CALCULATING PROBABILITIES

BY: Farhan Islam, Haroun Mohamed, Yousef Mohamed

INTRODUCTION:

The question we picked was number 7 from the senior category. The question is about how an anxious teacher is pacing up and down a hallway, with a probability '*p*' to move right and probability '*q*' to move left, and $p + q = 1$. Since the probability of him to move left or right is 1 (100%) we can safely assume that he never stops moving.
If we denote by $f(n, t)$ the probability that the teacher is at position n at time t, then at time $t + 1$ the probability is given in terms of the probabilities at time t via the relation:
$f(n, t + 1) = pf(n - 1, t) + qf(n + 1, t)$.

We will construct an explicit formula to calculate the probability of the teacher to be at a certain distance, given by the letter '*n*', from the origin at a certain time. '*n*' represents the displacement relative to the starting point (i.e., when he moves 3 units to the left n is –3). $t \geq 0$ (We assumed that the teacher takes exactly one step per second).

Rules set down:

We first established a list of rules for a better understanding of the problem at hand.
Firstly, the teacher always starts at the origin, and thus the condition at $t = 0$ is:

$$f(0,0) = 1, \qquad f(n,0) = 0 \quad n \neq 0$$

Moreover, when t is odd, n must be odd and vice versa, otherwise the probability will equal 0 (we consider 0 to be even).

Constructing the Formula:

We proceeded to calculate values for the probability with constants p and q. This was done by hand for more on hand knowledge of what is going on. A list of values is shown below:

$f(1,1) = p$	$f(-1,1) = q$	$f(0,2) = 2pq$	$f(2,2) = p^2$
$f(1,3) = 3p^2q$	$f(3,3) = p^3$	$f(0,4) = 6\ p^2q^2$	$f(2,4) = 4\ p^3q$
$f(4,4) = p^4$	$f(1,5) = 10\ p^3q^2$	$f(3,5) = 5\ p^4q$	$f(5,5) = p^5$
$f(-2,2) = q^2$	$f(-1,3) = 3q^2p$	$f(-3,3) = q^3$	$f(-2,4) = 4\ q^3p$
$f(-4,4) = q^4$	$f(-1,5) = 10\ q^3p^2$	$f(-3,5) = 5\ q^4p$	$f(-5,5) = q^5$

Finding patterns:

The values tell us that the probability alternates between *p* and *q* when *n* is positive and negative, respectively.
We also see that the powers of *p* and *q* go up/down by one when the value of *n* is increased.

Formula for the value of the powers of *p* and *q:*

The powers when *t* is even always start at (*n* = 0) *t/2*. When compared with the powers when *t* is odd, which always start at (*n* = 1) ½ over *t/2* or ½ under *t/2*, the following expression can be derived.

$$p^{\frac{t}{2}+\frac{n}{2}} \times q^{\frac{t}{2}-\frac{n}{2}}$$

In the case where *n* is negative and the powers switch, this formula still works as the (+) and (–) will switch when the (–) sign is introduced, and thus switching the powers.

Formula for the value of the coefficients:

The coefficient seems to come from Pascal's Triangle, where *t* is the horizontal row number. The diagonal column can be achieved by subtracting *n* from *t* and diving by 2. This is summarized in the following equation, where m is the coefficient:

$$m = {}^{t}C_{(t-n)/2}$$

Now we combine the last 2 equations we created to achieve our final formula:

$$f(n,t) = p^{\frac{t}{2}+\frac{n}{2}} \times q^{\frac{t}{2}-\frac{n}{2}} \times {}^{t}C_{(t-n)/2}$$

We can now compare our values from the formula with the values from the following simple python code (where *p* = *q* = *0.5*) to ensure that our formula works, which shown below shows that it does.

Execute | Share | main.py | STDIN

```python
def prob(number, time):
    if abs(number) > time:
        return 0
    elif time == 0:
        return 1
    else:
        return (0.5 * prob(number-1, time-1) + 0.5 * prob(number+1,time-1))
number = -1
time = 3
print(prob(number,time))
```

Result

```
$python main.py
0.375
```

Moving in 2 Perpendicular Directions:

After assuming that the teacher only moves towards North and East when expressed on a Cartesian Plane, the teacher convers the 1St Quadrant only. Since we are dealing with only 2-dimensions, we will represent his movements as a set of coordinates on the Cartesian Plane. Let ***f((x,y),t)*** be the probability that he is at point *(x,y)* at time *t*, where **x∈R and y∈R**. The probability that he walks North is *N* and East is *E*, and $N + E = 1$.

Since he is unable to walk back to a point after he has passed it, $x + y = t$.

The condition at $t = 0$ is: $f((0,0),0) = 1, f((x,y),0) = 0$ $\{ x \neq 0, y \neq 0 \}$

Some rules:

To start, we must first express the probability in a recursion relation like the one used in the first section. For him to be at a point, he either must walk east from a point to the west of the original point or walk north from a point to the south of the original point. This is shown in the relation below:

$$f((x,y),t) = Nf((x,y-1),t-1) + Ef((x-1,y),t)$$

Some rules to note:

- When *t* is odd, either *x* or *y* is odd.
- When *t* is even, *x* and *y* have the same parity.

Finding Patterns:

Here is a list of probabilities for different values of *x* and *y*:

$f((1,0),1) = E$	$f((0,1),1) = N$	$f((1,1),2) = 2NE$	$f((2,0),2) = E^2$
$f((0,2),2) = N^2$	$f((1,2),3) = 3N^2E$	$f((2,1),3) = 3NE^2$	$f((3,0),3) = E^3$
$f((0,3),3) = N^3$	$f((1,3),4) = 4N^3E$	$f((2,2),4) = 6N^2E^2$	$f((3,1),4) = 4NE^3$
$f((4,0),4) = E^4$	$f((0,4),4) = N^4$	$f((1,4),5) = 5N^4E$	$f((2,3),5) = 10N^3E^2$
$f((3,2),5) = 10N^2E^3$	$f((4,1),5) = 5NE^4$	$f((5,0),5) = E^5$	$f((0,5),5) = N^5$

There are obvious patterns that are quite like the first section. These include patterns in the powers of *N* and *E* and the fact that the coefficients are from Pascal's triangle.
The powers of *N* and *E* are clearly just either *x* or *y*, which is expected since *x* represents movement to the East and *y* represents movement to the North.
On that basis, the following expression is derived:

$$N^y \times E^x$$

Formula for the value of the coefficients:

This will not be as easy as the powers for N and E, since whilst we know that the coefficients are from pascal's triangle, we need to find a rule that relates *x* and *y* to the diagonal column number. The row number is clearly *t.* For the diagonal column number, we can use the value of *x, y* or both in the following expression:

$$m = {}^{t}C_{(t-x+n)/2}$$

Putting it all together:

The expressions used before can be joined together to form a formula for the probability, denoted by ***f((x,y),t)***, for any value of *x, y* and *t.*

$$f\big((x,y),t\big) = N^y \times E^x \times {}^{t}C_{(t-x+n)/2}$$

Four Directions:

We decided to expand our problem to deal with four directions represented on the Cartesian plane. The teacher is now able to move North, South, East, and West. Previous assumptions apply (exactly one step each second, etc.). Once again, his position will be represented as a coordinate, (*x,y*),
where **$x \in R$ and $y \in R$.**

f((x,y),t) will be the probability that the teacher is at position (x,y) at time *t.* The teacher covers all 4 Quadrants on the Cartesian Plane and the probability that he walks North is *N,* South is *S,* East is *E,* and West is *W,* where N + S + E + W = 1

Once again, the condition at *t = 0* is:

- *f((0,0),0) = 1, f((x,y),0) = 0* { x ≠ 0, y ≠ 0 }

A recursive relation is shown below:

- *f((x,y),t) = Wf((x+1,y),t-1) + Ef((x-1,y),t-1) + Nf((x,y-1),t-1) + Sf((x,y+1),t-1)*

Some points to note:

- *If t is even, x and y must have the same parity.*
- *If t is odd, x and y must have different parities.*
- *|x| + |y| ≤ t*

Finding Patterns:

Below is a list of values for the probabilities for different values of *x, y* and *t*:

$f((1,0),1) = E$	$f((0,1),1) = N$	$f((1,1),2) = 2NE$	$f((2,0),2) = E^2$
$f((0,2),2) = N^2$	$f((1,2),3) = 3N^2E$	$f((2,1),3) = 3NE^2$	$f((3,0),3) = E^3$
$f((0,3),3) = N^3$	$f((1,3),4) = 4N^3E$	$f((2,2),4) = 6N^2E^2$	$f((3,1),4) = 4NE^3$
$f((4,0),4) = E^4$	$f((0,4),4) = N^4$	$f((-1,0),1) = W$	$f((0,-1),1) = S$
$f((-1,-1),2) = 2SW$	$f((-2,0),2) = W^2$	$f((0,-2),2) = S^2$	$f((-1,-2),3) = 3S^2W$
$f((-2,-1),3) = 3SW^2$	$f((-3,0),3) = W^3$	$f((0,-3),3) = S^3$	$f((1,-1),2) = 2SE$
$f((-1,1),2) = 2NW$	$f((-1,2),3) = 3N^2W$	$f((1,-2),3) = 3S^2E$	$f((-2,1),3) = 3NW^2$
$f((2,-1),3) = 3SE^2$	$f((-1,3),4) = 4N^3W$	$f((1,-3),4) = 4S^3E$	$f((-2,2),4) = 6N^2W^2$
$f((2,-2),4) = 6S^2E^2$	$f((-3,1),4) = 4NW^3$	$f((3,-1),4) = 4SE^3$	$f((-4,0),4) = W^4$

Once again, the coefficients resemble the numbers that makeup Pascal's Triangle, and the powers of the constants *N, S, E* and *W* seem to be linked to the values of '*x*' and '*y*'.
When *x>0, W* is raised to the power of 0 and E is raised to the power of '*x*'.
When *x<0, W* is raised to the power of |*x*| and *E* is raised to the power of 0. The same applies for '*y*', *N* and *S*.
This can be expressed in the following formula:

$$\boldsymbol{N}^{\frac{(y+|y|)}{2}} \times \boldsymbol{E}^{\frac{(x+|x|)}{2}} \times \boldsymbol{S}^{\frac{|(y-|y|)|}{2}} \times \boldsymbol{W}^{\frac{|(x-|x|)|}{2}}$$

When you input *y* where *y < 0,* you will get 0 for the power of *N* and y for the power of *S* and vice versa for when *y > 0.* The same applies for *x, E* and *W*.

Formula for the value of the coefficients:

|x| and *|y|* are the diagonal column in pascal's triangle from which we can get the coefficients. Thus, our formula for the value of the coefficient is:

$$m = {}^{t}C_{|x|} \qquad or \qquad m = {}^{t}C_{|y|}$$

Where ‘m’ is the value of the coefficient.

Putting it all together:

Now we just combine the two expressions we develop in 4.1 and 4.2 to create our formula for the probability given any real value of *x, y* and t where $t \geq 0$.

$$f((x, y), t) = N^{\frac{(y+|y|)}{2}} \times E^{\frac{(x+|x|)}{2}} \times S^{\frac{|(y-|y|)|}{2}} \times W^{\frac{|(x-|x|)|}{2}} \times {}^{t}C_{|x|}$$

Real Life Uses of the Formula:

In this section, we will discuss uses of the formulae derived above. We will discuss some useful and sensible situations and fields that this can be useful in.

As stated above, we assumed that the math teacher will take one step in a random direction for every second that has passed. We can play with this value in certain ways such that we would be able to change the rate that he steps at or the distance that he travels at. This means that we can get values such as him moving one kilometer every hour. This can be used for anything, not just a teacher. So, this formula can be used to calculate the probability of the location of any object that is moving. Since the above formula was only used to give two directions, then expanded so that it applies for four directions. This is how we believe that our formula can be used to predict the movements of real living and moving objects.

Military:

A real-life use where this formula can be used is in military warfare. It can be used to predict where enemy soldiers and troops will move. We can change the speed such that it will match the speed of any available army. Since when finding this formula, we assumed that the man is moving on a Cartesian plane with four quadrants. Thus, we can label the axis in whatever directions we would like such as the positive y-axis be north. If an enemy group is relocating their troops elsewhere and there is no assured way of tracking them, then the above formula can be used to predict where they might be or at the very least shorten the range. As most troops would not move at night you can assume that they are only moving during the day and the speed can be predicted using common sense. If they are moving on foot, then assume they are moving at a speed of 5km/h.

There are obviously more variables considering this situation. Unlike the man where we assumed that he will always be moving and that no outside force is going to hinder his movements, this cannot be the case in an army situation. These variables can be weather, terrain and travel restrictions. If the weather is unfavorable to the moving army, we can assume that they might move slower or might not decide to move at all. This would make predicting where they will be even more difficult. But it can still give a more suitable range to where they might be.

Terrain, same as weather will hinder the group and slow them down. Terrain can also help the trackers limit the sample space of where they might be. If there are mountains in a certain direction and area, we can assume that they would not have nor will travel through there. Thus, making it easier to predict where they will be. This formula can be useful if the enemy is moving on ground, if they are moving via air for example using planes it would be more difficult to accurately predict where they. This is because an extra two directions are added to the four previously mentioned directions. These directions would be forwards, backwards, left, right, up or down.

Animal Control and Tracking:

In recent news, a herd of elephants in China has left their traditional home range habitats due to the habitat loss in that area. Since setting off on their journey on the 18th of April, they have journeyed more than 500km from their original home. They are still travelling now at the time of writing this. Although a wonderful job has been done in tracking and protecting these elephants, this was only possible due to China's powerful economy. If this were a weaker 3rd world developing country, it would not have been the case and a worse job would have been performed due to their low budget and economy. This is where the formula above can come in handy. It is a cheap method that requires a person to only plug-in the numbers. Although it would not be accurate it could give a rough estimate.

While studying the route that the elephants have taken, scientists found that the elephants have not gone back to a place they have already travelled to. This clears off a large amount of the sample space. We know places such as mountains, elephants will not travel due to them not being able to live in mountain areas which there are many of China. This rules out another large number of areas that they could travel to. This reduces the range significantly and makes the predictions more accurate in this cheap and easy to use method.

This method cannot be used to only track these elephants but also any other animal that is moving randomly. By using some common sense, we can limit where they will move and help us predict what they will do. This formula can be plugged into machines that can do the math for you and add more to it such that the results will be more reliable.

Improvements:

Though significant improvements to the formula were made during the process of its construction and development, there are still improvements that are possible.

Some improvements that could've been done are:

- We could have used a larger sample size like using larger numbers to see if the pattern still applies when creating the formula. When solving for $f(n, t)$ the sample space for n is $-5 \leq n \leq 5$. Then we tried to prove that $N+E=1$ by using a sample space up to 5.
- We could have thought of other methods and explore them since we looked at Pascals triangle and after seeing it was related, we researched on this and continued using it without thinking of other factors that could be used to form this formula.

References:

- Online Python Complier URL: https://www.tutorialspoint.com/execute_python_online.php
- CNN, Julia Hollingsworth and Zixu Wang. 'Millions of People in China Can't Stop Watching a Pack of Roving Elephants'. *CNN*, https://www.cnn.com/2021/06/09/china/elephants-china-yunnan-intl-hnk/index.html. Accessed 28 June 2021.