5512, Spring-2019
ASSIGNMENT 4:
Assigned: 04/06/19 Due: ~~04/17/19~~ 04/14/19 at 11:55 PM (submit via Canvas, you may scan or take a picture of your paper answers) <u>Submit only pdf or txt files (for non-code part), separate submission for code files</u>
<u>Show as much work as possible for all problems!</u>

Problem 0. (5 points)
Go on Canvas and fill out "are you going to do a project?".

Problem 1. (30 points)
(1) Assume job A will pay N(0,1) money and job B will pay N(0.5, 0.5) money. Does job B stochastically dominate job A?

(2) If two jobs play $N(\mu_x, \delta_x^2)$ and $N(\mu_y, \delta_y^2)$ for jobs X and Y. Under what conditions does job Y stochastically dominate job X?

(3) Is it possible for a uniform distribution to stochastically dominate a normal distribution? Is it possible for a normal distribution to stochastically dominate a uniform distribution? Justify.

Problem 2. (15 points)
Suppose you go to a casino and there are three slot machines. The three slot machines have rewards as follows (random variables):
Slot X = [(10%, $20), (30%, $5), (40%, $1), (20%, $0)]
Slot Y = [(5%, $40), (25%, $4), (30%, $2), (40%, $0)]
Slot Z = [(25%, $10), (25%, $5), (25%, $2), (25%, $0)]

It costs $4 to play any one of the slot machines and you are only allowed to play a single time (between all slot machines, not once per slot machine). You can see the three slot machines, but you have no idea which machine corresponds to which random variable. If you had the option to choose one of these slot machines and identify it (i.e. be able to determine if it is "Slot X", "Slot Y" or "Slot Z"), how much would you pay for to have this identification done?

Problem 3 & 4 use the following Markov Decision Process (MDP) with rewards as shown:

Assume that when moving there is a 70% chance to end up where you want to go and a 15% chance to end up 90 degrees left/right of where you want to go.

So for example, if you intend to go "up": there is a 70% chance you go up, 15% chance you go right and 15% chance you go left

You may assume that when you hit the 50 or -50, that you cannot move anymore and just get that reward then stop the "game".

	50	
	0	-3
-50	-1	-10
	-3	-2

Problem 3. (20 points)
For all parts of this problem, use the MDP given above and assume γ=0.8.

(1) Run value iteration until convergence and report the utilities for every state.

(2) If your initial guesses for the utilities are all zero, what is the least amount of iterations to find best policy?

(3) On the iteration you found in part (2), what is largest difference the estimated utility vs. actual utility (i.e. between parts (1) and (2)). How does this compare to the theoretical bound for the utility?

Problem 4. (15 points)
Use policy iteration to solve the MDP shown above. Start by assuming all actions are "Up" and γ=0.8.

Problem 5. (20 points)
Assume you have the following POMDP with rewards as shown below. Assume your initial guess of where you are in this POMDP is 20% in the top-left, 30% in the top-right and 50% in the bottom-right (also shown below). Assume the movement is the same as in Problems 3 & 4, but the only actions are moving "left" or "down". There is a boolean evidence variable, e, and P(e|s) is shown for all possible states in the third picture.

Find all possible belief states that result from taking two actions and their associated likelihoods. Which sequence of actions is best if you only take two actions?

Rewards:

-1	-4
2	-2

Initial guesses for states:

20%	30%
0%	50%

P(e|s):

0.3	0
0.9	0.2