Stochastic Simulation

Autumn Semester 2024

Prof. Fabio Nobile *Assistant: Matteo Raviola*

Project - 10

Submission deadline: 16 January 2025

Stochastic simulation of epidemic models

This mini-project concerns stochastic simulation of epidemic models. There are multiple advantages of considering stochastic models over their deterministic counterparts in epidemiology. First, the insurgence of a contagion when two or more individuals are in contact seems to be better described as a probabilistic rather than a purely deterministic process. Second, it is difficult in deterministic models to admit the possibility of the sudden extinction of the epidemic or the possibility of a minor epidemic outbreak. On the other hand, stochastic models can naturally estimate quantities such as the probability of extinction or the probability that the epidemic lasts longer than a certain time.

1 Introduction and background

Let us consider a closed population of size N partitioned into three non-overlapping classes: S (susceptible), I (infectious) and R (removed). Assume that individuals in the I-class are at the same time infected and infectious, i.e., they are able to transmit the disease to susceptible individuals. Contagious contacts only can happen when an individual from the S-class meets an infectious individual. Infected individuals recover after an exponentially distributed random time and gain immunity to the disease, thus becoming individuals of the R-class. In the model considered in this project we assume that immunity lasts forever; that is, once individuals enter the R-class, they are no longer susceptible to the infection. Since we are considering a closed population in this simple SIR model, the sum $S + I + R$ is constant in time and equals the total number N of individuals in the population. Therefore, we just need to keep track of the sizes of S and I classes. Denote by $(s, i) \in \mathbb{N}^2$ the number of susceptible (s) and infected (i) people at a given time. The (stochastic) SIR model can be then modeled by the following discrete-state, continuous time Markov process $\{X(t) = (S(t), I(t)) \in \{0, 1, \dots, N\}^2,\ t > 0\}$ with transition probabilities:

$$\begin{aligned} &\mathbb{P}\left(X(t + \mathrm{d}t) = (s, i) + (-1, 1) | X(t) = (s, i)\right) = \beta si\mathrm{d}t + o(\mathrm{d}t) \\ &\mathbb{P}\left(X(t + \mathrm{d}t) = (s, i) + (0, -1) | X(t) = (s, i)\right) = \gamma i\mathrm{d}t + o(\mathrm{d}t), \end{aligned} \tag{1}$$

where $\beta, \gamma \in \mathbb{R}^+$.

When $N, S(t), I(t) \geq 1$, are large, so that $(S(t), I(t)) + (-1, 1) \approx (S(t), I(t))$ and $(S(t), I(t)) + (0, -1) \approx (S(t), I(t))$, i.e., a single reaction changes very little the state, one can obtain the

so-called *mean field approximation of SIR* which can be written in differential form as

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= -\beta S(t)I(t), \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta S(t)I(t) - \gamma I(t), \end{aligned} \tag{2}$$

with $(S(0), I(0)) = (S_0, I_0)$, where we assume here $S(t), I(t) \in \mathbb{R}^+ \quad \forall t$. A better approximation is given by the following *diffusion approximation of SIR*:

$$\begin{aligned} \mathrm{d}S &= -\beta S(t)I(t)\mathrm{d}t - \sqrt{\beta S(t)I(t)}\mathrm{d}W_S \\ \mathrm{d}I &= (\beta S(t)I(t) - \gamma I(t))\mathrm{d}t + \sqrt{\beta S(t)I(t)}\mathrm{d}W_S - \sqrt{\gamma I(t)}\mathrm{d}W_I \end{aligned} \tag{3}$$

with $(S(0), I(0)) = (S_0, I_0)$, and where W_S, W_I are two independent standard Wiener processes.

A more realistic epidemic model, hereafter named SIR-d, takes into account demographic effects, i.e., the death and birth of the population. Its mean field approximation is given by

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= m(S(t) + I(t) + R(t)) - mS(t) - \beta S(t)I(t), \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \beta S(t)I(t) - \gamma I(t) - (m+v)I(t), \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \gamma I(t) - mR(t), \end{aligned} \tag{4}$$

where m is the host death/birth rate and v is the pathogen-induced mortality rate. The term that describes the birth of susceptible hosts, $m(S + I + R)$, ensures that deaths due to non-pathogen-related causes are balanced, and the total population $(S + I + R)$ remains constant over time, as long as there is no death due to the epidemic (expressed by $-vI$). Thus, at any given time t, one of the following 6 processes a_i can happen: (a_1) host birth, (a_2) death of susceptible host, (a_3) death of infected host, (a_4) death of recovered host, (a_5) infection, (a_6) Recovery.

Goals of the project

(a) We begin by comparing the dynamics generated by (1), (2) and (3). To that end, consider a small population of $N = 100$ individuals and initial condition $(S_0, I_0) = (99, 1)$. Implement both the stochastic SIR (1), its mean field (2) and diffusion (3) approximations for a time $T = 10$. In particular, simulate exactly the stochastic SIR, and use suitable time discretizations for the mean field and diffusion models. Plot the resulting trajectories. Run your simulations for $\beta = 0.02$ and $\gamma = 0.4$.

(b) In the stochastic SIR model, the disease is considered extinct whenever $I(t) = 0$ for some $t \in [0, T]$. Using the same setting as in the previous point, use the SIR model (1) and a Monte Carlo approach to estimate the probability of extinction of the disease at time T, that is, estimate $\mathbb{P}(I(t) = 0)$ for some $t \leq T$. Report your values for $T = 1, 2, 10$. Choose appropriately the sample size to guarantee a relative error of 5%.

(c) Propose and implement a variance reduction technique for your previous estimate. Discuss your results. Test now your proposed method to estimate the probability of extinction at $T = 2$, starting from the initial condition $(S_0, I_0) = (95, 5)$. Comment the results and suggest possible improvements, if needed.

(d) Derive the formulation of the stochastic SIR-d process associated to (4). **Hint:** What is the rate of each (a_i), $i = 1, 2, \ldots, 6$?

(e) Simulate then the SIR-d process and compare to its mean field approximation (4). Set $T = 10$, $\beta = 0.02$, $\gamma = 0.4$ and experiment with different values of $m \in [10^{-4}, 10^{-3}]$ and $v \in [10^{-3}, 10^{-2}]$. Compare your results.

(f) In a deterministic model an epidemic will go extinct (i.e., $I(t) \to 0$ as $t \to \infty$) if the basic reproduction number, R_0 of the infection, given by

$$R_0 = \frac{\beta(S(0) + I(0) + R(0))}{m + v + \gamma},$$

is less than one. Consider the stochastic SIR-d model with demographic effects. Is it possible that the epidemic will be extinct at a time T for $R_0 = 1.01, 1.05, 1.1, 1.5$? Consider a population of $N = 100$ individuals with $R(0) = 0, S(0) = 95, I(0) = 5, m = 10^{-4}, v = 10^{-2}$ and β chosen to obtain the previous values of R_0 and estimate the probability of extinction for $T = 1, 2, 10$.

References

[1] Priscilla E Greenwood and Luis F Gordillo. Stochastic epidemic modeling. In *Mathematical and statistical estimation approaches in epidemiology*, pages 31–52. Springer, 2009.

[2] Alvaro Moraes. *Simulation and statistical inference of stochastic reaction networks with applications to epidemic models, Chapter 2.* PhD thesis, King Abdullah University of Science and Technology, 2015.