Stochastic Optimisation - Problem Sheet 2

Dom Hutchinson

October 18, 2020

Question 3.

Question 3. (a)
Show that, if arm 2 is played by the above algorithm in round $s+1$ (i.e. $I(s+1)=2$) then one of the following statements must be true.

i). $N_2(s) < \dfrac{2\alpha\ln(s)}{\Delta^2}$

ii). $\hat{\mu}_{2,N_2(s)} \geq \mu_2 + \sqrt{\dfrac{\alpha\ln(s)}{2N_2(s)}}$

Answer 3. (a)
This is a proof by contradiction.
Suppose $I(s+1)=2$ but that none of the statements above hold. Then

$$\begin{array}{rrcll}
& \hat{\mu}_{2,N_2(s)} - \sqrt{\dfrac{\alpha\ln(s)}{2N_2(s)}} & < & \mu_2 & \text{by not ii)} \\
& & = & \mu_1 - \Delta & \text{by def. of } \Delta \\
& & \leq & \mu_1 - \sqrt{\dfrac{2\alpha\ln(s)}{N_2(s)}} & \text{by not i)} \\
\implies & \hat{\mu}_{2,N_2(s)} + \sqrt{\dfrac{2\alpha\ln(s)}{N_2(s)}} - \sqrt{\dfrac{\alpha\ln(s)}{2N_2(s)}} & < & \mu_1 & \\
\implies & \hat{\mu}_{2,N_2(s)} + \left(\sqrt{2}-\dfrac{1}{\sqrt{2}}\right)\sqrt{\dfrac{\alpha\ln(s)}{N_2(s)}} & < & \mu_1 & \\
\implies & \hat{\mu}_{2,N_2(s)} + \sqrt{\dfrac{\alpha\ln(s)}{2N_2(s)}} & < & \mu_1 & \\
\implies & i_2(s) & < & \mu_1 &
\end{array}$$

This means $I(s+1)=1$, which is a contradiction. Thus at least one of i) or ii) must be true.□

Question 3. (b)
Recall that $N_2(t)=\sum_{s=1}^t\{I(S)=2\}$. For an arbitrary positive integer u and any $t\in\mathbb{N}$ explain why

$$N_2(t) \leq u + \sum_{s=u+1}^t \big\{\{N_2(s-1)\geq u\} \text{ and } \{I(s)=2\}\big\}$$

Answer 3. (b)
Fix $t,u\in\mathbb{N}$. We have two possibilities

Case 1 $N_2(t) \leq u$ (i.e. Arm two has not been played u times yet). The result trivially holds in this case.

Case 2 $\exists\ s \in [1, t]$ such that $N(s) > u$ (i.e. Arm two has been played at least u times).
Let s^* denote the smallest such s. Then it must be true that $N(s^*-1) = u$ and $s^* \geq u+1$. Hence

$$\begin{aligned} N(t) &= \sum_{s=1}^{s^*-1} I(s) + \sum_{s=s^*}^{t} I(s) \\ &= N(s^*-1) + \sum_{s=s^*}^{t} I(s) \underbrace{\{N(s-1) \geq u\}}_{\text{true for all in sum}} \\ &\leq u + \sum_{s=u+1}^{t} \{N(s-1) \geq u\} \qquad \text{since } s^* \geq u+1 \end{aligned}$$

Thus the result holds in all cases. □

Question 3. (c)
Define $u = \lceil (2\alpha \ln(t))/\Delta^2 \rceil$. Using the answers to parts `(a)` and `(b)`, and relevant probability inequalities, show that

$$\mathbb{E}[N_2(t) \leq u + \sum_{s=u+1}^{t} e^{-\alpha \ln(s)}$$

Use this to show that $\mathbb{E}[N_2(t)] \leq u + \dfrac{1}{\alpha - 1}$.

Answer 3. (c)
We have

$$\mathbb{E}[N_2(t)] \leq u + \sum_{s=u+1}^{t} e^{-\alpha \ln(s)}$$

Taking expectations of both sides

$$\begin{aligned} \mathbb{E}[N_2(t)] &\leq u + \sum_{s=u+1}^{t} \mathbb{P}\left(\{N_2(s-1) \geq u\} \text{ and } \{I(s) = 2\}\right) \\ &\leq u + \sum_{s=u}^{t-1} \mathbb{P}\left(\{N_2(s) \geq u\} \text{ and } \{I(s+1) = 2\}\right) \end{aligned}$$

If $N_2(s) \geq u$ and $I(s+1) = 2$ then

$$\hat{\mu}_{2,N_2(s)} \geq \mu_2 + \sqrt{\frac{\alpha \ln(s)}{2N_2(s)}} \text{ by a)}$$

Thus

$$\mathbb{E}(N_2(t)) \leq u + \sum_{s=u}^{t-1} \mathbb{P}\left(\hat{\mu}_{2,N_2(s)} \geq \mu_2 + \sqrt{\frac{\alpha \ln(s)}{2N_2(s)}}\right) \quad (1)$$

Let $X_1,\dots,X_{N_2}$ be the random variables for each time arm 2 was played. Consider

$$\begin{array}{rrlr}
& \mathbb{P}\left(\hat{\mu}_{2,N_2(s)} \geq \mu_2 + \sqrt{\frac{\alpha\ln(s)}{2N_2(s)}}\right) & = \mathbb{P}\left(\frac{1}{N_2}\sum_{i=1}^{N_2} X_i \geq \mu_2 + \sqrt{\frac{\alpha\ln(s)}{2N_2(s)}}\right) & \\
& & = \mathbb{P}\left(\sum_{i=1}^{N_2}(X_i-\mu_2) \geq N_2\sqrt{\frac{\alpha\ln(s)}{2N_2(s)}}\right) & \\
& & \leq \exp\left(-2\cdot N_2\cdot\frac{\alpha\ln(s)}{2N_2(s)}\right) & \text{by Hoeffding's Ineq.}\\
& & = \exp(-\alpha\ln(s)) & \\
\implies & \mathbb{E}[N_2(t)] & \leq u + \sum_{s=u+1}^{t} e^{-\alpha\ln(s)} & \text{by (1)}
\end{array}$$

Further

$$\begin{array}{rcl}
\mathbb{E}[N_2(t)] & \leq & u + \sum_{s=u+1}^{t} e^{-\alpha\ln(s)}\\
& = & u + \sum_{s=u+1}^{t} s^{-\alpha}\\
& \leq & u + \int_u^\infty s^{-\alpha}ds \quad \text{since } \alpha > 1\\
& = & u + \left[\frac{s^{-\alpha+1}}{-\alpha+1}\right]_u^\infty\\
& = & u - \frac{u^{-\alpha+1}}{-\alpha+1}\\
& = & u + \frac{u^{-\alpha+1}}{\alpha+1}
\end{array}$$

By the definition of u, $u>1$ thus $u^{-\alpha+1}<1$ since $\alpha>1$. Giving us

$$\mathbb{E}[N_2(t)] \leq u + \frac{1}{\alpha-1}$$

Question 3. (d)
Use the answer to `(c)` to show that the regret of this algorithm is bounded above as

$$\mathcal{R}(T) \leq \frac{2\alpha\ln(T)}{\Delta} + \frac{\alpha}{\alpha-1}\Delta$$

Answer 3. (d)

$$\begin{array}{rcll}
\mathcal{R}(T) & := & \Delta\mathbb{E}[N_2(t)] & \\
& \leq & \Delta\left(u + \frac{1}{\alpha-1}\right) & \text{by 3. (c)}\\
& \leq & \Delta\left(\frac{2\alpha\ln(T)}{\Delta^2} + 1 + \frac{1}{\alpha-1}\right) & \text{by def. of } u\\
& = & \frac{2\alpha\ln(T)}{\Delta} + \Delta\left(1 + \frac{1}{\alpha-1}\right) & \\
& = & \frac{2\alpha\ln(T)}{\Delta} + \frac{\Delta\alpha}{\alpha-1} &
\end{array}$$