Homework-2

Capstone - Yuhua

1. implement Algorithm 1 (explore-then-commit) in the following two settings:

   - Arm 1 follows the normal distribution $N(\mu_1, 1)$, arm 2 follows the normal distribution $N(\mu_2, 1)$. Set $\mu_1 = 0$, $\mu_2 \in [0, 1]$, horizon $n = 10^3$. (Note that horizon is the total rounds of the experiment.)

   - Arm 1 follows the Bernoulli distribution, that is, it gets reward 1 with probability μ_1, and reward 0 with probability $1 - \mu_1$. Arm 2 follows the Bernoulli distribution with expectation μ_2. Set $\mu_1 = \frac{1}{2}$, $\mu_2 \in [0, 1]$, horizon $n = 1000$.

   Please show the performance of the algorithm in terms of regret in the above settings for $m = 25, 50, 75, 100$ and $m =$ "the optimal exploration rate according to equation (6.3)". You should plot two figures, one is for normal rewards and one is for Bernoulli rewards. For each figure, the y-axis is the regret and x-axis is the value of $\mu_2 - \mu_1$. There should be 5 lines in each figure. Show me your results and tell me your observations in the next class. Note that when $\mu_2 - \mu_1$ is small, the regret will be different for each experiment, then you should take at least 10^3 simulations and average them to get the expected regret.