TMA4265 Stochastic Modelling – Fall 2019
Project 1

Background information

- The deadline for the project is Sunday October 6 23:59.
- This project counts 10% of the final mark in the course.
- This project must be passed to be admitted to the final exam.
- A reasonable attempt must be made for each problem to pass this project.
- The project should be done in groups of **two** people. Please sign up as a group in Blackboard before submitting your report and code.
- The project report should include equations with calculations, plots and interpretation as text. The computer code should be submitted as a separate file. Make sure this code runs. We may test it.
- There is a **6 page limit** for the project report. If you submit a longer report, we will not read it. The 6 page limit does not include the computer code, which should be submitted as a separate file.
- Make your computer code readable and add comments that describe what the code is doing.
- You are free to use any programming language you want as long as the code is readable, but you can only expect to receive help with `R`, `MATLAB` and `Python`.
- Lectures on October 1 and October 2 will be used for work on the project, and you can get assistance in the lecture room these days. The exercise classes on September 26 and October 3 takes place as normal and you can receive help with the project.
- If you have questions outside the aforementioned times, please contact `susan.anyosa@ntnu.no` or `rasmus.erlemann@ntnu.no`.

Problem 1: Modelling the outbreak of measles

We use a simplified model where each individual only has three possible states: susceptible (S), infected (I), and recovered and immune (R). Measles is an infectious disease, and we assume that each day a susceptible individual can become infected or remain susceptible, and an infected individual can become recovered or remain infected. We model on a daily scale and let $n = 0, 1, \ldots$ denote time measured in days.

In the first stage of modelling, we assume that the individuals in the population are independent, and assume that each day, any susceptible individual has a probability $0 < \beta < 1$ of becoming infected tomorrow and any infected individual has a probability $0 < \gamma < 1$ of becoming recovered tomorrow.

a) Consider one specific individual, and let X_n be the state of that individual at time n. Let the states 0, 1 and 2 correspond to S, I and R, respectively, and assume that $X_0 = 0$. Explain why X_n is a Markov chain and explain why the transition probability matrix is

$$\mathbf{P} = \begin{bmatrix} 1-\beta & \beta & 0 \\ 0 & 1-\gamma & \gamma \\ 0 & 0 & 1 \end{bmatrix}.$$

b) Answer the following questions about X_n.

- Is this a reducible or irreducible Markov chain?
- Determine the equivalence classes and determine whether they are recurrent or transient.
- What is the period of each state?

c) Assume that $\beta = 0.05$ and $\gamma = 0.20$. Calculate (by hand) the expected time until a susceptible individual becomes infected and calculate (by hand) the expected time until an infected individual becomes recovered.

d) Estimate the expected values calculated in **c)** by writing computer code that simulates the Markov chain 1000 times until the recovered state is reached. How similar are the results achieved by simulation to the exact answers in **c)**?

Due to the highly infectious nature of measles, the proportions of susceptible, infected and recovered individuals in the population will change with time. This in turn means that it is highly unrealistic to assume that β does not change with time. Assume that the total population $T = 1000$ is constant through time and at each time step consists of S_n susceptible individuals, I_n infected individuals, and R_n recovered individuals.

Assume that for each time step n, the probability that a susceptible individual becomes infected is $\beta_n = 0.5 I_n / T$ and that the probability that an infected individual recovers is $\gamma = 0.20$. We assume that the $T = 1000$ individuals change state independently of each other at each time step given the values of β_n and γ.

e) Consider the Markov chain $Y_n = (S_n, I_n, R_n)$, and assume that $Y_0 = (950, 50, 0)$. Write code that simulates the Markov chain until time step $n = 200$. Choose one realization and show the temporal evolution of S_n, I_n and R_n together in one figure.

Hint: At each time step, the number of new infected individuals can be simulated as Binomial(S_n, β_n) and the number of new recovered individuals can be simulated as Binomial(I_n, γ) using `rbinom` in `R`.

f) Based on 1000 simulations of the outbreak for time steps $n = 0, 1, \ldots, 200$, estimate the expected maximum number of infected individuals, $\mathrm{E}[\max\{I_0, I_1, \ldots, I_{200}\}]$, and the expected time at which the number of infected individuals first takes its highest value, $\mathrm{E}[\min\{\arg\max_{n \le 200}\{I_n\}\}]$.

Problem 2: Insurance claims

Let $X(t)$ denote the number of claims received by an insurance company in the time interval $[0, t]$. We will assume that $X(t)$ can be modelled as a Poisson process. t is measured in days since January 1st, 0:00.00.

a) Assume the intensity is given by $\lambda(t) = 1.5$, $t \geq 0$. What is the probability that there are more than 100 claims before March 1 (59 days)? Verify your calculations by simulating 1000 realizations from the Poisson process. Make a figure that shows 10 realizations of $X(t)$, $0 \leq t \leq 59$.

Assume that the monetary claims are independent, and independent of the claim arrival times. Each claim amount (in mill. kr.) has an exponential distribution with rate parameter $\beta = 10$. This means that claim $C_i \sim \text{Exp}(\beta)$, $i = 1, 2, \ldots$. The total claim amount at time t is defined by $Z(t) = \sum_{i=1}^{X(t)} C_i$.

b) Compute the expected total claim amount and the variance of the total claim amount at March 1 (59 days) through the law of total expectation and the law of total variance. Estimate the expected value and the variance by 1000 computer simulations and compare to the true values.