Assignment 01

Name:Abhishek Patel
Enrollment:23565001
Dept:MFS-AI
Email:abhishek_p@mfs.iitr.ac.in

Q1.Generate a multi-armed bandit instance using the following code snippet: Please use the seed as the sum of digits in your enrollment number. In this problem, we implement -Greedy strategies and plot expected reward, E(r t), vs time, t. Let the total time horizon T be 10000, with exploration parameter = 0.1.
(a) Implement G1, G2 and G3 algorithms (5 points).
(b) Plot E(r t) vs t for each of the above algorithms. (5 points)

Solution:

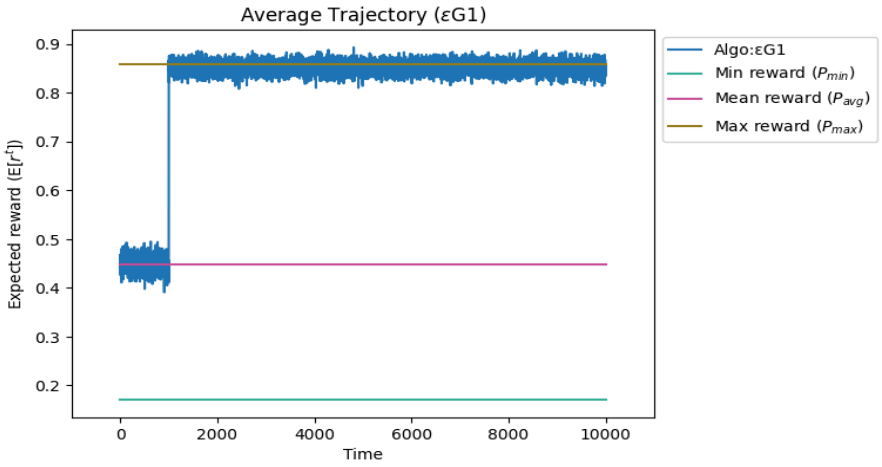
Probabilities of each bandit:

```
import numpy as np
# Enrollment:23565001
np.random.seed(22)
# arms=10
np.random.uniform(0,1,10)
```

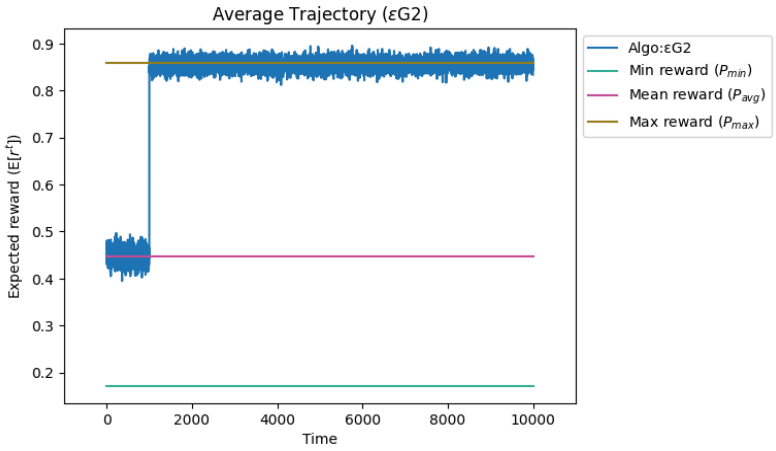
```
array([0.20846054, 0.48168106, 0.42053804, 0.859182  , 0.17116155,
       0.33886396, 0.27053283, 0.69104135, 0.22040452, 0.81195092])
```

Reward: Bernauli random variable

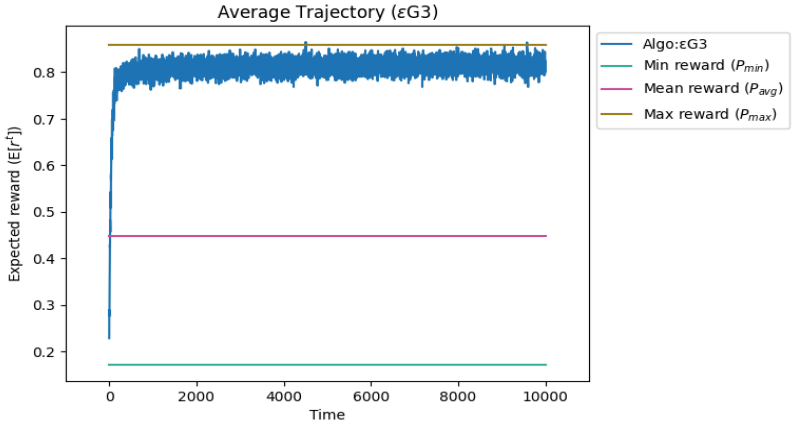
Epsilon Greedy 01



Epsilon Greedy 02



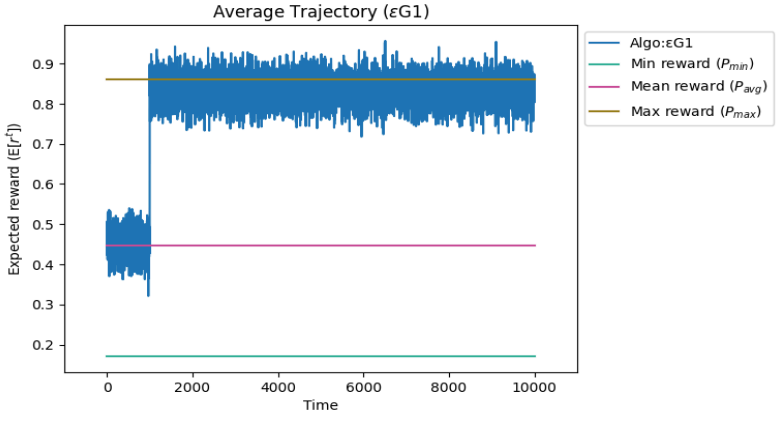
Epsilon Greedy 03



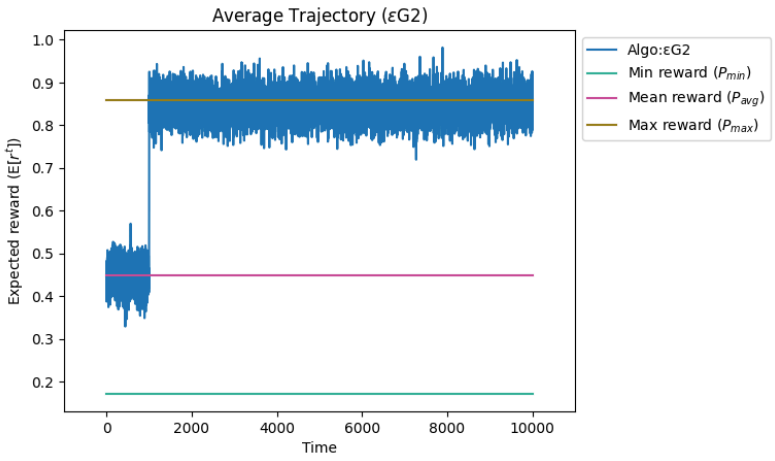
Q2 Repeat Q1 when each bandit gives reward from normal distribution with variance 1 and the mean same as above. (5 points)

Reward:Normal distribution

1. Epsilon Greedy 01



2. Epsilon Greedy 02



3. Epsilon Greedy 03