# Reinforcement Learning
# Weekly Written Assignment #2

Author:
CS17S011 Ajay Kumar Pandey

Submitted to-
Assoc. Professor: B.Ravindran

January 30, 2018

# 1 Solution:

Probability approximately correct(PAC) bound for Multi armed bandit(MAB):

Quarter elimination algorithm
Input:$\epsilon > 0, \delta > 0$
Output: An optimal arm
repeat
sample every arm a $\in \delta_l$ for $\frac{2}{\epsilon^2} \log \frac{7}{3\delta}$ times
find the median of sample means $Q_l(a)$ at round l, denoted by $m_l$
$s_{l+1} = s_l \diagdown \{a : Q_l(a) < m_l\}$
$\epsilon_{l+1} = \frac{3}{4}\epsilon_l; \delta_{l+1} = \frac{\delta_l}{2}; l = l+1$
$until |s_l| = 1;$
end
This algorithm runs $log_{4/3}K$ rounds K:number of arms

At every round there are two bad events, so summing up there bad probabilities should less than $\delta_l$
event $e_1 = \{Q(a_l^*) < q_*(a_l^*) - \epsilon_1/2\}$;
$P[e_1] \leq 3\delta_l/7$
In the 2nd bad event we get,
$P[\#Badarms \geq 3|s_l|/4 \diagup Q_l(a_l^*) \geq q_*(a_l^*) - \epsilon/2] \leq \frac{4\delta_l}{7}$

So happening of something bad together is $\frac{3\delta_l}{7} + \frac{4\delta_l}{7} = \delta_l$
Hence probability of good is $1 - \delta_l$

After all rounds of algorithm
$\sum_{l=1}^{logk} \delta_l = \delta/2 + \delta/4 + ....... + \delta/2^{log_{4/3}k}$
$\sum_{l=1}^{logk} \delta_l \leq \delta$
$\sum_{l=1}^{logk} \epsilon_l = (1/4)[\epsilon/(4/3) + \epsilon/(4/3)^2 + ....... + \epsilon/(4/3)^{log_{4/3}k}]$
$\sum_{l=1}^{logk} \epsilon_l \leq \epsilon$

showing sample complexity :
$\delta_1 = \delta/2; \delta_l = \delta_{l-1}/2 = \delta/2^l$
$\epsilon_1 = \epsilon/4; \epsilon_l = (3/4)\epsilon_{l-1} = (3/4)^{l-1}\epsilon/4$
$k_1 = k; k_l = (3/4)k_{l-1} = (3/4)^{l-1}k$

Therefore we have,
$\sum_{l=1}^{log_{4/3}k} k_l \frac{\log(7/3\delta_l)}{\epsilon^2/2} = \sum_{l=1}^{log_{4/3}k} \frac{2k \log \frac{2^l 7}{\delta 3}}{(4/3)^{l-1}((3/4)^{l-1}\epsilon/4)^2}$
$= 32 \sum_{l=1}^{log_{4/3}k} k(4/3)^{l-1}[\frac{\log(7/3)}{\epsilon^2} + \frac{\log(1/\delta)}{\epsilon^2} + l\frac{\log(2)}{\epsilon^2}]$
$\leq 32k\frac{\log(1/\delta)}{\epsilon^2} \sum_{l=1}^{\infty}(4/3)^{l-1}(lc_1 + c_2)$
$\leq O(k\frac{\log(1/\delta)}{\epsilon^2})$
Bound for both Median and Quarter elimination algorithm is same.