Project 1 - TMA4265 Stochastic Modelling

Jorstad, E., Nilsen, F., Sundhaug, P. R.

September 2020

Problem 1

a)

A Markov chain is a stochastic process that satisfies the Markov property, meaning that the transition probability is only dependent on the current state and not any earlier ones. We can see that the transition probabilities in step $t+1$ only depends on what state you are in at time t, meaning that it satisfies the Markov property. This seems to be logical in terms of the infection development we aim to model, since the probability of an infected individual to recover does not depend on previous states, e.g. how long the individual was susceptible before becoming infected is irrelevant. The transition probability matrix is given as

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

Figure 1 shows a visualization of the Markov chain. We observe that the probability of a susceptible becoming infected is β, and it makes sense that you cannot become directly recovered from being susceptible. Therefore that probability is zero. Furthermore, each row needs to have total probability of 1, so therefore the probability of staying susceptible is $1-\beta$. The probability of transitioning from infected to recovered is γ. Likewise, it is logical that you cannot become susceptible directly from being infected, thus the probability is 0. The same arguments can be made for the probabilities for transitioning from the recovered state. Hence, the transition probability matrix satisfies the Markov property, and $\{X_n : n = 0, 1, \dots\}$ is a Markov chain.

b)

Firstly it is practical to look into whether or not the states of our transition probability matrix communicates or not. Note that state 1 and 2 is trivially accessible from 0. Furthermore, 2 is trivially accessible from 1 and 0 is accessible from 1 through 2. Thus, $0 \sim 1 \sim 2$ and we only have one equivalence class. Note that, if there were no way of becoming susceptible (state 0) after being recovered (state 2), we would have several equivalence classes.

- By the definition of irreducibility, since only one equivalence class is induced, the Markov chain $\{X_n : n = 0, 1, \dots\}$ is irreducible.
- We recall that if i is recurrent, then

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

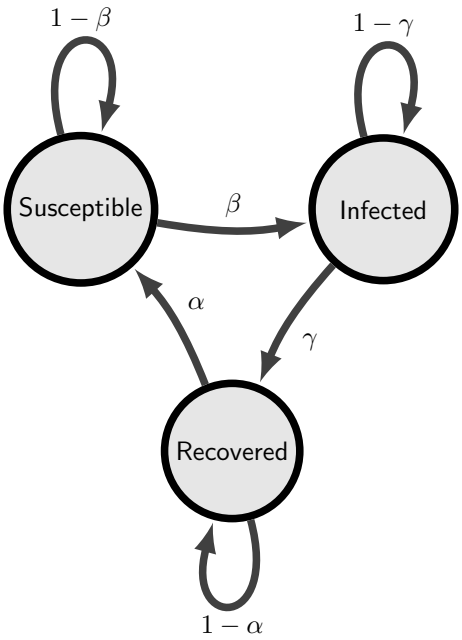

Figure 1: Visualization of the Markov chain with transition probability matrix P.

When discussing periodicity below, we show that for this Markov chain, $P_{ii}^{(n)} > 0 \ \forall \ n \geq 1$. Summing positive nonzero terms infinitely many times must result in an infinite number. Thus, the equivalence class must be recurrent.

- According to the Markov chain, it it will always be possible to stay in each state, e.g, $P_{00} = 1-\beta$ for state 0. Likewise, the probability of staying in infected and recovered state is $1-\gamma$ and $1-\alpha$, respectively. Thus, $P_{ii}^{(n)} > 0 \ \forall \ n \geq 1$. By the definition of aperiodicity,

$$d(i) = \gcd\{1, 2, \dots\} = 1, \quad i = \{0, 1, 2\},$$

meaning that all states are aperiodic.

c)

- A susceptible person can either be infected, with probability β or stay susceptible, with probability $1-\beta$. Let T_{SI} be the number of days until a susceptible person gets infected. Since T_{SI} is a Bernoulli process, $T_{SI} \sim Geo(\beta)$, and thus

$$E\left[T_{SI}\right] = \frac{1}{\beta} = \frac{1}{0.05} = 20$$

The expected time is 20 days.

- Because it is not possible to transfer directly from infected to susceptible, the expected time to go from susceptible to recovered, T_{SR}, must be the sum of the expected time to go from

State		Expected time (days)	
From	To	Simulated	Analytical
S	I	21.066176	20
S	R	30.161765	30
S	S	134.154412	130

Table 1: Comparison between the analytical and simulated values from problem 1d). Compares expected time from transitioning from susceptible (S) to infected (I), thereafter recovered (R), and finally completing an entire cycle and becoming susceptible again (S).

susceptible to infected and the expected time to go from infected to recovered.

$$E\left[T_{SR}\right] = \frac{1}{\beta} + \frac{1}{\gamma} = 30$$

The expected time is 30 days.

- By the same logic, the expected time to complete a full cycle in the states is

$$E\left[T_{SS}\right] = \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\alpha} = 130$$

The expected time is 130 days.

d)

Using Python to preform the simulations, we get the results summarized in Table 1. The code from the simulations is available from the attached code.
For this specific realization, we observe a 5.3%, 0.5% and 3.2% deviation between simulated and analytical expected time. Hence, our simulated results seem to be in accordance with the analytical time.

e)

I_n counts the number of infected individuals in a population. Because the probability of I_n increasing/decreasing is dependent on the number of susceptible individuals, it is not enough to know I_n at time t to determine the transition probabilities at time $t+1$. Thus, $\{I_n : n = 0, 1, ...\}$ is not a Markov chain.

As for $\{Z_n : n = 0, 1, ...\}$, if we know the state of Z_n at time t, this gives us the number of susceptible and infected individuals. Furthermore, because the size of the entire population is N, we also know the number of recovered individuals. Therefore, we have enough to determine the transition probability for $t+1$. Because the transition probabilities is only dependent on the current state, the chain satisfies the Markov property. Hence, $\{Z_n : n = 0, 1, ...\}$ is a Markov chain.

f)

For each day, we calculate the number of how many individuals transition between states by a binomial process. Individuals transitioning between states can be modeled as a binomial process,

State	Number	Proportion
S	235	23.5%
I	71	7.1%
R	694	69.4%

Table 2: Summary of long-run infection numbers from one realization modeled after 100 years.

e.g. Binomial(I_n, γ) for the number transitioning from infected to recovered. In Figure 2, we plot the realizations as a function of time step n.

We observe a big difference in behavior for the time intervals 0-50 and 50-300. At the start of the epidemic, we have a lot of susceptible individuals and little to no recovered individuals. As time progresses and more individuals become recovered, the curve will stabilize and somewhat flatten out. This is due to the low probability of transitioning from recovered to susceptible, $\alpha = 0.01$. Thus, we have less mobility between the states and the number of individuals in their state will stabilize.

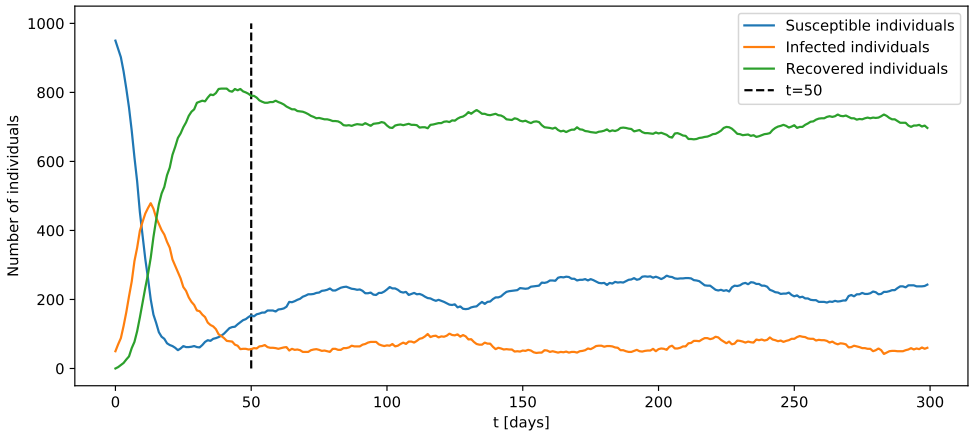

Figure 2: Temporal evolution of epidemic growth according to the SIR-model, with a β dependant on amount of infected individuals. The vertical line on time step 50 illustrates the difference in the model's behavior before and after time step 50. Note that the temporal evolution is represented as continuous variables and not discontinuous stochastic processes solely for the sake of visualization.

g)

Using the same functions as described in **1f)**, we model the epidemic growth for 36,500 time steps. The proportion of individuals in the states are shown in Table 2.

h)

Based on 1,000 simulations, the expected maximum number of infected was 489.961, and the expected first day of the highest number of infected was 13.123 in one realization. This means that the first day of the most infected is around day 13 and that around 490 individuals were infected. Both

of these values seem to match Figure 2 from the realization in problem **g)**. The smaller the value of $E[\min\{\arg\max_{n\leq 300}\{I_n\}\}]$, the severity of the epidemic increases, as there will be little time to prepare for the outbreak. In addition, a larger value for $E[\max\{I_0,\ldots,I_{300}\}]$ will lead to a more severe epidemic. If this value becomes too large, the governmental institutions and health care services will not be able to take care of all infected individuals.

Problem 2

a)

We have that

$$X(t) - X(0) \sim Poisson(\lambda t) = Poisson(88.5)$$

and thus

$$Pr\{X(59) - X(0) > 100\} = 1 - \sum_{x=0}^{100} \frac{88.5^x}{x!} e^{-88.5} \approx 0.103$$

Simulation also gives 0.103. The code can be found in the attached files.

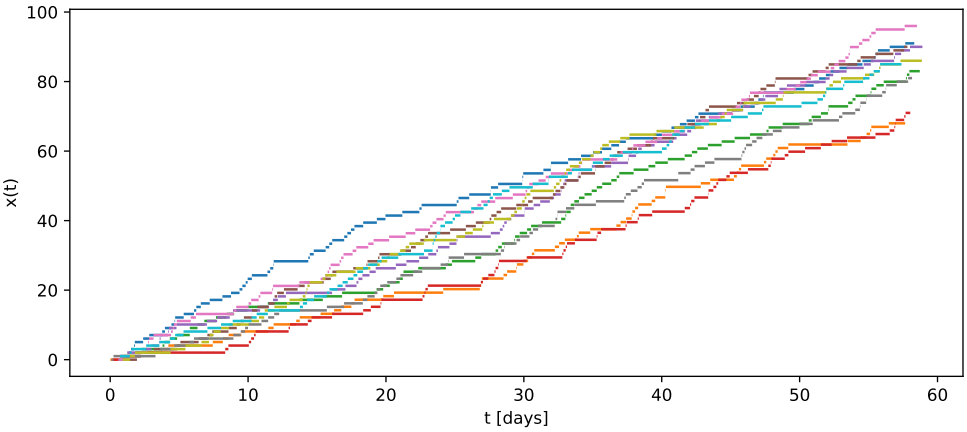

Figure 3: Number of claims as a function of $t, 0 \leq t \leq 59$. The ten realizations all have a unique color. From these realizations, we observe that one of them has over 100 claims.

b)

We know that the sum of n independent exponentially distributed stochastic variables follows a gamma distribution, so $Z(t)|X(t) \sim Gamma(X(t), \gamma)$ [1]. This gives

$$E\left[Z(59)\right] = E\left[E\left[Z(59)|X(59)\right]\right] = E\left[X(59)\frac{1}{\gamma}\right] = \frac{1}{\gamma}E\left[X(59)\right]$$

Because $X(59) \sim Poisson(59\lambda)$, we have

$$E\left[Z(59)\right] = 59\lambda\frac{1}{\gamma} = 8.85$$

The expected total claim amount is 8.85 mill. kroner. As for the variance

$$Var[Z(59)] = E\left[Var\left[Z(59)|\, X(59)\right]\right] + Var\left[E\left[Z(59)|\, X(59)\right]\right] = E\left[X(59)\gamma^2\right] + Var\left[X(59)\gamma\right]$$

$$= \frac{1}{\gamma^2}(E[X(59)] + Var[X(59)]) = \frac{1}{\gamma^2}(59\lambda + 59\lambda) = 118\lambda\frac{1}{\gamma^2} = 1.77$$

The variance for the claim amount is 1.77 (mill. kroner)^2. By comparison, the average of 1000 realizations of $Z(59)$ resulted in a simulated average $E[Z_{SIM}(59)] \approx 8.857563$ and simulated variance $Var[Z_{SIM}(59)] = 1.7360972$. This deviates very little from the analytical value.

References

[1] Institutt for matematiske fag. *Tabeller og formler i statistikk*. Fagbokforlaget, 2000.