COMP 138 RL: K-armed Bandit

Jun Sung Tak

September 23, 2021

1 Goal

The goal of this assignment is to explore the inadequacies of the sample average method in non-stationary problems. Additionally, multiple feedback techniques will be used (ie, e-greedy, greedy, optimistic initial value, and Upper bound selection) on both stationary and non-stationary settings in order to illuminate the differences between the feedback techniques.

2 Conclusion

"I always thought something was fundamentally wrong with the universe" [?]

References