

# Epidemics Evolution

Luca Graziano

November 2021

## Objective definition

The purpose of this project is understanding the evolution of a stochastic process representing the spread of a COVID-19 epidemics in a population composed of 10'000 individuals. The scenario simulation presents three possible states for an individual to be in:

- S = susceptible = not infected and not immune (then s/he can be infected)
- I = infected
- R = removed = recovered and immune, or dead

In order to construct the transition matrix, it is assumed that at each time step each susceptible individual becomes infected with probability 0.18, each infected individual is removed with probability 0.53 and each removed individual returns back to be susceptible with probability 0.1.

## Implementation

The problem we are facing can be represented as an homogeneous Markov chain, whose transition matrix is the one pictured in Figure 1. The rows represent the state distributions; each entry in the matrix is the conditional probability of being in a state at time  $n+1$  (columns) given that we are in a certain state at time  $n$  (rows).

$$P = \begin{matrix} & \begin{matrix} S & I & R \end{matrix} \\ \begin{matrix} S \\ I \\ R \end{matrix} & \begin{pmatrix} 0.82 & 0.18 & 0 \\ 0 & 0.47 & 0.53 \\ 0.10 & 0 & 0.90 \end{pmatrix} \end{matrix}$$

Figure 1: Transition matrix 3x3

The first row of the matrix represents the distribution probability for the Susceptible (S) state; the second row is the distribution for the Infected (I) state, and the third row is the distribution for the Removed (R) state. This means that if you are in the susceptible state you have an 82% probability of remaining in that state, and an 18% probability of becoming infected. If you are in the infected state, you have a 47% probability of remaining infected and a 53% probability of becoming removed. Lastly, if you are in the removed state, you have a 10% probability of becoming susceptible again, and a 90% probability of remaining in the removed state.

This Markov chain is irreducible, meaning that every state intercommunicate with each other. Formally, we can write this as

$$P(X_{m+n} = S_j \mid X_m = S_i) > 0 \quad \forall S_i, S_j \in \text{state space } S$$

The following is the transition graph of our Markov Chain:

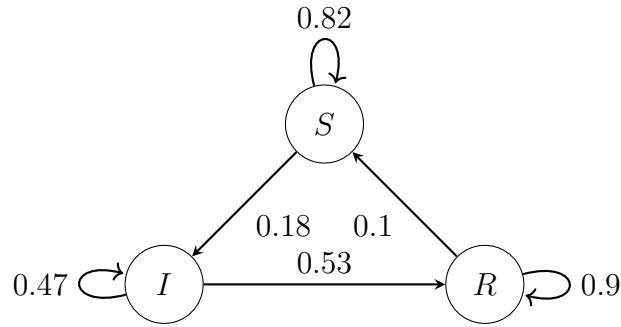


Figure 2: Transition graph

It conveys the same information contained in the transition matrix, but via a graph representation. It is not a simple graph, since at each state we have a non-zero probability of remaining in the same state, and thus we have loops; it is an oriented graph and the degree of all its vertices is four.

Since for every state we have a loop (a self-transition), we can state that every state of the Markov chain is aperiodic, and thus that the Markov chain itself is aperiodic.

$$d(s_i) = 1 \quad \forall s_i \in S$$

The initial distribution of our Markov chain is assumed to be

$$\underline{\mu}^{(0)} = [1 \ 0 \ 0]$$

meaning that the whole population of individuals starts in the susceptible state.

Since we know that our Markov chain is irreducible and aperiodic, we can use the following theorem:

**Theorem 1:** for any irreducible and aperiodic Markov chain there exists one stationary distribution.

We also know that as the number of iterations  $n$  grows to infinity, the distribution of the Markov chain tends in total variation to the stationary distribution, and that this stationary distribution is unique.

Reminding that the distribution of an homogeneous Markov Chain at a time-step  $n$  can be calculated as

$$\underline{\mu}^{(n)} = \underline{\mu}^{(0)} P^n$$

where  $P$  is the transition matrix.

In Figure 3 we can observe a representation of the evolution of the epidemics. Here, it is shown how the distribution becomes stationary after the 30th iteration. The stationary distribution of the Markov Chain after the 30th iteration is given by the following vector  $\underline{\mu}^{(n)}$

$$\underline{\mu}^{(n)} = [0.319 \quad 0.108 \quad 0.573]$$

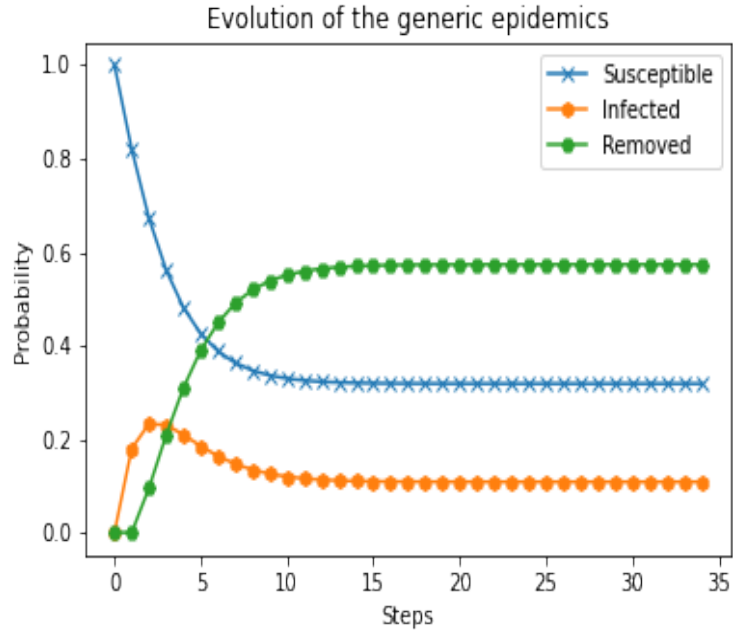


Figure 3: Evolution of the Markov chain

Assuming the population of our Markov chain was  $N = 10,000$ , this distribution would correspond to

$N_s = 3203$  (number of susceptible subjects at an arbitrary large time-step)

$N_i = 1095$  (number of infected subjects at an arbitrary large time-step)

$N_r = 5702$  (number of infected subjects at an arbitrary large time-step)

In Figure 4, it is represented a realisation of the markov chain, which is a simulation of the evolution of an actual epidemics. The simulation scenario is done on a vector of 10'000 people for 75 iterations, but we can see the re-alisation becoming stationary at around the 20th iteration. Contrary to the generic simulation of markov chain, the realisation never really becomes perfectly stationary and will always present a randomic component that makes it slightly volatile even once stationarity has been reached.

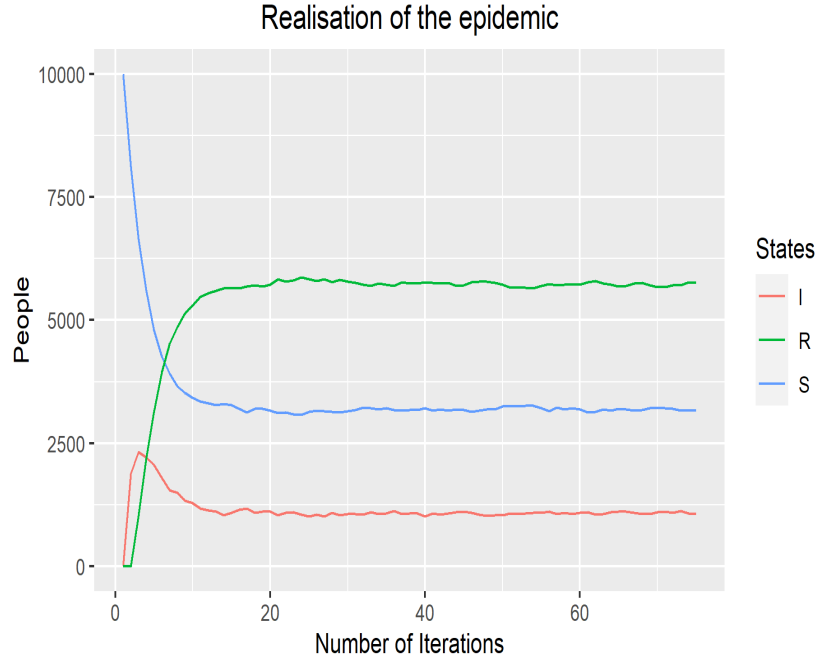


Figure 4: Realisation graph

Due to the randomic component, the number of individuals at each state cannot be perfectly forecasted but we can see that it follows the generic distribution. This realisation at the 75th iterations presents a distribution of:

$N_S = 3180$  (number of susceptible subjects at an arbitrary large time-step)  
 $N_I = 1061$  (number of infected subjects at an arbitrary large time-step)  
 $N_R = 5759$  (number of infected subjects at an arbitrary large time-step)

Finally the variation distance, which is an indicator of how dissimilar two stochastic vectors are, between the vector(g), representing the stationary distribution of the generic Markov chain, and the vector(r), representing the realisation of the epidemic, can be computed as follows:

$$\text{distance variation} = \frac{1}{2} \sum_{i=1}^3 |\vec{v}_i^{(g)} - \vec{v}_i^{(r)}| = 0.00258$$