MTL106 Assignment 1

Puran Mayur
2023MT10366

20 September 2024

(For entry number 2023MT10366, $T1 = 9$, $T2 = 3$, $T3 = 6$, $T4 = 6$, $T1T2 = 93$, $T3T4 = 66$, $T = 159$)

1 Question 1

1.1 Part A

For part a, we need to find the probability that after T, rounds Alice wins $T1T2$, and Bob wins $T3T4$. For this, first we define a matrix dp that contains the probabilities of Alice winning i times and Bob winning j times at the $(i, j)th$ index (which is $P(i, j)$). We then can consider the following recursion:

$$P(i, j) = dp[i][j] = \frac{j}{i + j - 1} dp[i - 1][j] + \frac{i}{i + j - 1} dp[i][j - 1]$$

Now this recursion is valid, because if we know the probability of Alice and Bob having $i-1$, j and i, $j-1$ wins, we can get to the i, jth value of the matrix by multiplying individual probabilities of Alice and Bob. Also there is no other way to get to i, j wins, so the sum of these two gives us the answer.

Probability of Alice winning at the $(i + j - 1)$th round $= \frac{n_b}{n_a+n_b} = \frac{j}{i+j-1}$
Probability of Alice winning at the $(i + j - 1)$th round $= \frac{n_a}{n_a+n_b} = \frac{i}{i+j-1}$

Now to solve the recursion, we just need a base case, and that is already given, as $dp[0][1] = 1, dp[1][1] = 1$

Hence, we solve this recursion using dynamic programming in our python code, and the value of dp[93][66], which is our answer, and the corresponding value of x comes out to be

834691793

where $x.q \equiv p(modM)$, and $\frac{p}{q}$ is the required probability.

1.2 Part B

For this part, let us first define a random variable $Y = \sum_{i=1}^{T} X_i$, for a given T. We can see that the random variable Y gives us the difference of points between Alice and Bob after round T, as for all the rounds from 1 to i, Y will increase by 1 if Alice wins, and decrease by 1 if Bob wins. (draw isn't possible as probability of draw is 0 if both attack)

$$E(Y) = E\left(\sum_{i=1}^{T} X_i\right) = E\left(E\left(\sum_{i=1}^{T} X_i \mid Z = z\right)\right) = \sum_{z=0}^{T}(2z - T) \cdot P(T, z)$$

where Z is the random variable for the number of points of Alice.

Now, E[Y] $= \sum_{i=1}^{T}(P(i,j))(i-j)$

where j is fixed as $j = T - i$

Using the recursion in part a, we find values of $P(i,j)$, and find the expectation of Y, which comes out to be **0**.

We do a similar calculation for Var(Y). Since E[Y] $= 0$,

$$Var(Y) = E\left[Y^2\right] - (E\left[Y\right])^2 = E\left[Y^2\right]$$

So the Var(Y) will just be equal to

$$E\left[Y^2\right] = \sum_{i=1}^{T} P(i,j)(i-j)^2$$

$j = T - i$
Again using the recursion in part a, we get:

Var(Y) = 22

2 Question 2

2.1 Part A

In the given play style of Bob, we always know what he will do next based on the previous game, and hence the optimal strategy for Alice will be to play that move in response that has the probability of increasing her points the most.

The following table shows the expected values of increase in points for Alice against every move of Bob.

	Bob		
	Attack	**Balanced**	**Defence**
Alice Attack	$\frac{n_B}{n_A+n_B}$	$\frac{7}{10}$	$\frac{5}{11}$
Alice Balanced	$\frac{3}{10}$	$\frac{1}{2}$	$\frac{11}{20}$
Alice Defence	$\frac{6}{11}$	$\frac{9}{20}$	$\frac{1}{2}$

where these evaluates can be calculated from the expression:
P(Alice Wins).(1 point) + P(Alice Draws).($\frac{1}{2}$ point)

From this table, we can conclude what move will be optimal for Alice in response to a move by Bob.

If Bob plays *balanced*, Alice should always attack as that maximises her probability of increasing points.

If Bob plays *defensive*, Alice should still always attack as attack still gives her the best expected value.

If Bob *attacks*, we can either attack or defend (not balanced as $\frac{3}{10} < \frac{6}{11}$). The optimal strategy would be to attack when ($\frac{n_B}{n_A+n_B} > \frac{6}{11}$), and defend otherwise.

Plugging all of this into the given starter code, and running the Monte Carlo simluation for 10^5 iterations, we can see that Alice has an advantage in points over Bob.

2.2 Part B

Yes it is indeed optimal for Alice to play the greedy strategy from part A in every round to maximise the total points, because by using greedy algorithm we are basically calculating the maximum expected points increase per round, and using greedy will maximise that over the whole rounds.

2.3 Part C

To calculate the expected number of rounds (τ) for Alice to win T times, we can simulate a function that runs multiple trials (defaulting to 100,000). In each trial, we can initialized instances of Alice and Bob, then conduct rounds until Alice achieves T wins, counting the total rounds taken. After completing all trials, we can average the number of rounds across all simulations to estimate E[T], providing an estimation of the expected rounds required for Alice to reach her winning goal.

3 Question 3

3.1 Part A

Similar to question 2 part a, for maximum points in the current round for Alice, we need to consider the expected increase in points for each play style of Alice.

Firstly, since Bob plays all three play styles at random, the probability of him playing any specific play style is $\frac{1}{3}$.

Next, we need to calculate of increase in points for each play style of Alice.

If Alice attacks:

$$expected\ points = \frac{n_B}{n_A + n_B} \cdot \frac{1}{3} + \frac{7}{10} \cdot \frac{1}{3} + \frac{5}{11} \cdot \frac{1}{3} = \frac{n_B}{n_A + n_B} \cdot \frac{1}{3} + \frac{127}{330} \tag{1}$$

If Alice plays balanced:

$$expected\ points = \frac{3}{10} \cdot \frac{1}{3} + (\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2}) \cdot \frac{1}{3} + (\frac{3}{10} + \frac{1}{2} \cdot \frac{1}{2}) \cdot \frac{1}{3} = \frac{9}{20} \tag{2}$$

If Alice defends:

$$expected\ points = \frac{6}{11} \cdot \frac{1}{3} + (\frac{1}{5} + \frac{1}{2} \cdot \frac{1}{2}) \cdot \frac{1}{3} + (\frac{1}{10} + \frac{4}{5} \cdot \frac{1}{2}) \cdot \frac{1}{3} = \frac{329}{660} \tag{3}$$

Since $\frac{329}{660} > \frac{9}{20}$, Alice is never better off playing balanced on average, so we should either attack or defend.
Therefore, we will attack when $\frac{n_B}{n_A+n_B} \cdot \frac{1}{3} + \frac{127}{330} > \frac{329}{660}$, i.e. $\frac{n_B}{n_A+n_B} + \frac{127}{110} > \frac{329}{220}$, and defend otherwise.
Putting this into the python code, we can see that with this strategy Alice indeed finds herself in a better off position than Bob in terms of points.

3.2 Part B

For Alice's strategy to maximise the expected number of points in T (=T3T4) future rounds, we can use an approach that assumes the expected values of points of Alice in the next round, and then uses backtracking to get to the current answer.

Let dp[i][j] represent the expected points gained by Alice after T rounds when she has i/2 points initially, and Bob has j/2 points initially.

Now we can define the recursion as follows:

a = $dp[i+2][j] \cdot (\frac{n_B}{n_A+n_B} \cdot \frac{1}{3} + \frac{7}{10} \cdot \frac{1}{3} + \frac{5}{11} \cdot \frac{1}{3})$

b = $dp[i+1][j+1] \cdot (\frac{3}{10} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + \frac{3}{10} \cdot \frac{1}{3})$

c = $dp[i][j+2] \cdot (\frac{6}{11} \cdot \frac{1}{3} + \frac{1}{5} \cdot \frac{1}{3} + \frac{1}{10} \cdot \frac{1}{3})$

dp[i][j] = max(a,b,c)