The Deterministic SIR Model

Class exercises
September 2, 2024

Solving differential equations using `ode()` in R

In this exercise, we will show you how to program differential equations in R, and how to solve them numerically, which is a key concept in mathematical modeling of infectious diseases. We will walk you through a basic example of a deterministic SIR model and how to solve the corresponding system of differential equations using the `ode()` function from the `deSolve` package, whereas the packages `reshape2`, `tidyverse`, and `ggplot2` will be used for data manipulation and plotting. The .R codes for the class exercises are available at GitHub repository.

PART 1: The basic deterministic SIR model

Construct a deterministic SIR (Susceptible-Infectious-Recovered) model for the following scenario: In a closed population of 1,000,000 individuals, a single person is initially infected with a novel pathogen at time t_0, with no individuals initially recovered. The basic reproduction number R_0 of this pathogen is 2, and the recovery rate ν corresponds to an average infectious period of 7 days. We aim to analyze the progression of the epidemic over a period of 730 days (i.e., 2 years). Implement the necessary code, representing the subpopulations in raw numbers and proportions.

1. *Determine how the proportions of susceptible, infectious, and recovered individuals evolve over time by plotting the trajectories of the S, I, and R compartments over the 730-day period.*
2. *What is the proportion of the population that is recovered at the end of the period?*
3. *Identify the peak proportion of infectious individuals and the day at which this peak occurs.*
4. *(Optional) How would the epidemic dynamics change if we keep the recovery rate (i.e., $\nu = 1/7$) constant but vary the basic reproduction number from 1 to 6?*

PART 2: Extensions of the SIR model: the SIRS and SIS models

Using the materials from PART 1, extend your SIR model to an SIRS and SIS model with vital dynamics, knowing that the transition rate from the R compartment to the S compartment is equal to $\sigma = 1/7$, and the life expectancy is considered to be 80 years. Implement the necessary codes, representing the subpopulations in terms of proportions.

PART 3: Adding more compartments in the SIR model: the MSEIR model

In a constant population of 1,000,000 individuals, a single person is initially infected with a novel pathogen at time t_0, and with no individuals initially recovered. Develop an MSEIR model and run it over a period of 1000 years (with a step size of 0.01 units), using both absolute numbers and proportions, and with the following parameters: life expectancy is 80 years, duration of maternal immunity is 3 months, duration

of the latent period is 2 days, duration of the infectious period is 1 week, and the basic reproduction number is equal to 2.

5. *(Optional) What would happen if one switches to a resolution of days for the time scale?*

PART 4: The basic SIR model with vaccination

In this part, we will extend the SIR model to incorporate vaccination and analyze its impact on the population. Reuse the R code from PART 1 with the following parameters: recovery rate $\nu = 1/7$, mortality rate $\mu = 1/80$, the basic reproduction number $R_0 = 3$, and the vaccination coverage $p_{vac} = 0.4$. Assuming that the vaccine has 100% efficacy, we aim to analyze the progression of the epidemic over a period of 730 days (i.e., two years). Write the codes in both numbers and proportions.

6. *What is the vaccination coverage that should prevent an epidemic in the population?*
7. *If the efficacy of this vaccine drops to only 70%, what proportion of the population would have to be vaccinated to prevent an epidemic?*

PART 5: The basic SIR model with multiple sub-populations

In this task, you will learn how to contruct an age-structured SIR model for two groups, i.e., children (group 1) and adults (group 2). For simplicity, we assume that there is no aging within the population, meaning that individuals do not transit from the children group (representing 20% of the total population) to the adult group. We also assume that both groups share the same rate of recovery $\nu = 1/7$. Given the assumption of age-specific mixing patterns, the force of infection differs between the age groups. In this model, we assume the mortality rate is equal to $\mu = 1/80$, the total population size is 1 million, and the probability of getting infection per contact is $b = 0.05$. The average number of contacts per day, retrieved from the Belgian social contact study in 2006 (SOCRATES tool) are: $c_{11} = 5.43$, $c_{12} = 6$, $c_{21} = 1.57$, and $c_{22} = 10.05$. Using this information, write code to simulate the spread of an infection over a 100-day period and plot the model output.

8. *What are the proportions of children and adults who became infected?*
9. *Extend your model to three (or more) subgroups. The social contact matrix can be obtained from the above-mentioned SOCRATES tool.*

PART 6: Serological data analysis

Consider the Belgian hepatitis A serological survey data from 2002. In the dataset the age of the individual (Age) and the immunological status (HAVres) with regard to hepatitis A is reported.

11. *Depict graphically the seroprevalence of hepatitis A in Belgium anno 2002 by age.*
12. *Estimate the age-dependent force of infection by using at least three different statistical models.*
13. *Show the fit of these three models to the observed data. How would you compare their performance in describing the observed seroprevalence data?*