The SIR Model for Spread of Disease

Part 5. The Contact Number

In Part 4 we took it for granted that the parameters **b** and **k** could be estimated somehow, and therefore it would be possible to generate numerical solutions of the differential equations. In fact, as we have seen, the fraction **k** of infecteds recovering in a given day *can* be estimated from observation of infected individuals. Specifically, **k** is roughly the reciprocal of the number of days an individual is sick enough to infect others. For many contagious diseases, the infectious time is approximately the same for most infecteds and is known by observation.

There is no direct way to observe **b**, but there is an indirect way. Consider the ratio of **b** to **k**:

b/k = b x 1/k

= the number of close contacts per day per infected x the number of days infected

= the number of close contacts per infected individual.

We call this ratio the **contact number**, and we write **c = b/k**. The contact number **c** is a combined characteristic of the population and of the disease. In similar populations, it measures the relative contagiousness of the disease, because it tells us indirectly how many of the contacts are close enough to actually spread the disease. We now use calculus to show that **c** can be estimated after the epidemic has run its course. Then **b** can be calculated as **c k**.

Here again are our differential equations for **s** and **i**:

$$\frac{ds}{dt} = -b\, s(t)\, i(t),$$

$$\frac{di}{dt} = b\, s(t)\, i(t) - k\, i(t).$$

We observe about these two equations that the most complicated term in both would cancel and leave something simpler if we were to *divide the second equation by the first* -- provided we can figure out what it means to divide the derivatives on the left.

1. Use the Chain Rule to explain why

$$\frac{di}{ds} = -1 + \frac{1}{cs}$$

The differential equation in step 1 determines (except for dependence on an initial condition) the infected fraction **i** as a function of the susceptible fraction **s**. We will use solutions of this differential equation for two special initial conditions to describe a method for determining the contact number.

Three features of this new differential equation are particularly worth noting:

- The only parameter that appears is **c**, the one we are trying to determine.
- The equation is *independent of time*. That is, whatever we learn about the relationship between **i** and **s** must be true for the entire duration of the epidemic.
- The right-hand side is an explicit function of **s**, which is now the *independent* variable.

2. Show that **i(s)** must have the form

$$i = -s + \frac{1}{c}\ln s + q$$

   where **q** is a constant.

3. Explain why the quantity

$$i + s - \frac{1}{c}\ln s$$

   must be independent of time.

There are two times when we know (or can estimate) the values of **i** and **s** -- at **t = 0** and **t = infinity**. For a disease such as the Hong Kong flu, **i(0)** is approximately **0** and **s(0)** is approximately **1**. A long time after the onset of the epidemic, we have **i(infinity)** approximately **0** again, and **s(infinity)** has settled to its steady state value. If there has been good reporting of the numbers who have contracted the disease, then the steady state is observable as the fraction of the population that *did not* get the disease.

4. For such an epidemic, explain why

$$c = \frac{\ln s_\infty}{s_\infty - 1}$$

   [Hint: Use the fact that the quantity in step 3 is the same at **t = 0** and at **t = infinity**.]

5. Use one of your numerical solutions in Part 4 to estimate the value of **s(infinity)**. Use this value to calculate the contact number **c** for the Hong Kong flu. Compare your calculated value with the one you get by direct calculation from the definition, **c = b/k**.