Basic and effective reproduction numbers Part II*

Marc R. Roussel
Department of Chemistry and Biochemistry
University of Lethbridge

May 10, 2021

Mass-action models are not only used in the chemical sciences. They are also used extensively in ecological and epidemiological modeling. One of the most famous models in epidemiology is the **SIR** model, which describes a disease in which the population can be divided into Susceptible, Infectious and Recovered individuals. While this is a very simple model, it does capture the essential features of many diseases. And it goes without saying that epidemiologists use a range of models as appropriate, some much more complicated than a basic SIR model. But the SIR model remains a basic entry point for a lot of modeling in epidemiology.

The SIR model consists of the following set of mass-action events:

1. If a susceptible individual comes in contact with an infected individual, there is a chance that the disease will be transmitted from one individual to another. At the population level, this corresponds to a transmission process

$$\mathrm{S} + \mathrm{I} \xrightarrow{k_t} 2\,\mathrm{I} \tag{1a}$$

with rate $k_t SI$, using capital letters to represent either the number of individuals in a given "compartment" (category) or their density (e.g. in units of invididuals per unit area). The constant k_t depends on how

*These notes are enrichment. I don't intend to examine this material although clearly some parts of it (writing rate equations, etc.) fall squarely within the examinable material of this course.

transmissible the disease is, and on the average number of contacts per unit time that could lead to transmission. In the case of a pandemic such as the current Covid-19 pandemic, public health measures can reduce the value of k_t by reducing the number of potentially hazardous contacts, so k_t is not necessarily constant. In the basic SIR model however, we treat it as such.

2. Infectious individuals recover (or die—the simple SIR model makes no fundamental distinction between these cases, although clearly they matter a lot to us).

$$\mathrm{I} \xrightarrow{\mathrm{k}_r} \mathrm{R}. \tag{1b}$$

And that is the whole model. We can write the mass-action differential equations for this model as follows:

$$\frac{dS}{dt} = -k_t SI, \tag{2a}$$

$$\frac{dI}{dt} = k_t SI - k_r I = (k_t S - k_r)I, \tag{2b}$$

$$\frac{dR}{dt} = k_r I. \tag{2c}$$

The SIR model assumes a constant population size. Note that

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0,$$

and since, by the addition rule for derivatives,

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = \frac{d}{dt}\left(S + I + R\right),$$

it follows that $S + I + R$ is a constant. This is not a bad approximation provided a disease doesn't hang around for too long. Obviously, for an endemic disease (such as, e.g., chicken pox), we might need a model that accounts for changes in population size from various causes.

Let's look a little harder at equation (2b). An outbreak will spread or die out depending on the sign of the term in parentheses: if this sign is positive, the number of infectious individuals grows, and this number decreases if the quantity in parentheses is negative. Thus, the disease spreads if

$$k_t S - k_r > 0$$

or

$$\frac{k_t S}{k_r} > 1,$$

and dies out otherwise. This is exactly what we want for an R number. If we take $S = S_0$, the size of the susceptible population at time zero (essentially, the entire population), then we get the basic reproduction number:

$$R_0 = \frac{k_t S_0}{k_r}.$$

On the other hand, if we take S at time t, we get the effective reproduction number[1]

$$R_e = \frac{k_t S}{k_r}. \tag{3}$$

In our original treatment of reproduction numbers, we had τ, the average time that a person remains infectious. There is a connection between population-level statistics and individual-level statistics that we won't go into here (but that you can learn about in my Foundations of Chemical Kinetics course, which will be offered next Fall if all goes well). This allows us to make an important connection between the rate constant k_r, which gives the rate at which infectious individuals recover, and τ, the average time of infection:[2]

$$k_r = \frac{1}{\tau}.$$

If we substitute this relationship into equation (3), we get

$$R_e = k_t S \tau. \tag{4}$$

Equation (4) has a simple interpretation: $k_t S$ is the rate at which the infection is transferred per infectious person, while τ is the average time a person remains infectious. If I multiply the rate of transfer of infection by the interval of time that a person is infectious, I get the approximate number of individuals to which the disease is passed on.

[1] You can probably see now why I use R_e for the effective reproduction number. If I didn't, we would have a problem distinguishing it from the number of recovered individuals in an SIR model.

[2] Something to ponder for the mathematically inclined among you: the average time isn't the same as the half-life. If you have taken a statistics course, it may be obvious to you why this is.

Finally note that the two definitions of R_0 and of R_e we have seen, in this and the previous notes, are slightly different and need not agree. However, there is a relationship between them, provided $k\tau$ is not too large. Recall that we had

$$R_e = e^{k\tau},$$

where k is the slope of the tangent to the graph of $\ln I$ vs t. Now take a Taylor series[3] of this equation for R_e at small values of $x = k\tau$, stopping at the first "interesting" term (the first one that depends on $k\tau$):

$$R_e \approx 1 + k\tau. \tag{5}$$

If the two definitions of R_e are to agree, we should have

$$\begin{aligned} k_t S\tau &= 1 + k\tau, \\ \therefore k &= \frac{k_t S\tau - 1}{\tau} = k_t S - \frac{1}{\tau} = k_t S - k_r. \end{aligned}$$

So the k determined from data on the time dependence of the infections is exactly the pseudo-first-order rate constant that appears as the coefficient of I in equation (2b)! This shows the two approaches to be consistent: $k > 0$ and thus R_e obtained from data is greater than 1 exactly when R_e obtained from the SIR model [equation (3)] is also greater than 1.

So far, we have assumed that we would be given τ, and therefore focused on calculating k. We can however estimate R_e directly from an SIR-based analysis. Since people recover after being infected, we would expect there to be a simple relationship between R and I. In fact, we can solve the SIR equations exactly and write down an explicit relationship between the two, but it turns out to be more convenient to think about this relationship in a more abstract way. If $R = R(I)$, then differentiating this relationship with respect to time and applying the chain rule gives

$$\frac{dR}{dt} = \frac{dR}{dI}\frac{dI}{dt}$$

or

$$\frac{dR}{dI} = \frac{dR/dt}{dI/dt}.$$

[3]Easily the most underappreciated and most useful bit of applied mathematics taught in introductory calculus courses.

Using equations 2b and 2c, we get

$$\frac{dR}{dI} = \frac{k_r I}{(k_t S - k_r)I} = \frac{k_r}{k_t S - k_r} = \frac{1}{\frac{k_t S}{k_r} - 1} = \frac{1}{R_e - 1}.$$

We decided that, logically, R should be a function of I. Looking at the expression above though, it would be more convenient to express the relationship the other way: $I = I(R)$. This is possible provided the relationship between the two variables is locally invertible, which should be the case given that R is a strictly increasing function of time. Thus,

$$\frac{dI}{dR} = R_e - 1.$$

The procedure for calculating R_e is therefore as follows: First, get the slope of a plot of I vs R. Then we can get R_e simply by calculating

$$R_e = 1 + \frac{dI}{dR}.$$

Note that the sign of dI/dR determines whether R_e is larger than 1 or smaller than 1.

If we also get k using the techniques of the previous set of notes, then since $k\tau = R_e - 1$ from equation (5), we get

$$\tau = \frac{1}{k}(R_e - 1) = \frac{1}{k}\,\frac{dI}{dR}.$$