COT 3100 Fall 2020 Homework #9 Solutions

1) (5 pts) Two dice are rolled. Both have six sides labeled 1 through 6, inclusive, but the second die is not fair. (The first one is.) The probability of the second die landing on the side with k dots is $\frac{k}{21}$. What is the probability of rolling a sum of 9 when these two dice are rolled?

Solution

Let a dice roll be represented as (a, b), where a is the value shown on the first (fair) die, and b is the value shown on the biased die. The four outcomes that lead to a sum of 9 are:

(6, 3), probability = $\frac{1}{6} \times \frac{3}{21} = \frac{3}{126}$

(5, 4), probability = $\frac{1}{6} \times \frac{4}{21} = \frac{4}{126}$

(4, 5), probability = $\frac{1}{6} \times \frac{5}{21} = \frac{5}{126}$

(3, 6), probability = $\frac{1}{6} \times \frac{6}{21} = \frac{6}{126}$

The probabilities of rolling each particular ordered pair is shown above. The sum of these probabilities is $\frac{18}{126} = \boldsymbol{\frac{1}{7}}$.

2) (8 pts) Let X be a continuous random variable described below:

$$p(X) = x/4, \text{ for all x in the range } 0 \le x \le 2$$
$$= x^2/8, \text{ for all x in the range } 2 < x \le k$$

(a) What is the value of k?

(b) What is E(X)?

Solution

(a) The area under the curve described by p(X) must equal 1. The first portion of the function from x = 0 to x = 2 represents a right triangle with base 2 and height 1/2, which has an area of 1/2. Thus, in order for the function to be valid, the area under the other portion of the function must also equal 1/2. Here is the desired integral that must be satisfied:

$$\int_2^k \frac{x^2}{8}\,dx = \frac{1}{2}$$

$$\frac{x^3}{24}\Big|_2^k = \frac{1}{2}$$

$$\frac{k^3 - 8}{24} = \frac{1}{2}$$

$$k^3 - 8 = 12$$
$$k^3 = 20$$
$$\boldsymbol{k = \sqrt[3]{20}}$$

(b) Now, plug into the definition of expectation and solve. Since the function is a step function, the integral has to be split into two different integrals:

$$\int_0^2 x\left(\frac{x}{4}\right)dx + \int_2^{\sqrt[3]{20}} x\left(\frac{x^2}{8}\right)dx =$$

$$\int_0^2 \left(\frac{x^2}{4}\right)dx + \int_2^{\sqrt[3]{20}} \left(\frac{x^3}{8}\right)dx =$$

$$\left.\frac{x^3}{12}\right|_0^2 + \left.\frac{x^4}{32}\right|_2^{\sqrt[3]{20}} =$$

$$\frac{8}{12} - 0 + \frac{20\sqrt[3]{20}}{32} - \frac{16}{32} =$$

$$\boldsymbol{\frac{1}{6} + \frac{5\sqrt[3]{20}}{8}}$$

3) (7 pts) A binary search tree is created by iteratively inserting elements into the tree. To insert an element, into an existing tree, put the element to the left side of the tree if it is less than the value stored at the root (top) of the tree, and put the element to the right side of the tree if it is greater than the value stored at the root of the tree. Repeat this process until there is an opening to insert the item. For example, for the tree shown below:

```
      5
    /   \
   2     8
```

the value 1 would be inserted to the left of the 2, the values 3 and 4 would be inserted to the right of the 2, the values 6 and 7 would be inserted to the left of the 8 and anything greater than 8 would be inserted to the right of 8, for the next insertion. Where elements end up in the tree depends on the order that they are inserted.

If the values 1, 2, 3, 4, 5, 6, and 7 were inserted in random order into a binary search tree, what is the probability that the final tree structure would look like this:

```
         4
      /     \
     2       6
    / \     / \
   1   3   5   7
```

Solution

There are 7! orders in which we can insert the elements. This is our sample space.

We know that 4 must be inserted first, so we have 1 choice for the first insertion. This is followed by 6 more insertions. Here is a picture of our ordering:

4 , ___ , ___ , ___ , ___ , ___ , ___

We must choose 3 slots out of 6 for inserting 1, 2, and 3. We can do this in $\binom{6}{3}$ ways. Consider one possible option highlighted in yellow:

4 , ___ , ___ , ___ , ___ , ___ , ___

Notice that 2 must go in the first slot, but 1 and 3 can go in either order for the second slot. Thus, for this, we multiply by 2. Namely, there are $2\binom{6}{3}$ ways to select slots for 1, 2 and 3. Once these slots are selected, there are only 2 ways left to select slots for 5, 6 and 7, since 6 must go in the earliest slot while 5 and 7 can be inserted in either order. Thus, the total number of insertion orders that result in the given tree is $2 \times 2\binom{6}{3} = 80$ and the desired probability is $\frac{80}{7!} = \frac{80}{5040} = \mathbf{\frac{1}{63}}$**.**

4) (5 pts) A box contains 5 apples and 6 oranges. Four children each receive a fruit from the box, one after the other, randomly chosen, without replacement. What is the probability that all four children receive the same fruit?

Solution

Just add up the probability of choosing 4 apples in row and the probability of choosing 4 oranges in a row. Multiplication principle can be used for both, noting the number of apples or oranges left after each draw. The result is:

$$\frac{5}{11} \times \frac{4}{10} \times \frac{3}{9} \times \frac{2}{8} + \frac{6}{11} \times \frac{5}{10} \times \frac{4}{9} \times \frac{3}{8} = \frac{(5)(4)(3)(2+6)}{11(10)(9)(8)} = \mathbf{\frac{2}{33}}$$

5) (5 pts) A biased coin, when tossed 7 times, has the same probability of obtaining 2 heads out of 7 as it does of obtaining 3 heads out of the 7 tosses. What is the probability the coin lands heads on a single toss?

Solution
Let p be the probability of obtaining heads on a single toss. Using the binomial probability distribution, we have:

$$\binom{7}{2}p^2(1-p)^5 = \binom{7}{3}p^3(1-p)^4$$

$$21(1-p) = 35p$$

$$21 - 21p = 35p$$

$$21 = 56p$$

$$p = \mathbf{\frac{3}{8}}$$

6) (5 pts) What is the probability that a randomly chosen divisor of 30^{49} is a multiple of 60^{19}?

Solution
$30^{49} = 2^{49}3^{49}5^{49}$. Thus, a randomly chosen divisor of this number has the form $2^a3^b5^c$, where $0 \le a,b,c \le 49$. This is a sample space of size 50^3. We need to determine the number of these values that are multiples of $60^{19} = (2^2)^{19}3^{19}5^{19} = 2^{38}3^{19}5^{19}$. To be a multiple of this value we have the following ranges for a, b and c, respectively:

$38 \le a \le 49$
$19 \le b \le 49$
$19 \le c \le 49$

There are 12 x 31 x 31 such ordered triplets (a, b, c) that satisfy the inequalities above. It follows that the desired probability is $\frac{12\times 31\times 31}{50\times 50\times 50} = \mathbf{\frac{2883}{31250}}$.

7) (10 pts) A major TV network would like to maximize the number of games a best of seven sports series is played. Unfortunately, the teams, team A and team B, are not evenly matched. If team A plays at its home stadium, it has a 70% chance of winning. If team A plays at team B's stadium, it still has a 45% chance of winning. Luckily, there is a clause in the contract that the TV network gets to choose which stadium the games are played in, so long as no more than 4 games are scheduled for either stadium. But, the TV network must schedule the location of all games before the series starts. (So, they can't see who actually wins a game before deciding where to schedule the next one.) How should the network schedule the games to maximize the expected number of games played, and what is the expected number of games played for that schedule. Note: a complete schedule is an ordering of seven game locations, where each location is either team A or team B, and there aren't more than 4 As or 4 Bs. For example, the schedule ABBAAAB, represents that we play the first game, fourth game, and if necessary the fifth and sixth games at team A's home stadium. Also note that the teams stop playing games as soon as one team reaches four wins, and that each game always ends in a win for exactly one team. (There are no draws.) **Note: This is a challenging question, don't feel bad if you don't get it!**

Solution

Since team A has a heavy advantage. It seems fairly clear that we want 3 home games for team A and 4 home games for team B. There may be situations where the probabilities are closer where this isn't the case, but for the big difference here, this is a safe assumption.

Once we know how many home games there are in the first four games, the actual order of those games doesn't matter, since each ordering will lead to the same probability distribution for wins of team A, and all four of these games are guaranteed to be played. So, we can split up our work into four cases for the first four games:

1) 0 home games
2) 1 home game
3) 2 home games
4) 3 home games

Case 1: AAAAHHH

Since we are playing all 4 games away, that means the last three must be at home. We can work out the expectation exactly for this one case.

Probably the easiest way to do this is to make a table with each entry representing the number of wins for team A and team B. The entry in row a, column b, prob[a][b], will represent the probability that after a+b games, we have a wins for team A and b wins for team b. This table is relatively easy to make because we can build it from previous entries. In particular, we have

prob[a][b] = prob[a-1][b]*pA + probl[a][b-1]*(1-pA)

where pA is the probability of team A winning the $(a+b)^{th}$ game.

Here is the table:

B A	0	1	2	3	4
0	1	.55	$.55^2$	$.55^3$	$.55^4$
1	.45	(2)(.55)(.45)	$(3)(.55)^2(.45)$	$(4)(.55)^3(.45)$	$(4)(.55)^3(.45)(.3)$
2	$.45^2$	$(3)(.55)(.45)^2$	$(6)(.55)^2(.45)^2$	$(6)(.55)^2(.45)^2(.3)+$ $(4)(.55)^3(.45)(.7)$	$[(6)(.55)^2(.45)^2(.3)+$ $(4)(.55)^3(.45)(.7)](.3)$
3	$.45^3$	$(4)(.55)(.45)^3$	$(4)(.55)(.45)^3(.3)\ +$ $(6)(.55)^2(.45)^2(.7)$		
4	$.45^4$	$(4)(.55)(.45)^3(.7)$	$[(4)(.55)(.45)^3(.3)\ +$ $(6)(.55)^2(.45)^2(.7)](.7)$		

We can use this partially filled out table to calculate the probability the series ends in 4, 5, 6 or 7 games. Namely, the table has the answers for 4, 5 and 6, and we can subtract from 1 to get the probability of getting to 7 games.

P(4 games) = $.45^4 + .55^4 = \frac{9^4+11^4}{20^4} = \frac{21202}{160000} = \frac{10601}{80000}$

P(5 games) = $(4)(.55)(.45)^3(.7) + (4)(.55)^3(.45)(.3) = \frac{4(11\times9^3\times7+11^3\times9\times3)}{(10)20^4} = \frac{92070}{400000} = \frac{9207}{40000}$

P(6 games) = $[(4)(.55)(.45)^3(.3) + (6)(.55)^2(.45)^2(.7)](.7) + [(6)(.55)^2(.45)^2(.3)+(4)(.55)^3(.45)(.7)](.3)$

$= \frac{4\times11\times9^3\times3\times7+6\times11^2\times9^2\times(7\times7+3\times3)+4\times11^3\times9\times7\times3}{10^2 20^4} = \frac{5090580}{16000000} = \frac{254529}{800000}$

P(7 games) = $1 - \frac{106010}{800000} - \frac{184140}{800000} - \frac{254529}{800000} = \frac{255321}{800000}$

The corresponding expectation for the number of games in the series is:

E(X) = $4\left(\frac{106010}{800000}\right) + 5\left(\frac{184140}{800000}\right) + 6\left(\frac{254529}{800000}\right) + 7\left(\frac{255321}{800000}\right) = \frac{4659161}{800000} = 5.82395125$

Case 2: HAAAAHH or HAAAHAH or HAAAHHA

Of these cases, it turns out the one that gives the maximum expectation for the length of the series is the sequence **HAAAAHH**, which has the same exact expectation as the sequences AHAAAHH, AAHAAHH, and AAAHAHHH.

Here is the corresponding chart for this case:

B A	0	1	2	3	4
0	1	.3	(.3).55	$(.3).55^2$	$(.3).55^3$
1	.7	(.3)(.45)+(.7)(.55)	$2(.3)(.55)(.45)+$ $(.7)(.55)^2$	$3(.3)(.55)^2(.45)+$ $(.7)(.55)^3$	$3(.3)(.55)^3(.45)+$ $(.7)(.55)^4$
2	(.7).45	$2(.7).45(.55)+$ $(.3)(.45)^2$	$3(.3)(.55)(.45)^2+$ $3(.7)(.55)^2(.45)$	$6(.3)(.55)^2(.45)^2+$ $4(.7)(.55)^3(.45)$	$6(.3)^2(.55)^2(.45)^2+$ $4(.7)(.3)(.55)^3(.45)$
3	$(.7).45^2$	$3(.7)(.55)(.45)^2+$ $(.3)(.45)^3$	$4(.3)(.55)(.45)^3+$ $6(.7)(.55)^2(.45)^2$		
4	$(.7).45^3$	$3(.7)(.55)(.45)^3+$ $(.3)(.45)^4$	$4(.3)(.7)(.55)(.45)^3+$ $6(.7)^2(.55)^2(.45)^2$		

Let's work out the probability of achieving each number of games for the series:

P(4 games) = $(.7).45^3 + (.3).55^3 = \frac{7\times 9^3+3\times 11^3}{(10)20^3} = \frac{9096}{80000} = \frac{1137}{10000}$

P(5 games) = $3(.7)(.55)(.45)^3+(.3)(.45)^4 + 3(.3)(.55)^3(.45)+(.7)(.55)^4$

$$= \frac{3\times 7\times 11\times 9^3+3\times 9^4+3\times 3\times 11^3\times 9+7\times 11^4}{10(20^4)} = \frac{398380}{1600000} = \frac{19919}{80000}$$

P(6 games) = $4(.3)(.7)(.55)(.45)^3 + 6(.7)^2(.55)^2(.45)^2 + 6(.3)^2(.55)^2(.45)^2 + 4(.7)(.3)(.55)^3(.45)$

$$= \frac{4\times 3\times 7\times 11\times 9^3+6\times 7^2\times 11^2\times 9^2+6\times 3^2\times 11^2\times 9^2+4\times 7\times 3\times 11^3\times 9}{10^2 20^4} = \frac{5090580}{16000000} = \frac{254529}{800000}$$

P(7 games) = $1 - \frac{90960}{800000} - \frac{199190}{800000} - \frac{254529}{800000} = \frac{255321}{800000}$

E(X) = $4\left(\frac{90960}{800000}\right) + 5\left(\frac{199190}{800000}\right) + 6\left(\frac{254529}{800000}\right) + 7\left(\frac{255321}{800000}\right) = \frac{4674211}{800000} = \mathbf{5.84276375}$

When we look at the other arrangements of the games in this case, case 3 and case 4, their best expectation is below this. To verify this, I wrote a program called playoffs.c which goes through and calculates the aforementioned table for all 35 orderings of the games. This file is posted along with this solution.

Thus, networks should schedule the games in one of the four following orders:

Home, Away, Away, Away, Away, Home, Home
Away, Home, Away, Away, Away, Home, Home
Away, Away, Home, Away, Away, Home, Home
Away, Away, Away, Home, Away, Home, Home

For each of these schedules, the network can expect $\mathbf{5.84276375}$ games.

8) (5 pts) Please give a summary of the life and mathematical contributions of Pafnuty Chebyshev. Please aim for a length of roughly 200 - 400 words. ***<u>Your summary must be typed.</u>*** Please state the sources you used in writing your summary.

<u>Sample Summary</u>

Pafnuty Chebyshev was born in Russia in 1821, one of nine children to wealthy landowners. Chebyshev had Trendelenburg's gait and walked with a stick, so being an officer, as was the family tradition was out of the question. Instead, Chebyshev took an interest in mathematics very early in his life. When he was 11, his family moved to Moscow, where he had excellent mathematics teachers. In 1837, he started attending Moscow University. In 1841, he graduated, earning praise for his work on finding roots of polynomials of the n^{th} degree, building on Newton's work. As Chebyshev continued his studies, he discovered probability in relation to the insurance industry and wrote his master's thesis on the theory of probability. It's his work in probability for which he is best known. Chebyshev transitioned to St. Petersburg University to earn his doctorate, and did so in number theory, building upon rediscovered works of Leonard Euler. From 1850 through 1882, Chebyshev worked as a professor at St. Petersburg University.

Chebyshev proved what is known as the Chebyshev inequality, which states that the probability of a random variable being further than a standard deviations from the mean is less than $1/a^2$. This inequality is used to prove the weak law of large numbers. He also discovered what is known as the Chebyshev bias. Although prime numbers of various forms are equally distributed, for all values less than 26,681 there are either an equal number or more primes of the form 4k+3 than there are primes of the form 4k+1. He is also known for Chebyshev polynomials, which are two sequences based on sine and cosine. More specifically, these polynomials are involved in expression sin (nθ) and cos (nθ) in terms of sin (θ) and cos (θ).

Chebyshev is known as the father of Russian mathematics, and taught many students. To date, he has over 10,000 mathematical descendants.

<u>Sources</u>

https://en.wikipedia.org/wiki/Pafnuty_Chebyshev