This worksheet demonstrates a Susceptible-Infectious-Recovered (SIR) infectious disease model, allowing you to investigate properties of the model as disease-related parameters are changed. You will also be able to compare the deterministic and stochastic versions of the model.

Using Microsoft Excel, we can formulate the model as either a set of deterministic difference equations, where the outputs are the same every time (the deterministic model), or as a set of difference equations where we incorporate stochasticity (or randomness) into the infection and removal processes. The model structure can be drawn as a schematic diagram as in the figure below.

A diagram of a number of infection

AI-generated content may be incorrect.

**The spreadsheet**

In the top left of spreadsheet, there are parameters defining the model output that you can adjust. These include the population size (N), the basic reproduction number (R0), the infectious period in days (1/gamma), and the initial number of people infected.

Below this are parameters that are derived from these input parameters. These have been locked for editing but are visible so you can see how they change with the input parameters. These include the effective contact rate (beta, equal to R0 × gamma) and the recovery rate (gamma, 1/infectious period).

On the right are two charts – the top chart shows the epidemic curve (of prevalence, not incidence) using a deterministic model, and the bottom chart shows the epidemic curve (of prevalence, not incidence) using a stochastic model. Both use the same colours: blue for susceptible people, S(t), green for infectious people, I(t), and pink for recovered people, R(t), through time.

As discussed in the lectures, the stochastic model will look different every time you run it. ***You can rerun the stochastic model at any time by hitting the “F9” button, or by selecting “Formulas”, then “Calculate Sheet”.***

**Practical:**

1. Start with a population size of 1000, a R0 of 2.5, an infectious period of 5 days, and 10 people initially infected. This is the base scenario that we will return to throughout this sheet.

First, look at the resulting parameters. What is the value of beta? What is the value of gamma?

Next look at the model formulas in columns O:U, rows 5 and 6. Can you work out how the formulas relate to the schematic model diagram above?

1. Next, look at the deterministic model output. What is the final epidemic size (i.e. proportion of the population infected during the epidemic)? How long before the epidemic has died out?

***Note: Hover your cursor over the S(t), I(t), and R(t) curves to read off specific values.***

1. Now look at the stochastic version of this model. Does it give similar results? How much does the final epidemic size and the timing of the epidemic change as you re-run the model by hitting F9? (Add rows to the table as necessary).

|  |  |  |  |
| --- | --- | --- | --- |
| Run # | Final epidemic size | Peak timing | End of epidemic |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |

1. Increase R0 to 4 (beta = 0.8) and explore what effect this has on the final epidemic size and timing in both the deterministic and stochastic models. What has changed from the scenario where the R0 was 2? Does the stochastic model change more/less with R0 = 4?

Deterministic model:

Stochastic model:

|  |  |  |  |
| --- | --- | --- | --- |
| Run # | Final epidemic size | Peak timing | End of epidemic |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |

1. Now decrease R0 to 1.25 and the infectious period to 2.5 days (beta = 0.5 and gamma = 0.4). What is the final epidemic size? How long does the epidemic last using the deterministic model? What about the stochastic model? Run the stochastic model a few times by pressing F9 and compare findings.

Deterministic model:

Stochastic model:

|  |  |  |  |
| --- | --- | --- | --- |
| Run # | Final epidemic size | Peak timing | End of epidemic |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |

1. We will now investigate the effect of the infectious period for a fixed value of R0. Increase and decrease the duration of the infectious period (to keep R0 fixed). What happens to the timing of the epidemic and the final epidemic size in the deterministic model?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| R0 | Infectious period | Beta | Gamma | Peak timing | Peak size | Final epidemic size |
| 4 | 5 days | 0.8 | 0.2 |  |  |  |
| 4 | 4 days | 1.0 | 0.25 |  |  |  |
| 4 | 2.5 days | 1.6 | 0.4 |  |  |  |
| 4 | 8 days | 0.5 | 0.125 |  |  |  |
| 4 | 10 days | 0.4 | 0.1 |  |  |  |
| 4 | 20 days | 0.2 | 0.05 |  |  |  |

1. Return to the scenario where R0 = 2.5 and the infectious period = 5 days (beta = 0.5 and gamma = 0.2). Now we are going to explore the effect of the population size. Keep the initial number of people infected at 10 but decrease the population size from 1000 to 100. What happens to the deterministic model? What about the stochastic model?
2. Now we will explore the effect of the initial number of people infected. Start with our base scenario of R0 = 2.5, infectious period = 5 days, a population size of 1000, and 10 initial infected people. Now decrease the number of initial people infected to one. What happens to the deterministic model? Run the stochastic model a few times by hitting F9. How do things change? What is happening here? Try gradually increasing the number of initial cases and see what happens.
3. Now we will investigate the relative timing and size of epidemics using parameters consistent with common human pathogens.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Pathogen | R0 | Infectious period | Peak timing | Peak size | Final epidemic size |
| Influenza | 1.5 | 2.5 |  |  |  |
| COVID-19  (ancestral) | 4 | 5 |  |  |  |
| Measles (lower bound) | 16 | 8 |  |  |  |
| Measles (upper bound) | 20 | 8 |  |  |  |

1. Throughout this worksheet, we have focused on prevalence *I(t)*. Of course, for many diseases we usually (only) have knowledge of *incidence*, the number of newly recorded infections in a specified time period (perhaps also per 1,000 population). For example: the number of new cases of influenza reported over a 24hr period (the daily incidence). *Note this would also typically be reported daily, although there is no strict requirement to have the same period of time for calculating incidence and for the reporting frequency.*

Firstly, try to write down an equation for the incidence from the SIR equations. *Note: this is not straightforward!*

Secondly, consider how to compute this quantity in the spreadsheet. *Note: this is also rather tricky, and certainly tedious to do in Excel.*