Sapienza University of Rome, Italy
Master in Artificial Intelligence and Robotics
Machine Learning (2018/19)

Exercise 5. RL k-Armed Bandit

Luca Iocchi

Exercise 5. Reinforcement Learning k-Armed Bandit

Compare the following two strategies for the stochastic k-Armed Bandit problem (with Gaussian distributions), by plotting the reward over time.

1. For each of the k actions, perform 30 trials and compute the mean reward; then always play the action with the highest estimated mean.
2. ϵ-greedy strategy (with different values of ϵ) and training rule from previous slide.

Note: realize a parametric software with respect to k and the parameters of the Gaussian distributions and use the following values for the experiments: $k = 4$, $r(a_1) = \mathcal{N}(100, 50)$, $r(a_2) = \mathcal{N}(90, 20)$, $r(a_3) = \mathcal{N}(70, 50)$, $r(a_4) = \mathcal{N}(50, 50)$.

Exercise 5. Reinforcement Learning
k-Armed Bandit

Learning goals

- Understanding RL strategies
- Comparing learning strategies with deterministic solutions