Project 1

TMA4265

Clara Panchaud
Silje Anfindsen

October 4, 2019

This is the project report for Project 1 in Stochastic Modelling. The computer code for all the problems can be found in the file: "Project1.ipynb".

1 Problem 1: Modelling the outbreak of measles

In problem 1 we look at a discrete sample space consisting of three possible states, $X_n \in [0, 1, 2]$, where X_n denotes the state of an individual at time n. The states 0,1,2 corresponds to being susceptible, infected, and recovered/immune. We let $n = 0, 1, 2, 3...$ denote time measured in days.

1.1 a)

A Markov Chain is a stochastic process which satisfies the Markov Property. In other words, the process is memoryless which means that the probability of being in a state tomorrow depends only on the present state, not on the past states. In this problem we are only aware of the probability of being in a state tomorrow given today′s state. Even if we knew about the past states, it would not affect the future states in this model. Since we do not care about the past, we can assume that X_n is a Markov Chain.

The transition probability matrix consists of the transition probabilities defined as:

$$P_{ij} = P\{X_{n+1} = j \mid X_n = i\}. \tag{1}$$

Each day, any susceptible (0) individual has a probability β of becoming infected tomorrow (1). This gives us $P_{01} = \beta$. Since it is not possible for an individual to go from being susceptible to recovered without first being infected, we have that $P_{02} = 0$. We know that the transition probability matrix require each row sum to be one, $\sum_j P_{ij} = 1$ for all states i. This gives us the last state in row 0, $P_{00} = 1 - \beta$. The same strategy follows for row 1, where the individual can go from being infected to recovered, but not the opposite way. In the last row, the individual can only go from being recovered to staying recovered, this is called an absorbing state.

$$\begin{bmatrix} 1-\beta & \beta & 0 \\ 0 & 1-\gamma & \gamma \\ 0 & 0 & 1 \end{bmatrix}$$

1.2 b)

Let's start by finding whether the Markov chain is reducible or irreducible. We recall that in a Markov chain we have an equivalence relation $\sim$ such that two states are in the same equivalence class if they are accessible from each other. The Markov chain is irreducible if $\sim$ induces exactly one equivalence class. From the matrix we can build a visualisation of the states in Figure 1. It is clear that once we leave any of the states there is no coming back. Therefore we have three equivalence classes: $\{0\}, \{1\}$ and $\{2\}$. This implies that we have a reducible Markov chain.

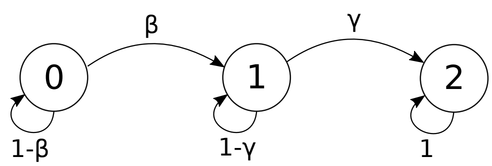

Figure 1: Markov Chain

Remember that a state is recurrent if $f_{ii} = 1$ where f_{ii} is the probability of ever returning to state i. If $f_{ii} < 1$ there is a positive probability of never returning and the state is called transient. A theorem from the lecture states that a state is recurrent if and only if

$$\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty \tag{2}$$

The period of a state is defined by

$$d(i) = gcd\{n \geq 1 : P_{ii}^{n} > 0\} \tag{3}$$

we will now go through each state to figure out if it is recurrent or transient, and then find its period.

- $\{0\}$ For this state we can use Figure 1 to see that once we leave state 0 there is no way of coming back. This implies that there is a probability β of leaving and never coming back. $f_{00} \leq 1 - \beta < 1$ which makes 0 a transient state.

  Since once we leave state 0 we never come back, the only way to be in state 0 is to stay there, so we have $P_{00}^{(n)} = (1-\beta)^n > 0$. Then $d(0) = gcd(1, 2, 3...) = 1$

- $\{1\}$ We use the same argument as for state 0 to say that state 1 is transient. Indeed $f_{11} \leq 1-\gamma < 1$.

  We follow the same procedure as for state 1 to say that $P_{11}^{(n)} = (1-\gamma)^n > 0$ and $d(1) = gcd(1,2,3...) = 1$.

- $\{2\}$ From state 2 we stay in state 2 with probability 1 so $P_{22}^{(n)} = 1 \quad \forall n \geq 1$. To be rigorous we are going to use the theorem stated previously. We can calculate $\sum_{n=1}^{\infty} P_{22}^{(n)} = \sum_{n=1}^{\infty} 1 = \infty$ so 2 is a recurrent state.

  Since $P_{22}^{(n)} = 1 \quad \forall n \geq 1$ we have that $d(2) = gcd(1,2,3...) = 1$.

1.3 c)

We will now assume that $\beta = 0.05$ and $\gamma = 0.20$ We will calculate two cases.
- Case 1: Expected time until a susceptible (0) individual becomes infected (1)
- Case 2: Expected time until an infected (1) individual becomes recovered (2).

Let

$$T_{ij} = \min\{n > 0 : X_n = j, X_0 = i\} \tag{4}$$

denote the time to move from state i to state j. We want to find the expected time for case 1.

$E[T_{01}] = \sum_{k=0}^{2}(E[T_{k0}]+1)P_{0k} = P_{00}(1+E[T_{01}]) + P_{01} = \frac{P_{00}+P_{01}}{(1-P_{00})} = 1/\beta = 20$

Now we want to find the expected time for case 2. We follow the same procedure as for case 1.

$$E[T_{12}] = P_{11}(1+E[T_{12}]) + P_{12} = 1/\gamma = 5$$

1.4 d)

We will now write computer code that simulates the Markov chain 1000 times for both cases to be able to confirm the theoretical answers in 1c).

When simulating the process 1000 times, the expected time for case 1 is $E[T_{01}] = 21,26$ for one run of the code. The result from the simulation is clearly very close to the theoretical result which is $E[T_{01}] = 20$.

For case 2 the expected value by computer simulation is $E[T_{01}] = 5,82$ for one run. This is also very close to the result from 1c) which is $E[T_{01}] = 5$.
We conclude that the results from c) seem to be correct.

1.5 e)

We will now consider the Markov chain $Y_n = (S_n, I_n, R_n)$ and simulate it for $n = 200$ days. The population is constant, $T = 1000$. Assume that for each time step n, the probability that a susceptible individual becomes infected is

$\beta_n = \frac{0,5 I_n}{T}$, and that the probability that an infected individual recovers is $\gamma = 0.20$.

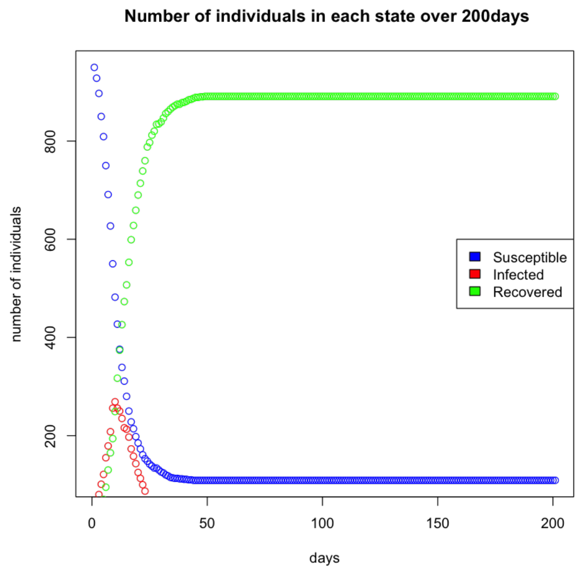

Figure 2: One realization of the Markov Chain Y_n

To consider how the individuals change states during 200 days, we will look at one realization, see Figure 2. As there are more and more infected people in the population there is an increasing probability that a susceptible individual becomes infected. Therefore it is natural that the number of infected people increases as the number of susceptible people decreases during the first days. After a while the infected individuals becomes recovered such that they can not infect the susceptible people. Therefore the number of infected people start to decrease as there are more recovered people. We can see that after about 50 days the states in the population is stable, then people are either recovered or susceptible.

1.6 f)

We will now estimate the expected maximum number of infected individuals and the expected time at which the number of infected individuals first takes its highest value. Look at the computer code to see how this is done. The results based on 1000 simulations are:

$$E[\max\{I_0, I_1, ..., I_{200}\}] = 275.457$$

$$E[\min\{\max_{n\leq 200} I_n\}] = 12.348$$

2 Problem 2: Insurance claims

In problem 2 we model a Poisson process $X(t)$ that denotes the number of claims in the time interval $[0, t]$. We start on January 1st and look at the interval $[0, 59]$, ending on March 1st. The intensity of the Poisson process is $\lambda = 1,5$. The amount of money for each claim is denoted by C_i, which follow the exponential distribution with rate parameter $\beta = 10$. We assume that the C_i are independent, and independent of the claim arrival time. The total claim amount is $Z(t) = \sum_{i=1}^{X(t)} C_i$.

2.1 a)

In this task we calculate the probability that there are more than 100 claims before March 1st. This means we are looking for $P(X(59) > 100)$.

Recall that for a Poisson process $X(t) = X(t) - X(0) \sim Poisson(\lambda t)$. Therefore $P(X(59) > 100) = 1 - P(X(59) - X(0)) = 1 - \sum_{x=0}^{100} e^{-\lambda} \frac{\lambda^x}{x!} = 0.10282$

Doing a 1000 simulations of this Poisson process we get a probability of 0.119 which is pretty close to what we calculated.

10 realizations of the process are found in Figure 3. Indeed most of those processes are below a 100 claims by March 1st.

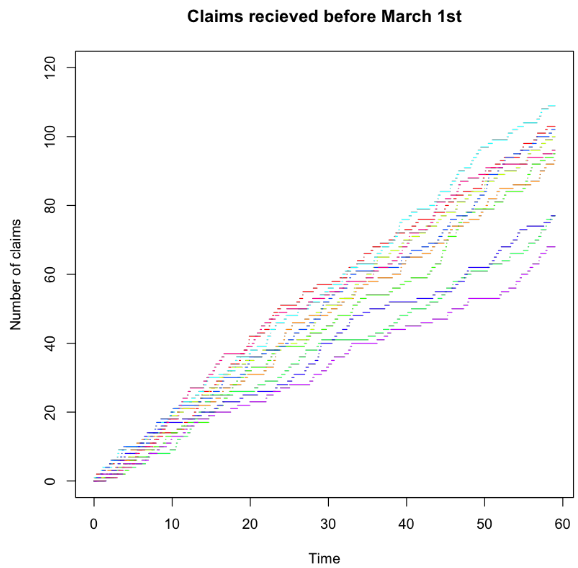

Figure 3: 10 realizations of $X(t)$, $0 \leq t \leq 59$

2.2 b)

Now we use the law of total expectation to calculate the expected total claim amount.

$E[Z(t)] = E[E[Z(t)|X(t)]] = E[E[\sum_{i=1}^{X(t)} C_i]] = E[\sum_{i=1}^{X(t)} E[C_i]] = E[\sum_{i=1}^{X(t)} \frac{1}{\beta}] =$

$E[X(t)\frac{1}{\beta}] = \frac{1}{\beta}E[X(t)] = \frac{1}{\beta}\lambda t = 8.85$

Now we use the law of total variance for the variance of total claim amount.

$Var(Z(t)) = E[Var(Z(t)|X(t))] + Var(E[Z(t)|X(t)]) = E[Var(\sum_{i=1}^{X(t)} C_i)]$

$+Var(E[\sum_{i=1}^{X(t)} C_i]) = E[\sum_{i=1}^{X(t)} Var(C_i)] + Var(\sum_{i=1}^{X(t)} E[C_i]) = E[X(t)\frac{1}{\beta}^2]+$

$Var(X(t)\frac{1}{\beta}) = \frac{1}{\beta}^2 E[X(t)] + \frac{1}{\beta}^2 Var(X(t)) = \frac{1}{\beta}^2 \lambda t + \frac{1}{\beta} a^2 \lambda t = 2\frac{1}{\beta}^2 \lambda t = 1.77$

When calculated over 1000 simulations, we get 8.87 for the expected value and 1.65 for the variance. Both values are close to the real values.