2) i) X is σ - subgaussian.

$\therefore \mathbb{E}(\exp(\lambda x)) \leq \exp(\lambda^2 \sigma^2/2)$, $\lambda \in \mathbb{R}$

Now $\exp\left(\frac{\lambda^2 \sigma^2}{2}\right) = \sum_{n=0}^{\infty} \left(\frac{\sigma^2 \lambda^2}{2}\right)^n \frac{1}{n!}$

Similarly,

$\exp(\lambda x) = \sum_{n=0}^{\infty} \frac{(\lambda x)^n}{n!}$

$\therefore \mathbb{E}(1) + \mathbb{E}(\lambda x) + \frac{\mathbb{E}(\lambda^2 x^2)}{2!} \ldots \leq \mathbb{E}(\cancel{\sigma}1) + \frac{\sigma^2 \lambda^2}{2} + \frac{\lambda^4 \sigma^4}{2 \times 2!} + \ldots$

~~Taking~~

Dividing by $\lambda > 0$ we get

$\mathbb{E}(x) + \frac{\mathbb{E}(\lambda x^2)}{2!} \ldots \leq \frac{\sigma^2 \lambda}{2} + \frac{\lambda^3 \sigma^4}{2 \times 2!} \ldots$

Now, limit λ tends to 0^+ ($\because \lambda > 0$), we get,

$\cancel{\mathbb{E}(x) + \mathbb{E}(\lambda x^2}$

$\lim_{\lambda \to 0^+} \left(\mathbb{E}(x) + \frac{\lambda}{2} \mathbb{E}(\lambda x^2) + \ldots\right) \leq \lim_{\lambda \to 0^+} \left(\frac{\sigma^2 \lambda}{2} + \ldots\right)$

$\mathbb{E}(x) \leq 0$

Similarly if divided by $\lambda < 0$ we get,

$\mathbb{E}(x) + \frac{\mathbb{E}(\lambda x^2)}{2!} \ldots \geq \frac{\sigma^2 \lambda}{2} + \frac{\lambda^3 \sigma^4}{2 \cdot 2!} + \ldots$

As λ tends to 0^- we get,

$\mathbb{E}(x) \geq 0$

$\therefore \mathbb{E}(x) = 0$

Now, using $\mathbb{E}(x) = 0$, we get

$$\mathbb{E}(\exp(\lambda x)) \leq \exp\left(\frac{\lambda^2\sigma^2}{2}\right)$$

$$\Rightarrow \frac{\mathbb{E}(\lambda^2 x^2)}{2!} + \frac{\mathbb{E}(\lambda^3 x^3)}{3!} \dots \leq \frac{\lambda^2\sigma^2}{2} + \frac{\lambda^4\sigma^4}{2\times 2!} + \dots$$

Dividing by λ^2, $\lambda \neq 0$, we get

$$\frac{\mathbb{E}(x^2)}{2} + \frac{\lambda \mathbb{E}(x^3)}{3!} \dots \leq \frac{\sigma^2}{2} + \frac{\lambda^2\sigma^4}{2\times 2!} + \dots$$

Now as λ tends to zero we get

$$\lim_{\lambda \to 0}\left(\frac{\mathbb{E}(x^2)}{2} + \frac{\lambda \mathbb{E}(x^3)}{3!} \dots\right) \leq \lim_{\lambda \to 0}\left(\frac{\sigma^2}{2} + \frac{\lambda^2\sigma^4}{2\times 2!} + \dots\right)$$

$$\Rightarrow \frac{\mathbb{E}(x^2)}{\cancel{2}} \leq \frac{\sigma^2}{\cancel{2}}$$

$$\Rightarrow \mathbb{E}(x^2) \leq \sigma^2$$

ii) x is subgaussion.

$$\therefore E(\exp(\lambda x)) \leq \exp\left(\frac{\lambda^2\sigma^2}{2}\right), \lambda \in \mathbb{R}$$

Let $\lambda = \lambda' c x$, where $\lambda' \in \mathbb{R}$

$$\Rightarrow \mathbb{E}(\exp(\lambda'(cx))) \leq \exp\left(\frac{\lambda'^2(|c|\sigma)^2}{2}\right) \text{ where } \lambda' \in \mathbb{R}$$

cx is $|c|\sigma$ - subgaussion

iii) $\mathbb{E}(\exp(\lambda(x_1+x_2)))$

$= \mathbb{E}(\exp(\lambda x_1)\cdot\exp(\lambda x_2))$

Now as x_1 and x_2 independent, we can take the expectation inside,

$\therefore \mathbb{E}(\exp(\lambda x_1))\cdot\mathbb{E}(\exp(\lambda x_2))$

$\leq \exp\left(\frac{\lambda^2\sigma_1^2}{2}\right)\cdot\exp\left(\frac{\lambda^2\sigma_2^2}{2}\right)$ [$\because x_1$ & x_2 are σ_1 & σ_2 subgaussian respectively]

$= \exp\left(\frac{\lambda^2(\sigma_1^2+\sigma_2^2)}{2}\right)$

$\therefore x_1+x_2$ is $\sqrt{\sigma_1^2+\sigma_2^2}$ - subgaussian.

3) i) Given $\mathbb{E}(X)=0$, $P(x\in[a,b])=1$,

Now, any $x\in[a,b]$ can be written as, $x=\alpha b+(1-\alpha)a$, $\alpha\in[0,1]$

Since, e^{sx} is a convex function, we have,

$e^{sx}=e^{s(\alpha b+(1-\alpha)a)} \leq \alpha e^{sb}+(1-\alpha)e^{sa}$,

For a given x, $\alpha=\frac{x-a}{b-a}$

$e^{sx} \leq \left(\frac{x-a}{b-a}\right)e^{sb}+\left(\frac{b-x}{b-a}\right)e^{sa}$

Now, $E(e^{sx}) \leq \mathbb{E}\left(\left(\frac{x-a}{b-a}\right)e^{sb}+\left(\frac{b-x}{b-a}\right)e^{sa}\right)$

$\Rightarrow \mathbb{E}(e^{sx}) \leq \frac{-a}{b-a} e^{sb} + \frac{b}{b-a} e^{sa}$ $\left[\because \mathbb{E}\left(\frac{x-a}{b-a}\right) = \frac{\mathbb{E}(x)-a}{b-a} = \frac{-a}{b-a}\right]$

let $\theta = \frac{-a}{b-a}$, now, $a < 0 < b$ as $\mathbb{E}(x) = 0$,

$\therefore \theta > 0$.

$$\therefore \mathbb{E}(e^{sx}) \leq \theta e^{sb} + (1-\theta) e^{sa}$$

$$= e^{sa}\left(1-\theta+\theta e^{s(b-a)}\right)$$

Let, $u = s(b-a)$

$$\Rightarrow \mathbb{E}(e^{sx}) \leq e^{sa}(1-\theta+\theta e^{u})$$

$$= e^{-s\frac{(-a)}{(b-a)}\cdot(b-a)}(1-\theta+\theta e^{u})$$

$$= e^{-\theta u}(1-\theta+\theta e^{u})$$

Taking log on both sides

$$\log(\mathbb{E}(e^{sx})) \leq -\theta u + \log(1-\theta+\theta e^{u})$$

Note that, as (a, b) are fixed θ is a constant, hence only u is a variable.

Let $\psi(u) = -\theta u + \log(1-\theta+\theta e^{u})$

Now $\psi(0) = 0$

$$\psi'(u) = -\theta + \frac{\theta e^{u}}{1-\theta+\theta e^{u}}$$

$$\psi'(0) = -\theta + \theta = 0$$

$$\psi''(u) = \frac{\theta e^{u}(1-\theta+\theta e^{u}) - \theta^2 e^{2u}}{(1-\theta+\theta e^{u})^2}$$

$$\psi''(u) = \frac{\theta e^u}{(1-\theta+\theta e^u)} \times \left(1 - \frac{\theta e^u}{1-\theta+\theta e^u}\right)$$

Now, $e^u > 0$, $\theta > 0$ and,

$$1-\theta+\theta e^u = 1 + \frac{a}{b-a} + \theta e^u$$

$$= \frac{b}{b-a} + \theta e^u > 0 \quad [\text{as } b > 0 > a]$$

$\therefore \psi''(u) = t(1-t)$ where $t = \frac{\theta e^u}{1-\theta+\theta e^u} > 0$

$\leq \frac{1}{4}$ [as $t(1-t)$ is maximum at $t = 1/2$]

$$\therefore \psi(u) = \psi(0) + u\psi'(0) + \frac{1}{2}u^2\psi''(v), \quad v \in [0, u]$$

$$= 0 + u \times 0 + \frac{1}{2}u^2\psi''(v)$$

$$\leq \frac{1}{2}u^2 \times \frac{1}{4} = \frac{u^2}{8}$$

$$\therefore \log(\mathbb{E}(e^{sx})) \leq -\theta u + \log(1-\theta+\theta e^u) = \psi(u) \leq \frac{u^2}{8}$$

$$\Rightarrow \mathbb{E}(e^{sx}) \leq e^{u^2/8}$$

$$= e^{\frac{s^2(b-a)^2}{8}}$$

$$= e^{\frac{s^2}{2}\left(\frac{b-a}{2}\right)^2}$$

$\therefore x$ is $\left(\frac{b-a}{2}\right)$ subgaussian.

ii) $P\left(\sum_{t=1}^{n}(X_t - \mathbb{E}(X_t)) \ge \varepsilon\right)$, $\varepsilon > 0$

Let $\sum_{t=1}^{n} X_t = S_n$

$\therefore P(t(S_n - \mathbb{E}(S_n)) \ge t\varepsilon)$ ($\because X_1, X_2 \ldots X_t$ are independent) $(t>0)$

~~$= P(\exp(S_n - \mathbb{E}(S_n)) \ge \exp(\varepsilon)) \le \exp(-\varepsilon)\,\mathbb{E}(\exp(S_n - \mathbb{E}(S_n)))$ [Markov inequal]~~

~~$= \exp(-\varepsilon)\,\mathbb{E} \prod_{t=1}^{n} \exp(X_t - \mathbb{E}(X_t))$~~

~~$\le e$~~

$= P(\exp(t(S_n - \mathbb{E}(S_n))) \cancel{\ge \exp(t\varepsilon)})$

$\le \exp(-t\varepsilon)\,\mathbb{E}(\exp(t(S_n - \mathbb{E}(S_n))))$ [Markov inequality]

$= e^{-t\varepsilon}\,\mathbb{E}\left(\prod_{t=1}^{n} \exp(t(X_t - \mathbb{E}(X_t)))\right)$

$= e^{-t\varepsilon} \prod_{t=1}^{n} \mathbb{E}(\exp(t(X_t - \mathbb{E}(X_t))))$ [$\because X_i$ are independent]

$\le e^{-t\varepsilon} \prod_{i=1}^{n} e^{\frac{t^2 (b_i - a_i)^2}{8}}$ [From previous proof]

$= \exp\left(-t\varepsilon + \frac{1}{8} t^2 \sum (b_i - a_i)^2\right)$

Substituting ~~$t = 4t$~~ $t = 4\varepsilon / \sum (b_i - a_i)^2$ we get,

$P(S_n - \mathbb{E}(S_n) \ge \epsilon) \le \exp\left(\frac{-2\epsilon^2}{\sum (b_i - a_i)^2}\right)$ //

4) $R_T \leq \min\left\{ T\Delta, \Delta + \frac{4}{\Delta}\left(1 + \max\left\{0, \log\left(\frac{T\Delta^2}{4}\right)\right\}\right)\right\}$

$m = \max\left\{1, \left\lceil \frac{4}{\Delta^2} \log \frac{T\Delta^2}{4} \right\rceil\right\}$

Let $\Delta \geq 2/\sqrt{T}$, for a sufficiently large T.

Note that $\frac{4}{\Delta} \log \frac{\Delta^2}{4} = \frac{8}{\Delta} \log \frac{\Delta}{4}$ decreases with increasing Δ.

$\therefore \frac{4}{\Delta}\left(1 + \max\{0, \log \frac{T\Delta^2}{4}\}\right) \leq 2\sqrt{T}\left(1 + \max\left\{0, \log \frac{T \times 4}{T 4}\right\}\right)$

$= 2\sqrt{T}$

$\therefore R_T \leq \min\left\{ T\Delta, \Delta + 2\sqrt{T} \right\}$

For T sufficiently large $T\Delta > 2\sqrt{T}$

$\therefore R_T \leq \Delta + 2\sqrt{T}$

$\therefore$ $\underline{C = 2}$

6) Let all rewards be 1- subgaussian

$\therefore \Pr(x_i - u_i > \varepsilon) \leq \exp\left(\frac{-\varepsilon^2}{\cancel{2}}\right) \leq \cancel{\exp}$

$\therefore \Pr(R_T > E(R_T) + \varepsilon)$

$= \Pr(T\mu^* - \Sigma x_t - T\mu^* + \Sigma u_{A_t} > \varepsilon)$

$= \Pr(\Sigma(u_{A_t} - x_t) > \varepsilon)$

$\leq \exp\left(\frac{-\varepsilon^2}{1+1+\dots}\right)$ [Sum of sub-gaussians is a sub-gaussian]

$= \exp\left(\frac{-\varepsilon^2}{K}\right)$

$\therefore$ If $\exp\left(-\frac{\varepsilon^2}{K}\right) = \delta$ we have,

$\varepsilon = \sqrt{K \log \frac{1}{\delta}}$

$\therefore \Pr\left(R_T < E(R_T) + \sqrt{K \log \frac{1}{\delta}}\right) = 1-\delta$

7) 1) $\bar{R}_{nV}(T) = \sum_{a=1}^{K} \Delta_a \mathbb{E}(N_a(T))$

$N_a(T) = \sum_{t=1}^{T} 1_{\{I_t = a\}}$

$= 1 + \sum_{t=K+1}^{T} 1_{\{I_t = a\}}$ [$\because$ Each arm is played once]

$\Rightarrow \mathbb{E}(N_a(T)) = 1 + \sum_{t=K+1}^{T} P(I_t = a)$

$= 1 + \sum_{t=K+1}^{T} P(I_t = a \mid \text{exploitation})(1-\varepsilon) + \sum_{t=K+1}^{T} P(I_t = a \mid \text{Explore})\,\varepsilon$

$= 1 + \sum_{t=K+1}^{T} P(a = \text{argmax}_i\, \hat{u}_{t,i})(1-\varepsilon) + \frac{\varepsilon \cdot T}{K}$

Now, $\sum \Delta_a \mathbb{E}(N_a(T))$

$= \sum_{a=1}^{K} \Delta_a + \sum_{i=1}^{K} \Delta_i \sum_{t=K+1}^{T} P(i = \text{argmax}_j\, \hat{u}_{t,j})(1-\varepsilon) + \frac{\varepsilon}{K} T \sum_{i=1} \Delta_i$

Let $\Delta = \max \Delta_i$, i^* is the optimal arm, $\Delta_{min} = \min(\Delta_i)$

$\therefore \sum_a \Delta_a \mathbb{E}(N_a(T)) \leq \sum \Delta_i + \sum_{\substack{i=1 \\ i \neq i^*}}^{K} \Delta \sum_{t=K+1}^{T} P(i = \text{argmax}\, \hat{u}_{t,j})(1-\varepsilon) + \frac{\varepsilon}{K} T \sum \Delta_i$ [Not considering i^* as $\Delta_{i^*} = 0$]

$= \sum \Delta_i + \Delta \sum_{t=K+1}^{T} \sum_{\substack{i=1 \\ i \neq i^*}}^{K} P(i = \text{argmax}\, \hat{u}_t)(1-\varepsilon) + \frac{\varepsilon T}{K} \sum \Delta_i$

Without loss of generality, $i^* = 1$, therefore,

$\sum_{i=2}^{K} P(i = \text{argmax}_j\, \hat{u}_{t,j}) = \sum P(\hat{u}_{i^*,t} < \max_j \hat{u}_{j,t})$

$\therefore$ We have,

$$\sum \Delta_a \mathbb{E}(N_a(T)) \leq \Delta_i + \sum_{t=K+1}^{T} \Delta(1-\varepsilon) P\left(\hat{u}_{i^*, N_i(t)} < \max \hat{u}_{j, N_j(t)}\right) + \frac{\varepsilon T \sum \Delta_i}{K}$$

$$\leq \Delta_i + \sum_{t=K+1}^{T} \Delta(1-\varepsilon) \exp\left(\frac{-t \Delta_{min}^2}{K\, 4}\right) + \frac{\varepsilon T}{K} \sum \Delta_i$$

Note that, $P\left(\max_j \hat{u}_{j, N_j(t)} > \hat{u}_{i^*, N_i(t)}\right) \leq P\left(\hat{u}_{j, N_j(t)} > \hat{u}_{i^*, N_i(t)}\right)$

where $\Delta_{min} = u_{i^*} - u_j$

and $N_{i^*}(t) \geq t/K$ was used above.

$\therefore$ We have,

$$R_{\pi,\nu}(T) \leq \Delta_i + \Delta(1-\varepsilon) \sum_{t=0}^{T} \exp\left(-\frac{t}{K} \frac{\Delta_{min}^2}{4}\right) + \frac{\varepsilon T}{K} \sum \Delta_i$$

$$= \Delta_i + \Delta(1-\varepsilon) \frac{1}{1 - \exp\left(\frac{-\Delta_{min}^2}{K\, 4}\right)} + \frac{\varepsilon T}{K} \sum \Delta_i$$

$$\frac{R_{\pi\nu}(T)}{T} \leq \frac{\Delta_i}{T} + \frac{\Delta}{T} \frac{(1-\varepsilon)}{\left(1 - \exp\left(-\frac{\Delta_{min}^2}{4}\right)\right)} + \frac{\varepsilon}{K} \sum_{i=1}^{K} \Delta_i$$

$$\lim_{T \to \infty} \frac{R_{\pi,\nu}(T)}{T} \leq 0 + 0 + \frac{\varepsilon}{K} \sum_{i=1}^{K} \Delta_i$$

But $\displaystyle \frac{R_{\pi\nu}(T)}{T} = \frac{1}{T} + \sum_{i=1}^{K} \Delta_i \sum_{t=K+1}^{T} P\left(a = \arg\max \hat{u}_{N_i(t), i}\right) \frac{1}{T} + \frac{\varepsilon}{K} \sum_{i=1}^{K} \Delta_i$

$$\geq \frac{\varepsilon}{K} \sum_{i=1}^{K} \Delta_i$$

$$\therefore \lim_{T \to \infty} \frac{R_{\pi,\nu}(T)}{T} = \frac{\varepsilon}{K} \sum_{i=1}^{K} \Delta_i$$

8) We know that, if $I_t = K$ in U.C.B then of the conditions hold,

a) $\hat{\mu}_{K^*, N_{K^*}(t-1)} + \sqrt{\frac{\alpha \log t}{N_{K^*}(t-1)}} \leq \mu^*$

b) $\hat{\mu}_{K, N_i(t-1)} - \sqrt{\frac{\alpha \log t}{N_K(t-1)}} > \mu_K$

c) $N_i(t-1) \leq \frac{4\alpha \log t}{\Delta_K^2}$

Now in the modified algorithm the arm selection is done in phases that is at, $t = 1, 2, 4 \ldots$ or that is same as,

$\log_2 t = 0, 1, 2, \ldots$

Therefore let $\tilde{t} = \log_2 t$

$$\therefore N_K(T) = \sum_{t=1}^{T} 1_{\{I_t = K\}}$$

$$= \sum_{\tilde{t}=0}^{\log_2 T} 2^{\tilde{t}} \{I_{\tilde{t}} = K\}$$

[$2^{\tilde{t}}$ because that is the number times the K will be played after selection]

$$N_K(T) \leq u + \sum_{\tilde{t}=0}^{\log_2 T} 2^{\tilde{t}} 1_{\{I_{\tilde{t}} = K, N_K(t-1) > u\}}$$

$$= u + \sum_{\tilde{t}=0}^{\log_2 T} 2^{\tilde{t}} \cdot 1_{\{I_{\tilde{t}} = K, (a) \text{ or } (b)\}}$$

[let $u = \frac{4\alpha \log T}{\Delta_K^2}$]

$\therefore$ Following the proof of U.C.B's regret bound we have,

$$E[N_k(T)] \leq u + \sum_{\tilde{t}}^{\tilde{T}} \left(\Pr\{I_{\tilde{t}} = k, (a) \text{ holds}\} + \Pr\{I_t = k, b \text{ holds}\}\right) \times 2^{\tilde{t}}$$

Now, $2^{\tilde{t}} \Pr\{(a) \text{ holds}\} \leq \cancel{\sum_{\tilde{t}} \frac{2^{\tilde{t}}}{\tilde{t}^{2\alpha}}} \quad \frac{2^{\tilde{t}}}{\tilde{t}^{2\alpha-1}}$

and $2^{\tilde{t}} \Pr\{(b) \text{ holds}\} \leq \frac{1}{\tilde{t}^{2\alpha-1}} = \frac{2^{\tilde{t}}}{\tilde{t}^{2\alpha-1}}$

$$\therefore E(N_k(T)) \leq u + \cancel{2\alpha 2}\, 2\alpha \sum_{\tilde{t}=1}^{\log_2 T} \frac{2^{\tilde{t}}}{\tilde{t}^{2\alpha-1}}$$

Let $\alpha = 3/2$

$$E[N_k(T)] \leq u + 2 \sum^{\log_2 T} \frac{2^{\tilde{t}}}{\tilde{t}^2}$$

Now $\sum \frac{2^{\tilde{t}}}{\tilde{t}^2} \leq \sqrt{\sum_{\tilde{t}=0}^{\log_2 T} (2^{\tilde{t}})^2 \times \sum_{\tilde{t}=0}^{\log_2 T} \frac{1}{\tilde{t}^4}}$ Cauchy-Schwartz

As $\sum_{\tilde{t}}^{\log T} \frac{1}{\tilde{t}^4} \leq \cancel{\log} \sum_{\tilde{t}}^{\infty} \frac{1}{\tilde{t}^4} = \frac{\pi^4}{90}$

$$\sum^{\log_2 T} 4^{\tilde{t}} = \frac{4^{\log_2 T} - 1}{4-1} = \frac{T^2 - 1}{3} < \frac{T^2}{3}$$

$\therefore 2\sum 2^{\tilde{t}}/\tilde{t}^2 \leq \frac{2\alpha T \pi^2}{3\sqrt{30}}$, $\therefore E[N_k(T)] \leq \frac{6\log T}{\Delta_k^2} + \frac{2T\pi^2}{3\sqrt{30}} + 1$

$$\therefore R_T(U.C.B) = \sum \Delta_k E(N_k(T))$$
$$\leq 6 \log T \sum \frac{1}{\Delta_k} + K\left(\frac{2T\pi^2}{3\sqrt{30}} + 1\right)$$