

Epidemic Modeling

NONLINEAR SYSTEMS

# Part I: A Simple Model

#### THE SI MODEL

The SI model is considered to be the most basic model of an epidemic in a closed population, and contains only two categories of individuals: those who are susceptible and those who are infected. Births and deaths are not considered in this model, meaning that the total population in the model is a constant. This model is best used for non-lethal diseases that may be caught repeatedly, such as the common cold. The SI model is shown below:

$$\frac{dS}{dt} = \frac{\rho I}{N} - \frac{\beta SI}{N}$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \frac{\rho I}{N}$$

In these equations S and I are variables representing the number of susceptible and infected individuals. The constants  $\rho$ ,  $\beta$ , and N represent the recovery rate of those infected, the contact rate between susceptible and infected persons, and the total number of individuals in the population, respectively. We may assume in this model that none of the three populations (infected, susceptible, or total) will be negative, and we may similarly assume that neither the contact rate nor the recovery rate will be negative. A negative population, recovery rate, or contact rate would all be illogical occurrences with no practical application.

# **EQUILIBRIA OF THE SYSTEM**

There are two equilibrium points in this system, with the first equilibrium point being found when there are zero infected individuals and the entire population is comprised of susceptible individuals. The second equilibrium point could be described as the "carrying capacity" of the system for infected individuals in the population, and occurs when the following equation holds true:

$$S = \frac{\rho}{\beta}$$

This means that the population of infected and susceptible individuals will remain stable when the rate of recovery is equal to the contact rate multiplied by the number of remaining susceptible individuals. This is because the rate at

which individuals are infected and recover may be modeled by the following equations:

$$Infection = \frac{\beta SI}{N}$$

$$Recovery = \frac{\rho I}{N}$$

The first equilibrium point occurs when the rates of infection and recovery are both zero (which occurs when the number of infected individuals is 0), but the second equilibrium point occurs when they are equal to each other and non-zero. The equilibrium of  $S=\frac{\rho}{\beta}$  is obtained by factoring  $\frac{I}{N}$  from the infection and recovery equations.

In terms of an epidemic, this second equilibrium is the point at which the growth of a disease is capped. In the case of the common cold this is why there are always people who have a cold but there is never a time when every person has a cold without special circumstances. A school environment, for example, can act to dramatically increase the rate of contact while simultaneously slowing the rate of recovery to almost zero, allowing the infected population to rise to nearly the entire total population as a result of each individual being confined in a room with at least 25 other infected people for eight hours of the day. Once safety precautions are implemented to counter the epidemic the contact rate will plummet and the recovery rate will improve, allowing the school's population to return to a much lower equilibrium value for the infected population (potentially even the disease free equilibrium).

## STABILITY OF THE EQUILIBRIA

The stability of an equilibrium point of a linear system depends upon the eigenvalues of that system at the equilibrium point in question. If both eigenvalues are less than zero, then the equilibrium point is stable. If both eigenvalues are greater than zero the equilibrium point is unstable, as solution curves are repelled from both eigenvectors. If one eigenvalue is greater than zero and one is less than zero you will have a saddle point where the solution trajectories are drawn towards one eigenvector but repelled from the second, resulting in an unstable equilibrium. If one eigenvalue is zero then the stability of the equilibrium is determined solely by the sign of the second eigenvalue.

Finally, if both eigenvalues are zero then the solution trajectories will take the form of lines parallel to the single eigenvector, generating a neutrally stable equilibrium.

These various statements, however, only apply to systems of linear differential equations, which ours is not. To solve this issue we found the Jacobian of our system of equations at both equilibria, providing us with a linearization of our model at the necessary points (when the infected population is zero and when  $S = \frac{\rho}{\beta}$ ). The two Jacobian matrices are presented below:

Jacobian at the Disease Free Equilibrium

$$\begin{bmatrix} 0 & \frac{\rho}{N} - \beta \\ 0 & \beta - \frac{\rho}{N} \end{bmatrix}$$

Jacobian at the Equilibrium When  $S=rac{
ho}{eta}$ 

$$\begin{bmatrix} \frac{\rho}{N} - \beta & 0 \\ \beta - \frac{\rho}{N} & 0 \end{bmatrix}$$

These linearizations of our equation allow us to find eigenvalues and eigenvectors that will apply to the specific equilibrium points that we wish to inspect. The specific eigenvalues and their associated eigenvectors are shown below:

Eigenvectors and Eigenvalues at the Disease Free Equilibrium

$$\lambda = 0, \left(\beta - \frac{\rho}{N}\right)$$

$$\underline{V}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \ \underline{V}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Eigenvectors and Eigenvalues when  $S = \frac{\rho}{\beta}$ 

$$\lambda = 0, \left(\frac{\rho}{N} - \beta\right)$$

$$\underline{V_1} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \, \underline{V_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Since the only non-zero eigenvalues are functions of the constants in the equation, specifically the rate of contact, the rate of recovery, and the total population, this means that the behavior at each of the equilibria will depend upon what the initial conditions of the scenario are. If  $\beta - \frac{\rho}{N}$  is a value less than zero, meaning that when  $\beta < \frac{\rho}{N}$  the disease free equilibrium will be a stable equilibrium. If the reverse is true, such that  $\frac{\rho}{N} < \beta$ , then it becomes clear that the stable equilibrium will be the equilibrium found when  $S = \frac{\rho}{\beta}$ .

#### APPLICATION OF STABILITY

To translate this finding into terms of the basic reproduction rate  $(R_0 = \frac{\beta N}{\rho})$  simple algebraic manipulation may be employed to find the following equivalencies:

$$\beta < \frac{\rho}{N}$$
 is equivalent to  $\frac{\beta N}{\rho} < 1$ 

$$\frac{\rho}{N} < \beta$$
 is equivalent to  $\frac{\beta N}{\rho} > 1$ 

This means that the disease free equilibrium will be a stable equilibrium when the basic reproduction rate is less than one, while the second possible equilibrium will be stable (and the disease free equilibrium unstable) when  $R_0$  is greater than one.

This is a finding that makes sense when you look at what  $\beta$ ,  $\rho$ , and N are, in terms of our model.  $\beta$  is the rate of contact between different individuals and N is the total number of individuals, meaning that  $\beta N$  would be the total number of contacts between individuals per time.  $\rho$  is the rate at which afflicted individuals recover from their illness per time. When the rate at which people recover over time is greater than the rate at which there is a potential for infection (since each contact is a potential infection), then you will see the disease begin to naturally die out. Likewise when the rate at which people contact each other (and potentially infect each other) outstrips the rate at which people are recovering you will see the population of infected individuals grow

until it reaches its eventual maximum. This maximum is dictated by the rate of recovery and the number of uninfected individuals, because it becomes harder and harder to infect new people if few are left uninfected while simultaneously more people are cured every day when the number of infected individuals is high.

# Part II: A Less Simple Model

### THE SIR MODEL

One of the primary limitations to the SI model is its inability to account for individuals that cannot be infected by the disease. While this makes for a perfect modeling tool when studying diseases such as the common cold, which are almost never fatal and can be caught time and time again, it falls short of modeling the behavior of a majority of diseases in the world. This is because it has no way to represent immunization to diseases or the death of infected individuals (both populations that would then be unable to catch the disease even though they may have interactions with the infected). For diseases like chicken pox and measles it makes more sense to have a third category of people, those who are either immune through prior infection or dead. This is where the SIR model comes into play.

The SIR model provides this third category of individuals in the form of an "R" group that will represent individuals who are either immune or dead. The equations for the SIR model are shown below:

$$\frac{dS}{dt} = -\frac{\beta SI}{N}$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \frac{\gamma I}{N}$$

$$\frac{dR}{dt} = \frac{\gamma I}{N}$$

As in the previous equation the symbols S and I represent the populations of infected and susceptible individuals, while the symbol R represents the number of individuals who cannot be infected due to death or prior exposure.

 $\beta$  now represents the rate of infection, while  $\gamma$  is the symbol for the rate of removal for the system or the rate at which individuals are added to the R population.

This system, however, is rather unwieldy since it deals with the number of individuals in each population. Instead of modeling the system in terms of individuals, it is easier to model it in terms of percentages of the total population. This can be achieved by using the terms s, i, and r, to represent the percentage of the total population contained by the susceptible, infected, and removed groups respectively. The equations, once transformed into this format, are shown below:

$$\frac{ds}{dt} = -R_0 s i$$

$$\frac{di}{dt} = R_0 s i - i$$

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

In this system the total of s+i+r is equal to 1, since the number of individuals contained within should total to 1 full population of individuals, and  $R_0$  is slightly modified to the form of  $\frac{\beta N}{\gamma}$ . The convenience of this system is that there are still three dynamic variables (s, i, and r) but there is only one parameter,  $R_0$ .

### **EQUILIBRIUM POINTS**

This model, unlike the SI model, has a single equilibrium point. This is shown below:

$$i = 0$$
  $s + r = 1$ 

The reasoning behind this is fairly simple, because there is only one variable in one of the equations. At an equilibrium point you know that  $\frac{di}{dt}$ ,  $\frac{dr}{dt}$ , and  $\frac{ds}{dt}$  will all be equal to zero, and  $\frac{dr}{dt}=i$  meaning that i must be 0. If i is equal to zero, then we know that both  $\frac{di}{dt}$  and  $\frac{ds}{dt}$  must also be zero. The only remaining constraint, that s+r=1, is a result of one of the bounds placed on the system initially (that s+i+r=1).

This equilibrium point shows that the behavior of the SIR model differs greatly from the behavior of the SI model. One of the equilibrium points in the SI model is a point at which there is a balance between the infected populace and those who are susceptible. This means that there can be a stable solution that allows for the populations of the infected and susceptible individuals to remain a non-zero constant. The SIR model does not allow for this, since the only equilibrium point is when the number of infected people is zero. As time passes you will always see a disease eventually die out in an SIR model with an unchanging population (no new births) since the population of infected individuals will always tend towards zero as time tends towards infinity.

#### SIZE OF AN EPIDEMIC

This distinct lack of equilibrium with a non-zero infected populace means that one of the best ways to measure an epidemic modeled by the SIR equations is its total size, or the number of people who are infected over the course of the epidemic. In this model anyone who is infected will eventually be removed from the infected population (and added to the removed population), meaning that the size of an epidemic with r<sub>o</sub> people who are initially in the removed population can be modeled by the following equation:

$$\lim_{t\to\infty}r(t)-r_0$$

This allows you to look "retrospectively" at the removed population once the epidemic has passed and subtract the number of people who were initially in the removed population to find the number of people who were added to it during the epidemic, and thus find the total number of people who had been infected over the course of the epidemic.

## EFFECTS OF THE BASIC REPRODUCTION RATE

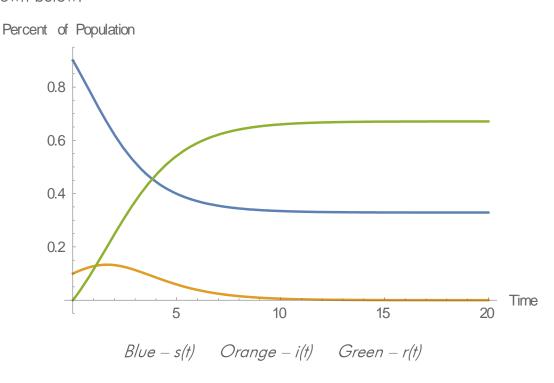
By changing the value of  $R_0$  we can study the effects of the various parameters on the overall size of the resulting epidemic. This is achieved by setting the basic reproduction rate to different values while keeping the initial conditions identical. Because  $R_0$  is equal to  $\frac{\beta N}{\gamma}$  we can know that an increase in  $R_0$  is equivalent to either a decrease in the rate of removal or an increase in the infection rate and/or population, while a decrease in  $R_0$  is equivalent to an

increase in the rate of removal or a decrease in the infection rate and/or population.

In our tests on the effects of the basic reproduction rate we will maintain the initial conditions of i(0) = 0. 1 and s(0) = 0.9, with r(0) = 0.

$$R_0 = 1.5$$

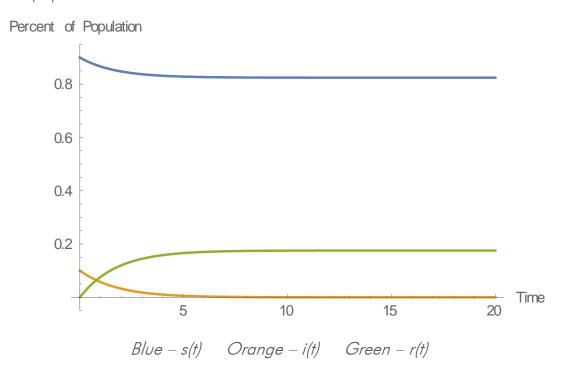
When the basic reproduction rate had a value of 1.5 the total size of the group that was infected over the course of the epidemic was 67.11% of the total population. The graph of the population of the three groups as time passes is shown below:



As you can see, the infected population grew for a short period of time until it attained its peak value, before petering off to zero as time passed. The maximum portion of the population that was infected at any given time was less than 0.15, but when you look at the total number of people that were infected (the graph of r(t)) that portion grows to be much larger since people are constantly being infected.

## $R_0 = 0.5$

This test was performed with the same initial conditions, but a basic reproduction rate of 0.5 rather than of 1.5. In this case the total percentage of the populace that was infected during the epidemic was only 17.57%, a reduction of nearly 50% as compared to when  $R_0$  was larger. The graph of the three populations can be seen below:



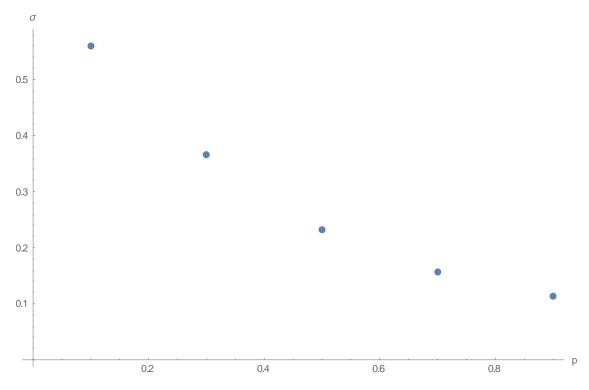
This graph behaves much differently than when  $R_0$  had a greater value. Rather than initially increase the number of infected individuals started off by decreasing before beginning to taper off towards zero. This is because the basic reproduction rate was less than one, meaning that the removal rate (the rate at which infected people either die or are cured) is larger than the rate at which new individuals become infected. This causes an negative initial slope for i(t), which only continues to trend lower towards the equilibrium solution as time passes.

### FFFFCTS OF VACCINATION

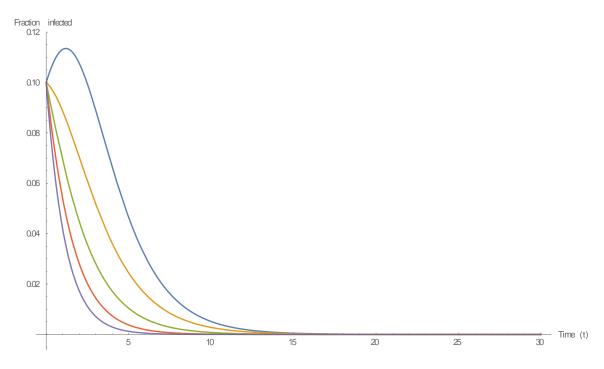
The one scenario that the most basic SIR model doesn't cover is vaccination, a practice commonly seen in developed nations across the globe. By exposing oneself to a harmless version of a virus, people are able to gain an immunity to this virus should the real version attack them in the future. The effects of vaccination are easily added, however, by moving the vaccinated individuals from the susceptible population to the removed population. This means that, for a percentage p of susceptible individuals that have been vaccinated, the initial conditions may be described as shown below:

$$s(0) = (1 - p)s_0$$
$$i(0) = i_0$$
$$r(0) = ps_0$$

This allows us to test the effects of vaccinating various percentages of the susceptible population at the start of an outbreak to see how effective it is at combating an epidemic. Below is a scatterplot that compares the size of an epidemic ( $\sigma$ ) to the percentage of the susceptible population ( $\rho$ ) that is vaccinated when s<sub>0</sub> is equal to 0.9, i<sub>0</sub> is equal to 0.1, and R<sub>0</sub> is equal to 1.5.



From this scatterplot we can clearly see a decrease in the overall size of the epidemic when the percentage of people in the susceptible population is increased. At the extreme end of the spectrum the size of an epidemic that starts with 10% of the population infected only ends up infecting a total of 11.42% of the population when 90% of susceptible individuals are vaccinated at the onset of the epidemic. This means that only 1.42% of the total population was infected beyond those who already had been infected at the start. A telling graphic, showing the percentage of the population that is infected over time, is included below:

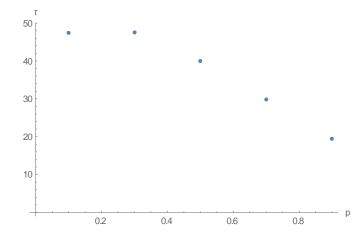


Blue - p = .1 Orange -p = .3 Green -p = .5 Red -p = .7 Purple -p = .9

When only 10% of the population is vaccinated it is clear that the number of infected individuals is on the rise when t = 0.

Similarly, the duration of an epidemic is affected by the percentage of the population that is vaccinated. An epidemic ends when there are zero infected individuals left. As you can see in the plot above, higher percentages of vaccinated individuals result in epidemics that reach zero infected individuals

faster. This trend is easier to visualize in the graph to the right, where  $\tau$  is the duration of an epidemic and p is the percentage of the population vaccinated before the infection begins. As you can see, the duration of an epidemic decreases as vaccination percentage increases.



We can take this graph a step further if we ask ourselves what percent of the population we need to vaccinate in order to prevent an outbreak from becoming an epidemic. In this case, we define an epidemic to be any disease outbreak that results in the number of infected individuals increasing before enough of them are removed from the population that the epidemic stops. Conversely, an outbreak is defined as a disease by which the number of infected individuals is decreasing at time t=0. Thus, in order to find the bifurcation value between these two states of infection, we simply set t=0 and allow the rate of infection to vary as a function of p, the percentage of individuals vaccinated.

$$i'[p] = R_0(1-p)s_0i_0 - i_0$$

Solving for i'[p] = 0, we find that the bifurcation value occurs at  $p \approx 0.26$ , or 26 percent of the population. If the percent vaccinated exceeds 26 percent, the epidemic is defanged, and becomes a simple outbreak. Below this number, any outbreak has the possibility of becoming an epidemic.

#### SHORTCOMINGS OF THE MODEL

Unfortunately no model is perfect, and this one is no exception. One of the notable limitations of this model is that it is very limited in scope. That is to say that it would be entirely inaccurate if it were used to model the spread of disease across an entire country, because the rate of infection is a constant. This means that it would only be able to be applied to a setting that is consistently urban throughout the entire population, or consistently rural. It cannot be used to model the disease in a city or area that has vastly different population densities.

A second issue with this model is the fact that the initial conditions that were picked were rather arbitrary and sometimes unrealistic. There will never be a situation in which 10% of the population is suddenly afflicted by an illness without any lead up to that point. This means that you would likely start with a number of infected individuals of around 50 in total, and it would spread from that point rather than from a point in which 10% of the population is afflicted.

The final issue with the model is that Wolfram Mathematica, when solving the differential equation, plots specific points before using an interpolating function to "connect the dots" between the points. This isn't an issue when the slope of the function is significant, but as it begins to level off it starts to behave erratically, particularly when looking closely at the graph to find the zeros of the function. This means that the values for  $\tau$  may not be as accurate as if an exact solution was used to find them.

# **Appendix**

Diff project Part I

$$x = sost = y = torrect$$
 $x' = \frac{p}{y} = \frac{p \times y}{N} = 0$ 
 $y' = \frac{p \times y}{N} = \frac{p}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}{N} = \frac{p \times y}{N}$ 
 $y' = \frac{p \times y}{N} = \frac{p \times y}$ 

$$\frac{1}{\sqrt{\frac{\beta \cdot N}{N}}} + \frac{\beta \cdot \frac{\rho}{\beta}}{\frac{\beta}{N}} = 0$$

$$\frac{1}{\sqrt{\frac{\beta \cdot N}{N}}} - \frac{\beta \cdot \frac{\rho}{\beta}}{\frac{\beta}{N}} = 0$$

$$\frac{1}{\sqrt{\frac{\beta \cdot N}{N}}} - \frac{\beta \cdot \frac{\rho}{\beta}}{\frac{\beta}{N}} = 0$$

$$\frac{1}{\sqrt{\frac{\beta \cdot N}{N}}} - \frac{\beta \cdot \frac{\rho}{\beta}}{\frac{\beta}{N}} = 0$$

$$\frac{1}{\sqrt{\frac{\beta}{N}}} - \frac{\rho}{N} = 0$$

$$\frac{1}{\sqrt{\frac{\beta}{N}}} - \frac{\rho}{$$

eq 2 
$$\begin{cases} \frac{\rho}{N} - \beta & 0 \\ \frac{\beta}{N} - \frac{\rho}{N} & 0 \end{cases} \xrightarrow{R_z + R_x} \begin{cases} \frac{\rho}{N} - \beta & 0 \\ \frac{N}{N} & 0 \end{cases}$$

$$\Rightarrow \begin{cases} 1 & 0 & 0 \\ 0 & 0 & 0 \end{cases} \xrightarrow{Y = K} \begin{cases} N_{z_1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{cases}$$

$$4 \quad R_0 = \frac{\beta N}{\rho}$$

$$\beta - \frac{\rho}{n} < 0 \qquad \beta < \frac{\rho}{n} \qquad \beta_n < \rho \qquad \frac{\beta N}{\rho} < 1$$

$$-\beta + \frac{\rho}{n} < 0 \qquad \frac{\rho}{n} < \beta \qquad \rho < \beta N \qquad 1 < \frac{\beta N}{\rho}$$

Diff project Part II

6. 
$$\frac{ds}{dt} = -R_0 si$$
 $\frac{di}{dt} = R_0 si - i$ 
 $\frac{dr}{dt} = i$ 
 $\frac{dr}{dt} = i$ 
 $\frac{S=0}{8} si = 0$ 
 $\frac{S=0}{8} si = 0$