

MB Assessed Practical 2 (Easter 2023)

1 Instructions

Your report should be written in the form of an Rmarkdown document in RStudio. Use the “BLANK_MB_AP2.Rmd” file provided, add your name and email after “author:” on line 3 and save the file as “CRSID_MB_AP2.Rmd”, where CRSID is the first part of your email address, for example or226. Once completed, you should “knit” the Markdown document in RStudio into an html file. Both the Rmd source file and the html file should be submitted electronically via Moodle, **before 17:00 UK time on Thursday 4th May 2023**.

Requests for extension on medical ground should be submitted by email to teach@path.cam.ac.uk via your Tutor or DoS.

Notes:

- This assignment does not require any external data files, so there is no need to specify a working directory (in other words, don't use the `setwd` function).
- **Pre-requisite:** you should complete Practical 15 before attempting this exercise, in particular to find out how to install and use the `deSolve` package in R.
- The code you have to write is very similar to that from the last two practicals (which are available on Moodle). Keep in mind that there may be different ways to write code to get correct results. The practical will be assessed primarily on the output (figures) and written answers where requested.
- Most (~90%) marks will be awarded to technically correct figures and answers, with a few extra marks awarded for higher quality outputs (with attention being paid to labels, colours, legends etc) and more detailed answers.

Assistance will be available **online** on the first Thursday of Easter term (April 27th) to answer specific questions.

2 Aims

This assessed practical makes use of the different programming methods you learned in the Lent term. The step-by-step questions will guide you to produce figures similar (but not necessarily identical) to those in Dr Restif's lecture notes on epidemic models (Lecture 5).

Specifically, you will use R to analyse the dynamics of the simple epidemic model (SIR) in a population of size N , defined by the following system of differential equations:

$$\begin{aligned}\frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \gamma I\end{aligned}$$

3 Questions

1. [~5% of marks] Why do we only need to analyse two differential equations even though the model has three variables (S , I and R)?

2. [~15%] Write a function `R0(beta,gamma,N)` which takes three arguments (β , γ and the population size N) and returns the basic reproductive number $\mathcal{R}_0 = \beta N / \gamma$. Then write code to generate a plot of \mathcal{R}_0 against the recovery rate γ ranging from 0.2 to 4, with fixed values $\beta = 0.02$ and $N = 100$. Using the `abline` function, add a horizontal line to show the threshold $\mathcal{R}_0 = 1$ on the same plot.
3. [~20%] Write a function `SIR_dydt(t,y,parms)` which takes three arguments (time t , current values of the variables, and values of the model parameters) and returns a list whose only element is a vector with the values of the derivatives dS/dt and dI/dt (see Practical 15). Before trying to solve the model, follow these steps test that your `SIR_dydt` function works correctly. According to the model definition, dS/dt should always be negative, whereas dI/dt will only be positive if $\beta S > \gamma$ and $I > 0$.
 - a. Setting $\beta = 0.02$, $\gamma = 1$, $I = 1$, write a for loop that prints the value of dI/dt for S ranging from 10 to 80 in steps of 10.
 - b. Based on the last output, what is the critical density of susceptible individuals required for an epidemic?
4. [~15%] Using the function `ode()` from the `deSolve` package, compute the solution of the SIR model with $N = 100$, $\beta = 0.02$, $\gamma = 1$, initial conditions $S(0) = 99.9$, $I(0) = 0.1$, and a vector of time points ranging from 0 to 20. Plot the solution in the (S,I) phase plane.
5. [~5%] Create a vector containing the number of recovered individuals $R(t)$ at the same time points as in the previous question. Plot all three variables (S, I and R) against time, using different colours.
6. [~30%] This question aims to estimate the final epidemic size f , defined as the *proportion* of the population that gets infected. Recall from lecture 5 (Epidemic models) that f is the non-zero solution of the following equation:

$$1 - f - e^{-\mathcal{R}_0 f} = 0$$

- a. Write a function `fes_eq(f,r0)` which calculates the left-hand-side of the equation for any values of f and \mathcal{R}_0 .
- b. Setting $\mathcal{R}_0 = 2$, plot the left-hand side of the equation against f , for f ranging from 0 to 1.
- c. Setting $\mathcal{R}_0 = 2$, write a `while()` loop that goes through values of f from 1 downwards in steps of 0.001, stopping when `fes_eq(f,r0) >= 0`. What is the final epidemic size for $\mathcal{R}_0 = 2$?
- d. Using your code from questions (a) and (c), write a function `fes_sol(r0)` which returns the value of the final epidemic size f for any given value of \mathcal{R}_0 .
- e. Plot the final epidemic size as a function of \mathcal{R}_0 , for \mathcal{R}_0 ranging from 1.1 to 6. You should obtain a graph similar to Figure 5.5 in the Lecture notes. (Note that function `fes_sol` will fail for $\mathcal{R}_0 = 1$).