# Exercise 4
[Avik Banerjee (3374885), Soumyadeep Bhattacharjee (3375428)]

*Text in italics are notes taken during the tutorial*

## 1   Monte Carlo Methods vs Dynamic Programming

a)
- Monte Carlo methods do not require the knowledge of the environment, particularly the transition probabilities and reward system.

- Monte Carlo methods are not biased by the initial conditions of the learning parameters (the initial values assigned to the value functions).

- *Less harmed by violations of Markov property, can be used with simulations, model-free, can focus on small subset of state space.*

b) Monte Carlo methods are used when it is not possible to have a complete knowledge of all possible state transitions and reward probabilities, such as when trying to teach an agent to play video games where there are millions of possible state and action combinations.

## 2   Monte Carlo ES for blackjack

a) Output plots: