Compartmental Models in Epidemiology

An application of differential equations to deterministic disease modeling

Hrishee Shastri

INTRODUCTION

Epidemiology studies the dynamics of diseases on populations. Quantitatively understanding these dynamics is difficult, so many scholars and public policy makers employ compartmental models to simplify the process. In a compartmental model, we divide the progression of a disease into a number of compartments. Individuals in a population move from one compartment to another based on a given transition rate between the two compartments. More specifically, at any point in time, an individual in a compartment can either stay in that same compartment or move to another compartment. We briefly gloss over the SIR (susceptible-infected-recovered) model in the section below to whet the reader's appetite. Then, we redirect our focus and dive deeper into the mathematics behind compartmental models by building and analyzing our own model for a made up disease: *hagoromoitis*.

THE SIR MODEL

The simplest compartmental model is the SIR model, an initialism for susceptible-infected-recovered. It comprises of three compartments : S, the population of those susceptible to the disease, I, the population of those infected with the disease, and R, the population of those who conferred immunity after being infected. The transition rates for the model usually depend on two quantities: the contact rate (average proportion of total population contacted per unit time by one individual) β and the recovery rate γ (on average, $\gamma I(t)$ people would recover per unit time). The system evolves over time, with $S(t), I(t), R(t)$ denoting the population at time t in each compartment. The total population, however, is constant (i.e. $S(t) + I(t) + R(t) = C$ for all $t \geq 0$). As a result, the population flows from susceptible to infected at a rate of $\beta S(t)I(t)$ (since on average each infected entity contacts and infects $\beta S(t)$ per unit time), and from infected to recovered at a rate of $\gamma I(t)$. This gives rise to the system of ordinary differential equations (ODEs)

$$\begin{aligned} S'(t) &= \frac{dS}{dt} = -\beta S(t)I(t) \\ I'(t) &= \frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t) \\ R'(t) &= \frac{dI}{dt} = \gamma I(t) \end{aligned}$$

with initial conditions $(S(0), I(0), R(0)) = (S_0, I_0, R_0)$ being the initial populations at time 0. See figure 1 for a diagram of the SIR model.

Despite the system's nonlinear nature, an exact analytic solution was found by Harko et al. in 2014. One could write this entire paper on the SIR model itself due to it being such a popular object of study. Instead, we build off of it and dive deeper.

CONSTRUCTING COMPARTMENTAL MODELS - HAGOROMOITIS

Let us examine the population of Reed's math department under a strange contagion: hagoromoitis. hagoromoitis is an affliction that renders the patient incapable of using any brand of chalk other than hagoromo. Clinical studies have identified four compartments in the progression of hagoromoitis (see fig. 2):

- **O**: where a patient is oblivious to hagoromo chalk (and is thus susceptible)
- **T**: where a patient is in the process of trying hagoromo chalk (and is infected)

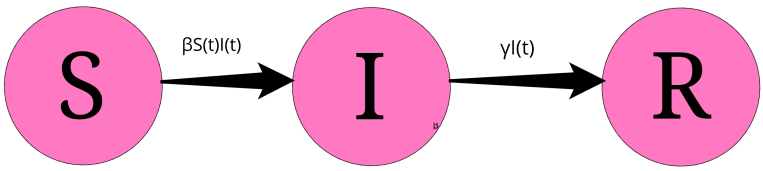

Figure 1. The SIR model

- **U**: where a patient is no longer attached to hagoromo chalk (and is thus immune)
- **E**: where a patient has become obsessed with hagoromo chalk (and is also infected)

We make some key assumptions for this simple model: the total population of the math department over time is constant (N) and the affliction is transmitted to someone in O through close contact with someone in compartments T or E. Further assume that an individual in T can develop more serious symptoms and move to compartment E (T $\to$ E). We also assume that someone in compartments T or E can eventually become immune to the affliction, since either they decided hagoromo chalk was not for them (E $\to$ U) or that its novelty wore off after being obsessed with it (T $\to$ U).

We denote $O(t), T(t), U(t), E(t)$ to be the number of people in compartments O, T, U, and E at time t, respectively. Since our total population is N for all t, note that $O(t) + T(t) + U(t) + E(t) = N$ for all t as well. Further assume that the average person comes into contact with βN others per unit time (β is the contact rate, discussed in the SIR model). Let γ be the rate at which people embrace hagoromo after trying it out (i.e. people trying the brand embrace it on average at a rate of $\gamma T(t)$ per unit time), and let δ be the rate at which people disregard the brand after trying it out (i.e. people trying the brand disregard it at a rate of $\delta T(t)$ per unit time). Finally, let ϵ be the rate at which the chalk's novelty wears off for those who were obsessed with it (i.e. people obsessed with it move on at a rate of $\epsilon E(t)$ per unit time).

Now we derive transition rates. The average person in O contacts βN people per unit time, and since the probability of contacting an afflicted is $\frac{T(t)+E(t)}{N}$, the rate of new afflictions (and consequently, the transition rate between O and T) is $O(t)\beta N\frac{T(t)+E(t)}{N} = \beta O(t)\left(T(t) + E(t)\right)$. Between T and U, the transition rate is simply $\delta T(t)$; between T and E, the transition rate is simply $\gamma T(t)$; and between E and U, the transition rate is simply $\epsilon E(t)$. See figure 2 for a diagram of our compartmental model.

Using the transition rates, we can derive a system of ODEs that describes the dynamics of hagoromoitis. In general, each differential equation is formulated by summing the transition rates incident to that compartment while subtracting the transition rates exiting the compartment. This matches our intuition about the flow of people in and out of a compartment, or in other words, the population change over time through that compartment. With-

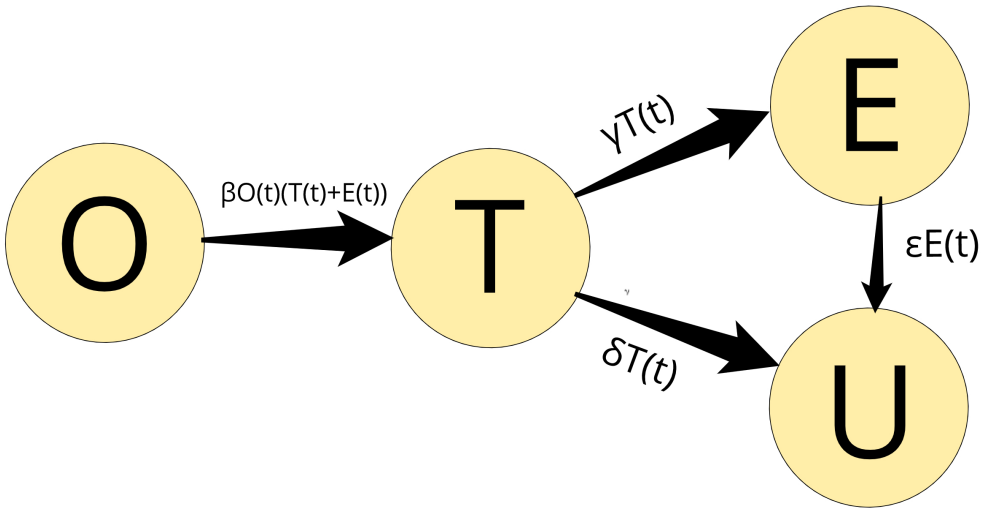

Figure 2. Compartmental model for hagoromoitis

out further ado, we have

$$\begin{aligned}
O'(t) &= \frac{dO(t)}{dt} = -\beta O(t)\left(T(t)+E(t)\right) \\
T'(t) &= \frac{dT(t)}{dt} = \beta O(t)\left(T(t)+E(t)\right) - \gamma T(t) - \delta T(t) \\
E'(t) &= \frac{dE(t)}{dt} = \gamma T(t) - \epsilon E(t) \\
U'(t) &= \frac{dU(t)}{dt} = \delta T(t) + \epsilon E(t)
\end{aligned}$$

with initial conditions $O(0) = O_0, T(0) = T_0, E(0) = E_0, U(0) = U_0$. The fact that this system is nonlinear makes it difficult to solve analytically. Nevertheless, we can examine the behavior of this system in other ways as well.

As as an aside, this model can be thought of as a generalization of some other models. Namely, if $\delta = 0$, we have the SEIR (susceptible→exposed→infected→recovered) model where both exposed and infected compartments are infectious. If $\epsilon = 0$, we have the carrier state model, where compartment E is a carrier state for the permanently infectious. And if $\gamma = \epsilon = 0$, we have the familiar SIR model (assuming $E_0 = 0$ as well).

BASIC REPRODUCTION NUMBER

Often when there are rumors of an epidemic, the term *basic reproduction number*, or the R_0 value, is tossed around. Qualitatively, the R_0 value for a disease is the expected number of infections created by a single infectious entity in a fully susceptible population. The value itself is model dependent and may be derived depending on the assumptions made. In terms of our running hagoromoitis example, the R_0 value would capture the impact a single person from compartment T or E (as those are our infectious compartments) would have on a wholly oblivious population.

One of the most important and basic principles derived from mathematical epidemiology is that the system of ODEs describing a disease exhibits threshold behavior. In other words, if $R_0 < 1$, the disease will likely die out, but if $R_0 > 1$ it will likely escalate into an epidemic.

DERIVING R_0 USING THE NEXT GENERATION MATRIX

The R_0 value can be easy to derive for simple compartmental models. For example, in the SIR model, R_0 can be found through simple reasoning: since the average time until an infectious entity recovers is $\frac{1}{\gamma}$ and the average

time between contacts is $\frac{1}{\beta}$, it follows that $R_0 = \frac{\frac{1}{\gamma}}{\frac{1}{\beta}} = \frac{\beta}{\gamma}$. For more complex models, such as hagoromoitis, this is not as easy – there are two infectious compartments, T and E, each with different characteristics.

The next generation matrix method

Diekmann et al. (1990) and van den Driessche and Watmough (2002) formulated the *next generation matrix* method to compute R_0 for a model with n compartments and $m, m < n$ infectious compartments. First, we establish some notation.

- Let x_i, $i \in \{1, ..., m\}$ be the population of the i^{th} infected compartment at time t.
- Let $F_i : \mathbb{R}^n \to \mathbb{R}$ be the rate of transfer of new infections into compartment i.
- Let $V_i : \mathbb{R}^n \to \mathbb{R}$ be defined by $V_i(x) = V_i^-(x) - V_i^+(x)$, where $V_i^+(x)$ is the rate of transfer of old infections into compartment i and $V_i^-(x)$ is the rate of transfer of entities out of compartment i.

Then we have, similar to how we derived our system of ODEs,

$$\frac{dx_i}{dt} = F_i(x) - V_i(x) \quad \forall i \in \{1, ..., m\}$$

or

$$\frac{dx}{dt} = F(x) - V(x)$$

where $\frac{dx}{dt} = [\frac{dx_1}{dt}, ..., \frac{dx_m}{dt}]^T$, $F(x) = [F_1(x), ..., F_m(x)]^T$, and $V(x) = [V_1(x), ..., V_m(x)]^T$. Next, Jacobian matrices of $F(x)$ and $V(x)$, $DF(x) = \left[\frac{\partial F_i(x)}{\partial x_j}\right]$ and $DV(x) = \left[\frac{\partial V_i(x)}{\partial x_j}\right]$ respectively, are evaluated at the *disease free equilibrium* point, x_0. The disease free equilibrium point is the steady state solution of the compartmental model for which there are no infections (the next section discusses this in greater detail). For example, in the SIR model, the disease free equilibrium would be $x_0 = (S*, I*, R*) = (N, 0, 0)$ where $*$ denotes the equilibrium solution and N is the total population.

The next generation matrix is defined as

$$DF(x_0) \cdot DV(x_0)^{-1}$$

where $\cdot$ denotes standard matrix multiplication. R_0 for the compartmental model is precisely the largest eigenvalue of this matrix.

We can now derive R_0 for hagoromoitis. For our $m = 2$ infectious compartments, we have

$$\begin{aligned}
\frac{dT(t)}{dt} &= F_T(x) - V_T(x) \\
&= \beta O(t)(T(t) + E(t)) - ((\gamma + \delta)T(t) - 0) \\
\frac{dE(t)}{dt} &= F_E(x) - V_E(x) \\
&= 0 - (\epsilon E(t) - \gamma T(t))
\end{aligned}$$

and so the Jacobians are

$$DF(x) = \begin{bmatrix} \beta O(t) & \beta O(t) \\ 0 & 0 \end{bmatrix} \qquad DV(x) = \begin{bmatrix} \gamma + \delta & 0 \\ -\gamma & \epsilon \end{bmatrix}$$

$$DV(x)^{-1} = \begin{bmatrix} \frac{1}{\delta+\gamma} & 0 \\ \frac{\gamma}{\epsilon(\delta+\gamma)} & \frac{1}{\epsilon} \end{bmatrix}$$

The disease free equilibrium where the entire population is susceptible is $(O*, T*, E*, U*) = (N, 0, 0, 0)$, so the next generation matrix is

$$DF(x_0) \cdot DV(x_0)^{-1} = \begin{bmatrix} \beta N & \beta N \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\delta+\gamma} & 0 \\ \frac{\gamma}{\epsilon(\delta+\gamma)} & \frac{1}{\epsilon} \end{bmatrix} = \begin{bmatrix} \frac{\epsilon\beta N+\beta N\gamma}{\epsilon(\gamma+\delta)} & \frac{\beta N}{\epsilon} \\ 0 & 0 \end{bmatrix}$$

This matrix has eigenvalues 0 and $\frac{\epsilon\beta N+\beta N\gamma}{\epsilon(\gamma+\delta)}$, and since all the rates must be positive, it follows that the maximal eigenvalue is $\frac{\epsilon\beta N+\beta N\gamma}{\epsilon(\gamma+\delta)}$. By definition this is the R_0 value, but usually we divide it by the total population N to gain a better picture of how the affliction would behave irrespective of initial population conditions. Thus, $R_0 = \frac{\beta(\epsilon+\gamma)}{\epsilon(\gamma+\delta)}$ for hagoromoitis.

Considerations for the next generation matrix method

The next generation matrix method may look like a panacea for computing R_0. However, there are instances of compartmental models for which the above method fails. The most glaring example is when $DV(x_0)$ is not invertible, which often occurs when a model allows a compartment where individuals perpetually stay infectious. Another example is if the model allows external infectious populations to infect susceptibles within the model (e.g. bats in a human only model). In any case, there are numerical (simulation based) approaches to compute an R_0-like metric, if not R_0 itself.

EQUILIBRIUM POINTS

The equilibrium points of a model can tell us useful information about its behavior. Specifically, they tell us the points at which the overall flow in the model is such that the populations in each compartment are constant over time. For hagoromoitis, we can compute the equilibrium points $x_0 = (O*, T*, E*, U*)$ such that $f(x_0) = 0$, where

$$f(x) = \begin{bmatrix} -\beta O(T+E) \\ \beta O(T+E) - \gamma T - \delta T \\ \gamma T - \epsilon E \\ \delta T + \epsilon E \end{bmatrix} = 0$$

where for simplicity we write O, T, E, U for $O(t), T(t), E(t), U(t)$. Recall N is the total population, which is constant for all t. Therefore, from inspection, we see that the equilibrium points are

$$x_0 \in \{(N-i, 0, 0, i) \mid 0 \leq i \leq N\}.$$

Epidemiologically this makes sense – the equilibrium points stipulate that the infectious compartments should be empty. It would take just one infectious entity to destabilize the model, either by infecting susceptibles or by gaining immunity. These types of equilibrium points are *disease free equilibrium points*, as mentioned earlier when we computed R_0.

Endemic equilibrium points

It should be mentioned that there is another type of equilibrium point in epidemiology: *endemic equilibrium points*. These equilibrium points correspond to the state where the population in each compartment stays constant while there is at least one infectious compartment with a population greater than zero. Neither hagoramoitis nor the basic SIR model has these equilibrium points, because they do not replenish their supply of susceptibles (i.e. the infectious run out of entities to infect).

If we slightly modify our model for hagoromoitis so that the rate from going from U to O is α, we can derive the endemic equilibrium points. We can interpret this as non-attachment to hagoromo chalk gradually fading into forgetting it ever existed at all, essentially rendering the patient susceptible to hagoromoitis again (i.e. immunity is not permanent). Then we compute the equilibrium points by solving

$$f(x) = \begin{bmatrix} -\beta O(T+E) + \alpha U \\ \beta O(T+E) - \gamma T - \delta T \\ \gamma T - \epsilon E \\ \delta T + \epsilon E - \alpha U \end{bmatrix} = 0$$

Thankfully, this nonlinear system of equations can be reduced to a linear system of three equations by making the observation that the nonlinear term $\beta O(T+E) = \alpha U$. Solving for the equilibrium point x_0 gives us

$$x_0 = \left(\frac{\epsilon(\delta+\gamma)}{\beta(\gamma+\epsilon)}, \frac{\epsilon}{\gamma}E^*, E^*, \frac{\epsilon(\delta+\gamma)}{\alpha\gamma}E^* \right) \quad \text{such that} \sum x_0 = N$$

The $\sum x_0 = N$ condition can be used to solve for E^* directly. If the initial values $(O_0, T_0, E_0, U_0) = x_0$, the population across compartments would theoretically be constant for all time t, meaning hagoromoitis would perpetually circulate around Reed's math department.

In fact, remember that our R_0 value for the permanent-immunity hagoromoitis was found to be $R_0 = \frac{\beta(\gamma+\epsilon)}{\epsilon(\delta+\gamma)}$ – this is the same R_0 value for this temporary-immunity version since our infectious compartments did not change. As such, O^* in x_0 is actually $\frac{1}{R_0}$. This is no surprise, as the stability of equilibrium points generally depend on R_0. More specifically, whether R_0 is above or below some threshold value determines the stability of the equilibrium points. Finding this threshold value can be difficult.

Classifying equilibrium points

Classifying the equilibrium points in terms of being stable or unstable is also useful, as it tells us the expected behavior of the system when starting at a point close to the equilibrium points. However, due to the nonlinearity of these systems, we are generally concerned with local stability. If an endemic or disease free equilibrium point is locally stable, then a system starting in the neighborhood of those points will eventually reach that equilibrium and remain there for all time t. If they are locally unstable, then a system starting in its neighborhood will diverge from the equilibrium point.

For compartmental models, the system of ODEs are nonlinear and so many practitioners classify equilibrium points by finding the eigenvalues of the Jacobian of the system evaluated at each equilibrium point (i.e. linearizing the system). Especially with models that are symbolic in nature, such as our hagoromoitis model, this proves to be a difficult task. However, if it were able to be done, then we could give conditions on $\alpha, \beta, \gamma, \delta, \epsilon$ or R_0 for which the equilibrium points are locally stable/unstable.

AGENT BASED SIMULATION

To conclude our investigation, we design and implement an interactive agent based simulation for our hagoromoitis model. The source code repository is available here and the applet here. The user can specify $O_0, T_0, E_0, U_0, \alpha, \beta, \gamma, \delta, \epsilon$ and watch the simulation unfold on Reed's front lawn. The graph below the simulation concurrently plots the population in each compartment after every 100 timesteps (each timestep is 10 milliseconds), effectively approximating $O(t), T(t), E(t), U(t)$. While the whole design and implementation is beyond the scope of this paper, I would like to briefly discuss a few salient details.

Agent AI

Each agent is represented by a small circle of fixed radius, colored based on which compartment it is currently in. The goal is to have each agent mimic natural movement to an acceptable degree. There are many solutions to this problem, each at different levels of sophistication, and ours is relatively simple. Each agent chooses a random point within a larger fixed radius and travels to that point, repeating ad infinitum. The speed of movement is also inversely proportional to the total population (i.e. more agents mean slower movement and fewer agents mean faster movement).

Implementing the contact rate

Recall that the contact rate, β, is the average proportion of the total population that an individual contacts per unit time. There is no reasonable way to force an agent to contact a fixed number of other agents per unit time, especially when each agent is wandering randomly throughout the plane. What we do instead is introduce a contact radius r_c. In other words, if any agent is a distance of r_c away from another agent, the two are said to be in contact with each other (and thus an infection can be transmitted between the two). Letting w, h be the width and height of the simulation window, we can equate β as the ratio between the area of the contact circle and the area of the simulation window, obtaining an expression for r_c:

$$\beta = \frac{\pi r_c^2}{wh}$$
$$r_c = \sqrt{\frac{\beta wh}{\pi}}.$$

r_c was implemented with this formula in mind.

Implementing transition rates

The transitions between compartment O and T are taken care of by the contact radius r_c. That leaves the transitions between all other compartments, whose rates are of the form ρC where $\rho \in \{\alpha, \gamma, \delta, \epsilon\}$ and C is the compartment (either T, E, U). We borrow a result from continuous-time Markov chains, which lets us interpret these rates as exponentially distributed wait times - for example, if an agent is in compartment E, it waits for X units of time until it transitions to compartment U, where X is an exponentially distributed random variable with rate ϵ (i.e. $X \sim Exp(\epsilon)$). For the case of an agent in compartment T, it waits for $min\{X, Y\}$ units of time, where $X \sim Exp(\gamma)$ and $Y \sim Exp(\delta)$. And so if $min\{X, Y\} = X$, the agent moves to compartment E, otherwise it moves to compartment U. Implementing the transitions in this manner hopefully makes our simulation more realistic than if we had a fixed wait time between compartments.

EXTRAPOLATING TO THE REAL WORLD

Barring the fact that hagoromoitis does not exist in the real world (as far as I know), the old adage "all models are wrong, but some are useful" hopefully rings true here. Compartmental models are no doubt powerful tools to simplify the disease modelling process, but the conclusions they make should only supplement decision making. We could have added more and more assumptions to our model to make it more realistic, such as introducting births and deaths, accounting for age, vaccinations, etc. However, the more assumptions added, the more complex the model becomes and the harder it is to analyze.

Modeling diseases with differential equations come with trade offs. Our system of ODEs is inherently deterministic, which is nowhere near accurate when modeling real world systems. Stochastic models, which take into account the inherent diversity and variability in populations, may be more useful. However, they are mathematically harder to analyze when compared to a system of ODEs. I will leave this topic for another day.

REFERENCES

1. Brauer F. (2008) *Compartmental Models in Epidemiology*

2. Diekmann O, Heesterbeek J A, Roberts MG. (2002). *The construction of next-generation matrices for compartmental epidemic models*

3. Blackwood J, Childs L. (2018) *An introduction to compartmental modeling for the budding infectious disease modeler*