# CS6700 : Reinforcement Learning

# Short Assignment on Bandits

Sai Vinay G
CE17B019
Indian Institute of Technology Madras

February 9, 2019

**1. You have come across Median Elimination as an algorithm to get $(\epsilon, \delta)$-PAC bounds on the best arm in a bandit problem. At every round, half of the arms are removed by removing arms with return estimates below the median of all estimates. How would this work if we removed only one-fourth of the worst estimated arms instead ? Attempt a derivation of the new sample complexity.**

Given, to remove only one fourth of the worst arms present after each round.
$E1 = \{\hat{p}_1 < p_1 - \epsilon_1/2\}$
$Pr[E1] \leq e^{-2\epsilon^2 n}$
$Pr[\#bad \geq 3n/4 | \hat{p}_1 \geq p_1 - \epsilon_1/2] \leq 4/3(e^{-2\epsilon^2 n})$
As, the sum of both probabilities should be less than $\delta$ we get,

$n = 2/\epsilon^2 ln(7/3\delta_l)$
For the summation to be bounded initialize $\epsilon_1 = \epsilon/16$ and update it using $\epsilon_{\ell+1} = 15/16\epsilon_\ell$

Now the new modified median elimination algorithm becomes

New algorithm :

1. Set $S = A$
2. $\epsilon_1 = \epsilon/16, \delta_1 = \delta/2, \ell = 1$
3. Sample every arm $a \in S$ for $\frac{1}{(\epsilon_\ell)^2/2} ln(7/3\delta_\ell)$ times, and let $\hat{p}_a^\ell$ denote its empirical value.
4. Find the first quartile of $\hat{p}_a^\ell$, denoted by $q_\ell^1$.
5. $S_{\ell+1} = S \setminus \{a : \hat{p}_a^\ell < q_\ell^1\}$
6. if $|S| = 1$ Then output $S_\ell$,
   Else $\epsilon_{\ell+1} = \frac{15}{16}\epsilon_\ell; \delta_{\ell+1} = \delta_\ell/2; \ell = \ell + 1$ Go to 3.

As, in each round one fourth of the worst arms are removed. It takes $log_{4/3}n$ rounds, so that we are left with one arm.

Sample arm complexity is,

$$
\begin{aligned}
\sum_{\ell=1}^{log_{4/3}n} \frac{2n_\ell}{(\epsilon_\ell)^2} ln(7/3\delta_\ell) &= 2 \sum_{\ell=1}^{log_{4/3}n} \frac{\frac{n}{(\frac{4}{3})^{\ell-1}} ln(2^\ell 7/3\delta)}{((\frac{15}{16})^{\ell-1}\epsilon/16)^2} \\
&= 512 \sum_{\ell=1}^{log_{4/3}n} n(\frac{64}{75})^{\ell-1}(\frac{ln(1/3\delta)}{\epsilon^2} + \frac{7}{\epsilon^2} + \frac{\ell ln(2)}{\epsilon^2}) \\
&= 512 \frac{n ln(1/\delta)}{\epsilon^2} \sum_{\ell=1}^{log_{4/3}n} (\frac{64}{75})^{\ell-1}(\ell C' + C) \leq O\Big(\frac{n ln(1/\delta)}{\epsilon^2}\Big)
\end{aligned}
$$

Thus sample complexity is bounded by $O\Big(\frac{n ln(1/\delta)}{\epsilon^2}\Big)$
And also we here have less probability of rejecting a good arm.