Rounak Seksaria
20231130114

Introduction to Biology

Assignment - 1 : The SIR Model.

Some diseases have the potential to affect large segments of a population; they are called epidemics (from the Greek words 'epi' (upon) + demos (the people)). Epidemiology is the scientific study of these diseases.

An epidemic is a complicated phenomenon, that affects all strata of society drastically, as we have seen time & again during the various waves of the COVID-19 (or SARS CoV-2) virus that devastated & 'shook-stopped' (in some sense) the entire world. The dangers posed by the contagion compel scientists in general, & mathematical biologists in particular, to learn as much as we can about the nature of epidemics.

Using mathematics, firstly we can try to draw out of the situation its essential features & describe them mathematically. This is calculus as 'language'. We substitute an "ideal" mathematical world for the real one. This mathematical world is called a model. Second, we can use mathematical insights & methods to analyze the model. This is calculus as 'tool'. Any conclusions we reach about the model can then be interpreted to tell us various things about our physical reality.

One such example of a mathematical ~~world~~ model applied in the extremely important study of epidemiology is ~~of~~ the 'SIR' model, ~~etc~~ which stands for susceptible, infected, recovered, of an epidemics. Its basic purpose is to help us understand the way a contagious disease spreads through a population so as to enable us to predict what fraction of the populations falls ill, & when.

A major assumption of many mathematical models of epidemics is that the population can be divided into a set of distinct compartments. These compartments are defined wrt to disease status. This is exactly what happens in the SIR model, which ~~was~~ is one of the simplest models

in mathematical biology, first described ~~as~~ by Kermack & McKendrick in 1927.

Susceptible(S): Individuals that are susceptible have, in the case of the basic SIR model, never been infected, & they are able to catch the disease. Once they have it, they move into the infected compartment.

Infected(I): Infected individuals can spread the disease to susceptible individuals. The time they spend in the infected compartment is the infectious period, after which they enter the recovered compartment.

Recovered(R): Individuals in the recovered compartment, are, in the SIR model, are assumed to be immune for life.

The model also assumes that all people who have been infected go to recovered section, which means there are no deaths. In other words, we can also say that 'recovered' is a combination of patients who have recovered from the disease & those that are dead.

We also assume that the system is closed, ie, we neglect natural births & deaths. The total size of the population, ie the sum S+I+R remain constant throughout.

The SIR model is easily written using ordinary differential equations (ODEs), which implies a deterministic model (no randomness is involved, the same starting conditions give the same output), with continuous time (as opposed to discrete time). Analogous to the principles of reaction kinetics, we assume that encounters b/w infected & susceptible individuals occur at a rate proportional to their respective

nos. in the population. ~~The rate of new~~

Before we proceed further, notation:
~~I~~ I : # of infected invidivals
S : # of susceptible individuals
R : # of recovered individuals
S' : dS/dt
I' : dI/dt
R' : dR/dt

Let's begin ~~with~~ by addressing R'.
Suppose infection lasts for k days [In the SIR model, we also assume that everyone ~~is~~ stays infected for exactly the same amount of time].
We also assume that in ~~an~~ the absence of any definite information to the contrary, that the ~~infection~~ infected population is "uniform with respect to the duration of infection", at any given point in time. This ~~means~~ means that, there are, at any instant, just as many people who have been infected for 1 day, as there are for those ~~who~~ who have been infected for 2, and so on ~~to~~, up to k days.

Then on any day, 1/k th of the infected population will recover. ~~In other words,~~ Mathematically,

$$R' = \frac{dR}{dt} = bI \text{ , where } b = \frac{1}{k}$$

Here b is a constant scalar value that doesn't change over the course of time. The value of b varies ~~for~~ for each disease because of the total amount of days for infection [infection period] varying among different diseases.
Thus, $R' \propto I$.

For S', we suppose that :

(i) Each susceptible individual comes into contact with a proportion p of the infected population each day. This implies that each susceptible person has contact with pI infected persons per day. This means that pSI is the total contact of susceptible & infected persons every day.

(ii) q is the proportion of ~~infected per~~ the contacts above that actually cause infections.

Here again, first as before, we assume, unless there is dramatic evidence against, that q & p are constant.

Thus no. of new infections per day : $q \times (pSI)$.
That means S decreases by $qpSI$ every day.
Mathematically,

$$\frac{dS}{dt} = S' = -aSI \quad \text{where } a = qp$$

Here a can easily be converted from being the product of proportions on a 'daily' basis to a different unit depending on the kind of data that we have.
a here is a constant.
The minus sign denotes a negative rate of ~~the~~ change.

$\because$ we know $S + I + R = k$, k is a constant $\therefore$

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \quad \text{[differentiating]}$$

$$\Rightarrow S' + I' + R' = 0$$

$$\Rightarrow I' = aSI - bI$$

We can also understand this by the logic that aSI new people get infected every day while bI ~~get~~ people decrease from I [as they are recovered], which intuitively matches with our formula.

To summarize,

$S' = -aSI$

$I' = aSI - bI$

$R' = bI$

We also note that:

① a is a true parameter called the transmission coefficient, with unit 1/(person-day)

② b is a true parameter called the recovery coefficient, with unit 1/day.

③ The units of S, I, & R are persons, time is day. S', I', & R' are persons per day.

④ S is a ~~decreasing~~ non-increasing quantity, it is conceivable that S or I are zero at certain points, in which case, $S' = 0$. Thus $S' \geq 0$

⑤ Similarly, $R' = 0$ for $I = 0$
& $R' > 0 \ \forall \ I \neq 0$, ∵ I is a no. & thus non-ve.
∴ $R' \geq 0$

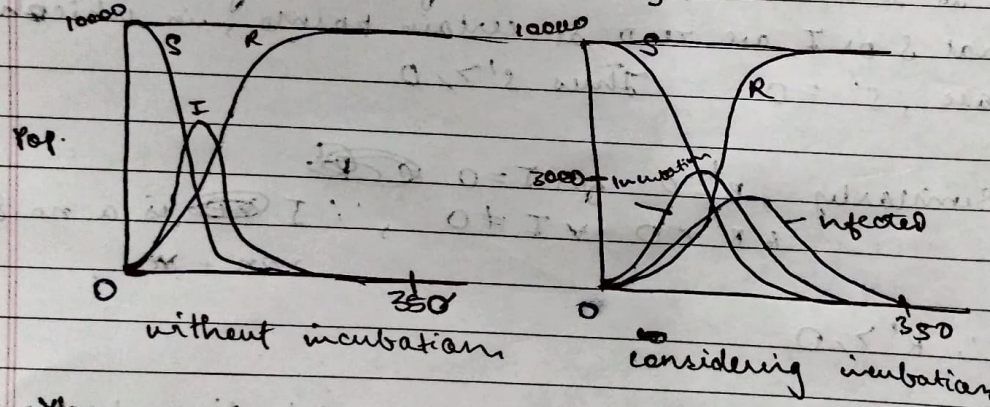
⑥ I' ~~is~~ $\in \mathbb{R}$ with no further specification, because it depends on relative sizes of aSI & bI.

During the last century, the SIR equations have been modified & extended to models a diverse range of epidemics including: Ebola, cholera, H1N1, tuberculosis, HIV/AIDS, influenza, malaria, Dengue fever, Zika virus, & most recently SARS-COV-2. In many of these models, additional parameters & terms are added to account for pathogen specific characters of transmission. Additional subcompartments may also be added to model different subpopulations (one such model is the SE/R model).

However, incorporating too many features can have subtle but important drawbacks including limited or unreliable inference of model parameters early in an epidemic. Nevertheless, SIR models do give us a very informative idea and basis for further predictions.

For SARS CoV-2 specifically, the large incubation period is a factor that needs to be taken into account for a more realistic prediction. The datasets received is extremely noisy, & in countries like India where people are reluctant to do tests until it is an extremely serious situation, & in places where healthcare isn't developed, data gathering is difficult. After the advent of self-test kits, the data gathering for no. of cases suffered another blow.

To display this, for SARS-CoV-2,



Thus, as far as I personally understand the data & the model, much like the prey predator model, it is a simplistic model that can be modified in order to better suit each disease based on its specifics. It acts as a great basis model that can help us really understand important insights about an epidemic, if modified sufficiently for that particular disease.