**Answer 8**

***5.2***

As per our inference in Lunar Lander, the learning rate will reduce in case of low epsilon value, as no new scenarios are being encountered.

**Case 1**: When epsilon is set at very low since the start of training

We get a graph as shown below. The orange line displays rewards when epsilon is set to 0.95 and the yellow line shows rewards, when epsilon is set to 0.05. In the latter case, convergence is not achieved quickly.

**Case 2**: When epsilon is set very high and never decreased.

We get a graph as shown below. The blue line displays the rewards with a decay function, while the orange line shows rewards over time without the decay function. The rewards converge after a period of time.

***5.3***

The Q values seem accurate when the epsilon is set high and isn’t decayed. However, the lunar lander does not land perfectly within the boundaries for every run. The reason lies with some properties that are left unconsidered.