IEOR 4575: Reinforcement Learning

Homework 0

Instructor: Shipra Agrawal

Note. *This homework is meant for assessing your mathematical background and pre-requisites for the course. You must work on this assignment by yourself. If you have mastered the prerequisites for the course, 50% or more of the problems should be readily solvable, and the rest with some self-study and help.*

Probability and Statistics

Problem 1. (10 points) There is 75% chance of rain on Monday. There is 50% chance of rain on Tuesday if that it rains on Monday, and 20% otherwise. If it rains on any given day, there is a 70% chance that Alice will get wet on her way to work. (She does not change her behavior based on what happened the earlier day). What is the probability that Alice will get wet from rain on both Monday and Tuesday?

Problem 2. (15 points) A standard 52-card deck [1] is randomly partitioned into four 13-element sets, which are dealt to players named Alice, Bob, Tom, and Harry.

(a) Calculate $\Pr(\text{Alice gets exactly 2 aces}|\text{Bob gets exactly 1 ace})$.
Hint: Use $\Pr(A|B) = \Pr(A, B)/\Pr(B)$.

(b) Let C and S denote the number of clubs and spades, respectively, dealt to Alice. Calculate $E(C|S)$ as a function of S.
Hint : Observe that given the number of spades, the number of clubs, hearts, diamonds dealt to Alice have the same conditional distribution.

Problem 3. (15 points) You are selling your bike. You get offers one by one. You have decided to stop and accept an offer as soon as you see one which is better than the first offer you got. (You always skip the first offer). What is the expected number of offers you will to wait for, including the first one, until you accept an offer? Mathematically, let's model this process as follows. Let X_1, X_2, ... denote an infinite sequence of independent uniformly-distributed random samples from the interval $[0, 1]$. (Interpretation: X_i is the i^{th} offer.) Let τ be the smallest $i > 1$ such that $X_i > X_1$. What is $E[\tau]$?
Hint: Use following formula for expected value of a non-negative integer valued random variable Z*:* $E[Z] = \sum_{n=0}^{\infty} \Pr(Z > n)$.

Problem 4. (30 points)

(a) You have multiple bikes to sell. For simplicity, assume that you get one offer every day, with X_i denoting the offer received on day i. Here X_1, X_2, . . . are a sequence of independent random variables uniformly distributed in $[0, 1]$. Now suppose that you accept *every* offer $X_i, i \geq 2$ such that $X_i > X_1$. You always reject the offer X_1 on day 1. Let Y_n be the number of accepted offers until and including day n. That is, $Y_n = \sum_{i=2}^{n} I(X_i > X_1)$, where $I(\cdot)$ is the indicator function. What is $E[Y_n]$?

[1] A standard 52-card deck is the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, king, ace\} \times \{clubs, diamonds, hearts, spades\}$.

(b) Consider the problem setting in part (a) again, and now let Z_n be the number of accepted offers until and including day n assuming $X_1 = 1/2$. That is, $Z_n = \sum_{i=2}^{n} I(X_i > 1/2)$, where $I(\cdot)$ is the indicator function. What is $E[Z_n]$? Also, prove that with probability at least $1-\delta$, for any $\delta \in (0,1)$, we have that $Z_n \leq \frac{n}{2} + O(\sqrt{n \log(1/\delta)})$. That is,

$$\Pr\left(Z_n \leq \frac{n}{2} + O(\sqrt{n \log(1/\delta)}\right) \geq 1-\delta$$

Hint: Use Chernoff-Hoeffding bounds (Look here or here) Also familiarize yourself with the Big O notation.

(c) Now suppose that you accept the offer on day i if it is better than the average offer from day $1, \ldots, i-1$. In more precise terms, let X_1, X_2, . . . be a sequence of independent random variables uniformly distributed in $[0,1]$ as before. Then you accept *every* offer $X_i, i = 2, 3, \ldots$, such that $X_i > \frac{1}{(i-1)} \sum_{j=1}^{i-1} X_j$. You always reject the offer X_1 on day 1. Let W_n be the number of accepted offers by day n. That is, $W_n = \sum_{i=2}^{n} I(X_i > \frac{1}{(i-1)} \sum_{j=1}^{i-1} X_j)$. What is $E[W_n]$?
Hint: Note that for two random variables A and B, $\Pr(A \geq B)$ can be written as $E_B[\Pr(A \geq B|B)]$.

Optimization

Problem 5. (15 points) Consider the following Linear Program (LP):

$$\begin{aligned}
\text{Minimize}_{v,x_1,x_2,x_3} \quad & v \\
s.t. \quad v + x_1 & \geq 2 + 0.5x_1 + 0.2x_2 + 0.3x_3 \\
v + x_2 & \geq 3 + 0.1x_1 + 0.9x_2 \\
v + x_3 & \geq 5 + 0.2x_1 + 0.95x_2 + 0.05x_3
\end{aligned}$$

Formulate the dual of the above LP.

Problem 6. (15 points) Consider the function $f(x) = \log(\sum_{i=1}^{n} e^{x_i})$.

1. What is the gradient $\nabla f(x)$ of f? What is the gradient at $x = (1, 1, \ldots, 1)$?

2. What is the Hessian $\nabla^2 f(x)$ of f? What is the Hessian at $x = (1, 1, \ldots, 1)$?

3. Consider the following optimization problem:

$$\begin{aligned}
\min_{x} \quad & \log(\textstyle\sum_{i=1}^{n} e^{x_i}) \\
s.t. \quad & \textstyle\sum_i x_i = 1
\end{aligned}$$

Argue that this is a convex optimization problem and solve it analytically using optimality conditions. State the optimal solution.