UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE
ISC2 - Introduction à la modélisation stochastique (Automne 2024)

Project
Simulation of an SIR stochastic process in continuous time

Instructions:

- Written report (4-5 pages) due by midnight on **Friday, November 15.**
  In your written report, please answer the questions below carefully, describing and explaining the methods and algorithms used.
- Oral project discussion **Friday afternoon, November 22** (30min per team: 10min oral presentation on ppt or pdf + questions)
  Precise chedules to be defined

Introduction to compartmental models in epidemiology

A mathematical model is a simplified description of a pheonomon of interest, enabling it to be studied using mathematical techniques. In the case of diseases that spread through a population, models enable us to better understand the transmission process, explain the observed data, and predict the future of the epidemic. These models can help, for example, to answer key questions for the Public Health System, such as the duration of an epidemic or the number of hospital beds that will be needed over time.
A common assumption in many transmissible disease models is to consider the population as a homogeneous group of individuals, randomly mixed and divided into separate epidemiological classes, according to their epidemic status. This seminal approach was first proposed in 1927 by W.O. Kermack and A.G. McKendrick, and is known as *Compartmental Models*.

There are many variants of these models, but in general we find the following 3 classes: S is used to represent individuals *susceptible* of contracting the disease; I designates individuals *infectious*, i.e. those capable of transmitting the disease; and R (*removed* in English) is the compartment used for those who have been infected but can no longer transmit (because they have recovered, died, etc.). Other classes can be added to this basic model to reflect the reality of a particular epidemic.
Transitions between different classes of individuals over time can be described in either a deterministic or a stochastic way. We are going to focus on the latter case, and in particular on the Markovian framework. See [1, 2, 3] for a detailed presentation of these models.
In [2] the usefulness of stochastic epidemic models is investigated, by comparing them with deterministic models. An important qualitative difference between so-called deterministic and stochastic epidemic models is asymptotic dynamics. In addition, stochastic models also take into account the possibility of disease extinction within a finite timeframe, so that the expected time to disease extinction can be calculated. It has also been observed that stochastic models better capture the

uncertainty and variability inherent in real epidemics, due to factors such as the unpredictability of person-to-person contact.

1 The SIRD Markovian model (*general stochastic epidemic model*)

As described above, we will consider 4 classes of individuals: S, I, R and D, where R stands for recovery (with acquired immunity, so that they are not susceptible anymore) and D for death. We will denote respectively $S(t), I(t)$, $R(t)$ and $D(t)$ the number of individuals in each class at time $t \geq 0$. We assume that the population size is constant over time, i.e. $N = S(t)+I(t)+R(t)+D(t)$ for all $t \in \mathbb{T}$. The population can be described by the continuous-time Markov process $(S, I, R) = \{(S(t), I(t), R(t)); t \in \mathbb{T}$ with values in the set $E = \{(s, i, r) \in \{0, 1, 2, \dots N\}^2 : s+i+r \leq N\}$. From now on, we will assume that each infected individual has a transmission rate β/N per contact with a susceptible individual, a rate γ to recover and a rate ν to die from the infection. Thus, the transition rates of the Markov chain (S, I, R), to go from state $(s, i, r) \in E$ to another state $(s + k, i + j, r + l) \in E$ are

$$Q_{(s,i,r),(s+k,i+j,r+l)} = \begin{cases} \dfrac{\beta si}{N} & \text{if } (k, j, l) = (-1, 1, 0), \\ \gamma i & \text{if } (k, j, l) = (0, -1, 1), \\ \nu i & \text{if } (k, j, l) = (0, -1, 0). \end{cases}$$

Notice that the state $(s, 0, r)$ is absorbing for all $s, r \in \{0, 1, \dots N\}$.

1. Write the complete transition probability matrix for this Markov chain for $N = 3$ for an infinitesimal time-step $\Delta t > 0$.

2. Simulate 100 trajectories of the SIR Markov process up to absorption, with $N = 1000$, $(S(0), I(0), R(0)) = (990, 10, 0)$, $\beta = 0.2$, $\gamma = 0.1$, and $\nu = 0.001$. Remember to save the simulated trajectories (values of (S, I, R) over time) so that you can answer the following questions. Explain the algorithm used for the simulations.

3. Calculate the empirical mean and variance of the epidemic duration (defined as $T_{\text{ext}} = \inf\{t \geq 0 : I(t) = 0\}$).

4. Calculate the empirical mean and the empirical variance of the final size of the epidemic, which is equal to $N - S_{T_{\text{ext}}}$ (corresponding to the number of individuals who were infected between the start of the epidemic at time 0 and the end of the epidemic at time T_{ext}).

5. Graph 10 simulated trajectories of (S, I, R, D) in the same graph (you can use different colors for each compartment).

For epidemic models, a major concern is to find the conditions under which a disease introduced into a community will develop into a large-scale epidemic, and if so, under what conditions the disease may become endemic. For stochastic models, all these questions are formulated in terms

of probabilities. A useful parameter in this respect is the *basic reproduction number (or ratio) of an epidemic*. Usually denoted R_0, this is the average number of individuals infected by a single infectious individual while he or she is contagious. The R_0 is calculated at the start of the epidemic, before the effects of saturation in the population (the number of susceptible individuals decreases over time in a population of fixed size). In the SIR model just described, $R_0 = \frac{\beta}{\gamma+\nu}$. If the number of basic reproductions is less than or equal to one, with a high probability the epidemic will be relatively small. For this reason, most studies of these examples focus on the complementary case. Without doubt, the most interesting case is the one where R_0 is close to one.

6. Propose a method for estimating the value of R_0 from contact real data, and support your choice with a discussion of data collection in the context of an epidemic.

In the SIR model described below, we assume that an individual can transmit the disease immediately after being infected. For many transmissible diseases this assumption is not realistic (e.g. Covid-19).

7. Propose a modification to the SIR model, by adding a class, to be able to take into account a delay between the moment when an individual is contagious and the moment from which he can transmit. Give the state space and transition rate matrix $\widetilde{Q}$ corresponding to this new model.

References

[1] H. Andersson and T. Britton. *Stochastic epidemic models and their statistical analysis*, volume 151 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2000.

[2] F. Brauer, P. van den Driessche, and J. Wu, editors. *Mathematical Epidemiology*. Springer Berlin Heidelberg, 2008.

[3] T. Britton and E. Pardoux, editors. *Stochastic Epidemic Models with Inference*. Springer International Publishing, 2019.