Project 1 : Lasse Uttian, Henrik Olaussen, and Alexander Laloi Dybdahl

Problem 1

a)

To model a specific individual X_n we have to know the transition probabilities for to and from each state.
We are given parameters β, γ, α. These give the probabilities that a susceptible individual becoming infected, a infected individual becoming recovered and a recovered individual becoming susceptible respectively. For every state an individual either changes state or continues as the same state.
Therefore we can complete the transition probabilities by incorporating the probability that an individual doesn't change state as 1 minus the respective parameter. This gives arise to a transition probability

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

Since the transition probability for an individual on any given day is only a function of the current state, and the sum of probabilities for every state's transition probabilities is 1, we can conclude that $\{X_n : n = 0, 1 \ldots\}$ is a Markov chain.

b)

Now we let $\beta = 0.01$, $\gamma = 0.10$ and $\alpha = 0.005$. This give the the transition matrix

$$\mathbf{P} = \begin{bmatrix} 0.99 & 0 & 0.01 \\ 0 & 0.90 & 0.10 \\ 0.005 & 0 & 0.995 \end{bmatrix}$$

Let's raise the matrix to the second power

$$\mathbf{P}^2 = \begin{bmatrix} 0.9801 & 0.0189 & 0.001 \\ 0.0005 & 0.81 & 0.1895 \\ 0.009925 & 0.00005 & 0.990025 \end{bmatrix}$$

Since all entries are non-negative in the above matrix we can conclude that the Markov chain is regular. Therefore we know there to be exist a limiting distribution. Let's find the limiting distribution

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} \tag{1}$$

$$\sum_{j=0}^{2} \pi_j = 1 \tag{2}$$

We can rewrite (1) as such

$$
\begin{aligned}
(I - \mathbf{P})\boldsymbol{\pi}^T &= 0 \\
\begin{bmatrix} -\beta & 0 & \alpha \\ 0 & -\gamma & 0 \\ \beta & \gamma & -\alpha \end{bmatrix} \boldsymbol{\pi}^T &= 0 \\
\Rightarrow \begin{bmatrix} -\beta & 0 & \alpha \\ 0 & -\gamma & 0 \\ 0 & 0 & 0 \end{bmatrix} \boldsymbol{\pi}^T &= 0
\end{aligned}
$$

$$\boldsymbol{\pi} = r(\alpha\gamma,\ \beta\alpha,\ \beta\gamma), \tag{3}$$

With (2) and (2) we end up with the limiting distribution

$$\boldsymbol{\pi} = \frac{1}{\alpha\gamma + \beta\alpha + \beta\gamma}(\alpha\gamma,\ \beta\alpha,\ \beta\gamma) \approx (0.3226,\ 0.03226,\ 0.6452), \tag{4}$$

We can calculate the long-run mean number of days spent in each state by taking the limiting distribution multiplied by the number of days in a year.

$$\overline{\boldsymbol{X}} \approx (118,\ 12,\ 235), \tag{5}$$

c)

Now let's simulate the Markov chain for 20 years for one individual starting as susceptible.

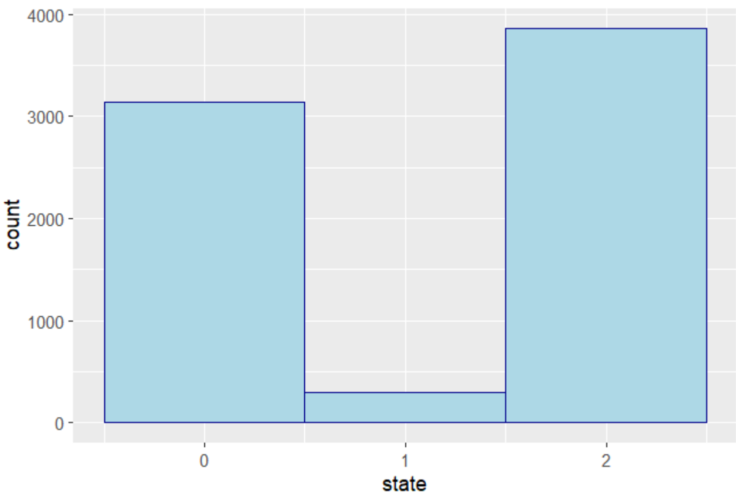

Figure 1: Histogram of the states which an individual occupied over a 20 year simulation

By running this simulation we can estimate the long-run mean number of days spent i each state, by finding the mean from the last 10 years of the simulation with

$$\overline{X}_j = \frac{1}{10}\sum_{t=3650}^{7300} \mathbb{1}\{X_t = j\}, \tag{6}$$

where j refers to the state. We compute the mean by (6) to be the following for data from figure 1 to be

$$\overline{\boldsymbol{X}} = (145.0,\ 12.4,\ 207.7)$$

Now let's run the experiment 30 times and compute a 95% confidence interval for the long-run mean number of days spent in each state.
Since we do not know the variance and expected value, we can assume the mean to be t-distributed with sample mean (SM_j) being the mean given by (6), of the mean of the 30 simulations and sample variance (S_j^2) given by

$$S_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\overline{X}_j^{(i)} - SM_j\right)^2, \tag{7}$$

with $n = 30$. Then with

$$T_j = \frac{SM_j - \mu}{S_j/\sqrt{n}} \tag{8}$$

$$PR(-c \leq T_j \leq c) = 0.95, \tag{9}$$

with c being the critical value for the 95% requirement of our confidence interval. This results in the confidence interval given by

$$\left[SM_j - \frac{cS}{\sqrt{n}}, SM_j + \frac{cS}{\sqrt{n}}\right] \tag{10}$$

Our code gives the following results

State	Lower bound	Upper bound
S	99.64	125.5
I	9.432	13.06
R	227.7	255.9

Table 1: The 95% confidence intervals for the long-run mean of states

Both our analytically and computed values for the long-mean run days spent in each state lies within their respective confidence interval given in Table 1.

d)

It is given that $Z_n = (S_n, I_n)$ for $n = 0, 1, ...$ In this task we are going to determine if $\{I_n : n = 0, 1, ...\}$, $\{Z_n : n = 0, 1, ...\}$, and $\{Y_n : n = 0, 1, ...\}$ are Markov Chains(MC). First we consider I_n. Let us say that we had 1000 infected yesterday in a population of 1000. Then we know that we can not have anyone that is in the state S (susceptible) today. Hence, the current state depends on the previous state. Therefore, I_n is not a MC. Secondly, we consider Y_n. We know that Y_n holds all the information S,I,R. The transition probability matrix for Y_n is therefore given by:

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

This implies that Y_n is a MC. Finally, we look at Z_n. The stochastic process Z_n holds exactly the same information as Y_n, since we know both S_n and I_n. This yields that $R_n = 1000 - S_n - I_n$. Because of this we can say that Z_n is a MC as well.

e)

Let's now simulate the Markov chain $\{Y_n : n = 0, 1, 2 \ldots\}$ until time step $n = 300$.

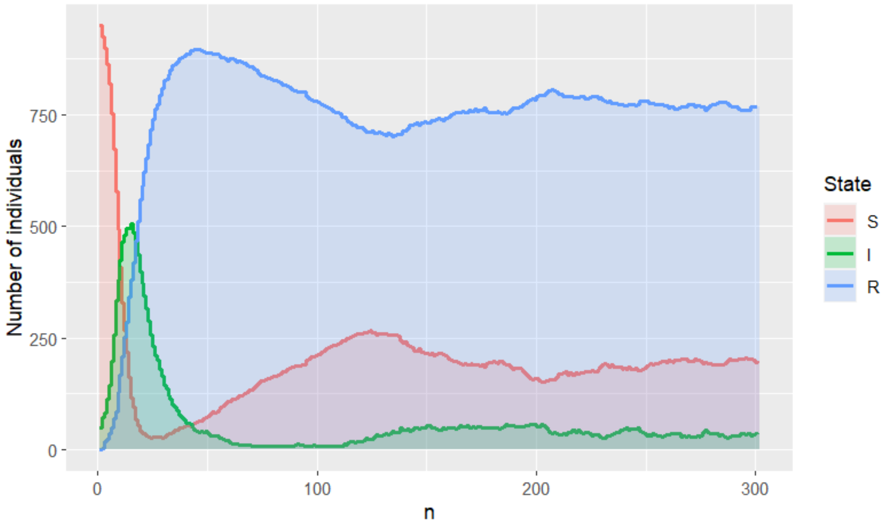

Figure 2: A realization of the dynamic Markov chain with initial conditions $Y_0 = (950, 50, 0)$

Since the infection parameter β_n is proportional to the number of infected divided by the population size, we observe a spike early when there are many susceptible individuals and a increasing number of infected. After the spike we can observe that the number of recovered dominates the population. This explains why the number of infected individuals plummets after the spike, since even with the high infection parameter, the number of susceptible is rather low.

f)

In this task we want to calculate the expected maximum number of infected individuals $E[max\left\{I0, I_1, ..., I_{300}\right\}]$ for n = 1,2,....,300, and the expected time it takes before we first reach such a maximum $E[min\left\{\underset{n\leq300}{argmax}\left\{I_n\right\}\right\}]$. Computing this in R we get:

$$E[max\left\{I_0, I_1, ..., I_{300}\right\}] = 524.538$$
$$E[min\left\{\underset{n\leq300}{argmax}\left\{I_n\right\}\right\}] = 12.89$$

Furthermore, we want to compute 95 present Confidence intervals (CI) for the expected values above. We did this in R. The resulting CIs are shown in the Table 2.

	Lower bound	Upper bound
$E[\max_{n\leq 300}\{I_n\}]$	523.2	525.8
$E[\min\{\arg\max_{n\leq 300}\{I_n\}\}]$	12.81	12.91

Table 2: The 95% confidence intervals

Looking at the CIs, we can say that a maximal number of infected often will occur at an early stage in our realization. To begin with the number of infected will increase rapidly until we reach a maximum of approximately 524 out of 1000 infected. From this point the number of infected decrease fast and remains low, according to our simulations in Figure 2. In conclusion the outbreak is severe to begin with, but not in the long-run because so many recovers.

g)

We can see from the graph below that vaccination is effective. When we have 600 or more vaccinated the number of infected individuals does not increase for time steps 150-300. Rather, the number of infected individuals stays 0 from here and onwards.

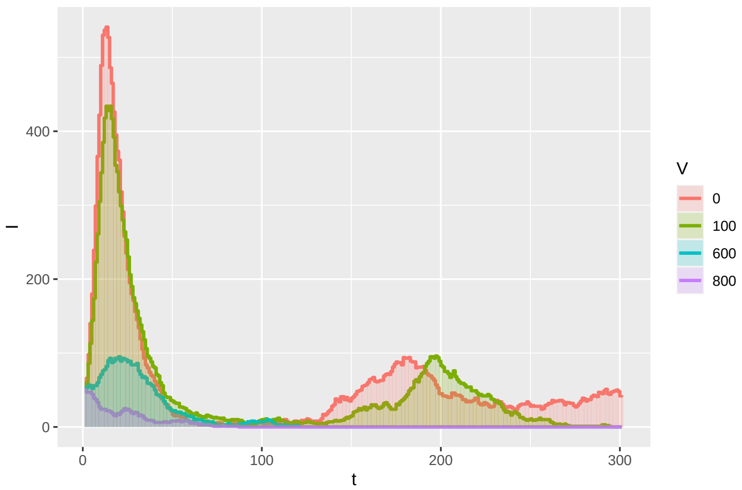

Figure 3: A realization of the model, now including vaccinated individuals

Problem 2

a)

We want to calculate the probability that the number of claims exceeds 100 after 59 days. In order to this we start by deriving an expression for $P(X(t) > 100)$:

$$P(X(59) > 100) = 1 - P(X(59) \leq 100)$$

We know that X(t) is Poisson distributed

$$\Rightarrow P(X(59) > 100) = 1 - \sum_{x=0}^{100} \frac{(\lambda t)^x}{x!} e^{-\lambda t}$$

We calculate the sum by using the command ppois($x, t\lambda$) in R with $x = 100$, $t = 59$ and $\lambda = 1.5$. We get:

$$P(X(59) > 100) = 0.1028222$$

Furthermore, we verify this calculation by simulating 1000 realizations of the Poisson process in R. Running the simulation once, we get that $P(X(59) > 100) = 0.099$. This result is not far from our exact calculations above. Therefore we conclude that the value of the exact calculation is reasonable.

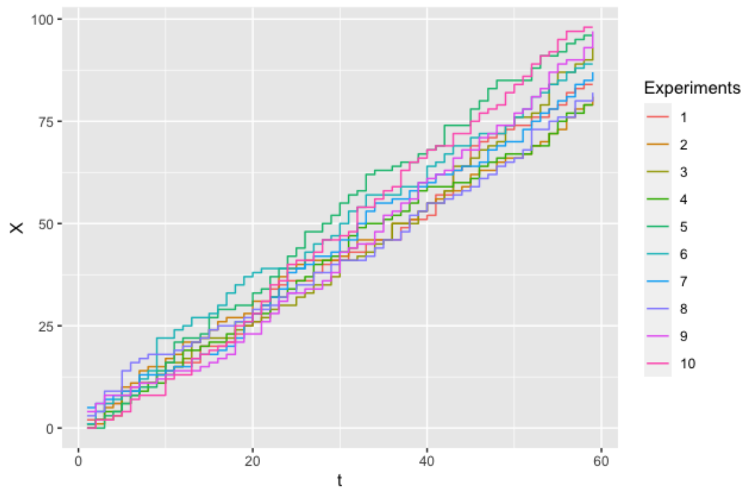

Figure 4: Realization of X(t), $0 \leq t \leq 59$, 10 simulations

b)

Assumes that the monetary claims are independent, and are independent of the claim arrival times. It is stated that each claim has an exponential distribution with rate parameter $\gamma = 10$. In this task we have written a code in R that simulates the total amount of monetary claims within a month 1000 times. We use this simulation to estimate the probability that the total claim amount exceeds 8 mill. kr. at March 1 at 00:00.00 (59 days). The resulting estimated probability:

$$P_{estim}(Z(t) \geq 8) = 0.722$$

Below is a Figur 5e that shows 10 realizations of Z(t), $0 \leq t \leq 59$:

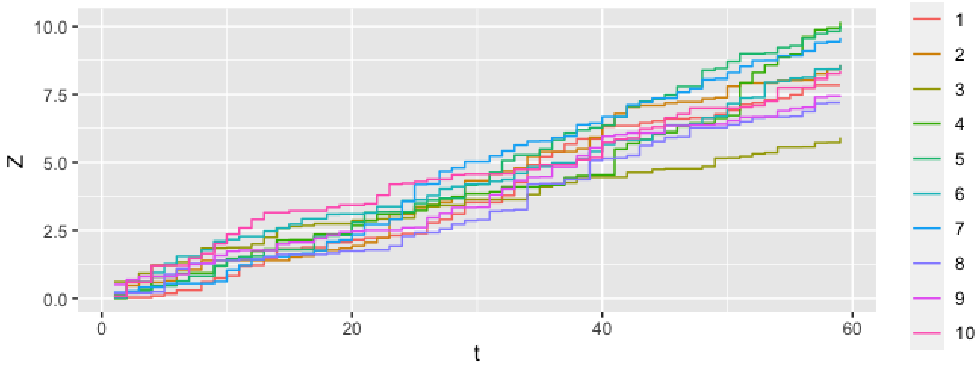

Figure 5: Realization of Z(t), $0 \leq t \leq 59$, 10 simulations

c)

$$Pr(C_i > 0.25) = e^{-\frac{\gamma}{4}} = \alpha$$

$$\Rightarrow \mathbb{1}\{C_i > 0.25\} \sim b(1; z)$$

$$\Rightarrow \sum_{i=0}^{x} \mathbb{1}\{C_i > 0.25\} \sim b(x; z)$$

$$\begin{aligned}
M_{Y(t)}(\tau) &= E\left[e^{\tau Y(t)}\right] \\
&= E\left[e^{\tau \sum_{i=0}^{X(t)} \mathbb{1}\{C_i > 0.25\}}\right] \\
&= E\left\{E\left[e^{\tau \sum_{i=0}^{X(t)} \mathbb{1}\{C_i > 0.25\}} \middle| X(t) = x\right]\right\} \\
&= \sum_{x=0}^{\infty} E\left[e^{\tau \sum_{i=0}^{x} \mathbb{1}\{C_i > 0.25\}}\right] \cdot Pr(X(t) = x) \\
&= \sum_{x=0}^{\infty} (\alpha(e^{\tau} - 1) + 1)^x \cdot \frac{(\lambda t)^x}{x!} e^{-\lambda t} \\
&= e^{-\lambda t} \sum_{x=0}^{\infty} \frac{(\lambda t \alpha(e^{\tau} - 1) + \lambda t)^x}{x!} \\
&= e^{-\lambda t} \cdot e^{-\lambda t \alpha(e^{\tau} - 1) + \lambda t} \\
&= e^{-\lambda t \alpha(e^{\tau} - 1)}
\end{aligned}$$

Since the moment generating function for $Y(t)$ is of the form of a Poisson distribution, $\{Y(t) : t \geq 0\}$ will therefore clearly be a Poisson process with rate

$$\mu = \lambda e^{-\frac{\gamma}{4}} \approx 0.1231 \tag{11}$$

We simulate to find the expected value of $Y(t)$, when using the mean as the estimator.

$$\mu \approx 0.12363 \tag{12}$$

Since the expected value of a Poisson process is the rate, the results from (12) is approximately the analytical value from (12).