edX

Exercise 2A Epsilon Greedy
Exercise 2.2A: Epsilon Greedy

In this exercise, you will examine the epsilon 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 "Ex2.2A Epsilon Greedy.ipynb" notebook under **Module 2** folder.

1. Examine the notebook. We have given you an implementation of the epsilon greedy algorithm.

2. Once you have studied the notebook, prepare a simulation. Don't change any other parameter, that is:

   - evaluation_seed = 5016

   - num_actions = 10

   - trials = 10000

   - distribution = "bernoulli"

3. Set the epsilon to zero.

4. Run the simulation, observe the results, and answer the following questions.

Lab Question

1.0/1.0 point (graded)

With epsilon set to zero, what do you observe?

With epsilon set to zero, what do you observe?

○ The epsilon greedy behaves randomly

○ The epsilon greedy behaves like the optimistic greedy algorithm

◉ The epsilon greedy behaves like the greedy algorithm
✔

| Submit | You have used 1 of 2 attempts |

Set the epsilon to one. Run the simulation again and observe the results.

Lab Question

1.0/1.0 point (graded)

With epsilon set to one, what do you observe?

◉ The epsilon greedy behaves randomly
✔

○ The epsilon greedy behaves like the optimistic greedy algorithm

○ The epsilon greedy behaves like the greedy algorithm

| Submit | You have used 2 of 2 attempts |

Now, try several different number of epsilons (0.05, 0.1, 0.15). 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 epsilon gives the most average reward?

- ◉ 0.05
  ✔

- ○ 0.1

- ○ 0.15

| 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 epsilons (0.05, 0.1, 0.15) and observe the results.

Lab Question

1.0/1.0 point (graded)

Which epsilon gives the most average reward?

○   0.05

◉   0.1
    ✔

○   0.15

Submit	You have used 1 of 2 attempts