

Project 1 - Stochastic Modelling

By: Sanne Jamila Razmara Olsen & Einride Brodahl Osland

Problem 1

For this problem we are analyzing an outbreak of measles. We use a simplified model where each individual can have in of the following states; susceptible (S), infected(I) and recovered/immune (R). Let $n = 0, 1, 2, \dots$ denote time measured in days and assume that each day

1. A susceptible individual can become infected or remain susceptible
2. An infected individual can become recovered or remain infected,
3. A recovered individual can lose immunity and become susceptible, or remain recovered.

Assume that the individuals in the population are independent, and assume that for each day, any susceptible individual has a probability of $0 < \beta < 1$ of becoming infected tomorrow, any infected individual has a probability $0 < \gamma < 1$ of becoming recovered tomorrow, and any recovered individual has a probability $0 < \alpha < 1$ of losing immunity tomorrow.

- a) Let X_n denote 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 $X_0 = 0$.

The Markov assumptions state that

1. The probabilities of moving from one state to all others sum to one.
2. The probabilities apply to all system participants.
3. The probabilities are constant over time.
4. The states are independent over time.

We have that for an individual that is in the state susceptible (S) has a probability of $0 < \beta < 1$ of becoming infected tomorrow, and thus a probability of $0 < 1 - \beta < 1$ of not becoming infected tomorrow. An individual of the state (S) can not become immune or recovered, as being susceptible is a contradiction to being immune and the susceptible individual has to become infected (I) before it can be recovered (R). Thus all the probabilities of moving from the state susceptible (S) sum to one. As the probability of becoming infected tomorrow is the same for any day as long as the individual is in state S, it is constant over time (not time-dependent). The same arguments hold for the other states as well. Thus the system $\{X_n : n = 0, 1, \dots\}$ is a Markov chain.

Mark that a recovered/immune (individual in state (R)) or an infected individual (individual in state (I)) can not become susceptible (S).

We have that for state 0 (being susceptible), there is a probability of $0 < \beta < 1$ of becoming infected tomorrow, and thus a probability of $0 < 1 - \beta < 1$ of staying susceptible (staying in state 0) tomorrow (not being infected). There is a 0 probability of becoming recovered tomorrow (moving to state 3).

Thus the first row of the transition probability matrix can be written as

$$\mathbf{P}_1 = \begin{bmatrix} 1 - \beta & \beta & 0 \end{bmatrix}$$

We have that for state 1 (being infected), there is a probability of $0 < \gamma < 1$ of becoming recovered tomorrow, and thus a probability of $0 < 1 - \gamma < 1$ of staying infected (staying in state 1) tomorrow (not being recovered). There is a 0 probability of becoming susceptible tomorrow (moving to state 0).

Thus the second row of the transition probability matrix can be written as

$$\mathbf{P}_2 = [0 \quad 1 - \gamma \quad \gamma]$$

For the second state (being recovered), there is a probability of $0 < \alpha < 1$ of becoming susceptible tomorrow, and thus a probability of $0 < 1 - \alpha < 1$ of staying recovered (staying in state 2) tomorrow (being recovered). There is a 0 probability of becoming infected tomorrow (moving to state 1).

Thus the third row of the transition probability matrix can be written as

$$\mathbf{P}_3 = [\alpha \quad 0 \quad 1 - \alpha]$$

Such that the full transition matrix can be written as

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

b) Assume $\beta = 0.01$, $\gamma = 0.1$ and $\alpha = 0.005$. We can thus write our transition matrix as

$$\mathbf{P} = \begin{bmatrix} 0.99 & 0.01 & 0 \\ 0 & 0.9 & 0.1 \\ 0.005 & 0 & 0.995 \end{bmatrix}$$

We have that

$$\mathbf{P}^2 = \begin{bmatrix} 0.9801 & 0.0189 & 0.001 \\ 0.005 & 0.81 & 0.1895 \\ 0.00995 & 0.005 & 0.990025 \end{bmatrix}$$

We know that the Markov chain is regular because $P_{ij}^2 > 0$ for all $i, j \in (0, 1, 2)$. Because $\{X_n : n = 0, 1, 2, \dots\}$ is a regular Markov chain with state space $\{0, 1, 2\}$ and transition probability \mathbf{P} , then there exists a limiting distribution $\pi = (\pi_0, \pi_1, \pi_2)$.

In a regular Markov chain $\{X_n : n = 0, 1, 2, \dots\}$ the limiting distribution $\pi = (\pi_0, \pi_1, \pi_2)$ gives the long-run mean fraction of time spent in each state.

We find these by solving the system $P\pi = \pi$.

We also have that $\pi_1 + \pi_2 + \pi_3 = 1$.

Choosing the following equations

$$0.99\pi_1 + 0.01\pi_2 = \pi_1$$

$$0.9\pi_2 + 0.1\pi_3 = \pi_2$$

$$0.005\pi_1 + 0.995\pi_3 = \pi_3$$

$$\pi_1 + \pi_2 + \pi_3 = 1$$

We obtain the solution $\pi_1 = \frac{10}{31}$, $\pi_2 = \frac{1}{31}$ and $\pi_3 = \frac{20}{31}$. Thus the long-run mean number of days per year spent in each state is 118 days susceptible, 12 days infected and 235 recovered.

c)

```
pi
```

```
## [1] 0.35113667 0.04053684 0.60832649
```

Assuming that the means of the simulated limiting probabilities are uniformly distributed, we consider the student-t distribution for computing our 95% confidence intervals.

Let $S_i^2 = \frac{1}{n-1} \sum_{j=0}^N (\pi_{ij} - \bar{X}_{ij})^2$ denote the sample variance and let \bar{X}_i denote the mean of the π_i for the i 'th limiting probability our confidence interval is then given by

$$\pi_i \in [\bar{X}_i - t_\alpha \frac{S_i}{\sqrt{n}}, \bar{X}_i + t_\alpha \frac{S_i}{\sqrt{n}}]$$

Where t_α is the critical value of the student-t distribution.

Implementing our solution in R (view document code.rmd) we obtain the simulated confidence intervals for our limiting probabilities.

```
confidence_interval(P,30,1)
```

```
## [1] 0.3579880 0.2921353
```

```
confidence_interval(P,30,2)
```

```
## [1] 0.03622791 0.02969011
```

```
confidence_interval(P,30,3)
```

```
## [1] 0.6818122 0.6290433
```

Which are close to our analytical solution of $\pi_1 = \frac{10}{31}, \pi_2 = \frac{1}{31}$ and $\pi_3 = \frac{20}{31}$, but noting that the intervals are quite high.

d) Let $Z_n = (S_n, I_n)$.

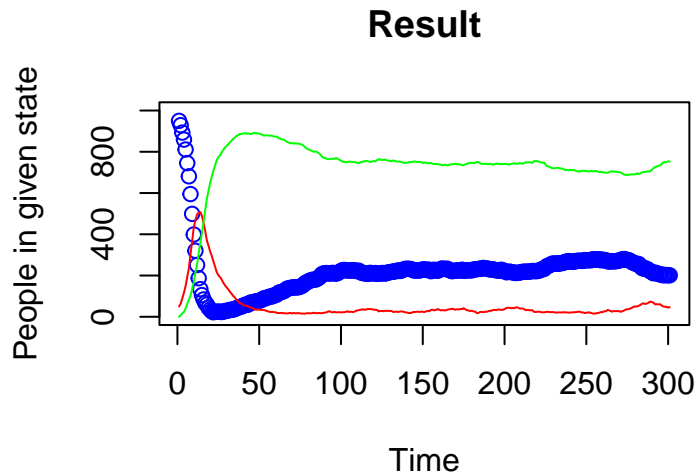
For I_n it will not be a markov chain, as it is missing a state space to transition to, if the probability of remaining in I was 1 it would be, though not a very interesting markov chain as it would only be an absorbing state. Furthermore the current state is not independent of the previous state, the next state depends on the number of infected individuals from the previous state

For Z_n we can look at it as a markov chain but the problem is that the transition probabilities dont add to 1, because we are missing a state, if it was possible to transition from I to S and skip R it could possibly be reduced to this Z, but that wouldnt be within the parameters of our problem, thus this is not a markov chain, by not being a complete state space.

For Y_n we can transition through the three states and the sum of the transition probabilities add to 1. We clearly see from the transition probability matrix that the states are independent of each other. Thus it satisfies the markov property and is therefore a markov chain.

e)

```
plot(step, S,col="blue", main="Result",xlab="Time", ylab="People in given state",ylim=c(0,1000))
lines(step, I,col="red")
lines(step, R, col="green")
```



As shown in the graph the difference between the interval 0-50 and 50-300 is that we start out with 950 Susceptible people, those people transition to other states, and theres a higher probability to stay in R so most people end up in that state.

f)

```
## [1] 522.611
```

```
## [1] 12.882
```

```
##      2.5%   97.5%
## 482.000 560.025
```

```
##      2.5% 97.5%
##      11   14
```

In 95% of cases at any time was between the maximum amount of infected was between 486 and 566. Which is over half the population in many cases, thus the situation would be quite severe. Furthermore it reaches this state after between 11-14 days which is a very fast growth rate, so there would be little time for response to the issue once an outbreak happens.

g) making a function to simulate with different number of vaccinated we assume that the vaccinated cant become sick, thus we just remove them from the S variable, and use the same code as in f

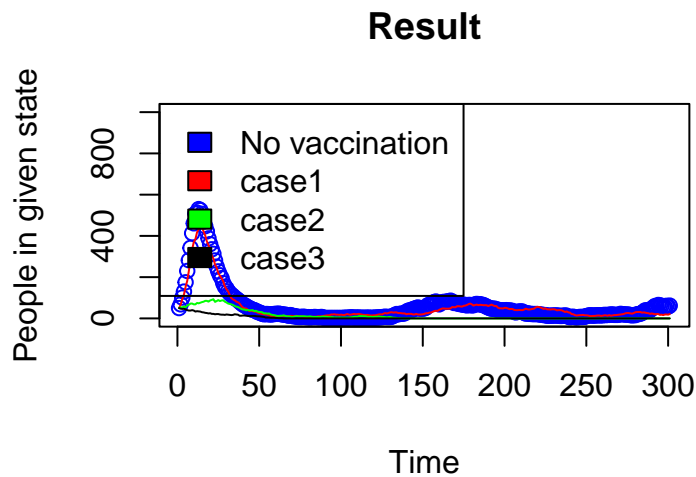
```
## [1] 522.61
## [1] 12.884
```

```
## [1] 439.697
## [1] 13.642
```

```
## [1] 97.619
## [1] 16.945
```

```
## [1] 50.033
## [1] 3.079
```

```
plot(step, No_vac,col="blue", main="Result",xlab="Time", ylab="People in given state",ylim=c(0,1000))
lines(step, I_case1,col="red")
lines(step, I_case2, col="green")
lines(step, I_case3, col="black")
legend("topleft",legend=c("No vaccination","case1","case2","case3"), fill= c("blue","red","green","black"))
```



As shown in the graph the number of infected drastically reduce with more vaccinated people, even just 100 people removed meant the max infected never reached 400 which is on the mean 150 less people infected.

As more people get vaccinated the initial hit of the outbreak also diminishes, which would give more time to respond to an outbreak, we can even see that when 80 percent are vaccinated the outbreak regresses immediately.

When testing the expected values we can see that the maximum infected timestep shifts further to the right as the timeline progresses, except for the 80% vaccinations where the maximum is in the first couple of timesteps.

The mean expected value is also reduced the more vaccinated people there are. Note that the mean value for case 3 is below 50 even though the simulation starts at 50, this is because the code doesnt account for the first timestep as a possible maxima, which is not ideal, but could be fixed

Problem 2

We have that $X(t)$ is the number of claims recieved by an insurance company in the time interval $[0, t]$. Assuming that $\{X(t) : t \geq 0\}$ can be modeled as a Poisson process, were t is measured in days since January 1st at 00:00:00, assume that the rate of the Poisson process is given by $\lambda(t) = 1.5, t \geq 0$.

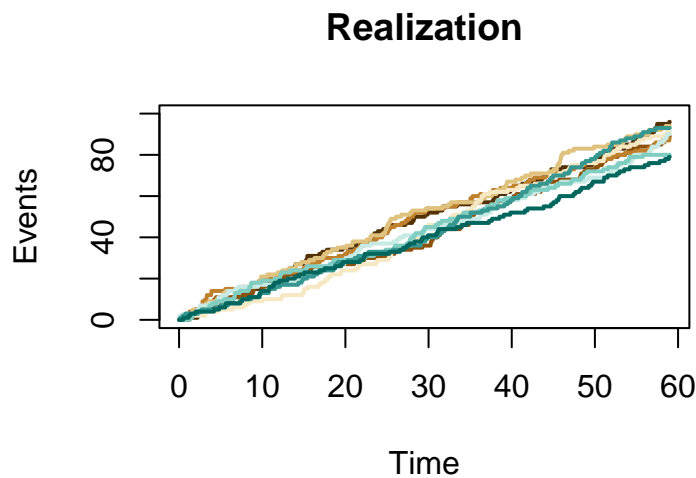
- a) We want to compute the probability that there are more than 100 claims at March 1 at 00:00:00. In other words, after 59 days.

We want to compute $P(X(59) > 100)$.

We have that $X(t) \sim \text{Poisson}(\lambda t)$ such that after $t = 59$ days the probability of $X_{59} > 100$ is given by

$$P(X_{59} > 100) = 1 - P(X_{59} \leq 100) = 1 - \sum_{n=0}^{100} P(X_{59} = n) = 1 - \sum_{n=0}^{100} \frac{(59\lambda)^n}{n!} \exp(-59\lambda) \approx 1 - 0.8777 = 0.1233$$

```
plot_poisson(t_value, lambda, N)
```



- b)

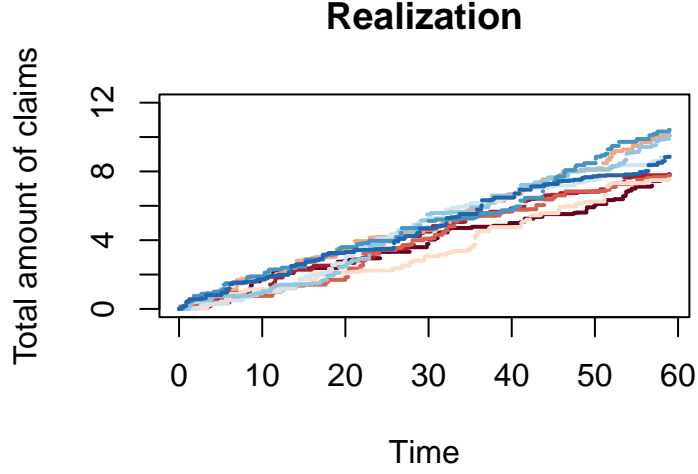
```
#Probability
mean(simulate_poisgam(N,gamma,lambda))
```

```
## [1] 0.709
```

The estimated probability of the total claim amount exceeding 8 mill kr. after 59 days is 0.971.

We now want to provide the estimated probability by making a figure that shows 10 realizations of $Z(t)$, for $t \in [0, 59]$ in the same figure.

```
final_simulation(lambda,gamma,t_value,N)
```



c)

Let Y_t denote the number of claims needing to be investigated recieved in the time interval $[0, t]$. We want to show that $\{Y(t) : t \geq 0\}$ is a Poisson process.

We have that

$$P(Y(t) = x) = \sum_{n=0}^{\infty} P(Y(t) = x \mid X(t) = n) \cdot P(X(t) = n)$$

From the law of total probability.

We have that our expression can be written as

$$\sum_{n=0}^{\infty} P(Y(t) = x \mid X(t) = n) \cdot P(X(t) = n) = \sum_{n=x}^{\infty} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \frac{(\lambda t)^n \exp(-\lambda t)}{n!}$$

Which can further be written as

$$\frac{\exp(\lambda t)(\lambda t p)^x}{x!} \sum_{n=x}^{\infty} \frac{\lambda t (1-p)^{n-x}}{(n-x)!} = \frac{\exp(\lambda t)(\lambda t p)^x}{x!} \exp(\lambda t(1-p)) = \frac{\exp(-p\lambda t)(\lambda t p)^x}{x!}$$

We thus have that $P(Y(t) = x) = \frac{\exp(-p\lambda t)(\lambda t p)^x}{x!}$ which is the equivalence of a Poisson distribution with rate λp .

We now want to find the value of p in order to compute the rate.

$$p = 1 - \int_0^{1/4} 1 - \exp(-\gamma t) dt = 0.082$$

Thus the rate is $\lambda p = 1.5 \cdot 0.082 = 0.123$