6

Between-Host Disease Models

Between-host models are often referred to as epidemiological models. This is one of the few chapters where we do not specify a particular biological context. The framework of these models is extremely general and include diseases such as the flu, HIV, Covid and other communicable and vector-born diseases. Sensitivity methods include visual screening methods such as spider plots, tornado plots and

6.1 Historical Background

Mathematics has been used to study disease progression within a population since at least the 1600s when John Graunt used death records from London to develop a method to estimate the risk of dying for different diseases [22]. In fact, it is arguable that alongside ecological models, epidemiological models have been among the earliest models developed as well as those with some of the most lasting affect on peoples daily lives. Insurance costs depends on factors that are, in part, determined by these models. The development of drugs from antibiotics, to antivirals, to vaccines (both in the development of the vaccine and in the use of the vaccine to prevent the spread of diseases) are motivated and assisted using mathematical models. Questions such as where should governments allocate resources for disease prevention such as HIV prevention (needle exchange? education?) and outbreak diseases like Ebola, West Nile virus are influenced by epidemiological models. Mathematicians from Daniel Bernoulli – who introduced a model of smallpox treatments in 1760 – to John Snow – who studied the spread of cholera and used his quantitative theory to locate the source of a cholera outbreak in London to current mathematicians including those working with governments throughout the world use

mathematical models to inform decision-making as well as preventative methods. During the development of modern theory, models have gone from highly disease specific (e.g. focusing on one disease in one region) to relatively general (focusing on theoretical and qualitative results) to back to more complex and specific models that use the general theory as a guide to understanding.

The outbreak and spread of COVID-19 is an example that is occurring at the time of this writing. Predicting the spread of the disease, excess mortality, impact of methods to control the pandemic and even economic cost of the disease all rely on models. Because models rely on specific assumptions each model analysis often produces different predictions. It is the job of modelers to separate out predictions that differ in magnitude, such as the percent of the population that will be infected, from those that do not, such as whether or not a disease will become endemic (i.e. permanently circulating in the population). It is also the job of modelers to assess the most viable targets for controlling the disease. One of the most controversial topics during the Covid pandemic was the extent masks helped prevent the spread of the disease. Part of this controversy was a consequence of evolving understanding of the spread of the disease. Masks help more if the disease is airborne than if it is not. Another source of different quantitative predictions was a lack of knowledge of specific measurements such as how many particles different types of masks can filter. These sorts of questions can be addressed by estimating the parameter sensitivities as well as analyzing aspects of mathematical models.

Another interesting outcome of the global pandemic is the prevalence of terms that are fundamental in epidemiology – terms like 'herd immunity' and 'R naught (R_0)'. It is defined as the number of new infections caused by one infected member of the population over the course of the infection. Intuitively, if R_0 is smaller than one the infection spreads too slowly to maintain or increase in number. This quantity is of fundamental importance in epidemiology since it provides the border between the spread and decline of a disease. Therefore estimates of R_0 and the effect policy decisions have on R_0 are central to community response to an infectious disease [26, 27, 5].

What is less understood is how to estimate this from data and from models. In demography this could be estimated by taking the total

number of offspring in a year and dividing this by the total adult population to arrive at the average offspring per adult produced in a year. However, there are clear issues with this including how to draw a line between adults and offspring and how to adjust for variations within a year. The other difficulty with this method is that it is not at all predictive. It may describe whether your population was increasing or decreasing *last* year, but it does not predict whether the population will grow in the following year. This is a particular deficiency when considering infectious diseases and epidemiology since the goal is to develop strategies to counter the spread of diseases including ones that are currently spreading. Having some predictive model helps guide policy decisions during disease outbreaks.

The concept of the basic reproduction number was actually described in the 1920's in the context of demographics about 60 years before it was well established in the study of the spread of diseases. In demographics R_0 is a measure of the number of offsprings per each female member of the population. Again, it is intuitive that larger R_0 implies faster growth. Demographics and epidemiology share many characteristics and many of the earliest research in both fields was performed by Lotka who was interested in population dynamics in a general sense.

There are two groups that dominate the development of modern epidemiology and provide an interesting insight into the role a quantitative theory plays in the development of science [61]. In the 1890's through the early 1900's Ronald Ross studied malaria and was the first to show that mosquitos transmit the malaria parasite. Although it had already been shown that certain parasites can live in the mosquito gut, Ross argued that the spread of malaria was caused by the mosquito as a disease vector and that controlling the mosquito population could prevent the spread of the disease within human populations. This idea was already 'in the air' in the early 1900's with several practical trials of mosquito control to reduced the spread of dengue fever; however, these empirical methods often failed inexplicably. Ross was motivated to use quantitative relationships between insect control and disease spread to develop estimates for the amount of mosquito population needed to be reduced and how large an area needed to be treated. Ross developed mathematical models that were combined with data that indicated that

it was not necessary to eradicate the entire mosquito population. Instead he estimated a threshold density below which the disease would not spread.

This work was extended and refined by George Macdonald in the 1950's. Macdonald extended Ross' theoretical model that had compartments for susceptible and infected humans and mosquito density, by including compartments that represented infected mosquitos and details about the mosquito life-stages and infectivity. Macdonald developed his theory at the same time that DDT was created and widely used. This lead Macdonald to focus on important aspects of control which extended the application of the theory. Additionally, Macdonald worked closely with field studies to refine estimates of key parameters. Together, the Ross-Macdonald provided a template, generally specific to the spread of malaria, for the interplay between theory and application.

In 1927 Kermack and McKendrick published one of the classic papers in mathematical biology – 'A Contribution to the Mathematical Theory of Epidemics' [38]. This paper generalized the Ross-Macdonald theory and presented a very general theory of epidemics. A special case developed in this paper is the version of epidemiological models that most students learn and the one presented below. It is one of the most fundamental models used in biology sharing overlap with ecological, demographic and population models.

Although Kermack and McKendrick are credited with modern epidemiological models, it must be noted that Ross, Macdonald and Kermack and McKendrick all focused on a population threshold that determines whether the disease spreads or not if it is introduced into a population. In some cases (as we will show below) the threshold and the basic reproduction number provide the same information. But in general, R_0 provides a more robust understanding about the spread of an infectious disease. It took almost 30 years for R_0 to replace the population threshold in epidemiology even though it was in widespread use in demography to estimate population decline. Epidemiology rests on incremental advancement in the science from observations that indicate the most important biological processes, for example the insect vector, to quantitative estimates based on the current knowledge that are compared to observations to determine next generation models. The theory

also needed cross-fertilization to develop practical methods to combat the spread of infectious diseases.

There are many methods to calculate R_0. One of the most intuitive is to determine parameters for which the infected population increases, $\frac{dI}{dt} > 0$ when the disease is introduced into a susceptible population. This is equivalent to determining whether the disease free state is stable or not. We note that this is not the only method used to estimate the propagation of the disease and is sometimes difficult to provide biological interpretations.

6.2 Two Compartment Models

6.2.1 Model

Epidemiological models are excellent examples of *compartmental* models. The dependent variables in these models are conceptualized as compartments with connections between compartments representing transitions between compartments. An example compartmental diagram is shown in Figure 6.1.

Kermack and Mckendrick

McKendrick was a medical doctor as well as a competent mathematician who published more than 50 papers on epidemiological theory. Kermack as trained as in mathematics and chemistry but an accident left him blind so he focused on theoretical studies. The 1927 paper written by Kermack and McKendrick is a relatively controversial paper but not because of its' content. This paper is often cited as the foundation of the SIR model; however, the differential equations as described here are only described as a special case of a more general framework. In fact, the differential equations were explored by Ross, McDonald and others. Kermack and McKendrick made important contributions – mainly by connecting the concept of per capita infection with a propagation threshold.

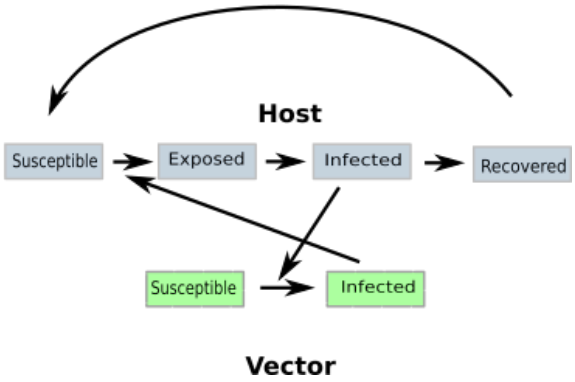

Figure 6.1: Schematic of a compartmental model. This example includes susceptible, exposed, infected and recovered host populations with susceptible and infected compartments for the disease vector. This schematic is relevant for diseases such as malaria, but not for diseases spread person to person such as the flu.

The basic ingredients to epidemiological models require some distinction between members of the population in order to describe how the disease propagates. The simplest model distinguishes members who are sick and those that are not sick but have the potential to become sick. In the simplest diseases, the illness is passed from infected individuals directly to susceptible members. We will begin with an illness that never kills the infected individual but the individual never recovers from. In this case our population has two compartments shown in Figure 6.2 – susceptible, S, and infected, I. A word equation that describes the dynamics is,

$$\begin{aligned} \text{Rate of Change of } S &= \text{Loss of } S \text{ to } I, \\ \text{Rate of Change of } I &= \text{Gain of } I \text{ from } S. \end{aligned}$$

If we denote the rate at which susceptible individuals become in-

Figure 6.2: Disease schematic with only two compartments.

fected as k, we can make some logical inferences about this rate. First it must depend on both S and I since there should be not transfer from susceptible to infected if either of these compartments is zero, so $k = k(S,I)$ and both $k(0,I) = 0$ and $k(S,0) = 0$. There are many forms of this rate but if we assume that the interactions between susceptible and infected individuals is random we can assume that the law of mass action approximates the rate of interaction and that transmission of the disease proportional to the product of S and I, $k(S,I) = kSI$. We note that this assumptions is definitely not universally true – during the COVID pandemic it became clear that front line workers (e.g. doctors and nurses) necessary workers have higher rates of contracting the disease implying that the transmission was not random. Another example where random interactions is not appropriate are diseases that pass from mother to child. This is an assumption that we should exam closely if we find the predictions from our analysis are unreasonable. Realistically, there are no important diseases which can be represented with this model. These are diseases for which there is no cure, but is never deadly. Instead, we can think of this as a step-wise model that provides a possible road-map to understanding the dynamics.

6.2.2 Analysis

Given our assumptions, we can translate the word equations into differential equations,

$$\frac{dS}{dt} = -kSI, \tag{6.1}$$

$$\frac{dI}{dt} = kSI. \tag{6.2}$$

Recall that we assumed the population was constant since there were not births or deaths included – another way to see that this assumption is consistent with this model is to add these two equations to find,

$$\begin{aligned} \frac{dS}{dt} + \frac{dI}{dt} &= 0, \\ \frac{d(S+I)}{dt} &= 0, \\ S+I &= N. \end{aligned}$$

Where we used the fact that the sum of derivatives is the derivative of the sum and that if the derivative of a function is zero the function is constant (denoted N). We see that the total population is always constant and we can use this to simplify the equations. Since $S+I=N$, $S=N-I$ and we can substitute this into Equation 6.2,

$$\frac{dI}{dt} = kI(N-I).$$

This should be familiar since it is the logistic equation. We see that $I=0$ and $I=N$ (that is $S=0$) are the only steady-states and that $I=0$ is unstable for all values of k. We can use the stability of the disease free state to understand how the disease spreads when initially introduced into the population. Because,

$$\begin{aligned} \frac{d\,(kI(N-I))}{dt}|_{I=0} &= kN \\ &> 0, \end{aligned} \tag{6.3}$$

the disease free state is always unstable. In this overly simplified model, there is no parameter choice where the disease does not propagate, so the concept of a particular parameter value that separates the spread and decline of the disease state is not really applicable. Instead, all parameters lead to an increase in the infected compartment.

6.2.3 Sensitivity Analysis: Spider Plot

There are many measures of the spread of a disease. For epidemics, one measure might be the maximum of the infected population since this gives some indication of the strain on hospitals. For this two-compartment model, we know that the maximum is N so there would be no variation in the maximum value. Therefore all parameters including the total population, N, transmission rate, k and initial condition would have the same sensitivity.

It also might be of interest to see how fast the disease is spreading. This could be quantified early in the disease process when the disease is first introduced or at different time points to see if the disease is increasing the rate of spread or not. This could be determined by the right-hand-side of Equation 6.3.

Another marker that provides information about the disease progression is the time the infected population reaches half of the maximum (N). This is a standard measure for sigmoidal curves referred to as the half-saturation constant . The workflow for this would be,

1. Select a parameter set
2. Solve the differential equation
3. Determine the time that $I = \frac{N}{2}$, $QoI = T_{half}$
4. Save this value.

We should note that for this example we could solve the equation analytically and find the half-saturation constant from the solution. We are using this as an example to motivate the sensitivity method. We will use a graphical method that provides information about the impact of parameter variations in a relative way. These methods are often referred to as 'ranking' methods since the goal is to determine which parameters have more impact than others rather than quantify the effect of varying parameters. Spider plots are relatively simple to understand. We change each parameter, one at a time, sweeping through percent changes (above and below the nominal value). Plotting the QoI's simultaneously provides a quick way to judge the relative importance taking into account the precent differences.

We will add some more interesting behavior to the equations that we will solve numerically since the SI model is completely solvable, so is much less interesting. For demonstration purposes, we can add vital dynamics. We will maintain the two classes, but assume that there is an immigration term (or a birth term) that enters the susceptible population and a death term for the infected population. The new model is,

$$\frac{dS}{dt} = rS(S-\kappa) - kSI, \tag{6.4}$$

$$\frac{dI}{dt} = kSI - \delta I. \tag{6.5}$$

This gives us six parameters to consider: r, κ, k, and δ and the initial populations. To decide on a QoI, we first solve the equations for

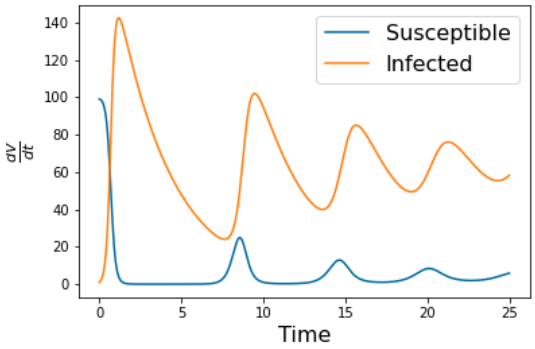

Figure 6.3: Susceptible/Infected dynamics including vital dynamics.

specific parameters. We will use qualitative parameters (e.g. not tuned for any specific observations): $r = .05$, $\kappa = 100$, $k = .05$, and $\delta = .3$. We will start with initial conditions, $S_0 = 99$ and $I_0 = 1$ so that the initial infected population is 1% of the initial populations. An example of the dynamics is shown in Figure 6.3. We see that adding vital dynamics can add oscillations, although this is parameter dependent.

We will use the ratio of I and the total population as the QoI – since this may oscillate, we will take the long-time solution. it can be shown that there is a stable equilibrium, so this is reasonable choice. We show the spider plots for a 50% variation above and below the nominal parameter set (see Figure 6.4).

We see immediately that r and κ lead to decreasing QoI while δ is positively correlated. Each of these has a stronger effect than varying any of the other parameters. In fact, it is impossible to see both the effect of varying S_0 and I_0 because they have exactly the same effect at all levels of variation.

6.3 Classical SIR

6.3.1 Model

We now move to the classic SIR (Susceptible, Infected, Recovered) model represented by the schematic diagram in Figure 6.5. This model

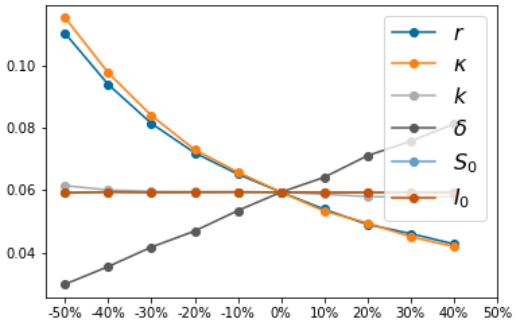

Figure 6.4: Spider plots comparing one-at-a-time variation of parameters in 10% increments. The QoI is $\frac{S_{steady}}{S_{steady}+I_{steady}}$.

can be used to represent the dynamics of a disease which is not deadly but everyone eventually recovers. One could imagine a very mild cold, for example. Although we should point out that there are segments of the population that could be challenged by almost *any* immune system challenge. Immune compromised individuals, elderly, or other subsets of the population need to be accounted for differently. We are also still neglecting births, so we have to keep this in mind when we examine our results.

If we assume that the infection has a fixed recovery rate, γ, we can

Figure 6.5: Disease schematic susceptible, infected and recovery states.

write the equations governing the disease dynamics,

$$\begin{aligned} \frac{dS}{dt} &= -kSI, && (6.6)\\ \frac{dI}{dt} &= kSI - \gamma I, && (6.7)\\ \frac{dR}{dt} &= \gamma I. && (6.8) \end{aligned}$$

It is worth noting that the road-map laid out by the SI model provides some insight into the next steps. If we add the equations we find that,

$$\begin{aligned} \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} &= 0,\\ \frac{d(S+I+R)}{dt} &= 0,\\ S+I+R &= N. \end{aligned}$$

So we can eliminate one of the compartments by replacing it with the difference between the total population and each of the other compartments – for example $R = N - S - I$. Because there are no births and deaths the total population is always N. For the SIR model, the population flows into the recovered compartment, which is not connected to the rest of the compartments in any other way so we can focus on the S and I dynamics and use the algebraic relationship, $R = N - S - I$, for R.

$$\begin{aligned} \frac{dS}{dt} &= -kSI, && (6.9)\\ \frac{dI}{dt} &= kSI - \gamma I, && (6.10)\\ R &= N - S - I. && (6.11) \end{aligned}$$

6.3.2 Analysis

Notice that for a steady-state either $I = 0$. This is can be seen in Figure 6.6, where the the rates of infection and recovery are different. We see different dynamics for the susceptible population. in either case, $S = 0$

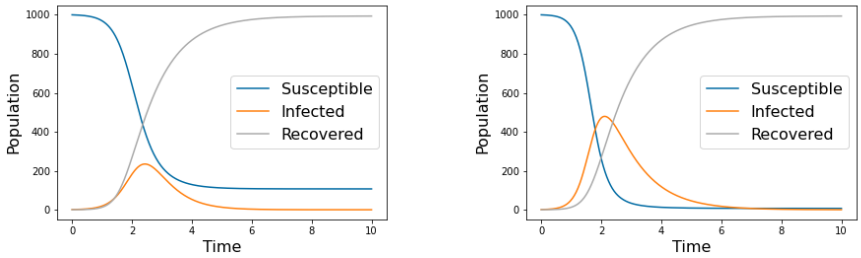

Figure 6.6: Dynamics for low infectivity, $N = 1000$, $k = 0.005$ and $\gamma = 0.5$.Comparison between two different disease examples. One is less infective($k = 0.005$) than the other ($k = 0.01$). These differences result in very different outcomes.

and $R = N$ as $t \to \infty$. In the first case, $k = 0.005$ and $\gamma = 1$. There is a wave of infection that recedes. The second case has a faster recovery time, $\gamma = 2$ and we still see a wave that passes. The second case has a residual of susceptible people who have not had the disease.

This disease wave that sweeps through the population is relatively simple here because the recovered population is immune and there are not variations in the susceptible population. Still, the peak of the wave depends on the characteristics of the disease – How easily is it transmitted? How fast is the recovery time? For many epidemics the goal is to reduce the peak infected population to reduce the impact on the infrastructure – this leads to methods to 'flatten the curve'.

There are several other aspects of the dynamics that are worth noting. The peak for the higher infectivity disease occurs earlier and is higher. This has policy implications since this means the number of people who have the disease and may require treatment is higher and more concentrated. This had a tremendous impact on the hospital infrastructure during the waves of COVID and was the focus of a lot of the modeling to predict resource needs.

These observations can be used to develop QoI's that are geared to answering specific policy questions such as the effectiveness of disease prevention – for example social distancing in terms of COVID or condom use for sexually transmitted diseases.

We can also try and understand the estimates of R_0. For the SIR model, one estimate for R_0 can be seen by determining when $\frac{dI}{dt} > 0$.

Going back to Equation 6.8, we can see that if $kSI - \gamma I > 0$, the infection grows. At the beginning of an epidemic most of the population is

R naught (R_0)

R_0 is notation for the basic reproductive ratio' and is a measure of whether a disease can spread throughout a population or not. Intuitively, if an infected individual is not able to infect at least one other person the disease will not spread since the secondary infections are fewer than the primary infections. In reality, this is far too simplistic of an idea. First of all this does not provide an algorithm for approximating the number of secondary infections. Second, diseases can be far more complicated with far different dynamics in the even of reinfection, for different members of the population and different behavior patterns. Therefore there has been widespread interest in how to calculate R_0 more accurately.

The most straightforward definition of R_0 is through the rate of change of the infected population. If this is positive, the disease is expanding and R_0 is large. Otherwise R_0 is small. The difficulty with this is that it is difficult to interpret this biologically in any general sense. It is the method that we rely on here, however.

Another method is the survival function approach [27] where there is a large population and $F(a)$ defines the probability that an individual remains infectious for a time interval of length a. The average number of secondary infections caused by an infected individual will be denoted $b(a)$. Then, we can define $R_0 = \int_0^\infty b(a)F(a)da$. This can be extended to complex models but rapidly becomes unwieldy.

The most general method is referred to as the 'next generation matrix'. This matrix is a statement of how the susceptible and infected populations change over one generation. This matrix defines the evolution of the population in discrete steps. The maximum eigenvalue of this matrix defines R_0. This is arguably the most general and accurate method to use. However, it is quite intricate in general so we will not focus on it.

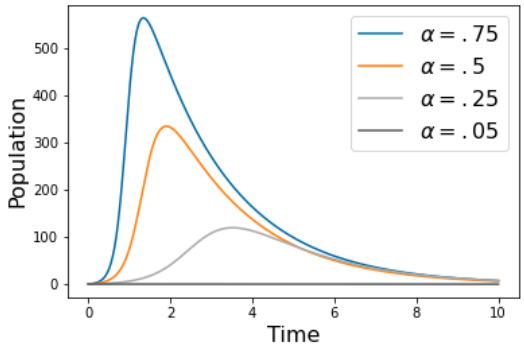

Figure 6.7: Comparison of the infected population for different values of α.

susceptible so that $S \approx N$ so,

$$\begin{aligned} kSI - \gamma I \approx I(kN - \gamma) &> 0 \\ \text{if} & \\ \frac{kN}{\gamma} &> 1. \end{aligned}$$

We can define $R_0 = \frac{kN}{\gamma}$ which ratio of the fraction of the population that is infected and the fraction that recovers. We can check whether this makes sense for our examples. The simulations shown in Figure 6.6 have $R_0 = \frac{0.005 \times 1000}{0.5} = 10$ and $R_0 = \frac{0.01 \times 1000}{0.5} = 20$, so the disease propagates. If we remove a portion of the population so that the susceptible population that has the potential of contracting the disease is αN, we can vary α and determine what percentage of the population needs to be quarantined to prevent the spread of the disease. Examining Figure 6.7, we find that as long as $\alpha < .05$ the disease essentially does not spread – that is an estimate of about 95% of the population must be removed from the susceptible population.

6.3.3 Sensitivity Analysis: Tornado Plots

We will again make the model a bit more complex so that there are a few more parameters to deal with. Just as in the SI model, we can include basic vital dynamics so that there are births of susceptible and

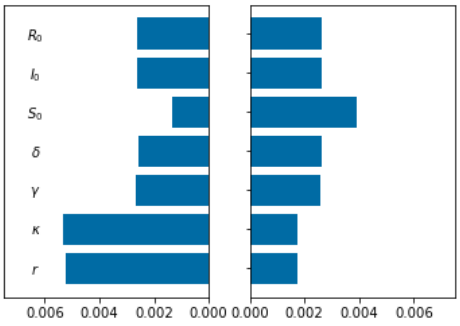

Figure 6.8: Tornado plot of the effect of a 50% increase or decrease in each parameter. We see that there is some asymmetry implying that decreasing κ and r has a stronger effect on the QoI than increasing these parameters. Also, this has a stronger, relative affect on the QoI.

deaths of infected. Our new model is,

$$\frac{dS}{dt} = rS(\kappa - S) - kSI, \tag{6.12}$$

$$\frac{dI}{dt} = kSI - \gamma I - \delta I, \tag{6.13}$$

$$\frac{dR}{dt} = \gamma I. \tag{6.14}$$

Our parameters are the growth rate ($r = .1$), the carrying capacity ($\kappa = 2 * N$), infectivity rate ($k = .01$), death rate ($\delta = .01$), recovery rate ($\gamma = .5$) and total initial population ($N = 1000$).

To consider the sensitivity, we can compare the QoI for the low, versus high values of the parameters. The specific QoI shown in is the same as the one for the SI model: $\frac{S_{steady}}{S_{steady}+I_{steady}}$. One way to visualize this is using tornado plots. These are horizontal bar plots that show differences in increasing (on the right) and decreasing (on the left) each parameter, one at a time. Just as spider plots give a visual and relative comparison, tornado plots quickly show whether increasing or decreasing a parameter has a larger effect. It also compares the variation in each parameter. A plot for this example is shown in Figure 6.8.

There are several caveats – this is still one-at-a-time sensitivity. A slightly more subtle issue is that just looking at the increase and decrease assumes that the behavior is linear. If the sensitivity is nonlinear (for example quadratic) it is possible to mis-represent the variation.

6.4 Waning Antigens

There are numerous examples of infectious diseases where the SIR model is not accurate because recovery does not grant permanent immunity as in diseases like measles. Instead, after a waiting time a recovered individual becomes susceptible so the model becomes an SIRS. It is important to note that the rate of transition between recovered and susceptible matters a lot as it modulates the speed of disease propagation. It is also a key measurable quantity used to estimate vaccination strategies. The goal of vaccination is to move members of the population from the susceptible category into the recovered category. There are many assumptions underlying this strategy that have varying degrees of correctness and depend on specific diseases. For example, we are assuming that recovery from the disease is equal to immunity from vaccination which is a tenuous assumption without proof. There are many diseases where there is a substantial difference between an induced immune response from a vaccination and the response from the active infection. However, all vaccine strategies aim to decrease R_0 by reducing the number of susceptible individuals that an infected individual to interact with. If a vaccine program is slower than the timescale of recovery to increased susceptibility, R_0 may never decrease enough to stop the spread of the disease.

For now, we will assume that there is no vaccine for the disease and that once an individual has been infected and recovered they move from the recovered compartment to the susceptible compartment at fixed rate (see Figure 6.9). We are again neglecting vital dynamics so we expect the total number of individuals in the population to remain constant – which will guide our analysis.

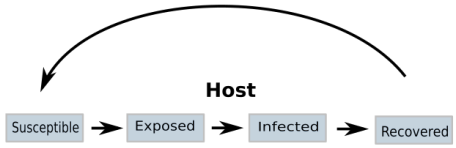

Figure 6.9: Disease schematic with four compartments.

6.4.1 Model: SIRS

Using the familiar compartmental model, we describe the dynamics of all compartments. The equations are the same as the SIR model except that we have added the transition from recovered to susceptible at a rate α.

$$\frac{dS}{dt} = -kSI + \alpha R, \tag{6.15}$$
$$\frac{dI}{dt} = kSI - \gamma I, \tag{6.16}$$
$$\frac{dR}{dt} = \gamma I - \alpha R. \tag{6.17}$$

Again, we have a conservation law since there are no deaths in this model: $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$ so that $R = N - S - I$ and the system of three Equations 6.16 - 6.17 can be reduced to two equations,

$$\frac{dS}{dt} = -kSI + \alpha(N - S - I), \tag{6.18}$$
$$\frac{dI}{dt} = kSI - \gamma I. \tag{6.19}$$
$$\tag{6.20}$$

Interestingly, the estimate for R_0 we used for the SIR model, based solely on whether I grows or decays, is the same for the SIRS model. This suggests that this simple definition of R_0 is missing something – either that or our understanding of the spread of a disease is not

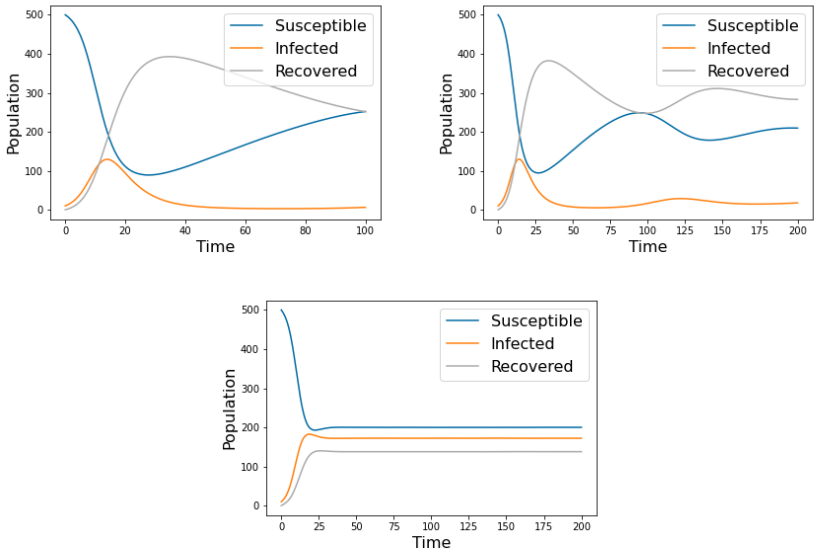

Figure 6.10: Comparison between three different disease examples where there is waning antigen so that recovered individuals can become susceptible. The parameters are: $k = 0.001$, $\gamma = .2$, $S_0 = 500$, $I_0 = 10$, $R_0 = 0$. The transition from recovered to susceptible is $\alpha = 0.01, 0.0125, 0.25$.)

sophisticated enough. The models are not the same, the behavior is also not the same so it seems hard to imagine our measure of disease propagation should be the same.

We can compare among different parameters and see important differences. In Figure 6.10, we show dynamics where the only parameter difference is the rate of antigen waning – that is how quickly recovered individuals enter the susceptible population. For low values, we see one isolated peak and as the rate of transition between recovered and susceptible increases we see repeated waves of infection and eventually a state where the infection never recedes and is endemic in the population. We will focus on the mathematical mechanisms for the waves of disease and consider how sensitivity analysis can provide insight into controlling secondary waves rather than digging deeper into the epidemiological aspect of R_0.

6.4.2 Analysis

How do we understand the differences in Figure 6.10? The first step is to return to our standard starting point: Steady-state analysis. Looking at Equations 6.19 and 6.20 to find steady-states, we can see that $I = 0$, $S = N$ and $R = 0$ is a steady-state. But there is another one with $S = \frac{\gamma}{k}$ and $I = \frac{\alpha(N-\frac{\gamma}{k})}{\alpha+\gamma}$, $R = N - \frac{\alpha(N-\frac{\gamma}{k})}{\alpha+\gamma}$. Linearization provides the reasons for the repeated waves as well as indication of the change between endemic and transient (See homework 6.8).

6.4.3 Sensitivity Analysis: Cobweb Diagrams

What can sensitivity analysis tell us about the behavior of the disease? Since the generic behavior – meaning the behavior that depends on a range of parameter values rather than one specific set of parameter values – is a decaying oscillation, it might be useful to know what controls the number of waves in a given time period. It also might be useful to know how the peaks of the waves depends on the parameters. This gives two QoIs that are easier to say in words than in equations. One way would be to determine the number of times the derivative of $I(t)$ is zero in a given time interval. Since this happens at the peaks and troughs it indicates the number of times the infectivity rate changes direction. We could also determine the value of I at these points to determine the amplitude of the oscillations. This type of QoI illustrates a key difference between sensitivity analysis that is useful in biological settings than other, more engineering based, settings. In engineering applications it is often more standard to have the dependent variables of a model be identified as the QoI. Often this is because engineering models are often developed with specific reliability issues in mind.

Oscillations are not simple to measure, numerically – especially when the shape of the curves can change. We will use a QoI that distinguishes between endemic and transient – namely, the value of the infected population is small or not (that is whether the infection is endemic or not). To indicate the sensitivity, we will use cobweb plots. The idea is to choose parameter sets where each parameter is perturbed from the basal state. The perturbations are randomly chosen and independent and are typically a percentage change from nominal. Once the parameter set is chosen, the solution is determined numerically. The

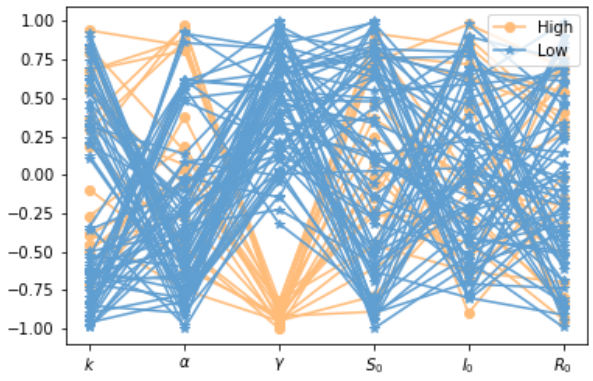

Figure 6.11: A variation of 100% indicating parameters that distinguish between high and low.)

QoI is sorted into bins. We are using two bins although others can be used. The parameters that lead to specific bins are connected visually. There are many different coloring options that help identify different aspects of the sensitivity. In Figure 6.11, we show a simple coloring identifying which parameters lead to high/low infected population.

There are a two parameters that appear to be well separated. When the infection tends to pass and not be endemic α is lower and γ is higher. There are other inferences that can be drawn from the shading – for example, γ is very certainly a parameter that moves the system between endemic and not while the initial conditions do not matter much at all.

We should also note that cobwebbing is the only global method that we have discussed in this chapter – the others are one-at-a-time. Global methods carry much more information, but are typically more complex to code up.

6.5 Caveats and State of the Art

We should make some mention about the scaling that we are using. There are two standard ways to implement the SIR-type models. These can be either dimensionalized or nondimensionalized – there are useful reasons for each, but it is important to understand the differences.

We can start with the dimensionalized version,

$$\frac{dS}{dt} = -kSI + \alpha R, \tag{6.21}$$

$$\frac{dI}{dt} = kSI - \gamma I, \tag{6.22}$$

$$\frac{dR}{dt} = \gamma I - \alpha R. \tag{6.23}$$

All the state (dependent) variables are measured in units of population (individuals, or appropriate units of individuals such as millions). We can measure the independent variable in units of time (for epidemics days is often appropriate). The units of the left-hand-sides must match the units on the right-hand-sides. Therefore

$$\left[\frac{dS}{dt}\right] = \frac{\text{population}}{\times\text{time}} = [kSI] = [k][S][I].$$

Therefore, $[k] = \frac{1}{\text{population time}}$ and k is a per capita rate. At the same time, α an γ are rates, $[*] = \frac{1}{\text{time}}$. This is important when using parameters relevant to specific diseases.

As before, the conservation law implies that the total population is N. This provides a scaling for the populations in terms of fractions of the total. Define $s(t) = \frac{S(t)}{N}$, $i(t) = \frac{I(t)}{N}$, an $r(t) = \frac{R(t)}{N}$. Then $\frac{dS}{dt} = N\frac{ds}{dt}$ and $[s] = 1$. We can re-write the equations in non-dimensional form,

$$\frac{ds}{dt} = -Nksi + \alpha r, \tag{6.24}$$

$$\frac{di}{dt} = Nksi - \gamma i, \tag{6.25}$$

$$\frac{dr}{dt} = \gamma i - \alpha r. \tag{6.26}$$

The parameter group, Nk has dimensions of $\frac{1}{\text{time}}$ since,

$$\begin{aligned}[Nk] &= \frac{[k]}{[N]}, \\ &= \frac{\text{population}}{\text{population} \times \text{time}}, \\ &= \frac{1}{\text{time}}.\end{aligned}$$

This means that the initial conditions for the two scalings must be consistent (non-dimensional scales are on the order of one, dimensional scale on the order of N). The parameters that involve nonlinear combinations of variables need to be scaled to either per capita or per time.

6.6 Problems

Problems 6.1 *Consider the SIRS model,*

$$\frac{dS}{dt} = -kSI + \alpha R, \tag{6.27}$$

$$\frac{dI}{dt} = kSI - \gamma I, \tag{6.28}$$

$$\frac{dR}{dt} = \gamma I - \alpha R. \tag{6.29}$$

(a) Find the equilibria

(b) Compare these with the two-variable version using the conservation of the population (Equations 6.16, 6.17 and $S+I+R=N$).

(c) Use any method to determine the stability of these equilibria

(d) Demonstrate the stability properties numerically using specific values of parameters that imply stability for the steady-states.

Problems 6.2 *(a) Show that the maximum of the two-compartment (SI) model is N*

(b) Demonstrate this using the numerical codes.

(c) What do you note about the approach to the equilibria?

Problems 6.3 *(a) Solve Equation 6.3 for $I(t)$*

(b) Write an expression for the analytic value of the half-saturation constant

(c) Use this to compare with the sensitivity method approximated numerically.

(d) Change the QoI to the value of the rate of change of the infected compartment evaluated at an increasing sequence of points, $(0,t_1,t_2,t_3,t_4)$. Compare the sensitivity plots for different times. What do you see as t_4 gets large?

Problems 6.4 *What do the sensitivity methods show for the sensitivity of the maximum of the two compartment model with respect to parameters? Are there differences in the methods?*

Problems 6.5 *Suppose the population is split into two different subsets: healthy and immune compromised.*

(a) Sketch a schematic where there are 4 compartments: Healthy susceptible, compromised susceptible, infected and recovered.

(b) Write a model that expands the susceptible compartment S in the SIR model to $S = S_h + S_c$.

(c) What happens if S_c *do not become infected but are removed from the population? Show this numerically.*

(d) Is this realistic? What happens to the total population? What would happen if there was a transition between S_h *and* S_c*?*

Problems 6.6 *Consider the SEIR model for these cases:*

- *Including a quarantine for sick people.*
- *Include vaccination*
- *Include age structured: The simplest way to do this is to consider SEIR models for some division of the population: Youth, Adults, Aged. Youth become Adults at a particular rate. We then have compartments for Susceptible, exposed, infected, recovered - youth (adults, aged). Just as susceptible can become exposed, youth can become adults at a specific rate.*

For all of these, provide representative simulations of the dynamics for some simple parameters. Consider the steady-state behavior. Can you find an analytic representation of the steady-states?

Problems 6.7 *Explore different methods to quantify the number of oscillations occur in a fixed interval. You might explore the command* `tspan[np.argmax(np.gradient(yp.y[1],tspan)<0)]` *as a starting place. Here* `np.gradient(yp.y[1],tspan)` *is the numerical derivative of* `yp_y[1]` *and* `np.argmax(a<0)` *provides the index of a that is the first occurrance of a negative value of a. The time this happens is in* `tspan`*.*

Problems 6.8 *Use sensitivity to determine the parameters that are mainly responsible for distinguishing between endemic and transient disease dynamics.*