Assignment 7

Question 3

- State Space $S_t = (x_t, s_t)$, where
  - $x_t = 1$ if employed at time t
    - 0 if unemployed at time t
  - $s_t \Rightarrow$ skill level at time t
- Action Space $A = \alpha \in [0,1]$ when $x_t = 1$
  - $= 0$ when $x_t = 0$
- Transition Probabilities

$$P((1,s), \alpha, (1, s(1-\alpha)g(s))) = 1-p$$

$$P((1,s), \alpha, (0, s(1-\alpha)g(s))) = p$$

$$P((0,s), 0, (0, se^{-\lambda})) = 1-h(s)$$

$$P((0,s), 0, (1, se^{-\lambda})) = h(s)$$

- Rewards

$$R((1,s), \alpha) = \alpha f(s)$$

$$R((0,s), 0) = 0$$

- If the horizon is long, then we would want to spend more time learning in the beginning so as to increase $f(s)$ & $h(s)$. We can then choose to earn in the later time steps.
- This trade-off on how much to learn & till when to learn will depend on the horizon and the discount factor gamma.
- For lower gammas, we have to balance earning & learning from the beginning. Same goes for shorter horizon
- Multiple jobs: • Each job i, will have its own $g_i()$ & $h_i()$ & $f_i()$
  - We will add an action c_t to choose the job at any time t.
- Consumption: To model this, we need to add e_t → accumulated earnings till time t in the state space.
  - Also, we will add action p_t → consumption of earnings at time t.
  - Modeling this would be interesting as it models the real world scenario.