CMPUT 397

Worksheet 2

September 10, 2020

1. Indepence and Correlation

(a) Show that if X and Y are independent, then $E[XY] = E[X]E[Y]$.

(b) The correlation between two variables X and Y is defined as :

$$\frac{E[(X - E[X]) \cdot (Y - E[Y])]}{\sigma_X \sigma_Y}.$$

Show that if X and Y are independent, then they are uncorrelated (correlation is zero). Provide an example of X and Y that are uncorrelated but not independent.

2. Suppose a game where you choose to flip one of two (possibly unfair) coins. You win \$1 if your chosen coin shows heads and lose \$1 if it shows tails.

(a) Model this as a K-armed bandit problem: define the action set.

(b) Is the reward a deterministic or stochastic function of your action?

(c) You do not know the coin flip probabilities. Instead, you are able to view 6 sample flips for each coin respectively: (T,H,H,T,T,T) and (H,T,H,H,H,T). Use the sample average formula (equation 2.1 in the book) to compute the estimates of the value of each action.

(d) Decide on which coin to flip next! Assume it's an exploit step.

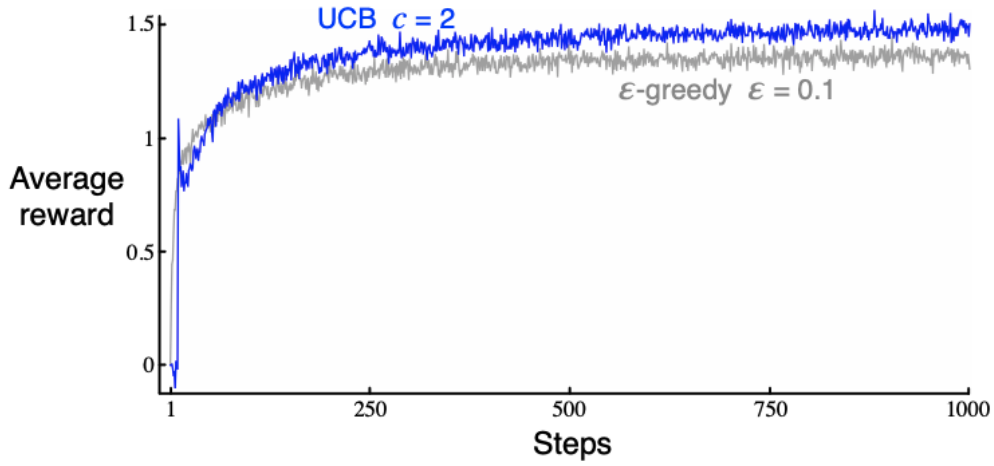

Figure 2.4: Average performance of UCB action selection on the 10-armed testbed. As shown, UCB generally performs better than ε-greedy action selection, except in the first k steps, when it selects randomly among the as-yet-untried actions.

3. **Discussion Question** : (Exercise 2.8 from Sutton & Barto, 2nd Edition) *UCB Spikes* : In Figure 2.4 the UCB algorithm shows a distinct spike in performance on the 11th step. Why is this? Note that for your answer to be fully satisfactory it must explain both why the reward increases on the 11th step and why it decreases on the subsequent steps. Hint: If $c = 1$, then the spike is less prominent, Hint 2: There were 10 bandit arms in this problem.