This handout includes space for every question that requires a written response. Please feel free to use it to handwrite your solutions (legibly, please). If you choose to typeset your solutions, the —README.md— for this assignment includes instructions to regenerate this handout with your typeset LaTeX solutions.

2.a

The distribution a_t matches the distribution of a^* and U_t is a deterministic function. The $L_t(a_t)$ cancel.

$$\mathbb{E}\left[U_t(a_t)\right] = \mathbb{E}\left[\mathbb{E}\left[U_t(a_t)\,|H_t\right]\right] = \mathbb{E}\left[\mathbb{E}\left[U_t(a^*)\,|H_t\right]\right] = \mathbb{E}\left[U_t(a^*)\right]$$

2.b

LHS $\leq$ RHS

$$\sum_{t=1}^{T}\left[L_t(a_t)-\mu_\theta(a_t)\right]+\left[\mu_\theta(a^*)-U_t(a^*)\right]\leq T\sum_{a}\mathbb{I}\left\{\bigcup_{t=1}^{T}\{\mu_\theta(a)\notin\left[L_t(a),U_t(a)\right]\}\right\}$$

If $L_t(a_t)-\mu_\theta(a_t)\leq 0$ and $\mu_\theta(a^*)-U_t(a^*)\leq 0$ for all $t\in[T]$. RHS ≥ 0, so LHS ≤ 0

If $t'\in[T]$ where $L_{t'}(a_{t'})-\mu_\theta(a_{t'})>0$, then $L_{t'}(a_{t'})>\mu_\theta(a_{t'})$

$\mu_\theta(a_{t'})\notin\left[L_{t'}(a_{t'})\right]$ so $\mathbb{I}\left\{\bigcup_{t=1}^{T}\{\mu_\theta(a_{t'})\notin\left[L_t(a_{t'}),U_t(a_{t'})\right]\}\right\}=1$

$L_t(a)$ and $\mu_\theta(a_{t'})$ is in $[0,1]$

$L_t(a_t)-\mu_\theta(a_t)\leq 1$

$$\sum_{t=1}^{T}\left[L_t(a_t)-\mu_\theta(a_t)\right]\leq T=T\mathbb{I}\left\{\bigcup_{t=1}^{T}\{\mu_\theta(a_{t'})\notin\left[L_t(a_{t'}),U_t(a_{t'})\right]\}\right\}$$

If $t'\in[T]$ where $\mu_\theta(a^*)-U_{t'}(a^*)>0$

$\mu_\theta(a^*)-U_{t'}(a^*)\leq 1$

$$\sum_{t=1}^{T}\left[\mu_\theta(a^*)-U_{t'}(a^*)\right]\leq T=T\mathbb{I}\left\{\bigcup_{t=1}^{T}\{\mu_\theta(a^*)\notin\left[L_t(a^*),U_t(a^*)\right]\}\right\}$$

If $\exists t$ where $\mu_\theta(a)\notin\left[L_t(a),U_t(a)\right]$, then RHS $\geq$ 2T, LHS $<$ 2T

If RHS = T and violated action is not a^*, then $\mu_\theta(a^*)-U_{t'}(a^*)\leq 0$ for all t, so LHS $\leq T$

IF RHS = T and violated action is a^* and $L_t\leq U_t$,

$L_t(a_t)-\mu_\theta(a_t)>0$ or $\mu_\theta(a^*)-U_t(a^*)>0$ so LHS $\leq T$

2.c

a gets picked up $n_T(a)$ times. When a is selected for the first time, $U_t(a_t) - L_t(a) = 1 - 0 = 1$
Each time a gets picked up, the difference is at most $2\sqrt{2 + 6logT}\frac{1}{\sqrt{n_T(a)}}$, $n_T(a)$ increment by one each time a gets picked up, so it yields the difference of $2\sqrt{2 + 6logT}\Sigma_{i-1}^{n_T(a)}\frac{1}{\sqrt{i}}$

2.d

$$\sum_{i=1}^{n_T(a)} \frac{1}{\sqrt{i}} \leq \int_0^{n_T(a)} x^{-\frac{1}{2}} dx = \left[2x^{\frac{1}{2}}\right]_0^{n_T(a)} = 2\sqrt{n_T(a)}$$

2.e

$$\mathrm{BR}(T) \le \mathbb{E}\left[\sum_{t=1}^{T}\left[U_t(a_t) - L_t(a_t)\right]\right] + K$$

$$= K + \mathbb{E}\left[\sum_{a}\sum_{t\in T_a}\left[U_t(a_t) - L_t(a_t)\right]\right]$$

$$\le K + \mathbb{E}\left[\sum_{a}\left(1 + \left(2\sqrt{2+6logT}\right)\left(2\sqrt{n_T(a)}\right)\right)\right]$$

$$= 2K + 4\sqrt{2+6logT}\,\mathbb{E}\left[\sum_{a}\sqrt{n_T(a)}\right]$$

AM-QM equality

$$\sum_{a} n_T(a) = T - 1 < T$$

$$\mathbb{E}\left[\sum_{a}\sqrt{n_T(a)}\right] = K\mathbb{E}\left[\frac{1}{K}\sum_{a}\sqrt{n_T(a)}\right] \le K\mathbb{E}\left[\frac{1}{K}\sum_{a} n_T(a)\right] = \sqrt{KT}$$

$$BR_T(\pi^{TS}) \le 2K + 4\sqrt{KT(2+6logT)}$$