21) To find $c>0$ such that for all $n \geq k > 1$ and for all policies π there exists a $\nu \in \mathcal{E}_N^K$ such that $R_n^s(\pi\, \nu) \geq c\sqrt{K/n}$

$$R_n^s(\pi, \nu) = E_{\nu\pi}[\Delta_{A_{n+1}}(\nu)]$$

$$= \sum_{i=1}^{K} P_{\nu\pi}(A_{n+1} = i)\Delta_i$$

Fix a policy π. Consider K-arm gaussian bandit, with variance 1 and mean $\mu \equiv (\Delta, 0, 0, \ldots 0)$. Denoting this environment as ν. A policy π and environment ν will give rise to $P_{\nu_\mu \pi}$ which will denoted as P_μ. 1 is the optimal arm.

Let $i = \underset{j>1}{\operatorname{argmin}}\, E_\mu(T_j(n+1))$. 1

As $\sum_{j=1}^{K} E_\mu(T(j)) = n+1$, we can say $E_\mu(T(i)) \leq \frac{n+1}{K-1}$ from (1) — (2)

Using this we define the second environment as with mean vector,

$\mu' = (\Delta, 0, 0, \ldots 2\Delta, 0, \ldots)$ and gaussian distribution.

Where 2Δ is at the i^{th} arm

This environment ν' will be rise to $P_{\nu'\pi}$ written as $P_{\mu'}$

Now,

$$R_n^S(\pi, \nu) = \sum_{j=2}^{K} P_{\pi\nu}(j)\Delta$$ [$\because$ For $j=1$, ~~so~~ regret $= 0$]

$$= (1 - P_{\pi\nu}(1))\Delta$$

$$= (1 - E_\mu(I_{A_{n+1}=1}))\Delta$$

Consider the event A: $E(I_{A_{n+1}=1}) \leq 1/2$ [Note that this is ~~a~~ a random event as no dependente on μ]

$$\text{Then } R_n^S(\pi, \nu) = (1 - E_\mu(I_{A_{n+1}=1}))\Delta$$

$$= (1 - E_\mu(I_{A_{n+1}=1}))P_\mu(E_\mu(I_{A_{n+1}=1}) \leq 1/2)\Delta$$

$$+ (1 - E(I_{A_{n+1}=1}))P_\mu(E(I_{A_{n+1}=1}) > 1/2)\Delta$$

$$\geq (1 - E(I_{A_{n+1}=1}))P_\mu(E(I_{A_{n+1}=1}) \leq 1/2)\Delta$$

$$\geq \frac{1}{2} P_\mu(E(I_{A_{n+1}=1}) \leq 1/2)\Delta = \frac{\Delta}{2} P_\mu(A)$$

For $R_n^S(\pi, \nu')$ we have,

$$R_n^S(\pi, \nu') = \Delta P_{\mu'}(1) + \sum_{j \neq i, 1}^{K} P(j) 2\Delta$$

$$\geq \Delta P_{\mu'}(1)$$

$$= \Delta E_{\mu'}(I_{A_{n+1}=1})$$

Consider the event A^c. $E(I_{A_{n+1}=1}) > 1/2$ [Again random as no depen environment is specified]

$$\therefore R_n^S(\pi, \nu') \geq \Delta \cdot E(I_{A_{n+1}=1}) \cdot P_{\mu'}((E_{A_{n+1}=1}) > 1/2)$$

$$\geq \frac{\Delta}{2} P_{\mu'}(A^c)$$

$$\therefore\ R_n^s(\pi,\nu) + R_n^s(\pi,\nu') \geq \frac{\Delta}{2} P_\mu(A) + \frac{\Delta}{2} P_{\mu'}(A^c)$$

$$= \frac{\Delta}{2}\left(P_\mu(A) + P_{\mu'}(A^c)\right)$$

From Bretagnolle-Huber inequality we know that,

$$P(A) + Q(A^c) \geq 1/2 \exp(-D(P,Q))$$

$$\therefore\ R_n^s(\pi,\nu) + R_n^s(\pi,\nu') \geq \frac{\Delta}{2} \times \frac{1}{2} \exp\left(-D\left(P_\mu, P_{\mu'}\right)\right)$$

Now in μ & μ' the only difference is the i^{th} term where $\mu_i' - \mu_i = 2\Delta$

$$\therefore\ D(P_\mu, P_{\mu'}) = E_\mu(A_n I_{A_{n+1}=i}) \cdot D[N(0,1), N(2\Delta,1)]$$

$$= P_\mu(i) \cdot (2\Delta)^2$$

But $P_\mu(i) \leq 1/K$

$$\therefore\ D(P_\mu, P_{\mu'}) \leq \frac{1}{K}(2\Delta)^2$$

$$\therefore\ R_n^s(\pi,\nu) + R_n^s(\pi,\nu') \geq \frac{\Delta}{2} \exp($$

$$\therefore\ D(P_\mu, P_{\mu'}) = E_\mu(T_i(n+1)) \cdot D[N(0,1), N(2\Delta,1)]$$

$$\leq \frac{n+1}{K-1} \frac{(2\Delta)^2}{2} \text{ from (2)}$$

$$R_n^s(\pi,\nu) + R_n^s(\pi,\nu') \geq \frac{\Delta}{4} \exp\left(-\frac{n 2\Delta^2}{K}\right)$$

Choose $\Delta = \sqrt{\frac{(K-1)}{4(n+1)}}$, as it is given that $n \geq K > 1$

$\Delta \leq 1/2$, [We need $\Delta \leq 1/2$ as we used 2Δ as mean and we have bounded means in $[0,1]$]

$$\therefore R_n^s(\pi, \nu) + R_n^s(\pi, \nu') \geq \frac{1}{4}\sqrt{\frac{(K-1)}{4(n+1)}} \exp(-1/2)$$

$$\Rightarrow 2\max(R_n^s(\pi, \nu), R_n^s(\pi, \nu')) \geq \frac{1}{8}\sqrt{\frac{(K-1)}{(n+1)}} \exp(-1/2)$$

Let $R_n^s(\pi, \nu) > R_n^s(\pi, \nu')$, then

$$R_n^s(\pi, \nu) \geq \frac{1}{16}\frac{1}{\sqrt{e}}\sqrt{\frac{(K-1)}{(n+1)}}$$

$$\geq \frac{1}{27}\sqrt{\frac{(K-1)}{(n+1)}}$$

Now, notice that, we need a λ such that

$$\frac{K-1}{n+1} \geq \lambda \frac{K}{n} \quad \text{where } n \geq K > 1$$

Choose $\lambda = 1/3$

$$\Rightarrow \frac{K-1}{n+1} > \frac{1}{3}\frac{K}{n}$$

$$\Rightarrow 3nK - 3n \geq Kn + K$$

$$\Rightarrow 2nK - 3n \geq K$$

$$\Rightarrow n(2K-3) \geq K$$

$$\Rightarrow (2K-3) \geq \frac{K}{n} > 0$$

Now $K \in [2, 3, \dots \infty)$

$\therefore 2K \in [4, 6, \dots \infty)$

$\therefore 2K-3 > 0$, Hence relation is satisfied for $\lambda = 3^{-1}$

$$\therefore R_n^s(\pi, \nu) \geq \frac{1}{27\sqrt{3}}\sqrt{\frac{K}{n}} \quad //$$

Q2) Choose ~~ν~~ the environment of K-arm gaussian bandit, with variance ~~off~~ of all arms as one and means as

$$\mu \equiv (\Delta, 0, 0, \ldots 0)$$

The policy is U.E

$$R_n^s(U.E, \nu) = \sum_{j=2}^{K} \Delta_j P_{\pi\nu}(j)$$

$$= \sum_{j=2}^{K} \Delta P_{\pi\nu}(j)$$

$$= (1 - P_{\pi\nu}(1))\Delta$$

Now $P_{\pi\nu}(1) = P_{\pi\nu}(\hat{\mu}_1 > \max_{j \neq 1} \hat{\mu}_j)$

$$\leq P_{\pi\nu}(\hat{\mu}_1 > \hat{\mu}_i) \quad \text{for some } i \neq 1$$

But as μ_1 and μ_i are gaussian and hence subgaussian by appropriately changing the means we know that

$$P_{\pi\nu}(\hat{\mu}_1 > \hat{\mu}_i) \leq \exp\left(-\frac{n}{K}\frac{\Delta^2}{4}\right) \quad (1)$$

$$\therefore R_n^s(U.E) = (1 - P_{\pi\nu}(1))\Delta$$

$$\geq \left(1 - \exp\left(-\frac{n}{K}\frac{\Delta^2}{4}\right)\right)\Delta \quad [\text{from } (1)]$$

let $\Delta = \sqrt{\frac{4K\log K}{n}}$

$$\therefore R_n^s(U.E) \geq \left(1 - \frac{1}{K}\right)\sqrt{\frac{4K\log K}{n}}$$

Now $K \in [2, 3, 4 \ldots \infty)$

$\therefore \; 1 - \frac{1}{K} \geq 1 - \frac{1}{2} = \frac{1}{2}$

$$\therefore \; R_n^s(U\text{-}B, U) \geq \left(1 - \frac{1}{K}\right)\sqrt{\frac{4K\log K}{n}}$$

$$\geq \frac{1}{2}\sqrt{\frac{4K\log K}{n}}$$

$$= \sqrt{\frac{K\log K}{n}}$$