Epidemics - Introduction and SIR Model

Udit Sajjanhar

03CS3011 {sajjanharudit@gmail.com}

I. Introduction

An epidemic by definition is generally a widespread disease that affects many individuals in a population. An Epidemic study on networks address the challenge to study how information flows on the network. Examples of epidemics include:

1. Computer Worms and viruses
2. Information flow in social networks, say news or rumors
3. Study of Asymptotic behavior of the population by Theoretical Biologists.

II. Models to Study Epidemics

We begin with a basic idea:

1. Group individuals into categories or States.
2. Calculate the rates at which various events occur (infection, recovery, death, etc.)
3. Move individuals between categories/change the population sizes.
4. Write down equations based on these rates. Use the equations to describe the system outcomes (analytically, statistically, or computationally).

All models have some details that are wrong but they help us to sort out the logical consequences of ideas and the quantitative consequences of our assumptions.

Box models

When we sit down to try to construct a model for a particular system, we first make a "box model" socalled because it divides the host population up into boxes according to epidemiological status. The most important categories are:

1. Susceptible: hosts that are susceptible to infection
2. Infective: hosts that are infected, and capable of infecting other hosts
3. Recovered (or Removed): hosts that have acquired immunity to the disease, or been quarantined, or vaccinated, or otherwise been made unavailable for infection
4. Exposed (or Latent): hosts that have been infected but are incubating the disease, not yet Infective.

We will refer to these as the State Variables of our system. The list of categories or states we have in a particular model depends on the biology of a particular disease: - Is there an incubation period? - Do any hosts recover? - If so do they acquire immunity or do they become susceptible again?

We refer to these models by their acronyms

1. SI (also known as the "simple epidemic") is an epidemic where hosts never recover,
2. SIS one where they recover but become susceptible again (e.g. Gonorrhea)

3. SIR where they recover and gain immunity (e.g. Measles),
4. SEIR where there is an incubation period, etc...

III. The SIR Model

We are interested in the spread of an infectious disease where individuals may be susceptible to the disease, may be currently infected with the disease, or may be recovered and immune from the disease. Thus we have three groups or states in which we can place individuals. In addition we see that our data is a time series where we have a number of infected individuals at each point in time. Similarly we also have a number of susceptible and recovered individuals at each point in time. Let us begin with some notation.

St = the number of susceptible individuals in the population at time t. It = the number of infected individuals in the population at time t. Rt = the number of recovered individuals in the population at time t. N = the population size.

Correspondingly define the three groups as fractions of the total population N in lower case.

st = St /N (the susceptible fraction of the population at time t.) it = It /N (the infected fraction of the population at time t.) rt = Rt /N (the recovered fraction of the population at time t.)

We will work with both of these sets of notation in the modelling process. Note that each individual in the population is in one of the three groups. Thus St + It + Rt = N and st + it + rt = 1.

A. Dynamics

If we think about the process of a disease that fits the SIR framework we have a flow of individuals from the susceptible group to the infected group and then to the recovered or removed group.

Susceptible -¿ Infected -¿ Recovered

An individual potentially moves from the susceptible to the infected group when she comes in contact with an infected person. What qualifies as a contact depends on the disease. For HIV a contact may be sexual contact or a blood transfusion. For influenza it may be walking within a few feet of an infected person that has recently coughed. Suppose that each infected person contacts γ individuals in each period of time on average. Now each contact may not result in transmission of the disease. Perhaps only α percent of the contacts result in transmission. Thus the potential number of transmissions may be at most $\gamma * \alpha$. Let us define this value as $\beta = \gamma * \alpha$. β is the average number of transmissions possible from a given infected person in each period.

Now, we must remember that there are three groups in the population. If we assume that individuals are mixed randomly then each potential transmission may be from an infected person to a susceptible person which results in a new infected person. Or a transmission may occur from an infected person to another infected person which results in nothing happening since the person is already infected. Or the potential transmission may occur from an infected person to a recovered or immune person. In this case again nothing changes. Since only s_t percent of the population is susceptible each infected person generates only βs_t new infections each period. Each infected person recovers (or is removed/ dies) at some rate. Let the fraction of the infected group that recovers be κ.

We are now ready to describe the SIR process.Given the current state of the population in period t described by S_t , I_t and R_t we can write a series of difference equations that describe the motion of the system. First let's describe the susceptible population. We begin period t with S_t individuals in the susceptible population. From this population we lose on average $\beta s_t I_t$ from the population. Thus in period $t + 1$ we have:

S_{t+1} = S_t - $\beta s_t I_t$ (1)

Through similar reasoning we see that:

R_{t+1} = R_t + κI_t (2)

and

I_{t+1} = I_t + $\beta s_t I_t$ - κI_t = $I_t(1$ + βs_t -κ) (3)

Similarly we could write each of these in terms of the population fractions:

s_{t+1} = s_t - $\beta s_t i_t$ (4)

Through similar reasoning we see that:

r_{t+1} = r_t + κi_t (5)

and

i_{t+1} = $i_t(1$ + βs_t - $\kappa)$ (6)

If we add up these equations we will find that s_{t+1} + i_{t+1} + r_{t+1} = s_t + i_t + r_t = 1. To solve the above equations we can use Euler's Methods starting with S_0=1, R_0=0 and I_0=0.0000001. However there is a problem here. We have assumed that there is a minor fraction of the population that was initially infected with the infection and thus we are not capturing the role of an external agent which brought this disease into the community.

B. Steady State

If there is a constant population there do not exist positive steady state levels of infected individuals in the SIR model. Let us see why this is the case. First suppose that i_t is equal to 0. Then from Equation 6 you can see that I_{t+1} = 0. Thus if we start with no people infected we stay there. There has to be an initial infected person in order for there to be more infected people. Suppose that i_t ¿ 0. I ask: will I_{t+1} be greater than, less than, or equal to I_t ? The answer is either greater than or less than but never equal to (except in a special case I will mention in a moment.)

So, in most cases it will either be increasing or decreasing. How do we know which? We see from Equation 6 that i_{t+1} = $i_t(1$ + βs_t - $\kappa)$. Let ρ_t = 1 + s_t - κ. This is the epidemic threshold for the SIR model with a constant population. If ρ_t is greater than 1 then we are multiplying it by a number greater than 1 so i_{t+1} ¿ i_t . The number of infected individuals is increasing. But if t is less than 1 we are multiplying it by a number less than 1 so i_{t+1} ¡ i_t . The number of infected individuals is decreasing. Clearly any time ρ_t ¿ 1 or ρ_t ¡ 1 we are not at a steady state.

Now let us consider the possibility of $\rho_t = 1$ and show why this can not persist. If $\rho_t = 1$ then either i_t = 0 or ρ_{t+1} ¡ ρ_t. The first case is not interesting. This just says that there are no infected individuals so there is no possibility of anyone else being infected. The second case is more tricky. Let us again look at the definition of ρ_t , $\rho_t = 1 + \beta s_t$ - κ. We see that there are three constants in the equation: 1,β and κ, as well as one state variable s_t.Let us take another look at the equation for st given by Equation 4. We see that this equation is decreasing whenever i_t ¿ 0. Thus if i_t ¿ 0 ρ_t is decreasing for any β ¿ 0. Thus there can never be a steady state where $\rho_t = 1$ and i_t ¿ 0. Further if i_t is increasing it is doing so at a decreasing rate. And if i_t is decreasing it is doing so at an increasing rate. Combining these results indicates that the infected population will always disappear in the long run if we have a constant population.

Now let us take a brief look at s_t and r_t . First let us look at s_t . We know that s_t is decreasing for i_t ¿ 0. And we also know that i_t equals 0 in steady state. So, from this we can deduce that s_t will also reach a steady state value since it is constant if i_t is equal to 0 (see Equation 4 to convince yourself.) Now, r_t is increasing for i_t ¿ 0 and at steady state when $i_t = 0$. Like s_t r_t will also reach a steady state value.