AV499-AVD871- Assignment 2

1) Suppose $\mathcal{S}$ is a discrete set and let $f(\cdot)$ be a function defined on $\mathcal{S}$ with $f : \mathcal{S} \to \mathbb{R}$. Let $g(\cdot)$ be another function such that $g : \mathcal{S} \to [0,1]$ and $\sum_{s \in \mathcal{S}} g(s) = 1$. Prove that

$$\max_{s \in \mathcal{S}} f(s) \geq \sum_{s \in \mathcal{S}} g(s) f(s) \quad \text{for any } f(\cdot) \text{ and } g(\cdot)$$

2) Suppose X, Y, Z are three random variables with joint PMF $P_{XYZ}(x, y, z)$. Let $g(\cdot)$ be a deterministic function of these three random variables (i.e $g(x, y, z) \to \mathbb{R}$). Write down a formula for $\mathbb{E}\, g(X, Y, Z)$. Show that $\mathbb{E}\, g(X, Y, Z)$ can be computed as $\mathbb{E}_X \left[\mathbb{E}_{Y|X} \left[\mathbb{E}_{Z|XY} \left[g(X, Y, Z) \right] \right] \right]$.
In $\mathbb{E}_X \cdot \mathbb{E}_{Y|X} \cdot \mathbb{E}_{Z|XY}\, g(X, Y, Z)$ for each expectation write down what is random and what the random quantity's distribution is.

3) In class, we had used the following theorem to argue that for the infinite horizon case, it is enough to consider current state dependent policies:
" For every history dependent policy π^h with a joint distribution $P_r^h(\mathcal{S}_t = s_t, A_t = a_t)$ at time t, $\exists$ a current state dependent policy π^s such that $P_r^s(\mathcal{S}_t = s_t, A_t = a_t) = P_r^h(\mathcal{S}_t = s_t, A_t = a_t)$."

Write down a proof of the following result using the above theorem.
- For the finite horizon total expected reward Markov decision process, there exists an optimal policy within the set of current state dependent policies.