An Epidemic Model

This is a simple model, due to Kermack and MacKendrick, of an epidemic of an infectious disease in a population. We assume the population consists of three types of individuals, whose numbers are denoted by the letters S, I and R (which is why this is called an SIR model). All these of course are functions of time.

S is the number of **susceptibles**, who don't have the disease but could get it.

I is the number of **infectives**, who have the disease and can transmit it to others.

R is the number of **removed**, who can't get the disease or transmit it: either they have a natural immunity, or they have recovered from the disease and are immune from getting it again, or they have been placed in isolation, or they have died. The mathematical model doesn't distinguish between these possibilities.

New infections occur as a result of contact between infectives and susceptibles. In this simple model the rate at which new infections occur is βSI for some positive constant β. When a new infection occurs, the individual infected moves from the susceptible class to the infective class. In our simple model, there's no other way individuals can enter or leave the susceptible class, so

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

On the other hand, infective individuals can enter the removed class, and we assume that this happens at a rate νI, so

$$\frac{dI}{dt} = \beta SI - \nu I$$
$$\frac{dR}{dt} = \nu I$$

Among the questions we want to answer:

(1) Suppose we start out with a population of susceptibles, and introduce a small number of infectives. Will the number of infectives increase, causing an epidemic, or will the disease fizzle out?

(2) Assuming there is an epidemic, how will it end? Will there still be susceptibles left when it is over?

(3) How long will it last?

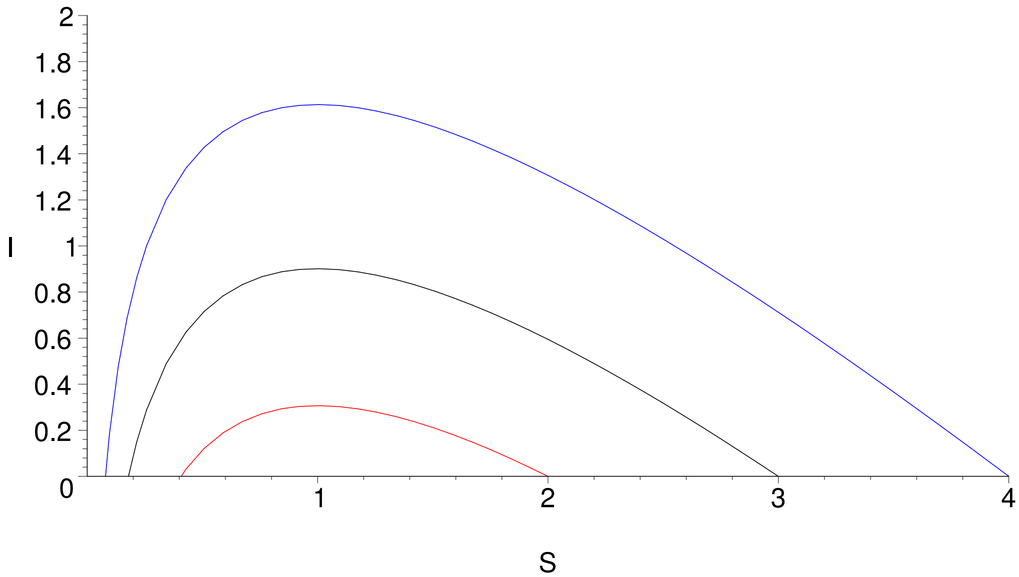
Now what we have is a system of differential equations, with three unknown functions. All we know about are single differential equations, with one unknown function. The trick is to consider I as a function of S. Then according to the chain rule

$$\frac{dI}{dS} = \frac{dI/dt}{dS/dt} = \frac{\beta SI - \nu I}{-\beta SI} = \frac{\nu}{\beta S} - 1$$

Since the right side here does not depend on I, all we need to do is integrate:

$$I = \int \left(\frac{\nu}{\beta S} - 1 \right) dS = \frac{\nu}{\beta} \ln S - S + C$$

Here are some plots of I as a function of S for various values of C. I'm taking $\nu = \beta = 1$.



Mathematically the curves extend to negative values of I as well as positive, but of course that doesn't make sense for populations, so I cut them off at $I = 0$. The line $I = 0$ really consists of equilibrium points: if $I = 0$ nothing will happen in our model. If you start out with $I > 0$ on one of these curves, as time goes on you travel along the curve to the left, eventually approaching $I = 0$ at some positive value of S. This must happen since on any of these curves, $I = (\nu/\beta)\ln S - S + C \to -\infty$ as $S \to 0$. So the answer to question (2) is that the epidemic will end as $I \to 0$ with S approaching some positive value: yes, there will always be some susceptibles left over.

For the answer to (1), note that $dI/dt < 0$ if $dI/dS > 0$, which happens if $\beta S < \nu$. In that case, the epidemic dies out. On the other hand, if $\beta S > \nu$, $dI/dt > 0$ and the number of infectives increases; the maximum of I will occur when S has decreased to the value ν/β. We can say that what determines whether the epidemic "takes off" is which occurs faster: new infections or removal of infectives. In SARS in 2003, despite some setbacks the health-care system managed to remove infectives (by isolating them) faster than new infections occurred, and the disease died out. If SARS had been more like influenza, β would have been much larger and the epidemic would inevitably have progressed until a large fraction of the population had been infected.

Now for question (3), we must go back to a differential equation involving time. Plugging in the formula for I as a function of S into the differential equation for dS/dt, we get

$$\frac{dS}{dt} = -\beta S I = -S\left(\nu \ln S - \beta S + \beta C\right)$$

To solve this differential equation using separation of variables, we write

$$-\int_{S(0)}^{S(T)} \frac{dS}{S\left(\nu \ln S - \beta S + \beta C\right)} = \int_0^T dt = T$$

Unfortunately we don't know an antiderivative for the integral on the left. There probably isn't a formula for it. However, we could use numerical methods (such as our left or right sums).

Let's say that at $t = 0$ we have $I = I_0$ (some small positive number) and $S = S_0 > \nu/\beta$. Then the constant

$$C = I_0 - \frac{\nu}{\beta} \ln S_0 + S_0$$

We'll declare the epidemic over when $I = I_0$ again, which will happen when $S = S_1 < \nu/\beta$, where

$$\frac{\nu}{\beta} \ln S_1 - S_1 = I_0 - C = \frac{\nu}{\beta} \ln S_0 - S_0$$

Again this can be solved using numerical methods. Then once we have S_1 we can say the epidemic is over in time T where

$$T = -\int_{S_0}^{S_1} \frac{dS}{S\left(\nu \ln S - \beta S + \beta C\right)}$$

For example, I tried $\beta = \nu = 1$, $S_0 = 3$, $I_0 = 10^{-6}$ (perhaps the unit for population is a million individuals, so think of a population of 3 million susceptibles with one infective). It turns out $S_1 \approx 0.17856$ (so when the epidemic dies out there will be only about 178,560 susceptibles left), and $C \approx 1.90139$. And the duration of the epidemic will be $T \approx 25.4711$. Here is a graph of I (in red) and S (in blue) as functions of t over the course of the epidemic, again produced by numerical methods.