

# Lecture 21: Compartment Models

36-467/36-667

15 November 2018

## Abstract

“Compartment models”, which track the transitions of large numbers of individuals across a series of qualitatively-distinct states (the “compartments”) are common in fields which study large populations, such as epidemiology and sociology. These notes cover such models as a particular kind of Markov chain, including some examples from epidemiology, evolutionary economics, and demography.

I made some mistakes in lecture; these notes correct them.

Epidemiology, ecology, sociology, demography, chemical kinetics, and some other fields are sometimes grouped together as “population sciences”, which study the behavior and evolution of large populations of individuals. All of these fields make extensive use of what are called (among other things) “compartment models”, where the members of the population are divided into a discrete set of qualitatively-distinct states (or “compartments”), and transition between them at rates which vary as functions of the current distribution over compartments<sup>1</sup>. They can be used whenever a population is divided into discrete, qualitatively-distinct types or classes, and transitions between classes are stochastic.

## 1 Compartment Models in General

### 1.1 Basic Definitions and Assumptions

Let’s begin with the case of a “closed” population of constant size  $n$ . (§5 will consider the modifications needed to make the population size vary over time.) This population is divided into  $r \geq 2$  qualitatively distinct classes, types or compartments<sup>2</sup>, and we are primarily interested in how many individuals are located in each compartment. We thus define the state of the population at time  $t$ ,  $X(t)$ , as a vector,

$$X(t) \equiv [X_1(t), X_2(t), \dots, X_r(t)] \quad (1)$$

where  $X_i(t)$  counts how many individuals are in compartment  $i$  at time  $t$ .

---

<sup>1</sup>Whether the ubiquity of these models reflects some underlying commonality across these fields, or just a tendency for scientists with similar mathematical training to borrow and/or re-invent similar models, is beyond the scope of this note.

<sup>2</sup>These words are all synonyms in the present context.

The state  $X(t)$  changes because individuals can move from one compartment to another. We assume that these moves or transitions follow a stochastic process, and we make two important assumptions about that process.

1. The  $i$ -to- $j$  **transition rate**  $p_{ij}$  is the probability that an individual in compartment  $i$  will move to compartment  $j$  at the next time step. ( $p_{ii}$  is thus the probability of staying in compartment  $i$ .) We assume that the transition rates are functions of the *current* population state, and not the past (or the future), so  $p_{ij}(X)$  is a well-defined function.

Note that this assumption allows some  $p_{ij}$  to be zero, either always or for certain states of the population. (These are sometimes called **forbidden** or **suppressed** transitions.) It also allows some  $p_{ij}$  to be constant, regardless of  $X$ .

2. All individuals in the population make *independent* transitions, conditional on  $X(t)$ . The moves of all the individuals in a given compartment are thus *conditionally* independent and identically distributed.

These assumptions are *not* implied by the mere fact that we've divided the population into compartments and are tracking the number of individuals in each compartment. They could be false, or true only for a different choice of compartments, etc.

— You may feel some scruples about applying these assumptions to people, at least to social processes which reflect human actions (as opposed to, say, models of disease spread). But modeling aggregated social processes as stochastic in this way does not commit us to saying that people act randomly, or don't have motives for their actions, or that all individuals in the same compartment are identical. What it does commit us to is that the differences between people in the same compartment which determine their motives and their actions are so various, so complicated, and so independent of the circumstances which landed them in that compartment that they can't be predicted, and can be effectively treated as random<sup>3</sup>.

## 1.2 Consequences of the Assumptions

These assumptions are enough to tell us that the sequence of random variables  $X(t)$  is a Markov chain. If it's a Markov process, it has to be a Markov *chain* because the state space is discrete, though large (Exercise 1), so what we really need to do is show that the Markov property holds. The easiest way to get there is through defining some additional variables.

---

<sup>3</sup>There are precise technical senses in which any sufficiently complicated function is indistinguishable from randomness. For a brilliant and accessible introduction to these ideas, and how they underlie probability modeling, I strongly recommend Ruelle (1991). The standard textbook on this subject is Li and Vitányi (1997). Salmon (1984, ch. 3) applies these ideas to the problem of defining "the" probability of an outcome for members of some class.

### 1.2.1 Flow or Flux Variables

For each pair of compartments  $i$  and  $j$ , define  $Y_{ij}(t+1)$  as the number of individuals who moved from compartment  $i$  to compartment  $j$  at time  $t+1$ . This is sometimes called the **flow** or **flux** from  $i$  to  $j$ .

$Y_{ij}(t+1)$  is a random quantity. By the assumptions above, in fact, it has a binomial distribution:

$$Y_{ij}(t+1) \sim \text{Binom}(X_i(t), p_{ij}(X(t))) \quad (2)$$

This implies that

$$\mathbb{E}[Y_{ij}(t+1)|X(t)] = X_i(t)p_{ij}(X(t)) \quad (3)$$

$$\text{Var}[Y_{ij}(t+1)|X(t)] = X_i(t)p_{ij}(X(t))(1-p_{ij}(X(t))) \quad (4)$$

There are two implications of these worth noting.

1. Both the expected flux and the variance in the flux are proportional to  $X_i(t)$ . This means that the standard deviation in the flux is only proportional to  $\sqrt{X_i}$ . Thus, if  $X_i(t)$  is large, the *relative* deviations in the flux are shrinking. This suggests that there should be nearly-deterministic behavior in large populations.
2. The assumption that each individual in compartment  $i$  has the *same* probability of moving to compartment  $j$  is crucial to calculating the variance. If there were randomly-distributed transition rates for all individuals in the compartment, we'd see a higher variance for the flux (Exercise 2).

If  $r > 2$ , then the set of fluxes out of  $i$ ,  $[Y_{i1}(t+1), Y_{i2}(t+1), \dots, Y_{ir}(t+1)]$  has as its joint distribution a multinomial, with the number of "trials" being  $X_i(t)$ , and the probabilities being  $(p_{i1}, p_{i2}, \dots, p_{ir})$ . (Remember that the marginal distributions of a multinomial distribution are binomial.)

### 1.2.2 Stock-Flow Consistency

The variables  $X_i(t)$  count the total number of individuals in each compartment, regardless of how or when they arrived there. Such variables are sometimes called **stock** or **level** variables. The  $Y_{ij}(t)$  variables, on the other hand, are (as already said) **flow** or **flux** variables. A basic principle is that the new level of the stock is the old level, plus flows in, minus flows out:

$$X_i(t+1) = X_i(t) + \sum_{j \neq i} Y_{ij}(t+1) - Y_{ji}(t+1) \quad (5)$$

$$\text{(new level)} = \text{(old level)} + \sum \text{(in flow)} - \sum \text{(out flow)} \quad (6)$$

This is sometimes called **stock-flow consistency**.

### 1.2.3 Markov Property

Because Eq. 5 says how the level in each compartment changes, it tells us how the state vector  $X(t)$  changes as well.

Notice that, from our assumptions,  $Y_{ij}(t+1)$  is independent of the past given  $X(t)$ . It is independent of earlier values of  $X$ , and also independent of earlier values of the  $Y$ 's. Since Eq. 5 tells us that  $X(t+1)$  is a deterministic function of  $X(t)$  and  $Y(t+1)$ , it follows that  $X(t+1)$  is independent of earlier  $X$ 's (and  $Y$ 's) given  $X(t)$ . Hence the state of the population follows a Markov chain.

Because the state space, while large, is finite, the Markov chain will eventually enter some closed, irreducible set, and stay there forever. In the long run, the distribution of  $X(t)$  will follow the invariant distribution for that irreducible set. These are properties of any finite Markov chain, and so apply to this one, even though it has special structure.

## 2 Epidemic Models

In epidemiology, compartment models are often known as “epidemic models”, and in that field they descend from pioneering work in mathematical epidemiology done in the early 20th century<sup>4</sup>, but they are applied much more broadly than just to epidemics. It's worth thinking through some of the most basic cases, to get a feel for how they work.

### 2.1 SI

Let's begin with a model with just two compartments, dividing the entire population into those who are infected with a certain disease — call this compartment  $I$  — and those who, not having caught the disease yet, are susceptible, in compartment  $S$ . The state of the population is  $X(t) = (X_I(t), X_S(t))$ , but since  $X_S(t) = n - X_I(t)$ , it is enough to keep track of  $X_I(t)$ .

With two compartments, we need to specify two rates,  $p_{SI}$  and  $p_{IS}$ . The former is the rate at which susceptible individuals become infected; the latter is the rate at which infected individuals are cured and become susceptible again.

A simple assumption is that  $p_{SI}(X_S, X_I) = \alpha X_I$ , because the more members of the population are infected, the more likely a susceptible individual is to encounter one and become infected. We might, however, simply set  $p_{IS} = 0$ , indicating that no one recovers from the disease, but rather the infected remain infected, and contagious, forever. There are a couple of ways to justify this latter assumption.

1. We're in a zombie movie. (This *is* Pittsburgh, after all.)
2. We're only interested in what happens on a comparatively short time scale. More precisely, whatever processes make infected individuals non-infectious (recovery, death, quarantine, etc.) happen much more slowly than the process

---

<sup>4</sup>It is conventional these days to attribute this to the work of Sir Ronald Ross on malaria, but early reviews (e.g., Lotka 1924, ch. VIII, pp. 79–82) make it clear that this was much more of a collective effort.

of contagion itself, and we're only trying to approximate the beginning of the epidemic, before the former processes can kick in.

3. The infected really do remain around, and contagious, forever (§2.1.1.)

Leaving aside how to justify the assumption that  $p_{SI} = 0$ , what are the consequences of these assumptions about the transition rates?

An obvious consequence is that if  $X_I(t) = n$  or  $X_I(t) = 0$ , then  $X_I$  will remain  $= n$  or  $= 0$  forever. That is, these are **absorbing** states of the population. You can convince yourself that these are the only absorbing states. You can also convince yourself that an absorbing state forms a closed, irreducible set of states, and that these are the only closed, irreducible sets for this chain (Exercise 3). There are thus two possible invariant distributions, one which gives probability 1 to  $X_I = 0$  and the other which gives probability 1 to  $X_I = n$ . The disease thus either dies out, or takes over the whole population. Finally, you can convince yourself that  $X_I(t)$  never decreases, so the disease must take over the population if it's ever present at all.

To see what happens on the way there, we'll need to use the flow variables. There is only one flux that we need to worry about,  $Y_{IS}$ :

$$X_I(t+1) = X_I(t) + Y_{IS}(t+1) \tag{7}$$

$$X_S(t+1) = X_S(t) - Y_{IS}(t+1) \tag{8}$$

Since  $Y_{IS}$  is binomially distributed, we can work out the expected change in the number of infected individuals.

$$\mathbb{E}[X_I(t+1)|X_I(t)] - X_I(t) = X_S(t)\alpha X_I(t) \tag{9}$$

$$= \alpha X_I(t)(n - X_I(t)) \tag{10}$$

Notice that this growth is very small when  $X_I \approx 0$ , and again when  $X_I \approx n$ . There is an intermediate range where the growth is very rapid.

Figure 1 plots a typical expected course of  $X_I(t)$ . This is what's known as **logistic** or **sigmoidal**<sup>5</sup> growth. This can be caricatured as “nothing seems to happen for a long time, then suddenly it explodes and grows exponentially, but eventually it levels off and saturates”.

Of course, Figure 1 is comes from iterating the *deterministic* equation 10. What happens if we actually follow the stochastic process? Figure 2 plots 100 realizations of the SI model, along with the deterministic approximation, and shows that while the *shape* is the same, the timing of the exponential take-off and of the eventual saturation becomes variable.

### 2.1.1 Diffusion of Innovations

A disease which, once caught, never goes away and remains infectious forever is a bit weird, biologically speaking. But the same ideas are also used to model the spread of new technologies and ideas, what is sometimes called the “diffusion of innovations”,

<sup>5</sup>Because somebody thought it that curve looked like an  $S$ , or perhaps the outline of  $\sigma$ .

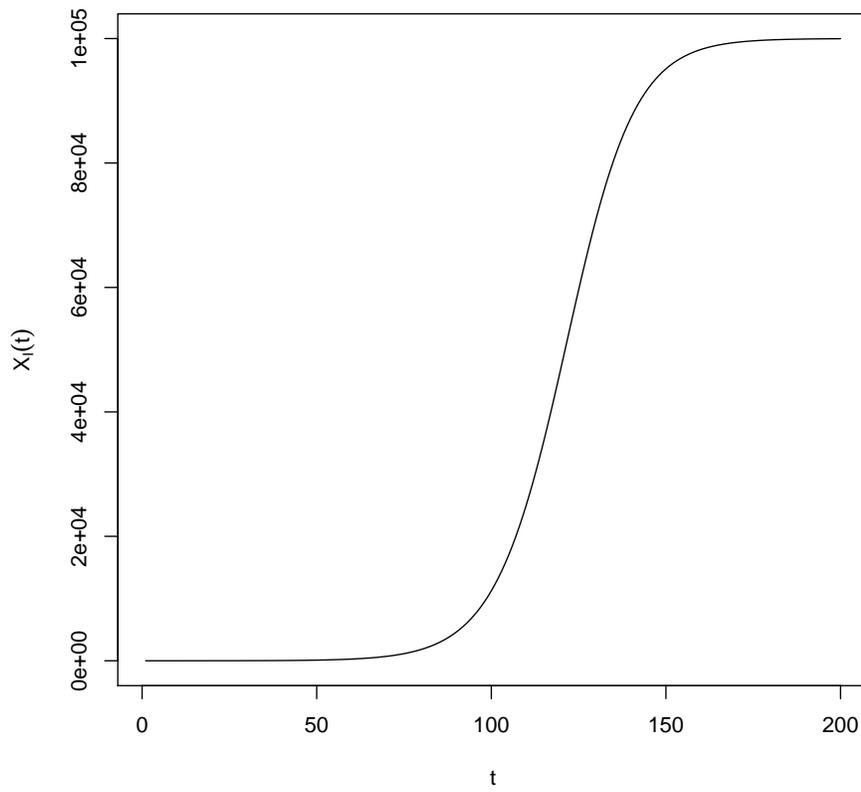


Figure 1: Illustration of the kind of behavior that results from Eq. 10. This is shown with  $n = 10^5$ ,  $\alpha = 10^{-6}$ , and  $X_I(1) = 1$ .

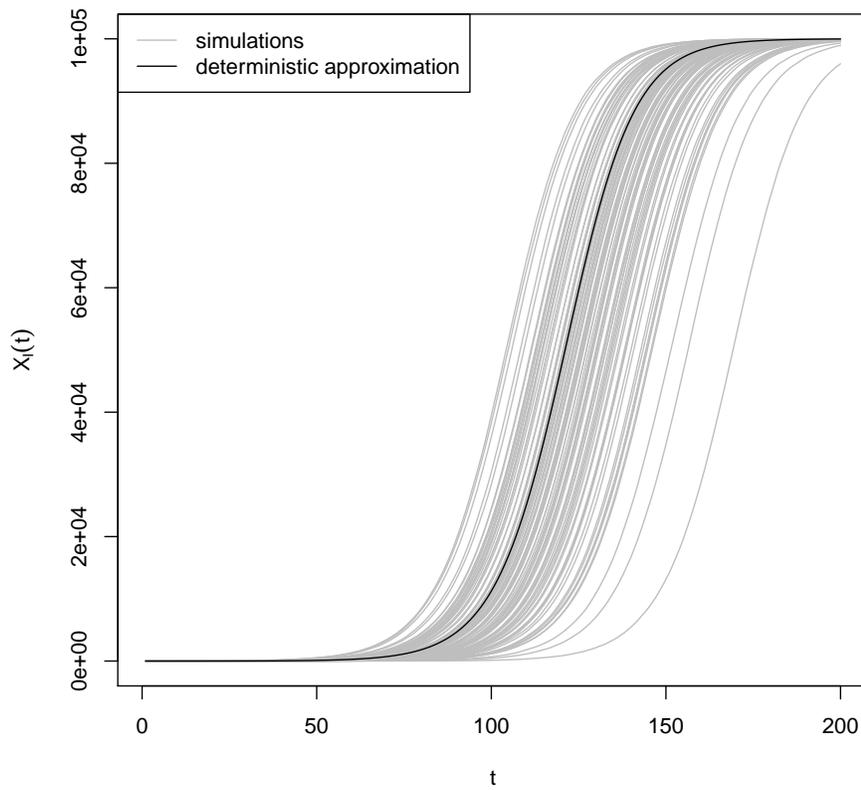


Figure 2: Comparison of stochastic simulations of the SI model (grey lines) to the deterministic approximation (black line). Parameters are as in Figure 1.

and then it makes sense to say that those who have “caught” an idea remain infected with it, and potentially contagious, for the rest of their lives.

This week’s homework explores this idea in the context of one of the classic datasets on the diffusion of innovations, about the spread of the use of a new antibiotic among doctors in four towns in Illinois in the 1950s. The same notions, however, have been used to model the adoption of a wide range of technologies, practices, art forms, and ideologies (Rogers, 2003).

It’s worth noting that the idea of analogizing the spread of ideas or behavior to the spread of disease is a very old one, but it’s usually been an analogy deployed when people *disapprove* of what is spreading. (It is very rare to talk about “infectious virtue”.) The oldest example I have found is the Roman writer and politician Pliny the Younger calling Christianity a “contagious superstition” (*superstitionis contagio*) in a letter to the Emperor Trajan in +110 (*Epistles* X 96.9), but the analogy has spontaneously re-occurred many, many times since then. It was also the basis for more systematic and scholarly studies, such as Siegfried (1960/1965), long before Dawkins (1976) coined the word “meme”.

## 2.2 SIS

A logical next step is to keep two compartments, but to allow for movement back and forth between them. That is, some infected individuals can recover and become susceptible again. Again, a simple assumption is that infected individuals recover at a constant rate,  $p_{SI} = \rho$ , while we keep  $p_{IS} = \alpha X_I$ . Because this allows for movement from each compartment to the other, we need to keep track of *two* fluxes:

$$Y_{SI}(t+1) \sim \text{Binom}(X_S(t), \alpha X_I(t)) \quad (11)$$

$$Y_{IS}(t+1) \sim \text{Binom}(X_I(t), \rho) \quad (12)$$

Now

$$X_I(t+1) = X_I(t) + Y_{SI}(t+1) - Y_{IS}(t+1) \quad (13)$$

$$X_S(t+1) = X_S(t) - Y_{SI}(t+1) + Y_{IS}(t+1) \quad (14)$$

To understand the dynamics here, start by considering what would need to happen for the expected change in  $X_I$  to be zero.

$$\mathbb{E}[X_I(t+1)|X_I(t)] - X_I(t) = \alpha X_I(t)(n - X_I(t)) - \rho X_I(t) \quad (15)$$

Let’s set this equal to 0 at  $X_I(t) = X_I^*$ . There are two solutions:

$$X_I^*(\alpha(n - X_I^*) - \rho) = 0 \quad (16)$$

$$X_I^* = 0 \text{ or} \quad (17)$$

$$\frac{\rho}{\alpha} + X_I^* = n \quad (18)$$

$$X_I^* = n - \frac{\rho}{\alpha} \quad (19)$$

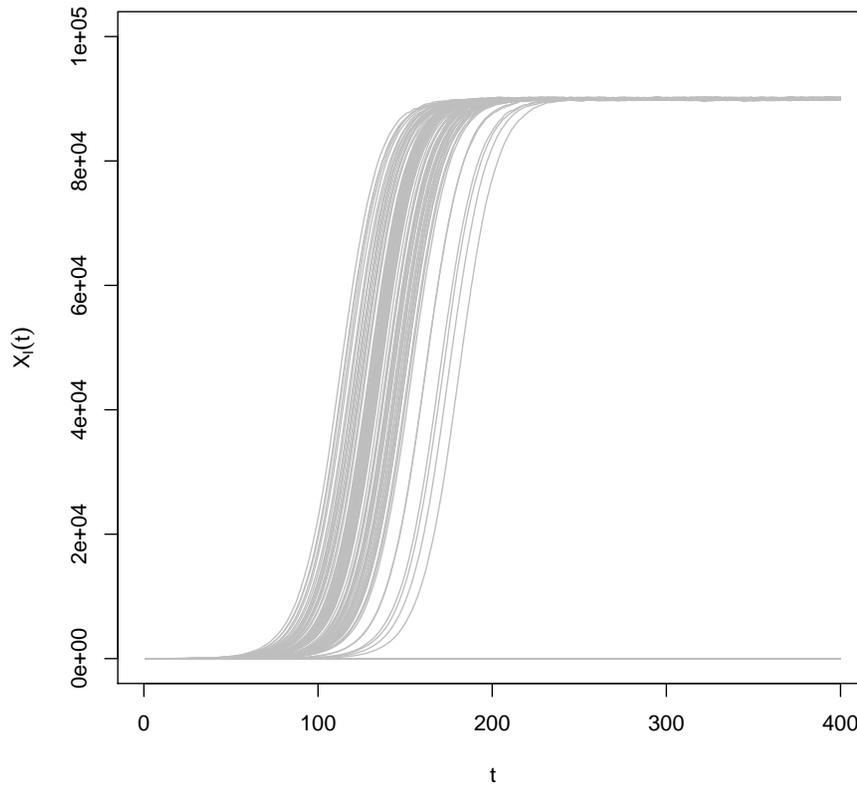


Figure 3: 100 trajectories of the SIS model, with  $n = 10^5$ ,  $\alpha = 10^{-6}$ ,  $\gamma = 10^{-2}$  (so  $X_I^* = 9 \times 10^4$ ) and  $X_I(1) = 1$ . Notice that there are *two* horizontal lines at large  $t$ .

The first solution is boring — if nobody’s infected, nobody gets infected. The other solution is more interesting: it gives the level of infection where new infections just balance recoveries (on average). Whether  $X_I(t)$  will tend to converge to this level is however another question. Obviously this is only possible if  $n > \rho/\alpha + 1$ , so let’s assume that going forward.

Suppose that  $X_I(t) = X_I^* + \epsilon$  for some small amount  $\epsilon$ . You can show (Exercise 4) that the expected change in  $X_I$  has the opposite sign from  $\epsilon$ , so that if the process is above  $X_I^*$  it will tend to be pulled down towards it, and vice versa. This suggests that the process should bounce around this level.

In fact, the situation is a little more subtle.  $X_I = 0$  is still an absorbing state, but now it is one which *can* be reached from other states of the population. If, after all,  $X_I(t) = 1$ , there is some probability that  $X_I(t + 1) = 0$ . (What is that probability?)

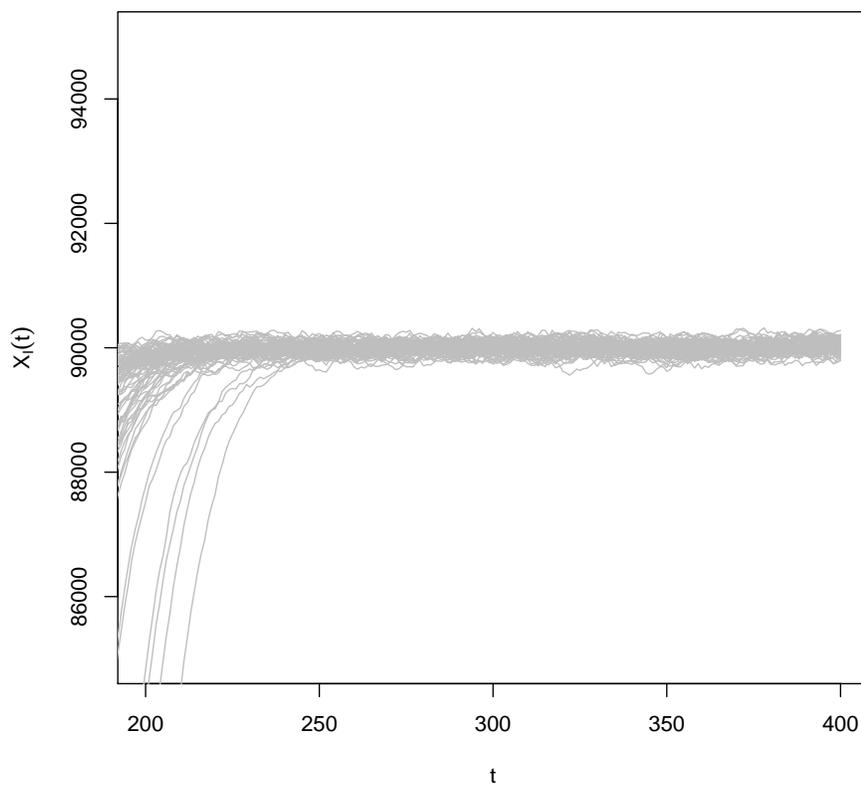


Figure 4: Close-up of Figure 3, showing fluctuations around  $X_I^*$  at large times.

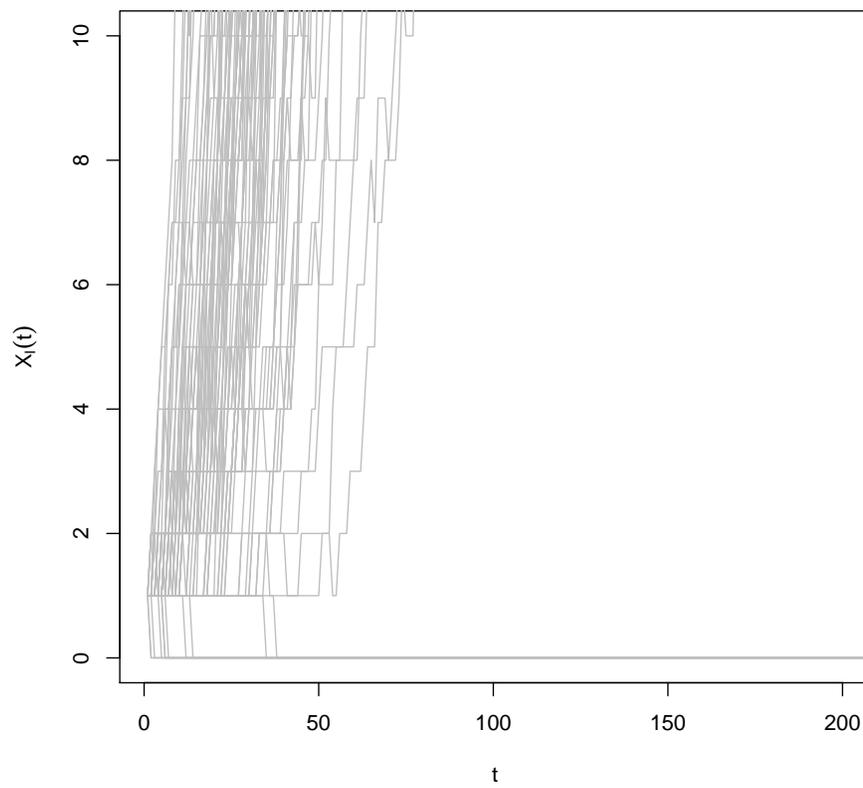


Figure 5: Close-up of Figure 3, showing that some simulations converge to zero.

And no matter how big  $X_I$  gets, there is always some probability that it will eventually reach 1, and then in turn 0. So  $X_I = 0$  is the only closed, irreducible set of states. Consequently, *the disease will always die out* eventually in an SIS model. It will, however, generally take a very long time to do so, and in the meanwhile it will spend most of its time oscillating around  $X_I^*$ .

### 2.3 SIR, SEIR, etc.

It is easy, and often convenient, to add more compartments to epidemic models of this sort. In an SIR model, for example, the third compartment,  $R$ , stands for “removed” or “recovered”, depending on whether we think of individuals in that compartment as having died off, or as having acquired immunity. The basic dynamics are

$$p_{SI} = \alpha X_I \tag{20}$$

$$p_{IR} = \rho \tag{21}$$

$$P_{SR} = 0 \tag{22}$$

$$p_{IS} = 0 \tag{23}$$

$$P_{RI} = 0 \tag{24}$$

$$p_{RS} = 0 \tag{25}$$

I will let you check that the only absorbing state is  $X_R = n$ , and also let you work out both the deterministic approximation and simulate some trajectories on the way there.

A related wrinkle is to add an intermediate compartment between  $S$  and  $I$ , say  $E$  for “exposed”, in which someone is not yet showing symptoms. Individuals in compartment  $E$  might or might not be infectious themselves (and, if so, might be more or less infectious than those in compartment  $I$ ). If one wants to have characteristic lengths of time for each phase of the disease, one can add even more compartments, say  $I_1, I_2, \dots, I_\tau$ , with automatic progression from one to the next.

## 3 Space

The easiest way to handle space in a compartment model is to say that you have  $r_1$  types or classes of individuals across  $r_2$  locations, so that the  $r = r_1 r_2$  compartments track combinations of types and locations. Migration from one location to another is then just another kind of transition. The transition rates can be made to be functions of the counts of types only in the same location (e.g.,  $S$ 's in location 1 can only be infected by  $I$ 's in location 1), or global aggregates calculated across all location, or some combination of both of these.

This approach will work poorly when the number of locations  $r_2$  will become comparable to the population size  $n$ , since then most locations will have a count of 0 for most types, and you would be better off using a different, more individual-based model.

## 4 Asymptotics

We haven't discussed what happens as  $n \rightarrow \infty$ . Basically, a compartment model converges on a system of ordinary differential equations.

Said a bit less baldly, suppose that the transition rates  $p_{ij}(X)$  are all of the form  $p_{ij}(X) = \rho_{ij}(X/n)h$ , where  $h$  represents the amount of time that passes from one value of  $X$  to the next. Notice that we assume that the transition rates are really functions of the *distribution* of individuals across the compartments, and not the *numbers* in compartments — that we get the same rate of infection when  $n = 20$  and  $X_I = 10$  as when  $n = 2000$  and  $X_I = 1000$ . Finally, we fix a finite, but perhaps large, interval of time  $T$ . Then what we find is that the trajectory of the compartment model between time 0 and time  $T$  approaches the solution of a differential equation. If we say that  $x(t)$  is the solution to

$$\frac{dx_i}{dt} = \sum_j \rho_{ji}(x(t))x_j(t) - \rho_{ij}(x(t))x_i(t) \quad (26)$$

then, as  $n \rightarrow \infty$  and  $h \rightarrow 0$ , the random sequence  $X(h)/n, X(2h)/n, X(3h)/n, \dots, X(T)/n$  comes closer and closer to the deterministic function  $x$  on the interval  $[0, T]$  (with high probability).

This is the trick that I used above, when I invoked “deterministic approximations” to the epidemic models. For instance, in the SI model, the limiting trajectory for  $(X_S(t)/n, X_I(t)/n)$  is given by

$$\frac{dx_S}{dt} = -ax_S(t)(1 - x_S(t)) \quad (27)$$

$$\frac{dx_I}{dt} = ax_I(t)(1 - x_I(t)) \quad (28)$$

$$x_S(t) + x_I(t) = 1 \quad (29)$$

$$(30)$$

and solving this differential equation does indeed give us sigmoidal growth curves. (Can you work out how the  $a$  here relates to the  $\alpha$  in the stochastic, finite- $n$  version of the SI model?)

It is important that the convergence only holds over finite times  $[0, T]$ . Basically, this is because infinite time gives a stochastic process infinitely many opportunities to do arbitrarily improbable things, so they will eventually. In particular, in the SIS model, the limiting behavior is

$$\frac{dx_I}{dt} = ax_I(t)(1 - x_I(t)) - gx_I(t) \quad (31)$$

and not extinction.

## 5 Demography: Birth, Aging, and Death

So far, we have considered models where the population size  $n$  is constant. Most real populations change in size, and an important application of compartment models is to demography, where what we really care about is how  $n$  changes over time. The three processes which change  $n$  are birth, death, and migration.

The SIR model shows how to handle death: we can treat “being dead” as just another compartment, say  $R$ <sup>6</sup>. Individuals in compartment  $i$  transition to this compartment at some rate  $\delta_i$ , and there are no transitions out of this compartment. Demographers typically define the **mortality rate** as the number of deaths per 1000 individuals per year, which you can translate into a probability per person per unit time (depending on the time-scale of your model).  $Y_{iR}(t)$  would be the random variable which counted the actual number of deaths among members of compartment  $i$  in time-period  $t$ . If this sort of formal simplicity isn’t needed, however, we can simply introduce a new variable  $D_i(t)$  for deaths in compartment  $i$  in period  $t$ .

Migration can be handled similarly, by introducing a new compartment for “leaving for the rest of the world”. If both in- and out- migration are possible, then we need to model the flow of immigrants.

Birth, however, requires a slightly special treatment in these models. The difficulty is that parents are (typically) still around. The easiest way to handle this is to introduce a new set of variables,  $B_{ij}(t)$ , which count the number of new births into compartment  $j$  from parents in compartment  $i$ <sup>7</sup>. If individuals can only be born one at a time, then we can make  $B_{ij}$  binomial reflecting a certain rate of parentage for individuals in compartment  $i$ . If there are a non-trivial number of twins, triplets, etc., it might be better to use something like a Poisson distribution.

The model thus looks like this:

$$Y_{ij}(t+1) \sim \text{Binom}(X_i(t), p_{ij}(X(t))) \quad (32)$$

$$D_i(t+1) \sim \text{Binom}(X_i(t), \delta_i(X(t))) \quad (33)$$

$$B_{ij}(t+1) \sim \text{Binom}(X_i(t), \beta_{ij}(X(t))) \quad (34)$$

$$X_i(t+1) = X_i(t) - D_i(t+1) + B_{ii}(t+1) + \sum_{j \neq i} B_{ji}(t+1) + Y_{ji}(t+1) - Y_{ij}(t) \quad (35)$$

In human demography, and to a lesser extent in population ecology, we need a fairly large number of compartments, which track combinations of age and sex. Demographers conventionally work with 5-year age brackets, so that there is a compartment for females age 0–5 and for males age 0–5, for females age 6–10 and for males age 6–10, and so on all down the line. The transition probabilities are very simple: every individual who does not die moves into the next age group (and the same sex), with probability 1. It is the age-specific death rates  $\delta_i$  which let us model realistic distributions of life-span; without them, we’d have geometric distributions of longevity.

<sup>6</sup>Saying that  $R$  stands for “removed” is euphemistic.

<sup>7</sup>I am writing as though individuals have only one parent, which of course isn’t true for (most) vertebrates, including human beings, but the reasons for this will become apparent shortly.

Births can only happen into the 0–5 age brackets, so  $B_{ij} = 0$  unless  $j$  is one of those two brackets. Because only females give birth,  $B_{ij} = 0$  unless  $i$  is a female age bracket of child-bearing age. In many applications, the age-fertility rates  $\beta_{ij}$  will be taken to be constant, but they can be allowed to vary with the number of males, or even with the ratio of males to females, depending on the customs of the population concerned. Note that these rates  $\beta_{ij}$  should not be confused with the **total fertility rate**, which is the *sum* of age-specific rates across child-bearing years.

## 6 Chemistry

Chemistry provides an important application for compartment models. Here the “compartments” are different types of molecules, or **chemical species**, and transitions are the consequences of reactions. We can write the SI model in this form, with one reaction,  $S + I \rightarrow 2I$ <sup>8</sup>. We can also write the SIS model as something with two reactions,  $S + I \rightarrow 2I$  and  $I \rightarrow S$ . The constant- $n$  case corresponds to all reactions having the same number of molecules on both the left- and the right- hand sides (as the SI and SIS cases do). In actual chemistry, though, it’s common for the number of molecules to change, as in  $2H_2 + O_2 \rightarrow 2H_2O$  (three molecules go in, two molecules come out). The way to handle this is to use  $Y$  variables to count the number of times each reaction happens, and then subtract the substrates (on the left) and add the products (on the right), so, e.g., if the hydrogen-and-oxygen-to-water reaction was the only one, we’d have

$$X_{H_2}(t+1) = X_{H_2}(t) - 2Y_{2H_2+O_2 \rightarrow 2H_2O}(t+1) \quad (36)$$

$$X_{O_2}(t+1) = X_{O_2}(t) - Y_{2H_2+O_2 \rightarrow 2H_2O}(t+1) \quad (37)$$

$$X_{H_2O}(t+1) = X_{H_2O}(t) + 2Y_{2H_2+O_2 \rightarrow 2H_2O}(t+1) \quad (38)$$

(We would also typically assume that the transition rate for this reaction is  $\propto X_{H_2}^2 X_{O_2}$ .)

## 7 Further reading

For a fine textbook treatment of compartment models in biology, see Ellner and Guckenheimer (2006), which includes a good chapter on epidemic models, and references to the further literature.

## 8 Exercises

To think through, as complements to the reading, not to hand in.

<sup>8</sup>A reaction like this, where the species  $I$  is involved in a process which makes more of the same species, is called **autocatalytic**.

1. *Compartment-model state spaces are very large* Fix  $n$ . A possible state  $X$  is a vector of non-negative integers  $[X_1, X_2, \dots, X_r]$ , with the constraint that  $\sum_{i=1}^r X_i = n$ . Prove that the number of such vectors is  $\binom{n+r-1}{n}$ .
  - (a) Imagine writing out a line of  $n$  dots, and inserting  $r-1$  vertical bars between the dots, with the possibility of bars being next to each other. Call the number of dots to the left of the left-most bar  $X_1$ , the number of dots between that first bar and the second  $X_2$ , and so on. Convince yourself that every arrangement of dots and bars corresponds to a valid value of  $X$ , and that every valid value of  $X$  can be represented in this way.
  - (b) Show that the number of arrangements of dots and bars is equal to choosing  $n$  locations for dots out of  $n+r-1$  places.
  - (c) Show that  $\binom{n+r-1}{n} = \binom{n+r-1}{r-1}$ .
  - (d) Use Stirling's approximation,  $\log n! \approx n \log n$ , to show that as  $n$  grows with  $r$  fixed, the size of the state space grows like  $n^{r-1}$ .
2. *Randomly-varying transition rates across individuals and over-dispersion* Suppose that each individual in compartment  $i$  has its own, random transition rate  $P_{ij}$  for moving to compartment  $j$ , with  $\mathbb{E}[P_{ij}] = p_{ij}$  and  $\text{Var}[P_{ij}] = v_{ij}$ . Show that  $\text{Var}[Y_{ij}(t+1)|X(t)] = X_i(p_{ij}(1-p_{ij}) + v_{ij})$ .
3. Show the following for the SI model of §2.1.
  - (a)  $X_I = 0$  and  $X_I = n$  are absorbing states.
  - (b) There are no other absorbing states.
  - (c) An absorbing state forms a closed, irreducible set of states. (This is a general fact about Markov chains.)
  - (d) This chain has no other closed, irreducible sets.
  - (e) With probability 1, either  $X_I(t) \rightarrow 0$  or  $X_I(t) \rightarrow n$ .
4. For the SIS model, calculate  $\mathbb{E}[X_I(t+1) - X_I(t) | X_I(t) = X_I^* + \epsilon]$  in terms of  $n, \alpha, \rho$  and  $\epsilon$ , and show that (for small  $\epsilon$ ) it has the opposite sign from  $\epsilon$ .
5. (a) Prove that if the probability of an event happening on each time step is a constant  $p$ , and successive time-steps are independent, the number of steps we need to wait for the event to happen has a geometric distribution. Find the parameter in terms of  $p$ .

## References

Dawkins, Richard (1976). *The Selfish Gene*. Oxford: Oxford University Press.

- Ellner, Stephen P. and John Guckenheimer (2006). *Dynamic Models in Biology*. Princeton, New Jersey: Princeton University Press.
- Li, Ming and Paul M. B. Vitányi (1997). *An Introduction to Kolmogorov Complexity and Its Applications*. New York: Springer-Verlag, 2nd edn.
- Lotka, Alfred J. (1924). *Elements of Physical Biology*. Baltimore, Maryland: Williams and Wilkins. Reprinted as *Elements of Mathematical Biology*, New York: Dover Books, 1956.
- Rogers, Everett M. (2003). *Diffusion of Innovations*. New York: Free Press, 5th edn.
- Ruelle, David (1991). *Chance and Chaos*. Princeton, New Jersey: Princeton University Press.
- Salmon, Wesley C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Siegfried, André (1960/1965). *Germs and Ideas: Routes of Epidemics and Ideologies*. Edinburgh: Oliver and Boyd. Translated by Jean Heanderson and Mercedes Clarasó from *Itinéraires de Contagions: Epidémies et idéologies* (Paris: Librairie Armand Colin).