Special Assignment in lieu of Semester End Practical Examination May-June 2020
B.Sc. (Hons.) Computer Science Semester-VI
Paper Code: 32341601
Paper Name: Artificial Intelligence

Epidemic Modelling

<u>Ans:</u>

<u>(a):</u> I have made the model from human perspective as to how the infection spreads. I have chosen to model this from human perspective because it can help us predict our present situation because of the COVID-19 Pandemic. Also, this model can help us understand our shortcomings in dealing with the Pandemic and can help us to be better prepared for future.

<u>(b.i & b.ii):</u> This model mainly includes how many people were initially infected and how the infection has spread. We find out that the spread of infection follows an exponential pattern and hence we can use an exponential model for calculation of rate of spread of infection.

Now let's consider the spread of infection in a city, let's say Delhi, with the population of 2 crore (2,00,00,000).

The city is divided into 9 regions based on the population density of those regions: -

15% in extreme congestion zone (ECZ)
30% in high congestion zone (HCZ)
35% in moderate congestion zone (MOCZ)
10% in mild congestion zone (MICZ)
7% in moderately spaced zone (MSZ)
2% in safe space zone (SSZ)
0.8% in ideal space zone (ISZ)
0.1% in sparse zone (SPZ)
0.1% in dispersed zone (DZ)

Zones have the following description: -

Zone	Average inter-person space (radial) in Dwelling Units (DU)
ECZ	<=0.5 meters
HCZ	More than 0.5 but < 1 meter
MOCZ	1 to 1.5 meters
MICZ	1.5 to 2 meters
MSZ	2 to 2.5 meters
SSZ	2.5 to 3 meters
ISZ	3 to 5 meters
SPZ	5 to 8 meters
DZ	>8 meters

Distributing the population according to the percentage given. So, the total number of people living in each region is:

ECZ:	$3x10^6$
HCZ:	$6x10^6$
MOCZ:	$7x10^6$

MICZ:	$2x10^6$
MSZ:	$1.4x10^6$
SSZ:	$4x10^5$
ISZ:	$1.6x10^5$
SPZ:	$2x10^4$
DZ:	$2x10^4$

Assuming a random distribution of initial carriers (i.e., carrier exists in a set of people with a very small probability, but uniformly – say there is one carrier per r persons (e.g., r = 1 million). Therefore, the number of infected people in each region given by (population of that region / 10,00,000) is:

DZ: 0.02 ≈ 0
SPZ: 0.02 ≈ 0
ISZ: 0.16 ≈ 0
SSZ: 0.4 ≈ 0
MSZ: 1.4 ≈ 1
MICZ: 2
MOCZ: 7
HCZ: 6
ECZ: 3

We have been given the conditions of transmission:
"The transmission of infection is possible with probability 0.7 if carrier is within 2-meter radius on another person for 5 seconds or more."

We have also been given the following three levels of daily Mobility of Population (MOP):
1. Static (S) – everyone stays at home
2. Dynamic (D) – normal movement for work etc.
3. Sensitive (SN) – only sick (possibly infected) stay at home/quarantine

Case 1: Static Mobility of Population (Best Case Scenario)
In this case, the infected people are quarantined at home and no one is outside. The pathogen can't spread to many people and actually dies out after a while being exposed to the outside environment. Since the probability of transmission of pathogen depends on two conditions:

1) if carrier is within 2-meter radius of another person
2) the duration is more than 5 seconds.

These four regions namely DZ, SPZ, ISZ and SSZ won't be affected by the pathogen since: -
1. The random distribution of initial carriers show there are no infected people in these regions.
2. The Average inter-person space (radial) for these four regions is more than 2.5 m and the city is already on complete lockdown which prevents people from meeting with each other.

As for the rest of regions, we have the count of initial carriers in each region. Given the conditions for the spread of the pathogen, and the fact that the *'Static Mobility Level'* ensures that everyone stays at home, we can say that the rate of transmission of pathogen is very slow.

Assumption: Let's assume that each initially infected person, infects 2 new people every day.

Justification: We can safely assume that this rate of spread of 2 new people every day is uniform throughout the city, even though the count of initial carriers in different zones is varied, because <u>our mobility level ensures that everyone is staying in their homes. And hence, even though the initial carrier count in different zones is varied, the rate of spread will be almost uniform.</u>

Now, calculating the total number of cases in each zone:

<u>Day 0:</u>

MSZ:	1
MICZ:	2
ECZ:	3
HCZ:	6
MOCZ:	7

Total cases on Day 0: 19

<u>Day 1:</u>

Now these 19 people will infect 2 people each. 19(2) = 38
Total cases on Day 1: 19+38=57

<u>Day 2:</u>

Now these 57 people will infect 2 people each. 57(2) = 114
Total cases on Day 2: 57+114=171

<u>Day 3:</u>

Now these 171 people will infect 2 people each. 171(2) = 342
Total cases on Day 3: 171+342=513

We get a series of total number of patients on each Day: 19, 57, 171, 513.....
On closer inspection we find that we have a geometric progression that is of the form:

$$a, ar^2, ar^3, ... , ar^{n-1}$$

where, a=19 and r=3

Using the formula to calculate the N^{th} term of a Geometric Progression, we can find on what day will a certain percentage of the population be under the infection, that indicates the rate of spread of infection.
The time needed for X% population to be infected can be calculated as:
(X / 100) x population affected

Note that since there are 4 zones that won't be affected by the pathogen, the Population affected is:

Population Affected = Total population – the population of zones not affected
= 2,00,00,000 – 6,00,000
= 1,84,00,000

Since the sequence is in GP with a = 19 and r = 3, n = days, can be calculated using the N^{th} term of GP

$a(r)^{n-1}$ = X% of population affected
$\Rightarrow 19(3)^{n-1} = (X/100) * 1,84,00,000$
$\Rightarrow (3)^{n-1} = (X/100) * (1,84,00,000/19)$
$\Rightarrow (3)^{n} = (X/100) * (1,84,00,000/19) * 3$

=> $\log_3(3^n) = \log_3(X * 29052.6315)$
=> $n = \log_3(X * 29052.6315)$

Since, we can't directly compute $\log_3$ prolog, we use $\log_{10}$ to obtain our answer using the formula:

$n = \log_{10}(X * 29052.6315) / \log_{10}(3)$

The value that we obtain is in decimals and we can round it off to the next number to give us an approx. number of days.

Case 2: Dynamic Mobility of Population (Worst Case Scenario)
In this case, the infected people are not in Quarantine. They are roaming freely and spreading the pathogen. Also, breaking the city into regions does not work here since there is no restriction on movement of people. Therefore, the whole city is considered as a single unit.

Number of people initially infected can be calculated as:
=> Total population / 10,00,000
=> 2,00,00,000 / 10,00,000
=> 20 people.

Assumption: Given the Mobility level, and the conditions of spread of the virus, we can assume that each person infects 20 people.

Justification: We can safely assume this value and consider it uniform throughout the city because <u>our mobility level ensures that infected people are not staying in their homes. And since we have already established that the city is considered a single unit, there is no scope of different rate of spread.</u>

Now, calculating the total number of cases each day:

<u>Day 0:</u>
Total cases on Day 0: 20

<u>Day 1:</u>
Now these 20 people will infect 20 people each. 20(20) = 400
Total cases on Day 1: 20+400=420

<u>Day 2:</u>
Now these 420 people will infect 20 people each. 420(20) = 8400
Total cases on Day 2: 420+8400=8820

<u>Day 3:</u>
Now these 8820 people will infect 20 people each. 8820(20) = 176400
Total cases on Day 3: 8820+176400=185220

We get a series of total number of patients on each Day: 20, 420, 8820, 185220 .....
On closer inspection we find that we have a geometric progression that is of the form:

$$a, ar^2, ar^3, \ldots\ldots, ar^{n-1}$$

where, a=20 and r=21

Using the formula to calculate the N^{th} term of a Geometric Progression, we can find on what day will a certain percentage of the population be under the infection, that indicates the rate of spread of infection.
The time needed for X% population to be infected can be calculated as:
(X / 100) x population affected

Since the sequence is in GP with a = 20 and r = 21, n = days, can be calculated using the N^{th} term of GP

$a(r)^{n-1}$ = X% of population affected
=> $20(21)^{n-1} = (X/100) * 2{,}00{,}00{,}000$
=> $(21)^{n-1} = (X/100) * (2{,}00{,}00{,}000/20)$
=> $(21)^{n} = (X/100) * (2{,}00{,}00{,}000/20) * 21$
=> $\log_{21}(21^n) = \log_{21}(X * 2{,}10{,}000)$
=> $\mathbf{n = \log_{21}(X * 2{,}10{,}000)}$

Since, we can't directly compute $\log_{21}$ prolog, we use $\log_{10}$ to obtain our answer using the formula:

$$\mathbf{n = \log_{10}(X * 2{,}10{,}000) / \log_{10}(21)}$$

The value that we obtain is in decimals and we can round it off to the next number to give us an approx. number of days.

Case 3: Sensitive Mobility of Population
In this case, only the possibly infected people are staying in. To explain this case, we have made a set of different assumptions than the previous two cases.

In this case, the city of Delhi having a population of 2 crores is considered to be divided in 4 regions that have an equal number of people living in that region. But, the Average inter-person space (radial) in Dwelling Units (DU) is different such that:

Zone	Average inter-person space (radial) in Dwelling Units (DU)
Zone 4	<=0.5 meters
Zone 3	More than 0.5 but < 1 meter
Zone 2	1 to 1.5 meters
Zone 1	1.5 to 2 meters

So, 2 crores/4= 50 lakh.
Each region is assumed to house 50 lakh people.
Considering r = 1 million (=10 lakh), initially infected people in all the regions equal 5.

Given that only possibly infected people are staying in and other people are freely going outside, and that there is difference in average inter-person space in the zones, we can assume that each zone has a different rate of spread of infection. The people can move from one zone to the other but considering that possibly infected people are staying in, we can consider the 4 zones as disjoint entities with their own rate of spread. The total number of people infected, however, is calculated by adding the number of infected people in that zone.

We know from the previous 2 cases that the spread follows a geometric progression, but since in this case we have different rate of spread, it is safe to assume that we will have 4 different geometric progressions, each of one zone.

According to the values assumed for the Average inter-person space and the initial count of infected people, we can assume that:-

For Zone 1, a=5 r=4
For Zone 2, a=5 r=6
For Zone 3, a=5 r=8
For Zone 4, a=5 r=10

Now to find out how long will it take for X% of the population to be infected, we can find the sum of the values obtained from the 4 G.P.s that gives us the total number of infected people in the city (since each individual G.P. gives us the number of infected people in that zone).

Note that we have 4 different G.P.s contributing to the total number of infected people, therefore, we can not use the method that has been used in the previous 2 cases.

Hence, the code for this Case is different from the other two cases and involves recursion to find out in how many days (rounded of value) will a certain percentage of the population be suffering from the disease caused by the pathogen.

We calculate the total population that has fallen ill because of the pathogen at each step of the recursion, the movement we get a value that is greater than that of the number required (according to the percentage), the count of days stop. **The last message printed tells us the final count of the number of days it takes for X% of the total population to be fallen ill.**

(b.iii): The purpose of this whole model is to understand our mistakes with dealing with the ongoing Pandemic COVID-19, so that we can refrain from making the same mistakes in future and can better prepare ourselves for a similar outbreak.

In my opinion, the situation can be made better by containment of infection and isolating the infected individuals as soon as they are identified. Special medical wards can be set up in remote places where infected people can stay during infection so that the infection does not spread from one person to another either directly or indirectly. And since we already know from this outbreak of COVID-19, the symptoms can take up to 20 days to be visible, therefore, everyone should take the disease seriously and practice best preventive measures.

Obviously, the rate of infection wouldn’t become 0 altogether but it will start to decrease and be withing a manageable range.

So, in this case the rate of spread becomes 0, that is no new people will be infected.
And as for infected people, they will stay under isolation until they are completely healthy.