

<u>Course</u> > <u>Bandits</u> > <u>Lab</u> > Exercis...

Exercise 2B Optimistic Greedy

Exercise 2.2B: Optimistic Greedy

In this exercise, you will implement the optimistic greedy policy.

Make sure that you have:

- 1. Completed the setup requirements as described in the Set Up Lab Environments section
- 2. Completed the previous exercises in this lab

Now, run jupyter notebook and open the "Ex2.2B Optimistic Greedy.ipynb" notebook under **Module 2** folder.

- 1. Examine the notebook.
- 2. Your task is to implement an optimistic greedy policy: that is initialize the \hat{r}_a to a large initial value R, which is implemented in the init() function, and then play the greedy algorithm.
- 3. We have given you some boiler plate code, you only need to modify the part as indicated.
- 4. Once you have done that, prepare a simulation. Don't change any other parameter, that is:
 - evaluation_seed = 5016
 - num actions = 10
 - trials = 10000
 - distribution = "bernoulli"
- 5. Set the R to zero.
- 6. Run the simulation, observe the results, and answer the following questions.

Lab Question

1.0/1.0 point (graded)

With R set to zero, what do you observe?

- The optimistic greedy behaves randomly
- The optimistic greedy behaves like the epsilon greedy algorithm
- The optimistic greedy behaves like the greedy algorithm

Submit

You have used 1 of 2 attempts

Set the $\it R$ to a very, very large number, let's use 10000, which is the same number with the number of trials. Run the simulation again and observe the results.

Lab Question

1.0/1.0 point (graded)

With R set to 10000, what do you observe?

The optimistic greedy behaves randomly



- The optimistic greedy behaves like the optimistic greedy algorithm
- The optimistic greedy behaves like the greedy algorithm

Submit

You have used 1 of 2 attempts

Now, try several different number of R (1, 3, 5). Make sure the other parameters stay the same, that is:

- evaluation_seed = 5016
- num_actions = 10
- trials = 10000
- distribution = "bernoulli"

Run the simulations and observe the results.

Lab Question

1.0/1.0 point (graded)

Which is the smallest R that allow the optimistic greedy to find the optimal arm?



Submit

You have used 1 of 2 attempts

Now let's prepare another simulation by setting a different seed, so your parameters should look like this:

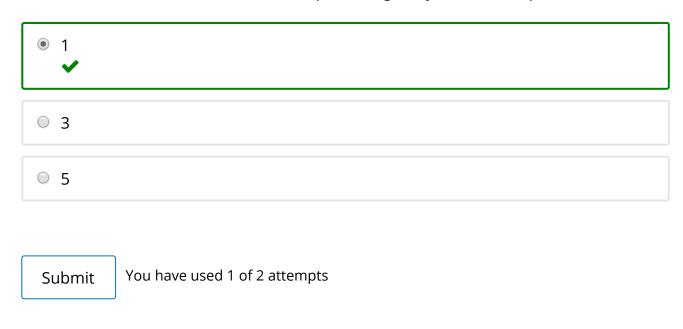
- evaluation_seed = 1239
- num_actions = 10
- trials = 10000
- distribution = "bernoulli"

Run the simulations with different number of R (1, 3, 5) and observe the results.

Lab Question

1.0/1.0 point (graded)

Which is the smallest ${\it R}$ that allow the optimistic greedy to find the optimal arm?



© All Rights Reserved