



USING CALCULUS TO MODEL EPIDEMICS USING S.I.R MODEL

SUBMITTED TO
DR.VIVEK KUMAR



Made by
Aaryan Rawat
2K20/B17/06
Parth Johri
2K20/B17/33

INTRODUCTION

Most of the human pandemics are mostly spread by contact with contagious people, often there are vectors, such as mosquitoes, rats and fleas involved in disease transmission. There are many kinds of contagion diseases, such as smallpox, polio and measles, that are easily spread through general contact. Other diseases, such as gonorrhoea, require more close contact. One of the major differences between the first group of diseases and gonorrhoea is that they “confer immunity” to someone who recovers from it, whereas gonorrhoea does not. In short, once you recover from chickenpox, you cannot catch it again. This characteristic offers the likelihood of control through vaccination. In this section we will devise our mathematical model of the spread of this disease known as COVID-19 to which people are susceptible, then infectious, and finally, recovered and immune

A scientific model is axiomatically a oversimplification. It is not possible to include contacts between individuals or even other ailments of the population, the amount of travel, climate and other factors that are responsible for small effects on the spread of the disease, because we would fail to analyze, compute and predict such a complicated model. However, by spotting the major effects, we can formulate a model that is manageable enough to use to compute and predict the overall spread of an illness.

The practicality of a mathematical model requires us to make accurate and precise scientific judgments about which effects are insignificant and which are not. We need to access our models against known data both to measure significant parameters and to recheck that we have neglected only minor factors and try to train our model with a larger data set to make the prediction more accurate. Once this is done, we can use our mathematical model to make predictions in new cases.



ASSUMPTIONS OF THE SIR MODEL

Since we already stated that since this is a scientific model, we cannot account of all the factors affecting our calculations. Thus, we have to ignore some insignificant factors.

The basic assumptions are as follows:

1) The entire sample size fits into one of the following categories:

- a.) S : Susceptible - those who can potentially catch the disease.
- b) I : Infectious - those who already have the disease and are potential carriers.
- c) R : Removed - those who are have recovered from the disease and are now immune to it.

This also include the people who have passed away due to the particular disease.

2) The samples size of people is fixed. The sample of people under study are well defined. A major part of this assumption is that the population growth rate/death rate is negligible; that is during the period of study, the population of the sample under consideration doesn't change by much.

3)The population is diverse. In theory, everyone will come in contact with the same portion of people in each category every single day. Again, it is not possible to keep a track of who a person meets every day especially when we have a large sample size under consideration.

4) Infected people recover or die at a constant rate.

PARAMETERS AND EQUATIONS

Let the constant population/sample size be 'N' By the assumptions, we know that the entire population can be divided in three parts.

Let the total number of susceptible individuals at any given time be given by the function:

$$s = S(t)$$

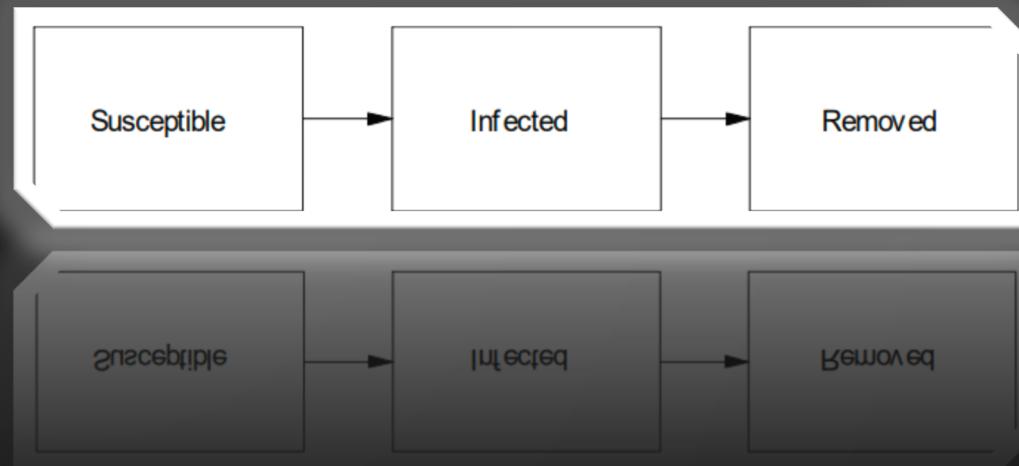
Let the total number of infected individuals at any given time be given by the function:

$$i = I(t)$$

Let the total number of recovered individuals at any given time be given by the function:

$$r = R(t)$$

This gives us: $(s + i + r) = N$ [at any time 't']



Now we will add a mathematical description of how individuals move between the compartments of ‘susceptible’, ‘infected’ and ‘recovered’ and the equations governing this movement.

- The number of people who move from susceptible to infected is directly proportional to the contact between the susceptible and the infected.
- The number of people who enter the infected compartment is equal to the number of people who leave the susceptible compartment minus the number of people who recover/die.
- The number of people who enter the recovered compartment is the number of people who get recovered or die

$$\frac{ds}{dt} = - \alpha * s * i$$

$$\frac{di}{dt} = \alpha * s * i - \beta * i$$

$$\frac{dr}{dt} = \beta * i$$

Here, ‘ α ’ represents rate of contact or rate of transmission and ‘ β ’, represents the recovery rate.

By our assumptions, both the parameters ‘ α ’ and ‘ β ’ are considered constants that is a constant transmission rate and a constant recovery rate.

At time $t = 0$, the entire sample size consists of many susceptible and a few infected. So, let us take

$$s(t=0) = 'S_0' \text{ initially}$$

$$i(t=0) = 'I_0' \text{ initially}$$

$$r(t=0) = '0' \text{ initially}$$

this gives us

$$s + i + r = S_0 + I_0 = N$$

WILL THE DISEASE SPREAD?

The equations we have till now are:

$$\frac{ds}{dt} = -\alpha * s * i \quad (1)$$

$$\frac{di}{dt} = \alpha * s * i - \beta * i \quad (2)$$

$$\frac{dr}{dt} = \beta * i \quad (3)$$

$$s + i + r = S_0 + I_0 = N \quad (4)$$

the condition of the disease spreading or not will depend on equation $\frac{di}{dt}$ (2)

(rate of change of infected people with time)

In the beginning of an outbreak like covid-19, almost everyone in the sample size lie in the ‘susceptible’ category. So, the term ‘s’ will always be decreasing in nature. So, we can say at any time ‘t’,

$$s \leq S_0 \quad (5)$$

We can use equation (5) and equation (2) to get
 $\frac{di}{dt} \leq \alpha * S_0 * i - \beta * i$

Here the term ‘ $\alpha * S_0 - \beta$ ’ is a constant and the nature of this constant {that is, is it positive or negative} will determine if ‘ $\frac{di}{dt}$ ’ will be positive or negative will then determine if number of infected people will increase or decrease.

$$\alpha * S_0 - \beta > 0$$

$$S_0 > \alpha / \beta \quad (6)$$

If equation (6) is valid, then the disease will spread.

In equation (6), the term ‘ α / β ’ is given by a parameter ‘ $1/q$ ’ $q = \beta / \alpha$ (7)

here, the constant ‘ q ’ is known as **contact ration**. This is the fraction of population that comes in contact with an infected subject during the period of consideration.

We have another parameter we can use to derive the condition of spread of a disease.

This is known as the “basic reproductive number”

The parameter ‘**Ro**’ is a much more standard term for measuring the “intensity of spread of disease”.

“If “ $R_0 > 1$ ” it means the disease is spreading and vice versa.

More the values of ‘Ro’, faster is the rate of spread of disease.

The basic reproduction number (R_0) for India was estimated at 1.379" ([1], 2021)

“In India, the ‘Ro’ at the first week from March 2–8, 2020 was 3.2. It remained around 2 units for three weeks, from March 9–29, 2020. After March 2020, it started declining and reached around 1.3 in the following week suggesting a stabilisation of the transmissibility rate.”[1] ((DOAJ, 2021)

“This was 1.450 for Maharashtra, 1.444 for Gujarat, 1.297 for Delhi and 1.405 for Tamil Nadu.

The study estimated a baseline R_0 of 1.379 for India. It also showed that the R was getting stabilised from first week of April (with an average R of 1.29), despite the increase in March. This decline in the value of ‘Ro’ indicated that in due course of time, there will be a reversal of epidemic”[1] ((DOAJ, 2021)

MAXIMUM NUMBER OF INFECTIVES AT ONE TIME

In order to find maximum number of subjects infected at one point of time during the epidemic, the following method is followed

We use the equations

$$\frac{ds}{dt} = -\alpha i s$$

$$\frac{di}{dt} = \alpha i s - \beta i$$

Divide the two given equations

$$\frac{di}{ds} = \frac{\alpha i s - \beta i}{-\alpha i s} = -1 + \frac{\alpha}{\beta s}$$

As we know that $q = \frac{\beta}{\alpha}$

$$\frac{di}{ds} = -1 + \frac{1}{qs}$$

Multiply both sides by ds and integrate

$$\int_{I_0}^i di = \int_{S_0}^s \left[-1 + \frac{1}{qs} \right] ds$$

$$i - I_0 = S_0 - s + \frac{1}{q} \ln \left| \frac{s}{S_0} \right|$$

$$i + s - \frac{1}{q} \ln |s| = I_0 + S_0 - \frac{1}{q} \ln |S_0|$$

In order to obtain i_{\max} , we differentiate ' i ' with respect to ' s ' and equate the acquired equation to zero.

But, we already have ' $\frac{di}{ds}$ '

$$\frac{di}{ds} = -1 + \frac{1}{qs}$$

By observation, when $s = \frac{1}{q}$, $\frac{di}{ds} = 0$

So, we put $s = \frac{1}{q}$ in the equation

$$i + s - \frac{1}{q} \ln |s| = I_0 + S_0 - \frac{1}{q} \ln |S_0|$$

Putting $s = \frac{1}{q}$ we get

$$i_{\max} = I_0 + S_0 - \frac{1}{q} (1 + \ln |q S_0|)$$

An interesting thing in this equation is that in the RHS,
 $q S_0 = R_0$ (basic reproduction number)

We derived the equation for i_{\max} as follows:

$$i_{\max} = I_0 + S_0 - \frac{1}{q} (1 + \ln(S_0))$$

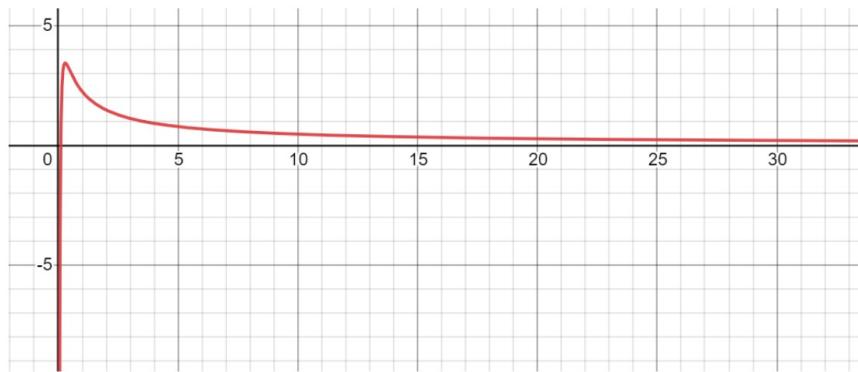
Since ' I_0 ' and ' S_0 ' were constants defined at time ($t=0$), we cannot change them. But, we can see the behaviour of the term $\left(\frac{1}{q}(1 + \ln(S_0))\right)$ at different values of ' q ' which we can control.

So, let us plot $y = \frac{(1 + \ln(qS_0))}{q}$

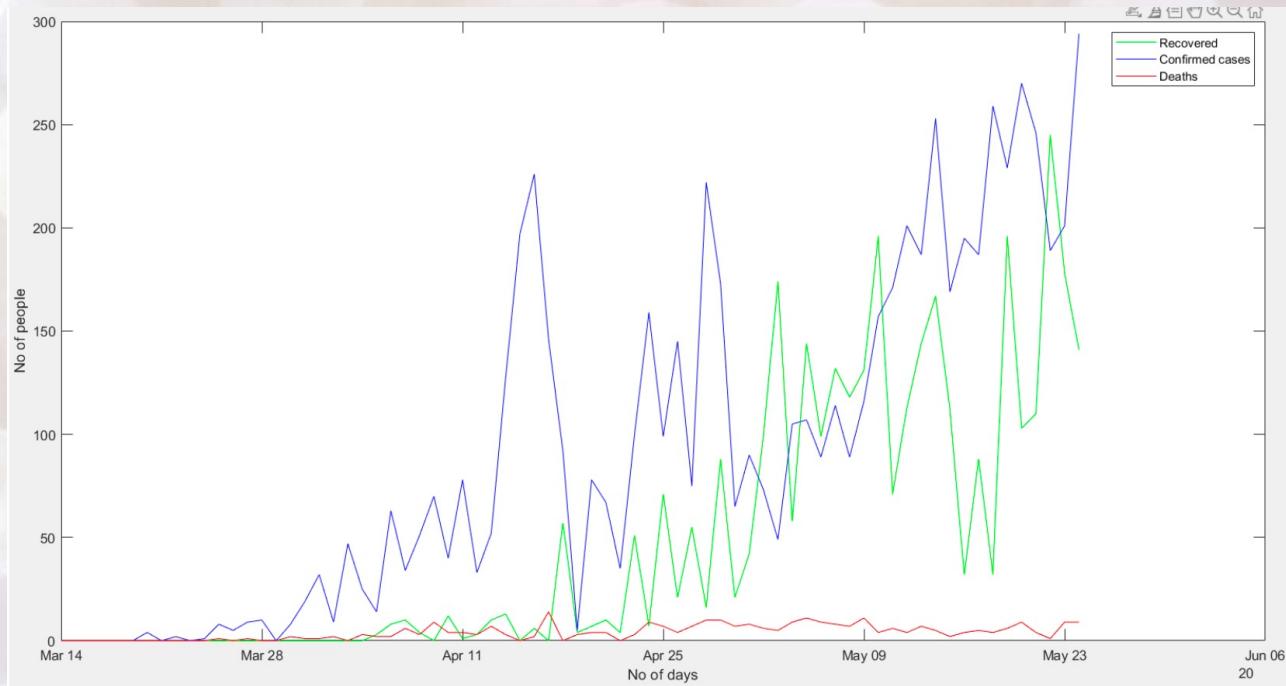
where $y = \frac{1}{q} (1 + \ln(qS_0))$

$$x = q$$

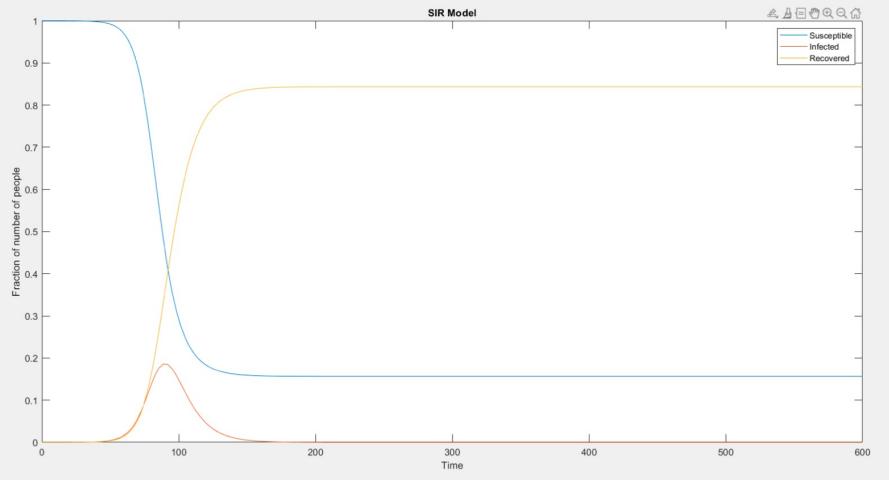
$$a = S_0$$



In the context of **Covid-19**, the contact ratio i.e., ‘q’ is very high, especially with the short incubation period of 4-5 days. This means that the part of curve we are at is somewhere in the right of the curve. This is a bad indicator since we are actually subtracting off ‘y’ from **‘Io+So’**. This represents almost the entire population who can be infected at some point of time.



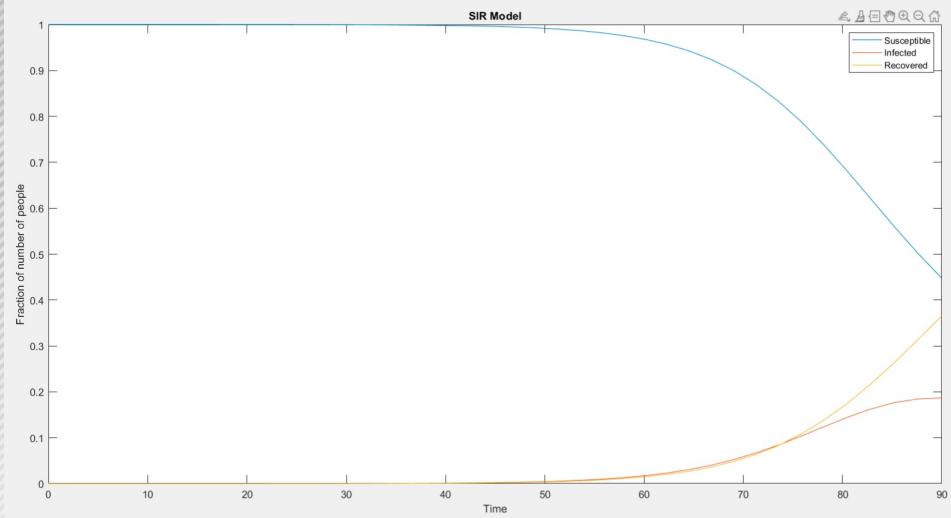
Graph plotted using actual data for 3 months (90 Days)



Calculated graph for coming 600 Days



Calculated graph for coming 90 Days



CODE

```
1 - close all
2 -
3 - %Reading the data from .csv file
4 - time = readtable('state_wise_daily.csv','range','A2:A217');
5 - time=table2array(time);
6 - time=(time(1:3:end,1));
7 - TN=readtable('state_wise_daily.csv','range','Y2:Y217');
8 - TN=table2array(TN);
9 -
10 - %Initializing the variables
11 - N=6.79e7;
12 - I=(TN(1:3:end,1)');
13 - R=(TN(2:3:end,1)');
14 - D=(TN(3:3:end,1)');
15 - S=N-I-D;
16 - 
17 - s=S/N;
18 - i=I/N;
19 - r=R/N;
20 - 
21 - 
22 - %Plotting the current data
23 - figure(1);
24 - plot(time,R,'g',time,I,'b',time,D,'r');
25 - xlabel('No of days');
26 - ylabel('No of people');
27 - legend('Recovered','Confirmed cases','Deaths');
28 - 
29 - %Finding b and k
30 - b=-1*([0 diff(s)]./(s.*i));
31 - b(isnan(b)) = [];
32 - b(isinf(b)) = [];
33 - b=abs(mean(b))
34 - 
35 - k=([0 diff(r)]./i);
36 - k(isnan(k)) = [];
37 - k(isinf(k)) = [];
38 - k=abs(mean(k))
39 - 
40 - 
41 - 
42 - %Solving the equations
43 - to = 0;
44 - tf =600;
45 - yo = [s(end) i(end) r(end)];
46 - [t y] = ode45('ypsisr',[to tf],yo);
47 - 
48 - 
49 - %Plotting the prediction
50 - figure(2)
51 - plot(t,y(:,1),t,y(:,2),t,y(:,3))
52 - title('SIR Model');
53 - xlabel('Time')
54 - ylabel('Fraction of number of people')
55 - legend('Susceptible', 'Infected', 'Recovered')
```

```
21 - figure(1);
22 - plot(time,R,'g',time,I,'b',time,D,'r');
23 - xlabel('No of days');
24 - ylabel('No of people');
25 - legend('Recovered','Confirmed cases','Deaths');
26 - 
27 - %Finding b and k
28 - b=-1*([0 diff(s)]./(s.*i));
29 - b(isnan(b)) = [];
30 - b(isinf(b)) = [];
31 - b=abs(mean(b))
32 - 
33 - k=([0 diff(r)]./i);
34 - k(isnan(k)) = [];
35 - k(isinf(k)) = [];
36 - k=abs(mean(k))
37 - 
38 - 
39 - 
40 - 
41 - 
42 - %Solving the equations
43 - to = 0;
44 - tf =600;
45 - yo = [s(end) i(end) r(end)];
46 - [t y] = ode45('ypsisr',[to tf],yo);
47 - 
48 - 
49 - %Plotting the prediction
50 - figure(2)
51 - plot(t,y(:,1),t,y(:,2),t,y(:,3))
52 - title('SIR Model');
53 - xlabel('Time')
54 - ylabel('Fraction of number of people')
55 - legend('Susceptible', 'Infected', 'Recovered')
```

COMPARISON



As we can see in the graph plotted using actual data, there are a lot of spikes, these are due to the variable rate of '**a**' and '**b**'.

Where '**a**' is the rate of contact between a susceptible and infected and '**b**' is the rate of change of population from infected to recovered.

But in our calculated graph there is a smooth curve, which is due to our assumption of this model as '**a**' and '**b**' being constants.

But as a similarity we can see the variation of infected and recovered, both infected and recovered are increasing. Also, if we talk about some parts of infected in our calculated curve, we can see the maximum value of infected people and observing that in future the infected persons are decreasing and recovered are increasing. This means the pandemic is coming to an end!!!

