Introduction to the Modelling of Epidemics - SIS Models

Troy Tassier

September 8, 2005

1 Introduction

Tonight we will begin to analyze some formal models of epidemics. In all of these initial lectures we will make some simplifying assumptions with which you may or may not be comfortable. Feel free to point out where you think these assumptions may or may not be true. One of the goals of the course is to prepare you to apply these models to empirical data and to make policy recommendations regarding the spread of epidemics. However I want to point out that just because we will make simplifying assumptions does not make the models invalid or of no use. As an example consider the model of supply and demand from your introductory economics course. There are lots of assumptions buried within the model, some of which you might not agree with, but that doesn't make the model useless.

Similarly, we will make some assumptions early on in this course that are not strictly true. Some of these we will modify later in the course to make more realistic models and attempt to make better predictions... some of them are still open topics on which researchers are working. So again, I stress that you should bring up questions about modelling assumptions. Some of your questions we will be able to quickly and directly address in class. Some of them may take a little or a lot of thought to answer. And I'm sure some of them will make great paper or thesis topics.

We have two goals tonight. One is to introduce you to difference equations and their use in modelling epidemics. And the other is to begin to develop formal models of epidemics that will be useful later in the course when we enter the applied realm.

2 The First Model

To begin let us start with the simplest possible model of an epidemic. It will be a simplified version of what is called an SIS model. The model we will develop will be a dynamic model. We want to monitor the number of individuals in the population who are infected with the disease in which we are interested. In some instances it will be useful to think in terms of the total number of people infected in the population. This will be something like 400 people or 4 million people. In other cases we will want to consider the fraction of the population infected. This will be something like 5 percent or 53 percent. Let us begin to develop some notation. We will use the variable t to denote the time period in most of the models we discuss. (Some papers we might read will use n to indicate the time period.) And let us use

N_t to denote the total size of the population in period t. In this first model the population will be divided up into two groups of people, those that have been infected by the disease and are infective, and those that are susceptible to being infected by the disease. We will label the number of those infected in period t as I_t and the number susceptible in period t as S_t. Where $I_t + S_t = N_t$. You may have guessed by now that the name SIS comes from the two population groups Susceptible and Infected. Individuals go from being susceptible to a disease to being infected. And then they recover and again become susceptible. Thus the progression of the disease from the standpoint of an individual is susceptible - infected - susceptible or SIS.

Individuals potentially move from the susceptible to the infected group when a susceptible person comes in contact with an infected person. What counts as a contact varies with the disease. Sometimes diseases are transmitted through sexual contact or are carried in blood as in HIV. In other cases one only needs to be near a person as the disease is carried via the air we breath as in SARS or influenza. In addition even if you come in contact with someone it is not always guaranteed that the diseased will be transmitted. For instance only a small fraction of sexual contacts result in the transmission of HIV. Finally once someone is infected it takes some amount of time, or some number of periods in our model for the infected individual to move back to the susceptible pool. (In some cases the infected individual never becomes susceptible again as they become immune to the disease after they recover. These models will be the topic of next week's discussion on susceptible - infected - removed, or SIR, models.)

Assume that $N_t = N$ for all t. In other words there is a constant population size. Further let us assume that all the people in the population are the same people from period to period. Note that these are two different assumptions. Can you think of an example where one does not imply the other? One example would be if there were equal birth and death rates in the population. Then the population size would be constant but some of the people in the population would be different each period. For now we will ignore this and assume that there is a constant unchanging population.

Suppose we begin in the initial period with I_0 infected individuals. In epidemics we are interested in how the disease will spread. Thus what we really want to know in many cases is how many infected individuals there will be in the next period. In other words we want to know I_1, and then I_2 and then .... We want to know how the spread of the disease will progress.

Let us begin by looking at the transition from period 0 to period 1. We begin period 0 with I_0 infected individuals. We can call this the state of the system at time 0. New individuals get infected by coming in contact with members of the infected population. Let us assume that each infected person contacts γ non-infected people in each period. Thus the number of possible new infections is γI_0. But not all of the contacts result in an infection. Suppose that only α percent of contacts result in an infection. Thus each infected individual results in $\gamma\alpha$ new infections in each period. Further to keep things simple let us assume that each infected person in period t moves back to the susceptible pool in the next period $t+1$. Let us now write out an equation that describes this process:

$$I_1 = I_0 + \gamma\alpha I_0 - I_0 \tag{1}$$

Note that since we assumed that each infected person recovers int he next period that the $I_t - I_t$ terms will cancel in each period. Therefore we ignore them in the following equations.

Now let us write the equation for period 2:

$$I_2 = \gamma\alpha I_1 \tag{2}$$

This is the same equation but with different time subscripts. Now if we substitute Equation 1 into Equation 2 we get:

$$I_2 = \gamma\alpha\gamma\alpha I_0 = (\gamma\alpha)^2 I_0 \tag{3}$$

Similarly, we would get the following for I_3:

$$I_3 = (\gamma\alpha)^3 I_0 \tag{4}$$

You should start to see a pattern here. And the general solution to this difference equation is:

$$I_t = (\gamma\alpha)^t I_0 \tag{5}$$

The number of people currently infected in period t is the product of the contact and transmission parameters raised to the power t multiplied by the initial size of the infected population.

Now, let us see if we can figure out how this system will behave. Let us look at the number of infected persons in period t where t is far into the future. In other words t is large, say 1,000. Suppose that there is one infected person in the initial period and that γ is 5 and α is 0.1. Will there be many people infected or a few at period 1,000? $\alpha\gamma = 0.5$. So we expect that there will be 0.5^{1000} people infected in period 1,000. You can check on your calculators if you like but this is a VERY small number, essentially 0. What if we increase α to .3? Now we get 1.5^{1000}, a VERY big number! What happens if α is 0.2? We get $1^{1000} = 1$. Lets try one more, let I_0 be 1,000,000 and α be 0.19. Thus in period 1,000 we would have $(0.19 * 5)^{1000} * 1,000,000 = 0.95^{1000} * 1,000,000$ which again is essentially 0.

What you have probably already noticed is that if $\alpha\gamma < 1$ the number of infected individuals decreases to 0 very rapidly; the disease disappears. If $\alpha\gamma > 1$ the number of infected individuals keeps increasing; the disease spreads throughout the population. This is sometimes called the *epidemic threshold*. Now what does this really mean? How can we interpret this result? If the number of contacts times the transmission rate is less than one this means that each infected person infects less than one person on average. So, the number of infected individuals will decrease. It is like the reproduction/ population models you may have studied, if each person has less than one offspring the population will die out. But if the average number of offspring is greater than one the population will grow. Just like our model when the average number of people infected is greater than one; the disease continues to spread to a larger and larger fraction of the population. Thus we reach the epidemic threshold whenever greater than one person is infected by each infected person.

The model of this section has some weaknesses that we will correct in the next section. But the main point of the model was that we can understand most of what is going on if we look at just a couple parameters in the model. And, from a public policy standpoint if

we can alter those parameters we can control an epidemic. As an example, if we can limit the number of contacts of infected people with non-infected people so that we are below the epidemic threshold we can end the epidemic.

3 A Full SIS Model

The model of the last section was kept overly simplistic in order to introduce you to some key ideas of the course. First, you were introduced to using difference equations to study a diffusion process. And second, you were introduced to some key parameters that we will use throughout the course. In this section we more fully develop the SIS model to a form that you are likely to encounter in policy and research discussions of diseases that fit the SIS framework.

First, in the last section we assumed that each contact of an infected person was with a non-infected, or susceptible, person. It is more realistic to assume that the number of susceptible contacts is a function of the number of susceptible persons in the population. Second, we assumed that each infected person was fully recovered after one time period. This may be true if we are measuring time in weeks or months, but it probably isn't true for some diseases if we are measuring time in days. Thus we would like our model to allow for the possibility that it takes multiple time periods for someone to move from the infected group back to the susceptible group.

We can do this in the following way. First let us define two new state variables that will measure the percentage of the total population that are susceptible and infected. Let $i_t = I_t/N_t$ be the percent of the population that is currently infected. Define $s_t = S_t/N_t$ as the percent of the population that is currently susceptible. We also will define a new parameter κ that measures the percent of the population that recovers from a disease each period. Thus if the time to recover is 3 time periods then $\kappa = 1/3$. This means that one-third of the population should recover each period on average. (It may seem that this is a weird assumption since it may be that there are different numbers of people infected in each period. Thus different numbers of people should recover each period. We will show in a moment that if we are in steady state this assumption will hold. If you still have a question on this ask after we have finished this lecture.)

We are now ready to write a system of equations that will describe our full SIS model of epidemics:

$$I_{t+1} = I_t - \kappa I_t + \alpha\gamma s_t I_t \tag{6}$$

$$S_{t+1} = S_t + \kappa I_t - \alpha\gamma s_t I_t \tag{7}$$

We have κI_t individuals who recover each period and thus leave the infected group and re-enter the susceptible group. And we have $\alpha\gamma s_t I_t$ individuals who enter the infected group and leave the susceptible group. Notice that the change in the infected group always equals the change in the susceptible group when we have a constant population. Thus $S_t + I_t = S_{t+1} + I_{t+1} = N$ if the population is constant.

Equivalently we can write these equations using our proportion state variables:

$$i_{t+1} = i_t - \kappa i_t + \alpha\gamma s_t i_t \tag{8}$$

$$s_{t+1} = s_t + \kappa i_t - \alpha\gamma s_t i_t \tag{9}$$

Now, with these equations written we can use them to understand the epidemic threshold in the SIS model. First let us ask the question: when will the number of infected individuals be increasing? Intuitively we can reason through this process by just looking at the equations and our box diagram. If more people flow into the infective state than flow out this means that the level of infectives is increasing. If the opposite is true (more flow out than in) the level of infectives is decreasing. Thus our epidemic threshold is determined by whether $\kappa i_t > \alpha\gamma s_t i_t$ or $\kappa i_t < \alpha\gamma s_t i_t$.

If $\kappa > \alpha\gamma s_t$ this means that more people leave the infective state than enter it. Thus the level of the disease is decreasing. If $\kappa < \alpha\gamma s_t$ this means that more people enter the infective state than leave it. Thus the level of the disease is increasing. Now, another way to write this inequality is: $\frac{\alpha\gamma s_t}{\kappa}$. This is the epidemic threshold for the SIS model. If the fraction is greater than one the level of infectives increases (more than one person is infected by each infective) And if it is less than one the level of infectives decreases (fewer than one person is infected by each infective.)

Now one thing different about this model compared to our previous one is that one of our state variables enters the equation for the epidemic threshold, s_t. If $\frac{\alpha\gamma s_t}{\kappa} > 1$ the disease spreads and the number of susceptible individuals decreases. Thus the fraction gets smaller in the next period. If $\frac{\alpha\gamma s_t}{\kappa} < 1$ the disease begins to die out and the number of susceptibles increases. Thus the fraction gets bigger in the next period. For those of you who have taken economics courses this reasoning may sound familiar. It may sound like a process that is working its way to equilibrium. And that is exactly what we will see in many cases here. The lack of susceptible individuals slows the epidemic when there are many infective individuals; and many susceptible individuals increases the spread of the epidemic when there are few infectives. The number of available hosts (or susceptible individuals) introduces negative feedback into the SIS model.

What would be the case where the disease reaches equilibrium? By this I mean the number of susceptibles and infectives is in steady state; both are constant proportions of the population. This would be the case if $\frac{\alpha\gamma s_t}{\kappa} = 1$ or $s_t = \frac{\kappa}{\alpha\gamma}$. This would define a steady state of the system where s_t and i_t are constant in all periods.

4 Steady State

We are now ready to analyze our first full model of epidemics a little more formally by looking at the steady state of our system of equations. Again, for those of you who have taken previous economics courses you may think of a steady state as an equilibrium: a situation where a system is not changing. In this model we will think of a steady state as a state where our state variables, I_t and S_t, do not change from period to period; they are constant. To find the steady state we need to find a solution to our system of equation above where $S_t = S_{t+1}$ and $I_t = I_{t+1}$. One way to do this is to drop the time subscript on our

equations above and solve for s and i. Thus we are looking for a solution to the following set of equations:

$$i = i - \kappa i + \alpha\gamma si \tag{10}$$

$$s = s + \kappa i - \alpha\gamma si \tag{11}$$

To solve these equations first rewrite Equation 10 as:

$$\kappa i = \alpha\gamma si \tag{12}$$

Which can then be written as:

$$s = \frac{\kappa}{\alpha\gamma} \tag{13}$$

This is the steady state value of the proportion of the population that is susceptible. It is the same equation as we found above from our intuitive understanding of the epidemic. And since we know that $s_t + i_t = 1$, we know that the steady state value of the infected proportion of the population is $i_t = 1 - s_t$ or :

$$i = 1 - \frac{\kappa}{\alpha\gamma} \tag{14}$$

Note the intuitive properties of these equations. As the number of contacts of an infected person or the transmission probability increases the number of susceptible individuals decreases (and the number of infected individuals increases). And as the time to recover increases (meaning κ decreases) the number of susceptible individuals decreases (the number of infected people increases.)

Next we will use Excel to view how parameter changes affect these outcomes and to check our analytical predictions of steady state values of susceptibles and infectives.

5 Computational Implementation

In this section we will use an Excel spreadsheet to view the behavior of this dynamic system. You can find the spreadsheet on our course web-page. To start note that the only time that α and γ appear in the equations above is as the product $\alpha\gamma$. So, let's just combine these into one parameter β. Let us start with $\kappa = 0.5$ and $\beta = 0.8$. This means that each infected person recovers in two time periods (again you may think of a time period as perhaps a day) and each infected person potentially contacts and infects an average of 0.8 persons per time period. We also need to specify an initial fraction of the infected population. Let us choose this value to be $i_0 = 0.01$. What happens when you input these parameters into the spreadsheet? You should see that the fraction of infected individuals in the population increases up to about 37% of the population in the figure. We can use our solution of the difference equations to figure out exactly what the level of infected individuals is. Recall that in steady state we expect that there will be $1 - \frac{\kappa}{\beta}$ infected individuals. In our example that means that there should be $1 - \frac{0.5}{0.8} = 3/8$ of the population infected in steady state.

And if we checked closely that is exactly what we find in our spreadsheet. Now leave the infection and recovery parameters the same but increase the fraction of the initial infected population to 0.8. Now you should see that the fraction infected steadily decreases down to the steady level of 3/8. It doesn't matter what the fraction of the infected you start with, your steady state is determined by the infection and recovery parameters of the model. We would say that the steady state infection level is independent of the initial state.

Now let us see how the parameters affect the steady state level. First increase the recovery rate, κ to 0.6. The steady fraction of infected individuals drops to 1/4 of the population. If we increase κ further to 0.7 the steady state fraction drops further. What should happen if we increase κ to 0.8? The disease should disappear. From our analytical results we should see that if $\kappa = \beta$ the susceptible fraction of the population goes to 1.0 and the infected fraction of the population goes to 0.

Set $\kappa = 1$ an leave the other parameters the same. This means that we are returning to the case where each infected person recovers in the period immediately following their infection. Why doesn't the steady state fraction of susceptible increase further? Because it can't be greater than 1.0. Also notice that even though we start with 80% of the population infected that the disease dies out very quickly.

Now let us view the effect of β. First lower β to 0.6 from 0.8. You should notice that the disease dies out even quicker now. As we decrease β we are lowering the number of contacts of each infected person. Thus the disease spreads more slowly. If we increase β to 1.2 we see that we again have a steady state level of the disease at $1/1.2 = 1/6$ of the population. As we keep increasing the number of contacts we see that the infected fraction of the population continues to increase. Try $\beta = 1.5$, $\beta = 2.0$, and $\beta = 2.5$.

Now before we increase the number f contacts further let us quickly revisit our system of equations with these parameters. We have $\kappa = 1.0$ and $\alpha * \gamma = \beta = 2.5$. So we have:

$$i_{t+1} = i_t - i_t + 2.5(s_t i_t) = 2.5(1 - i_t)i_t \tag{15}$$

Now what I want you to notice here is that this is a nonlinear equation. Thus the system might be messy for some parameter values. Now let us increase β to 2.8. You should notice in the picture that there no longer seems to be a stable steady state. The fraction of infecteds has begun to cycle be tween two values. Increase β to 3.0. Again you should observe cycles but this time the cycles are larger with the fraction of the infected population bouncing between 32% and 35% percent and back to 32% again on each consecutive period. If we continue to increase β we begin to see cycles of greater than two periods. And finally if we increase β nearer to 4.0 we see no discernable pattern in the time series. We have entered the chaotic region of the parameter space for this system of nonlinear equations. While we will not delve into the details of non-linear dynamics what you should know is that a chaotic system means that given any starting point you cannot predict the future state of the system without running through each step of the equations. In other words if you start the system with 80% infected and observe that there are 51% infected after 50 periods. And you start the system again at 79% infected or 81% infected there is no guarantee that in period 50 there will be anything close to 51% infected for these starting conditions. (in fact for $\beta = 4.0$ there are 16% infected in period 50 if you start the system at 79% infected and 38% infected if you start the system at 81% infected.) The important point here is that a time series

does you little good without an understanding of the underlying model that the universe has chosen. Looking at this data with almost any statistical procedure would return the answer that the system is random.

What is happening in the system is that each infected person causes a large number of contacts to become infected in the next period. But if there are a large number of infected individuals then there are very few people to infect. So, it is possible to get these wild fluctuations.

Wild fluctuations such as these are not the focus of this course. But I want you to be aware that many of the systems we will study are non-linear. Thus there is always the possibility that they will not be well behaved. Fortunately for you (and me) we will deal mostly with well behaved systems in this course where the parameters of interest will give us nice well behaved answers.

6 Homework

For homework I want you to solve the following problems assuming an SIS model of an epidemic. In order to simplify the problems you may assume that no death or birth occurs.

1. Suppose that $I_0 = 25\%$, $\kappa = 1/4$ and $\beta = 1/2$.
   a) What is the fraction of susceptible individuals in the next period (period 1).
   b) What is the steady state state level of infections in the population.
   c) Suppose that we lengthen the time it takes to recover from the disease. Will the steady fraction of susceptible individuals increase or decrease?

2. Suppose that there are two identical strains of a virus except for the fact that one strain can survive in the air for two minutes and one strain can survive for 5 minutes. Which strain do you think will cause there to be more infected individuals? Explain your answer using the equations and notation in our lesson today.

3. Suppose that $I_0 = 25\%$, $\alpha = .1$ and $\gamma = 3$. Individuals recover from the disease in five time periods.
   a) Write the equations describing the SIS model with these parameters.
   b) What is the steady state fraction of susceptible individuals?
   c) What is the steady state fraction of infected individuals?

4. Suppose that $I_0 = 25\%$, $\kappa = 0.4$, $\alpha = .1$ and $\gamma = 3$.
   a) Write the equations describing the SIS model with these parameters.
   b) What is the steady state fraction of susceptible individuals?
   c) What is the steady state fraction of infected individuals?