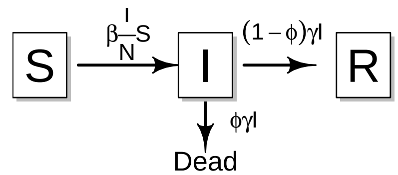
Demystifying the epidemic model

Mathematical models often seem like magic. As a consequence, when we see an equation we might think, *oh no! I can't do this!* or *skip to the good stuff* or similar. Today I would like to disabuse you of this notion that models are magic. Tedious? Sure! Indeed, that's what computers are for. But magic? No!

Today our goal is to understand what the computers are doing in the background when we "run" a mathematical model of an epidemic by pretending to be the computers. I promise, it will be a bit tedious, but you can do it[1] and you will learn a lot from doing it!

[1] And we will then let the computers take over more and more



The model

You will recall, perhaps vaguely, that we constructed a simple model of an epidemic, tracking the flow of individuals from the susceptible (S) to the infected and infectious (I) to the recovered and no long susceptible (R) categories. They looked all fancy:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta\frac{I}{N}S \\
\frac{dI}{dt} &= +\beta\frac{I}{N}S - \gamma I \\
\frac{dR}{dt} &= +(1-\phi)\gamma I
\end{aligned}$$

but they are just numbers. It helps to keep track of what the parameters (all the Greek letters) represent. First, β represents the number of contacts an average individual makes in a time step (i.e., contacts per day) *times*[2] the probability (from zero to 1) that a contact with an infected host caused a new infection. Second, γ is the average rate at which infections are lost (i.e., per day). It can be easier to think of $1/\gamma$ as the average duration of an infection (i.e., days). Finally, ϕ is the case fatality rate or the probability (from zero to 1) that an infections kills the individual, as opposed to the individual recovering, which happens with probability $1-\phi$.

[2] Meaning "and".

For fun, I'd like your group to come up with some values for these three parameters (write them down here).

$$\beta =$$

$$\gamma =$$

$$\phi =$$

We will use these parameters as well as the following starting values for this exercise. We will assume a population of $S = 100$ susceptible individuals at time $t = 0$ and no recovered and thus resistant individuals ($R = 0$). Further, let's assume that a single infected individual rides into town ($I = 0$).

Time	S	I	R	I/N	gain in I	loss of I	newly recovered
0	100	1	0	1/101			
1							
2							
3							

Simulating, slowly

The thing about our model, above, is that it does *not* tells us how many S, I, or R there are at any given time. Instead, it tells us about the *change* in S, I, and R per unit time. But, if we know how many of each we are starting with *and* we know how they change, then we know how many of each category (S, I, and R) we will have after a bit of time. We just have to add (or subtract) these changes to each class.

There is one little caveat: These are differential equations, which assume that everything happens all the time, instantaneously. We, however, have to deal in discrete time steps. We will revisit this, but for now let's just proceed in one-day increments. OK?

So if we start with:

we want to figure out what we have at time 1, then time 2, and time 3.

Let's start with the gain in infections, shall we? This happens at rate $\beta\frac{I}{N}S$. Well, we have all these numbers! You chose β and I gave you $I/N = 1/101$ and $S = 100$. So the new infections produced in the time step from time 0 to time 1 is $\beta\frac{1}{101} \times 100$. Calculate this number for whatever value of β you chose and write this in the "gain in I$ column.

We can take a similar approach to the loss of infections, which happens at rate γI. At time 0 I told you that $I = 1$ and you chose γ so their product is pretty easy to calculate, right? Write this number in the "loss of I" column.

Finally, we need to think about the rate at which animals recover. This is, overall, $(1-\phi)\gamma I$, but we can see that really, this is just a fraction $(1-\phi)$ of the infections that were lost. That is, those that didn't die, recover. So we can take our last value for "loss of I" and simply multiply by $(1-\phi)$, which is the "newly recovered". Go ahead and fill this out.

We now have the rate of change (over one day) in the S, I, and R categories. We simply need to add or subtract these amounts from the time 0 numbers to get the time 1 numbers. For instance, we know we have a certain "gain in I". These came from the initial $S = 100$, and so should be subtracted out, and go to, and thus should be added to, the initial $I = 1$. But we also lost some infections ("loss of I") that need to be subtracted out of the initial $I = 1$, too. And lastly, some fraction of these lost infections are now "newly recovered" that should be added to the initial $R = 0$.

Having done these additions and subtractions, we should have new values for S, I, and R at time = 1. Before moving on, make sure these numbers make sense[3].

[3] Should we have any zeros? Negative numbers? Numbers larger than 100? Is there anything else that looks squirrely?

To get the gains and loss in the *next* time step (from time = 1, to time = 2) we simply repeat the process with the new values of S, I, and R. Give it a try! And then repeat to get us to time = 3.

Enter the electronic computers

As I wrote above, none of this is magic, but it is pretty tedious. It turns out that computers are quite good at tedious tasks like this, so we rely on them quite a lot for "running" or "simulating" our models. Together we will put together a spreadsheet that does what you were just doing, but faster and more accurately. This will allow us to see how the simulated epidemic unfolds over time. We'll also construct it so that we change our parameter values and see how the epidemic changes. This is a nice halfway point to the prettier versions, like the Shiny app I made for you, that hides all of the calculations from you. With the spreadsheet we can easily go in and inspect any cell or value and, I hope, get a bit more intuitive sense of what's going on. But they all do the same thing.