

# POLITECNICO DI TORINO

---

*MSc in Data Science and Engineering*



**Politecnico  
di Torino**

Computer aided simulations and performance evaluation

**Laboratory #4**

---

*Professor:*  
Paolo Giaccone

*Student:*  
Paola Privitera

Academic Year 2020-2021

# Index

<i>Introduction</i> .....	2
<i>1 Numerical SIR model</i> .....	2
<i>2 Simulative SIR model</i> .....	3
2.1 Input parameters.....	3
2.2 Output metrics.....	3
2.3 Data structures .....	4
2.4 Algorithm .....	4
2.5 Replacement number $R_t$ .....	5

## Introduction

The scope of the laboratory is to study the evolution of an epidemic over time through the implementation of a SIR model, which is a compartmental model that divides the population into three different categories:

- Susceptible ( $S$ ): individual who is at risk of becoming infected by the disease;
- Infected ( $I$ ): individual who has the disease and can transmit it to another member of the population;
- Recovered ( $R$ ): individual who has recovered or died from the disease.

In the present work, a numerical and a simulative SIR model have been implemented.

### 1 Numerical SIR model

The numerical SIR model is a differential equation model characterized by an independent variable  $t$ , which in this case is the time measured in days, by some dependent variables and two main parameters. The dependent variables count people in each of the category defined above, as a function of  $t$ :

- $S(t)$  is the number of susceptible people at time  $t$ ;
- $I(t)$  is the number of infected individuals at time  $t$ ;
- $R(t)$  is the number of recovered people at time  $t$ .

The population size is  $N = S(t) + I(t) + R(t)$  for each  $t$ , in this case  $N = 10,000$  people.

The two main parameters are:

- $\beta$ : infection transmission rate per capita [ $\text{day}^{-1}$ ]; in this case,  $\beta = 0.2 \text{ day}^{-1}$  because it refers to the number of people with which an individual gets in contact each day;
- $\gamma$ : recovery rate, equal to the inverse of the infection duration [ $\text{day}^{-1}$ ]; since the infection period is 14 days,  $\gamma = \frac{1}{14} \text{ day}^{-1}$ .

The *reproduction number*  $R_0$  represents the average number of secondary infections that occurs when one infected individual is introduced into a completely susceptible population. For this reason, it is defined at the time of invasion.

Since the infective individual is expected to meet other people with a rate  $\beta$  during the infection period,  $R_0$  has been computed as  $R_0 = \frac{\beta}{\gamma} = 2.8$ .

The result is greater than 1, meaning that the invasion is possible, and the disease can spread through the population.

The numerical SIR model is defined through the following differential equations:

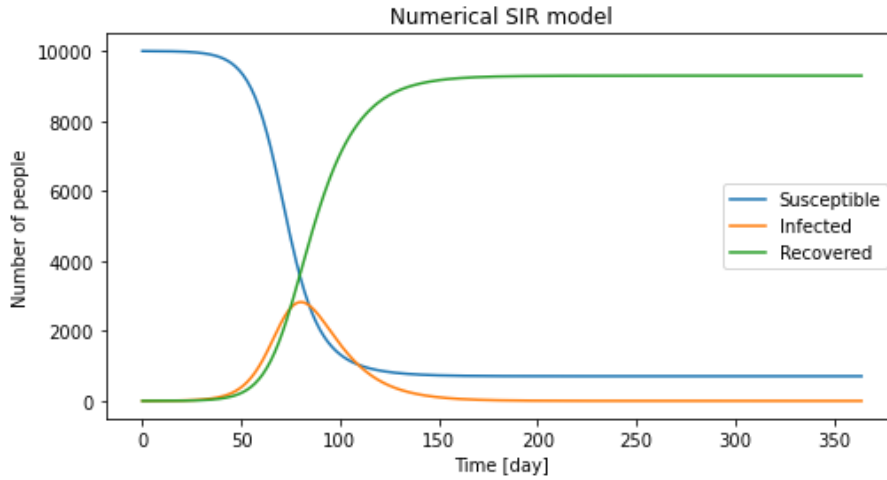
$$\begin{aligned}\frac{dS(t)}{dt} &= -\frac{\beta}{N}S(t)I(t) \\ \frac{dI(t)}{dt} &= \frac{\beta}{N}S(t)I(t) - \gamma I(t) \\ \frac{dR(t)}{dt} &= \gamma I(t)\end{aligned}$$

The first equation represents the variation of susceptible people with respect to time:  $\frac{\beta}{N}S(t)$  is the probability that an infected individual meets someone that is susceptible; this term multiplied by the number of infected people at that time gives the newly infected people. The negative sign comes from the fact that this portion of population moves from the set of susceptible to the one of infected. For this reason, the same term is also present in the second equation: the variation of infected people is given by the newly infected people minus the number of recovered individuals,  $\gamma I(t)$ . Finally, this last term characterizes the third differential equation.

The following assumptions have been made for the initial conditions, at day 0:

$$S(0) = N - 1 \quad I(0) = 1 \quad R(0) = 0$$

The values of  $S(t)$ ,  $I(t)$  and  $R(t)$  in function of the day during one year are plotted in the next picture:



As might be expected, the assumed value of  $S$  decreases over time, as susceptible individuals meet the infected ones. Such value cannot increase because the population size is fixed (no births are considered). The value for infected individuals first increases, until it reaches a peak, and then decreases as infected individuals recover from the disease or die. This implies that the value of recovered individuals rises over time, until there are no more infected, as confirmed by the graph.

The maximum number of infected people is equal to 2,823 and it occurs at day 80.

In the present work, the epidemics is assumed to end if, in a certain day, the number of infected people is less than 1, meaning that no one can be infected anymore. In this specific case, the epidemic ended after 229 days.

## 2 Simulative SIR model

The simulative SIR model consists in an agent-based simulator which is able to mimic the results of the numerical SIR model. An agent is the digital twin version of an individual of the population and each agent can virtually interact with other agents when moving around the world.

### 2.1 Input parameters

The input parameters used to run the simulator are:

- $N$ : number of people;
- $daysOfYear$ : number of days considered for the simulation;
- $\beta$ : infection transmission rate per capita [ $\text{day}^{-1}$ ];
- $\gamma$ : recovery rate [ $\text{day}^{-1}$ ];
- $runs$ : number of runs;
- $seed$ : initial seed;
- $confidence\ level$ .

### 2.2 Output metrics

The output metrics of the simulator are:

- $S(t)$ ,  $I(t)$  and  $R(t)$  for each day of the year;
- $R_t$ , replacement number, and related  $CI$ ;
- maximum number of infected people and corresponding time;
- day at which the epidemics ends.

## 2.3 Data structures

The main data structures for the simulator are:

- *people*: array of size  $N$  (= size of population) that represents the state of each individual. It takes value 0 if the individual  $i$  is susceptible, value 1 if  $i$  is infected, value 2 if recovered, with  $0 \leq i \leq N - 1$ ;
- *S*: array of length *daysOfYear* that indicates the number of susceptible people for each day of the year;
- *I*: array of length *daysOfYear* that indicates the number of infected individuals for each day of the year;
- *R*: array of length *daysOfYear* that indicates the number of recovered people for each day of the year;
- *recovery*: array of size  $N$  that takes into account how many days have passed since the beginning of the infection for each  $i$  person, with  $0 \leq i \leq N - 1$ ;
- *start\_infection*: array of size  $N$  that stores the day in which the infection started for each infected individual; it assumes value -1 in case the person does not get the disease;
- *num\_days\_for\_recovery*: array of size  $N$  that, for each person, memorizes the duration of the infection;
- *num\_contacts*: array of size  $N$  which indicates the number of people that each person meets during the current day;
- *rt\_matrix*: numpy array of shape  $(N, \text{daysOfYear})$  initialized at 0; for each infected, it memorizes the number of people that he or she infected in a specific day;
- *sums*: array of length  $N$  that stores, for each person, the total number of people that he or she infected during all the infection period; it takes value 0 if the person has never had the disease;
- *new\_rt\_matrix*: numpy array of shape  $(N, \text{daysOfYear})$  that, starting from the day on which the infection started for each infected, memorizes the number of people that he or she will infect (secondary infections) from a specific day to the end of the infectiousness; it takes value 0 in the days outside his or her infection period or if the person has never had the disease.

## 2.4 Algorithm

The algorithm starts by considering the same initial conditions of the numerical SIR model, for which the first day there is only one infected in the population, the number of susceptible is equal to  $N - 1$  and there are no recovered people. In case this individual is able to spread the disease, a meaningful instance of  $S(t)$ ,  $I(t)$  and  $R(t)$  in function of the day can be obtained. However, this is not always the case: for example, if the infected person recovers from the disease before infecting other people, degenerate outcomes occur.

For each day of the year, the number of people that each person meets (*num\_contacts*) is generated by using a geometric distribution with support that starts from 0 and mean  $\beta$ . To obtain compatible results with the numerical SIR model, the  $\beta$  value is halved ( $\beta = 0.1$ ) because there could be cases in which a person A meets a person B and, during the same day, a person C meets again A, doubling the average number of contacts of a person.

For the infected  $i$ , *num\_contacts*[ $i$ ] integer random numbers between 0 and  $N - 1$  are generated, indicating who are the people to meet that day, by keeping in mind that the considered infected cannot meet himself.

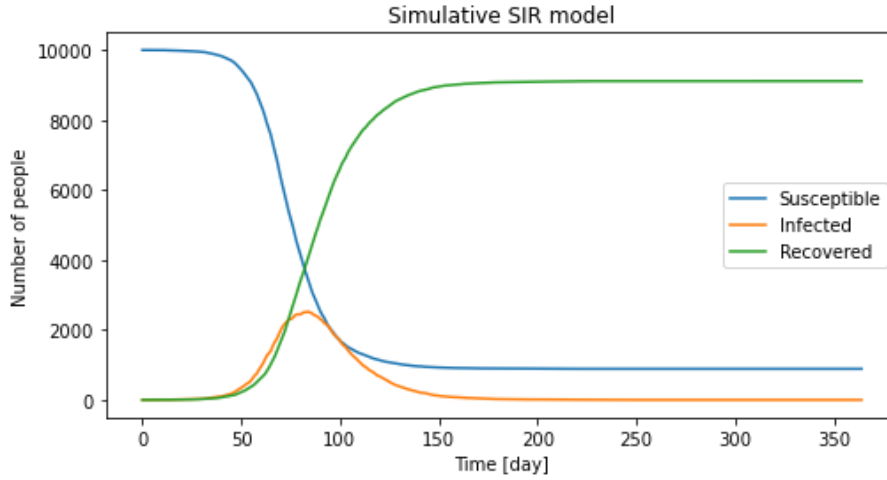
The infection occurs when the infected meets a susceptible: the respective counters are updated, the state of susceptible changes to infected and the day at which the infection starts (*start\_infection*) is stored. If the infected person meets another infected or a recovered individual, nothing happens.

The duration of the infection for each individual (*num\_days\_for\_recovery*) is estimated using a geometric distribution with parameter  $\frac{1}{\gamma}$ .

When, for a given person, the number of days passed since the beginning of the infection (*recovery*)

reaches the duration of the infection, the state of that person moves from infected to recovered and the related counters are updated.

The following graph shows a non-degenerate instance of  $S(t)$ ,  $I(t)$  and  $R(t)$  in function of the day during one year:



The time at which the maximum number of infected people achieved by simulation is similar to the one obtained with the numerical SIR model. Indeed, the peak of 2,523 infected people is reached at day 84.

By following the same approach used in the numerical model, it has been possible to observe that the epidemic ended after 245 days.

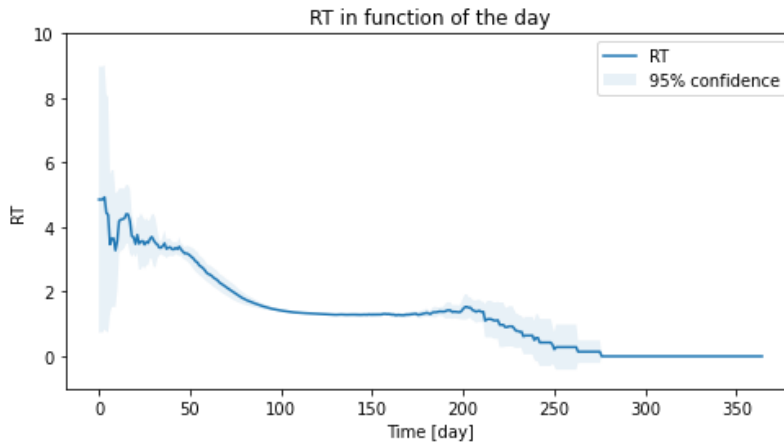
## 2.5 Replacement number $R_t$

The *replacement number*  $R_t$  is the average number of secondary infections produced by a typical infective during the entire period of infectiousness at time  $t$ . These values are stored in the *new\_rt\_matrix*.

If the total number of infected people at time  $t$  is different from 0, the  $R_t$  is given by:

$$R_t = \frac{\text{number of secondary infections from } t \text{ to the end of the infectiousness}}{\text{number of infected people at time } t}$$

The chart below shows the daily value of the  $R_t$  index over a one-year period:



Two main considerations can be drawn from the behavior of the curve: first, the  $R_t$  index decreases as time passes by, because the number of potential infections diminishes over time; secondly, the index variance is large at the beginning of the period, and then stabilizes up to around day 200, before increasing again for approximately 70 days, and eventually becoming negligible.