Maths and infectious diseases

By Elizabeth Buckingham-Jeffery, School of Mathematics at the University of Manchester
Website: `ebucksjeff.github.io` and Twitter: @ebucksJeff

Despite our advances in biology and medicine, infectious diseases like flu, Ebola, and malaria are still a major cause of death worldwide. We can use maths to support the medical response to an infectious disease outbreak by, for example, making a prediction of how many people will get ill and therefore will require medicine. We do this using a *mathematical model* of the disease spreading in the population.

A mathematical model is a simplified representation of a complex situation. A good model will be one that *replicates real life well enough*, but is also simple enough that we can solve the equations (either on a computer or by hand) to give predictions. In this worksheet you will be introduced to a simple version of one of the most popular infectious disease models, called the *SIR model*.

The 'SIR model' of infectious disease

We will develop a model that represents an infectious disease spreading in a population by writing down a set of equations. We *assume* that:

- while people are infected they can infect other people;
- after some time, people recover from infection;
- people do not die or leave the population;
- no new people are born or immigrate into the population;
- everyone *interacts equally* with everyone else;
- immunity from the disease lasts forever.

When developing a model, mathematicians must be aware of the assumptions they are making and how these relate to the real-world situation they are modelling. For example, the assumption that no one is born or dies is sensible when making predictions for a few weeks ahead, but certainly not when making predictions for a number of years.

We split the population into three compartments: people that are **S**usceptible to the disease, people that are **I**nfected and infectious with the disease, and people that are **R**ecovered from the disease.

Based on the *assumptions* written above, people can move between the classes in the following way:

Can you sketch the compartments and movements we would have if we were to change the assumptions in the following ways:

Immunity *wanes*; after some period of time you become susceptible to the disease again. What could we call this model?

New people are also born into the population; everyone is susceptible to the disease when they are born.

Challenge: New people are born into the population susceptible to the disease. A proportion of the population are vaccinated before they get infected. The immunity given by this vaccine lasts forever.

We use the notation $S(t)$ for the number of people in the susceptible compartment at time t, $I(t)$ for the number of people in the infected compartment at time t, $R(t)$ for the number of people in the recovered compartment at time t, and $N(t)$ for the total number of people in the population at time t.

Writing equations

Can you write an *equation* for $N(t)$ in terms of $S(t)$, $I(t)$ and $R(t)$?

Can you write an equation for $N(t)$ in terms of the number of people in the population at time zero, which we write as N_0?

Can you write an equation for $R(t)$ in terms of $S(t)$, $I(t)$, and N_0?

We know how many people are in each compartment currently, which we write as time t. We want to write down some equations that tell us how many people are in each compartment at a time which is shortly in the future, which we write as time $t + \Delta t$; Δt is a small length of time that we pick, for example a day, half a day, or an hour.

We will write down one *recurrence relation* for each compartment, this gives the SIR model:

$$\begin{aligned} S(t+\Delta t) &= S(t) - \left(\beta \frac{I(t)}{N_0} S(t)\right) \Delta t, \\ I(t+\Delta t) &= I(t) + \left(\beta \frac{I(t)}{N_0} S(t) - \gamma I(t)\right) \Delta t, \\ R(t+\Delta t) &= R(t) + \left(\gamma I(t)\right) \Delta t. \end{aligned}$$

β (the Greek letter 'beta') and γ (the Greek letter 'gamma') are *parameters of the model*. When we consider a specific disease β and γ have particular numerical values, such as $\beta = 0.47$ and $\gamma = 0.33$ for flu. When we write the equations with β and γ we mean a general model of diseases; when we write the equations with a specific number we mean a model of one specific disease. β is the parameter related to infection rate and γ is the parameter for recovery rate.

In the extension at the end of this worksheet, you can read about where these equations come from but, for now, we will consider some exercises to explore the *behaviour* of the SIR model for different parameter values.

Compute the number of people in each class

Consider a population with currently, at time 0, 99 susceptible people and 1 infected person for the infectious disease flu which has parameter values $\beta = \frac{1.4}{3}$ and $\gamma = \frac{1}{3}$. Can you write down the values of N_0, $S(0)$, $I(0)$, and $R(0)$?

Using the equations given above, can you write down the values of $S(1)$, $I(1)$, and $R(1)$?

What about the values of $S(2)$, $I(2)$, and $R(2)$?

As you can tell, it will get quite tedious to keep repeating this until day 100, or even just day 10. So we ask a computer to do the calculations for us.

Go to the web-page `ebucksjeff.shinyapps.io/basic_sir_model` to investigate the SIR model

What happens to the number of susceptible people, infected people, and recovered people as time progresses?

For which days is the epidemic worsening? When is it improving?

Experiment with changes in the value of the parameter β: keep γ fixed at 0.5 and look at plots with various values of β between 0.5 and 4. How do changes in β affect the graphs; in particular, the graph of I? Do larger values of β correspond to faster or slower epidemics?

Return β to 2 and look at plots with various values of γ between 0 and 0.7. How do changes in γ affect the graphs; in particular, the graph of I? Do larger values of γ correspond to faster or slower epidemics?

Challenge: Describe what happens when $\gamma = 0$. Why do you think this occurs?

We can consider values for the parameters β and γ that relate to specific diseases. Mathematicians have *derived* these parameter values by analysing data and by understanding the biology of diseases. Three diseases we consider here are:

- **Measles:** an infectious disease caused by a virus and spread from person to person by coughs and sneezes. It causes a red, itchy rash and in severe cases can lead to pneumonia which can cause death. The SIR model for measles has parameter values $\beta = 1.88$ and $\gamma = 0.13$.
- **Rubella:** an infectious disease caused by a virus and also spread from person to person by coughs and sneezes. It similarly causes a rash and fever but is considered less dangerous than measles except in pregnant mothers and newborn infants. The SIR model for rubella has parameter values $\beta = 0.62$ and $\gamma = 0.09$.
- **Influenza:** flu is again an infectious disease caused by a virus and spread from person to person by coughs and sneezes. Many cases are mild, with the main symptoms being a fever, runny nose, and muscles pains, however more severe cases can occur. The SIR model for influenza has parameter values $\beta = 0.47$ and $\gamma = 0.33$.

Exercises: Go to the "Compare Diseases" tab of the app

Which of these diseases is highly contagious, which is moderately contagious, and which is the least contagious?

For which disease does the largest proportion of the population avoid infection?

Roughly what proportion of the population avoids infection in each case?

Are there any other differences between the SIR model results for these diseases?

Aside: How was this app made?

I made this app with a *programming language* called R and, in particular, R Shiny. R is frequently used for statistics and data science and the software is completely free. The ability to program in R, or a similar programming language, is a fantastic, employable skill to have.

If you want to get started using R and R Shiny, I recommend the following:

- There are many guides online to introduce you to R and RStudio, a good reference is: `https://www.rstudio.com/online-learning/`.
- 'R for Data Science' is a fantastic online resource for learning R: `http://r4ds.had.co.nz`.
- If you feel comfortable using basic R, learn to make apps using R Shiny: `https://shiny.rstudio.com/tutorial/`.

Aside: Differential equations

The recurrence relations stated here give a simple version of the SIR model in *discrete time*. More often, mathematicians use a version of the model in *continuous time* which is made up of a set of three *differential equations*. If you haven't come across differential equations already, then it is likely you will by the end of your A-levels. The differential equations of the SIR model are:

$$\frac{dS}{dt} = -\frac{\beta}{N}SI, \quad \frac{dI}{dt} = \frac{\beta}{N}SI - \gamma I, \quad \frac{dR}{dt} = \gamma I.$$

If you are interested...

Extension: Where do the SIR equations come from?

Where do these equations of the SIR model come from? First consider the infected compartment. At time $t + \Delta t$ in the future the number of people that will be in the infected class is equal to the number of people currently infected plus a small change; we write this as

$$I(t + \Delta t) = \underbrace{I(t)}_{\text{number currently infected}} + \underbrace{\left(\beta \frac{I(t)}{N_0} S(t) - \gamma I(t)\right) \Delta t}_{\text{small change}}.$$

I'll explain where this particular formula for the small change comes from.

In the small time step Δt the number of people in the infected compartment can change in two ways: people can join from the susceptible compartment (infection) or people can leave to the recovered compartment (recovery). That gives the two separate *terms* in the expression:

$$\left(\underbrace{\beta \frac{I(t)}{N_0} S(t)}_{\text{Infection}} - \underbrace{\gamma I(t)}_{\text{Recovery}}\right) \Delta t$$

Recovery: We assume that there is a *rate* at which you recover from the disease, which we write as γ. Every individual in the infected compartment recovers at the same rate γ, there are $I(t)$ individuals currently in the compartment, so the total rate of people leaving the infected compartment at any time is $-\gamma I(t)$, and we consider the length of time Δt.

Infection: We assume that the rate at which each individual becomes infected by the disease is $\lambda(t)$. Every individual in the susceptible compartment becomes infected at the same rate and there are $S(t)$ individuals who could become infected. So the total rate of people entering the infected compartment at any time is $\lambda(t)S(t)$; we consider this for the length of time Δt.

However, we make one further step and break down $\lambda(t)$ into two parts. The risk of a susceptible person becoming infected depends on the proportion of people around them that are infected; one susceptible person mixing with nine infected people has a greater risk of being infected than one susceptible person mixing with one infected person and eight recovered people. So we rewrite $\lambda(t)$ to show it depends on the proportion of infected people $\frac{I(t)}{N_0}$ as well as other factors, such as contact rate, which we label as β. This gives $\lambda(t) = \beta(t)\frac{I(t)}{N_0}$ and then the equations above.

Derive other equations

Can you *derive* the equations for $S(t+\Delta t)$ and $R(t+\Delta t)$ in the same way we *derived* the equation for $I(t+\Delta t)$?

Can you write down equations for the model with waning immunity you considered on page 2?

References and extra reading

- More information on the SIR model and an introduction to the parameter R_0:
  `https://plus.maths.org/content/mathematics-diseases`
- Long introduction to the SIR model, including differential equations:
  `http://homepage.divms.uiowa.edu/~stroyan/CTLC3rdEd/3rdCTLCText/Chapters/Ch2.pdf`
- The Wikipedia page on the SIR model and extensions of it:
  `https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology`
- Interesting visualisation of measles and vaccination:
  `https://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated`