IE613: Online Machine Learning **Jan-Apr 2018**

Assignment 4: April 10

Instructions: *You are free to code in Python/Matlab/C/R. Discussion among the class participants is highly encouraged. But please make sure that you understand the algorithms and write your own code. Submit the code by* 19*th April.*

Question 1 (Pure exploration, best arm selection) *Consider a* K*-armed bandit where each arm is a Bernoulii random variable. We would like to identify the best arm with probability atleast* $1-\delta$*, i.e.,* $\Pr\{\hat{I} = I^*\} \geq 1-\delta$*, where* $\hat{I}$ *is the estimated best arm and* I^* *is the best arm. Numerically compare the expected sample complexity and the mistake probability of the following algorithms for different values of* K*. Set (confidence parameter)* $\delta = 0.1$.

- *KL-LUCB [1] with parameters* $\epsilon = 0, \alpha = 2, \kappa_1 = 4e + 4, \beta(t, \delta) = \log(\kappa_1 K t^\alpha/\delta) + \log\log(\kappa_1 K t^\alpha/\delta)$.
- *lil'UCB [2] with parameters* $\epsilon = 0.01, \beta = 1, \lambda = 9, \sigma = 0.5$.

Vary K between 10 to 50 in step size of 10. For each K, generate 1000 bandit instances with mean reward for each arm drawn uniformly at random from [0 1]. Average the sample complexities over all the instances to obtain the average sample complexity and the mistake probability should be calculated as the fraction of times non best arm is returned after stop signals are received.

Question 2 *Consider the same setting as in Question* 1*. But instead of generating different bandit instances, fix a bandit instance for each* $K = 10, 15, 20, 25, 30$ *as follows:*

$$\mu_i = \begin{cases} \frac{1}{2} \text{ if } i = 1 \\ \frac{1}{2} - \frac{i}{70} \text{ if } i = 2, 3, \cdots, K, \end{cases}$$

where μ_i *denotes the mean of* i*th arm. Numerically compare the expected sample complexity and the mistake probability for different values of* K *for both the algorithms. Run each algorithm atleast 200 times (more is better!) on the bandit instance to get estimates for sample complexity and mistake probability. Keep the parameters of algorithms same as in Question 1.*

References

[1] E. Kaufmann and S. Kalyanakrishnan. *Information Complexity in Bandit Subset Selection.* COLT 2013.

[2] K. Jamieson, M. Malloy, R. Nowak and S. Bubeck. *lil UCB : An Optimal Exploration Algorithm for Multi-Armed Bandits.* Journal of Machine Learning Research, December 2013