\max}\,\hat{\mu}_j(n-1)$, then the $\Delta_i = 0$, as per defination of Δ_i

Thus term ii would be equal to zero.

The regret thus simplifies to:

$$\bar{R}(\varepsilon\text{-greedy}, n) = \frac{\varepsilon n}{K}\sum_{i=1}^{K}\Delta_i$$

$$\Rightarrow \frac{\bar{R}_n}{n} = \frac{\varepsilon}{K}\sum_{i=1}^{K}\Delta_i$$

$\forall\, \varepsilon > 0$

Q.2 : $\hat{\mu} = \sum_{t=1}^{T} \frac{X_t}{T}$, $T : \Omega \to \{1, 2, 3, \dots\}$

T is independent from X_T for all T

To prove : $P\left(\hat{\mu} - \mu \geq \sqrt{\frac{2 \log(1/\delta)}{T}}\right) \leq \delta$

Proof : Let $T = n$

$$P(\hat{\mu} - \mu \geq t) \leq \exp\left(\frac{-nt^2}{2\sigma^2}\right) \quad \text{——— } (1)$$

Let $\exp\left(\frac{-nt^2}{2\sigma^2}\right) = \delta$

Solving for t

$$\frac{-nt^2}{2\sigma^2} = \log \delta$$

$$\Rightarrow t^2 = \frac{-2\sigma^2 \log \delta}{n}$$

$$\Rightarrow t^2 = \frac{2\sigma^2 \log(1/\delta)}{n}$$

$$\Rightarrow t = \sqrt{\frac{2\sigma^2 \log(1/\delta)}{n}}$$

Since X_t s are standard Gaussian random variables

$\therefore \sigma^2 = 1$

$$\therefore \quad t = \sqrt{\frac{2 \log(1/\delta)}{n}}$$

Substituting in eq (1)

$$P\left(\hat{\mu} - \mu \geq \sqrt{\frac{2 \log(1/\delta)}{n}}\right) \leq \delta$$

Hence Proved