SIR Model of Epidemics*

Troy Tassier

September 15, 2005

1 Introduction

In the last class we discussed a model of disease spread where individuals were susceptible to a disease, potentially contracted the disease, and then recover and are once again susceptible. we called this model the SIS model. Today we study a model of different diseases. Those for which the person contracting the disease becomes immune to future infections after recovery. This model is called SIR for Susceptible - Infected - Recovered. As a variant on this title SIR can also stand for Susceptible - Infected - Removed if we allow for people to die from a disease and thus leave the population we are studying. Examples of diseases which fit the first category include chicken pox and influenza. Others that fit the second category where we may consider removals from a population would be HIV or the Bubonic Plague where many or most individuals die from contracting the disease. In order to keep things simple we will concentrate on the first version of the model so that we may keep the population constant across time. This decision is entirely to keep things simple. And, you should be able to directly extend the model to one of a fluctuating population (with births and deaths) without much difficulty.

1.1 An Example

As an example consider the following data from an influenza epidemic in New York in the winter of 1968-69. This strain of influenza was dubbed the "Hong Kong flu" due to its place of discovery. Table 1 lists the number of "excess" pneumonia-influenza deaths that winter.

Of course not everyone dies from influenza thus the data do not count all instances of the influenza. But we might view the number of excess deaths as an indication of how wide spread the Hong Kong flu was in each week if we assume that the number of excess deaths is proportional to the number of total incidents of the disease in our population of interest. Our goal for today will be to attempt to devise a model that will describe this data.

*Data presentation on the Hong Kong Flu adapted from David Smith and Lang Moore, Duke University

Table 1: Excess Deaths

Week	Deaths
1	14
2	28
3	50
4	66
5	156
6	190
7	156
8	108
9	68
10	77
11	33
12	65
13	24

2 The SIR Model

In the introduction we mentioned that tonight 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.

S_t = the number of susceptible individuals in the population at time t.

I_t = the number of infected individuals in the population at time t.

R_t = 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.

$s_t = S_t/N$ (the susceptible fraction of the population at time t.)

$i_t = I_t/N$ (the infected fraction of the population at time t.)

$r_t = R_t/N$ (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 $S_t + I_t + R_t = N$ and $s_t + i_t + r_t = 1$.

2.1 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:

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

Through similar reasoning we see that:

$$R_{t+1} = R_t + \kappa I_t \tag{2}$$

and

$$I_{t+1} = I_t + \beta s_t I_t - \kappa I_t = I_t(1 + \beta s_t - \kappa) \tag{3}$$

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

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

Through similar reasoning we see that:

$$r_{t+1} = r_t + \kappa i_t \tag{5}$$

and

$$i_{t+1} = i_t(1 + \beta s_t - \kappa) \tag{6}$$

If we add up these equations we will find that $s_{t+1} + i_{t+1} + r_{t+1} = s_t + i_t + r_t = 1$.

3 Steady States

Unlike the SIS model studied in the last class there do not exist positive steady state levels of infected individuals in the SIR model if there is a constant population. 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. This was true in the SIS model too. There has to be an initial infected person in order for there to be more infected people. But in the SIS model there were other steady states as well. Let us see why those do not exist here. 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 i_t will either be increasing or decreasing. How do we know which? We see from Equation 6 that $i_{t+1} = i_t(1 + \beta s_t - \kappa)$. Let $\rho_t = 1 + \beta s_t - \kappa$. This is the epidemic threshold for the SIR model with a constant population. If ρ_t is greater than 1 then we are multiplying i_t by a number greater than 1 so $i_{t+1} > i_t$. The number of infected individuals is increasing. But if ρ_t is less than 1 we are multiplying i_t by a number less than 1 so $i_{t+1} < i_t$. The number of infected individuals is decreasing. Clearly any time $\rho_t > 1$ or $\rho_t < 1$ 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 $\rho_{t+1} < \rho_t$. 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, $\rho_t = 1 + \beta s_t - \kappa$ We see that there are three constants in the equation: $1, \beta$ and κ as well as one state variable s_t. Let us take another look at the equation for s_t 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 $\beta > 0$. 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. What remains is to calculate what those steady state values are and the dynamic path to steady state given a set of initial conditions.

We could calculate this using analytical methods but we can most likely proceed quicker and with more understanding if we use computational techniques. This also will allow us to further expand our toolbox for understanding the rise and fall of epidemics.

4 Computational Analysis

To get a better feel for how the model works we will use an Excel spreadsheet to view the dynamic path of the epidemic. (I will place a copy of the spreadsheet on the course web-page

for you to use in a HW assignment at the end of class.) People usually recover from influenza in about 3 three days. Note that what we are really interested in here is the length of time that it takes to become unable to pass on the disease. So, even if one still does not feel well after three days it may be that they are past the time that they are infective. With this in mind a good estimate for κ is 1/3. β is probably a little tougher to estimate. So, let us begin by looking at a range of parameter values for β and see what looks probable. What value of β creates a time series that looks like the data? Answering this question also will give us a chance to see how changing the value of β affects the course of the epidemic.

Let us start with $\kappa = 1/3$ and $\beta = .1$ and $I_0 = 1$. What happens? Essentially nothing. Similarly if we increase β to 0.2 and then to 0.3 nothing happens. And this makes sense if we look at the analytics of the model. Recall that I_t will only grow if $1 + \beta s_t - \kappa > 1$. So, since s_t is approximately 1.0 with only 1 infected person, ρ_t is approximately $1+\beta-\kappa$. Thus if $\beta < \kappa$ then an epidemic will not occur. But let us look at what happens when we increase β to 0.4. Now we get a small spike in infections around period 200. And about 32% of the people get the disease. Now note that while 32% of the people get the disease there are never more than about 1.5% that are infected at any one time. If we increase β further, to 0.5, we start to get something that resembles the data with which we began this lecture. Here we get a peak in infections around period 100 with about 6.5% of the population infected and just under 60% of the population has been infected by the end of the epidemic.

If we keep increasing β we find that the height of the spike in infections increases and a larger percentage of the population becomes infected over the life of the disease. For instance if $\beta = 1$ then about 96% of the population gets infected at some point.

Now let us return to parameters that look fairly similar to the Hong Kong flu outbreak: $\kappa = 1/3$ and $\beta = 0.6$. And let us use our spreadsheet to view the effect of the parameter κ which is the period of infectiousness. First let us decrease κ to 0.2. This means that 20% of the infected population recovers in each period. Or, it take 1/0.2=5 periods for an infected person to recover. When this occurs we get a bigger spike in infections at the height of the epidemic and a much larger percent of the population is infected over the course of the epidemic (95% compared to 75%.) If we lower κ further to 0.1 we see and even bigger spike and a larger percent of the population (essentially everyone) gets infected. In addition the height of the epidemic lasts longer. By comparison if κ is smaller (people recover from the epidemic faster) the epidemic is smaller. For instance if people recover in two periods instead of 3 ($\kappa = 1/2$ instead of $\kappa = 1/3$) only 32% of the population becomes infected instead of 75%. Thus a shorter period of infectiousness implies that an epidemic is of smaller magnitude and lasts for a shorter period of time. And as we saw above if we increased κ even further so that $\kappa = \beta$ the epidemic never occurs.

Recall that there are two ways that people leave the infected pool. One is that the individual recovers. The other interpretation of the model is that people move to the R pool because they are removed, or in other words, they die. One interpretation of our results above then is that epidemics that slowly kill people can be far more dangerous to the population as a whole than epidemics that kill people quickly. Thus if we compare some of the more sinister diseases of the past, epidemics such as the black plague (which usually resulted in death after less than one week) are potentially far less dangerous than something like HIV/ AIDS which can take years. Of course this is tempered by our current medical knowledge, treatment procedures, and understanding of the transmission of diseases. Yet, this is a result

that many find surprising. If a disease removes its carriers quickly the disease is not likely to have a long life itself.

5 Herd Immunity

In the last section we saw one surprising fact, if people are going to die from an infectious disease it is better that they die fast *for the purpose of ending the epidemic*. In this section we will see another potentially surprising fact: You do not need to immunize everyone in the population in order to prevent an epidemic.

To see this first recall that i_t is decreasing in the amount of susceptible individuals in the population. More exactly we saw above that if $1 + \beta s_t - \kappa$ is less than 1 the number of infected individuals will decrease. The social contact structure and the rate of removal, β and κ, are parameters of the social structure of the population and the disease of interest. There is nothing that we can do about κ since it is a function of biology and physiology. (Well, doctors and medical researchers can work on this problem but for the most part we cannot.) We do have some control over the β since it is a function of the population. We can make sick children and workers stay home when they are ill so that they do not infect others in the population. Recall the days of school closings if too many students are sick. (You probably recall those days fondly if you weren't one of those infected.) We also can quarantine people for certain diseases. This would probably happen if a small pox outbreak occurred. In the next section of the course we will spend a great deal of time discussing specifically the contact structure of social populations and the potential effects of public policies designed to deal with epidemics.

But for now we will do something more simple. Let us assume that the good doctors and medical researchers of the world have done their job and found a vaccine for our disease of interest. And that if we immunize a person they are forever immune from catching the disease and thus from passing it on to someone else. Now, from a modelling perspective this essentially means that when a person is immunized they shift directly from the susceptible pool to the recovered pool without going through the infected phase. Instead of the following picture:

Susceptible $\rightarrow$ Infected $\rightarrow$ Recovered

we have:

Susceptible $\rightarrow$ Recovered for those that are immunized.

Immunizing people from a disease is costly. There are the direct costs of producing and administering the dosage as well as indirect costs of providing information to the public. Thus we would like to be able to provide safety from disease at the lowest possible cost. What we will see in a moment is that we only have to immunize a fraction (many times a small fraction) of the population in order to succeed in our task. Specifically we need to move enough people from the susceptible pool to the recovered (or in this case immune) pool such that the disease disappears of its own accord. In terms of our model we need to move enough people such that: $1 + \beta s_0 - \kappa < 1$. In other words we need to lower the epidemic threshold below one. Or to write it more directly the fraction of people that needs to be immunized is such that $s_0 < \kappa/\beta$. Thus you need to immunize $1 - \kappa/\beta$ percent of the susceptible pool. We should see that this makes intuitive sense. If κ is small that means

that it takes longer to recover from infection and an infective person has more time to infect people. Thus as κ decreases $1 - \kappa/\beta$ increases; you need to inoculate a larger fraction of the population. And as β increases, each infected person contacts more people in a given period, $1 - \kappa/\beta$ increases. Thus again you need to inoculate a larger fraction of the population.

Let's return to our computational model of the epidemic with base parameters of $\beta = 6/10$ and $\kappa = 1/3$. We start the model with $I_0 = 1$; there is one infected individual. With these parameters we end up with about 75% of the population becoming infected. Now let's see what happens when we inoculate some fraction of the population. Start with 50,000. From looking at the graph not much changes. But we do drop to about 74% of the population becoming infected. Go to 500,000 inoculated. You should notice three things: First, the number of people who become infected drops to about 66% of the population. Second, the number of infected individuals at the peak of the epidemic drops. And third, the peak of the epidemic occurs later in the model. Go to 1,000,000 people inoculated. Again we see similar results. Go to 2,000,000. Here we see that the peak level of infections has dropped dramatically as to be almost imperceptible and only about 35% of the people become infected. When we go to 3,000,000 the peak of the epidemic is almost imperceptible and only about 12% of the population becomes infected. And at 4,000,000 the epidemic never occurs.

Let us see how these computational results compare to our analytical results above. We calculated that the number of people that you need to inoculate in order to stop an epidemic is κ/β. In our example this is $\frac{1/3}{6/10}$ or 5/9 of the population. And we saw that the disease disappeared when we inoculated just over 1/2 half of the population (4 million of 7.9 million.) So, in this exercise we have seen how one can use both analytical techniques and computational techniques to help gain a better understanding of a problem. The analytical techniques do a nice job of helping us to find the herd immunity threshold but do a poorer job of intuitively describing the dynamic story. The computational model does a great job of helping us easily see the dynamic story but we have to find the herd immunity threshold by trial and error.

6 Home Work: Non-Constant Populations

Before we close we should briefly consider our assumption of a constant population. In most societies (at least all that I know of!) individuals enter and leave the population either through immigration or even more simply by birth and death. We have not included any of these ideas in our model so far for the sake of simplicity. However, these can easily be incorporated into our equations above.

For homework let us suppose that each period n individuals are born into the population and that m individuals die each period of random causes. Further let us assume that all newly born individuals are born into the susceptible population and that individuals are equally likely to die in any of the population groups. Again let us make a simplifying assumption. Let us assume that $n = m$ so that the population size is constant.

The difference equations for S_t, I_t, R_t, s_t, i_t, and r_t with the birth and death rates included are:

Our equation for S_{t+1} is now:

$$S_{t+1} = S_t - bs_t I_t + n - ms_t \tag{7}$$

Our equation for R_{t+1} becomes:

$$R_{t+1} = R_t + kI_t - mi_t \tag{8}$$

and our equation for I_{t+1} becomes:

$$I_{t+1} = I_t + bs_t I_t - kI_t - mi_t = (1 + bs_t - k)I_t - mi_t \tag{9}$$

Or in terms of the proportion equations:

$$s_{t+1} = s_t - bs_t i_t + \frac{n - ms_t}{N} \tag{10}$$

Our equation for R_{t+1} becomes:

$$r_{t+1} = r_t + ki_t - \frac{mr_t}{N} \tag{11}$$

and our equation for I_{t+1} becomes:

$$i_{t+1} = i_t + bs_t i_t - ki_t - \frac{mi_t}{N} = (1 + bs_t - k)i_t - \frac{mi_t}{N} \tag{12}$$

Note that these equations are identical to those without birth and death (and those in the class spreadsheet) except for the last term in each case. Thus when you change the spreadsheet in question 2 below you only need to add this one term in each equation.

What you should be able to see with a little work is that if there are enough births and deaths in the population there can be steady state levels of the infected population greater than 0. We will use an Excel spreadsheet to see this in the homework.

1) Recreate our spreadsheet from class with parameters $\beta = 3$, $\kappa = .2$, $N = 7,900,000$ and $I_0 = 1$. Use Excel to graph 300 periods of each population group. Print the graph to hand in next class. (Be sure that your spreadsheet matches the spreadsheet from class before you move on!)

2) Add a birth rate and a death rate to your spreadsheet equations where there is a constant number of people born each period and a constant number of people that die each period. Each new person born enters the susceptible pool. Deaths occur with equal likelihood from each group. Hand in the answers to the following five questions. Print the graph for each of these to hand in with your answers. (Print them small so that we don't waste too much paper.) You can write your answer on the back of your graph or on a separate sheet of paper.

a) Let $n = m = 1,000$. What happens?
b) Let $n = m = 2,500$. What happens?
c) Let $n = m = 7,000$. What happens?
d) Let $n = m = 10,000$. What happens?
e) Let $n = m = 100,000$. What happens?

Answer these in an intuitive fashion in two to three sentences. I do not need to know exactly what the equilibrium level of, for example, s is. I want you to tell me something like "s is higher than it was in the base case." Or, "s is increasing at the end of period 300." Tell me about the interesting phenomena that you observe as you change parameters. If nothing substantial changes tell me "nothing substantial changed"!

3) In question 2e) above you should have found a steady state level of infected individuals. Do you expect the disease to affect the old, the young, or all age groups evenly? Explain in two or three sentences. (Think about Chicken Pox and Measles!)