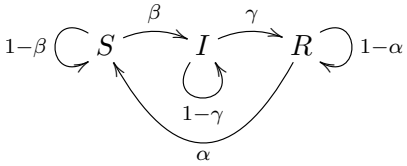
Project 1 TMA4265

Jasper Steinberg, Maximilian Leskovszky Rønseth, and Henrik Knudsen

Problem 1

a)

In this scenario, we have that $\Omega := \{0, 1, 2\}$ is the state space. We can justify using a Markovian model: Indeed, in a heavily simplified model, one could argue, that the result of day n is only dependent on what happened at day $n-1$. Indeed, the probability of transitioning from one state of health to another can be thought of as independent of whatever state one was in, in the states far into the past. We can also assume that our model is stationary. We indeed obtain that the one-step transition probabilities, are independent of time; i.e. $P_{ij}^{n,n+1} = P_{ij}$. Since $P(\Omega) = 1$, then for any event $A \in \mathcal{F}$, we have by σ-additivity of the probability measure that $P(\Omega \backslash A) = P(\Omega) - P(A) = 1 - P(A)$. Thus, the TPM is given by the $\mathbf{P} \in M_3(\mathbb{R})$ given in the exercise text, due to the following diagram:



The entry $\mathbf{P}_{02}$ is zero, because one cannot go from susceptible to recovered without being infected; $\mathbf{P}_{10}$ is zero, as one can not go from infected to susceptible without going through recovered (by definition); lastly, $\mathbf{P}_{21}$ is zero, as one cannot go from recovered to infected without first becoming susceptible.

b)

From Markovian theory, we know that a regular discrete-time Markov chain with $\Omega := \{0, \ldots, N\}$ and TPM $\mathbf{P}$, has a limiting distribution $\boldsymbol{\pi} \in [0, 1]^{N+1}$ exists and is unique. We observe that $\mathbf{P}$ is regular, as $\mathbf{P}_{\mathbf{ij}}^{\mathbf{2}} > 0$ for all i, j. Calculating the long-run mean number of days spent in each state is equivalent to calculating the limiting distributions; we do this through the Chapman-Kolmogorov equations. Through some linear algebra, we find by Cramer's rule, that $\boldsymbol{\pi} = \left(\frac{10}{31}, \frac{1}{31}, \frac{20}{31}\right)$ (notice that these sum to 1). We find the long-run number of days per year spent in each state by multiplying by 365; thus $365\boldsymbol{\pi} \simeq (118, 12, 235)$, which makes sense intuitively, given our choices of $\alpha, \beta, \gamma \in (0, 1)$.

c)

To compute the CI's, we used the t.test() function in R applied on a vector of the 30 sampled values for S, I and R, respectively. Afterwards we computed the confidence intervals through the conf.int function. When we scale the vector $\boldsymbol{\pi}$ to our time scale we approximately get $(1177, 118, 2355)$, which within what we would expect considering the confidence intervals in Table 1

Table 1: Confidence Intervals

$$CI_S = [1010.719, 1325.148]$$
$$CI_I = [96.9071, 125.5596]$$
$$CI_R = [2206.177, 2537.490]$$

d)

I_n is not a Markov chain as we do not have all information needed at step n to calculate step $n+1$. To see this, let I_n have k infected individuals. If R_n is $N-k$, we have no susceptible individuals and $I_{n+1} \leq I_n$. However, if $R_n = 0$, no people are immune and clearly I_{n+1} can be greater than I_n.
Z_n is a Markov chain, since if we have the amount of susceptible and infected people, we can calculate $R_n = N - I_n - S_n$ and thus have all information needed to calculate Z_{n+1}.
Y_n is just Z_n with even more information, so it is also a Markov chain.

e)

A realization is shown in figure 1. As we can see, during the first 50 days we have a large amount of susceptible people, which leads to a large outbreak and then most of the population becomes recovered. The disease then stays latent as the population immunity slowly decreases, causing a smaller second wave around day 175-200.

f)

Our computed 95% confidence intervals were

$$CI_{\text{Day}} = [11.71365, 11.98635]$$
$$CI_{\text{Peak}} = [518.9858, 528.0742]$$

From this we could conclude that any outbreak that peaked after day 12 could be classified as severe, and any that peaked on day 11 or before would be relatively mild. We could also expect that any severe outbreak has upwards of 525 concurrently infected people.

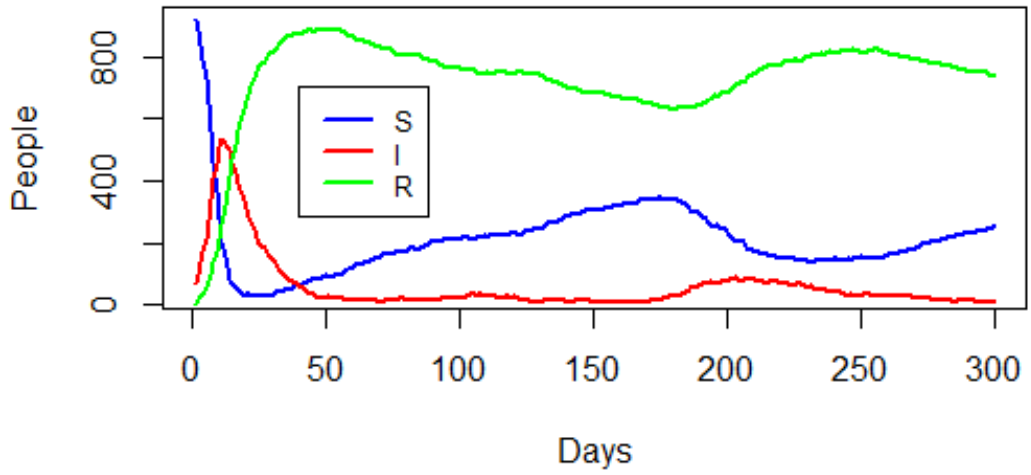


Figure 1: A realization of Y_n

Table 2: Expected values for each case

Vaccinateds	Peak day	Peak
0	11.849	522.798
100	12.618	438.996
600	15.606	96.858
800	2.012	50.033

g)

In figure 2 we can see four realizations with different amounts of vaccinated people. We can see second waves both in the 0 and the 100 cases, but about halved in the latter case. It also shows that the 600 people gives a small but significant peak at over 100 concurrent infected people, while the strictest case (800) is strictly decreasing, since only 100 people are susceptible. The expected values can be seen in table 2. There we can see that the peak number of infected people decreases as the number of vaccinated people increase. The day the peak happens on gets delayed as the number of vaccinated people increases, until enough people are immune and the peak is day 0.

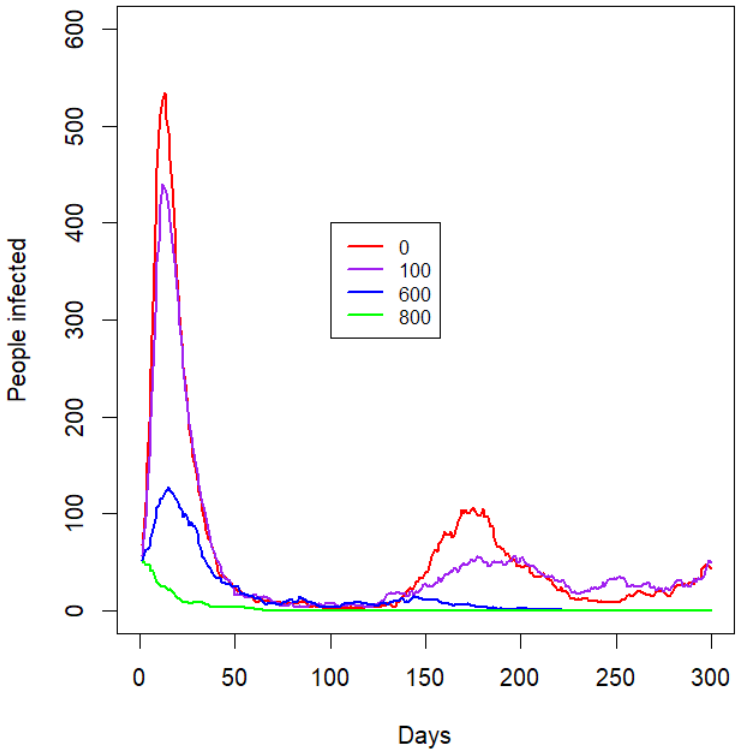


Figure 2: Number of people infected under different vaccination amounts

Problem 2

a)

By standard theory for Poisson processes, we know that, time increments are independent, so we can rewrite our probability in the following way: $\Pr(X(59) > 100) = \Pr(X(59) - X(0) > 100)$, where we know by definition that $X(0) := 0$. This way, we can use the PMF for the Poisson distribution, to calculate the probability:

$$\Pr\{X(59) > 100\} = 1 - \Pr\{X(59) \leqslant 100\} = 1 - \sum_{n=1}^{100} \Pr\{X(59) = n\}$$

where we know that $\Pr(X(t) = n) := \frac{(\lambda t)^n}{n!} e^{-\lambda t}$, where $\lambda(t) \equiv 1.5$ is our rate parameter. We calculate this by hand, and we get that $\Pr(X(59) > 100) \simeq 0.1028$. This is reasonable compared to our simulated value of 0.116. In Figure 3 we have plotted 10 realizations against the mean.

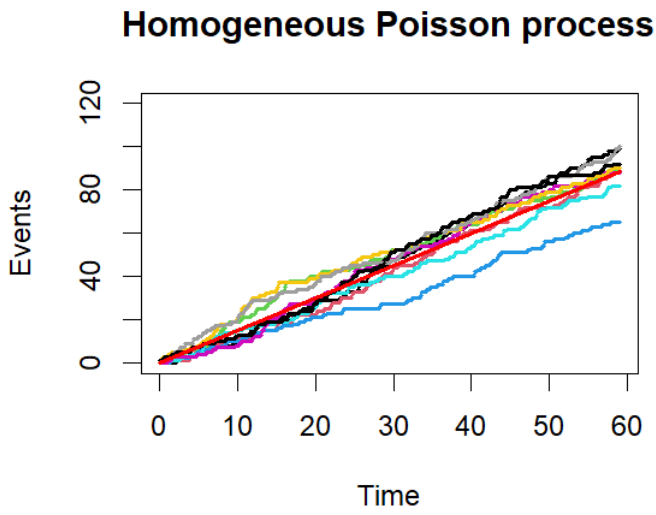


Figure 3

b)

After running 1000 simulations we found an estimated probability of 0.697 that the total claim amount exceeds 8 million. In Figure 4 we simulate 10 realizations.

c)

We start by finding the distribution of Y_t, before verifying that it is a Poisson process. If we condition on $X(t) = x$, Y_t can be thought of as the number of successes of an Bernoulli process with success probability $p = \Pr(C_i > 1/4)$ and x number of trials. Thus $Y_t|X_t = x \sim \text{Binom}(x, p)$. We find the distribution of Y_t by using the law of total probability and calculating in large steps

$$\Pr(Y_t = k) = \sum_{x=k}^{\infty} \Pr(Y_t = k|X_t = x) \Pr(X_t = x) = \sum_{x=k}^{\infty} \frac{x!}{k!(x-k)!} p^k (1-p)^{x-k} \frac{(\lambda t)^x}{x!} e^{-\lambda t}$$
$$= \frac{(p\lambda t)^k e^{-\lambda t}}{k!} \sum_{x=k}^{\infty} \frac{(1-p)^{x-k}(\lambda t)^{x-k}}{(x-k)!} = \frac{(p\lambda t)^k}{k!} e^{-\lambda t} e^{(1-p)\lambda t} = \frac{(p\lambda t)^k}{k!} e^{-p\lambda t}.$$

Thus we can conclude that $Y_t \sim \text{Pois}(p\lambda t)$. Now we can show that $\{Y_t : t \geq 0\}$ a Poisson process with rate $p\lambda$. Firstly the process increments are independent. This is clear since the claim amounts are independent. Furthermore,

$$Y_{t+s} - Y_s \sim \text{Pois}(p\lambda(s+t) - p\lambda s) \sim \text{Pois}(p\lambda t).$$

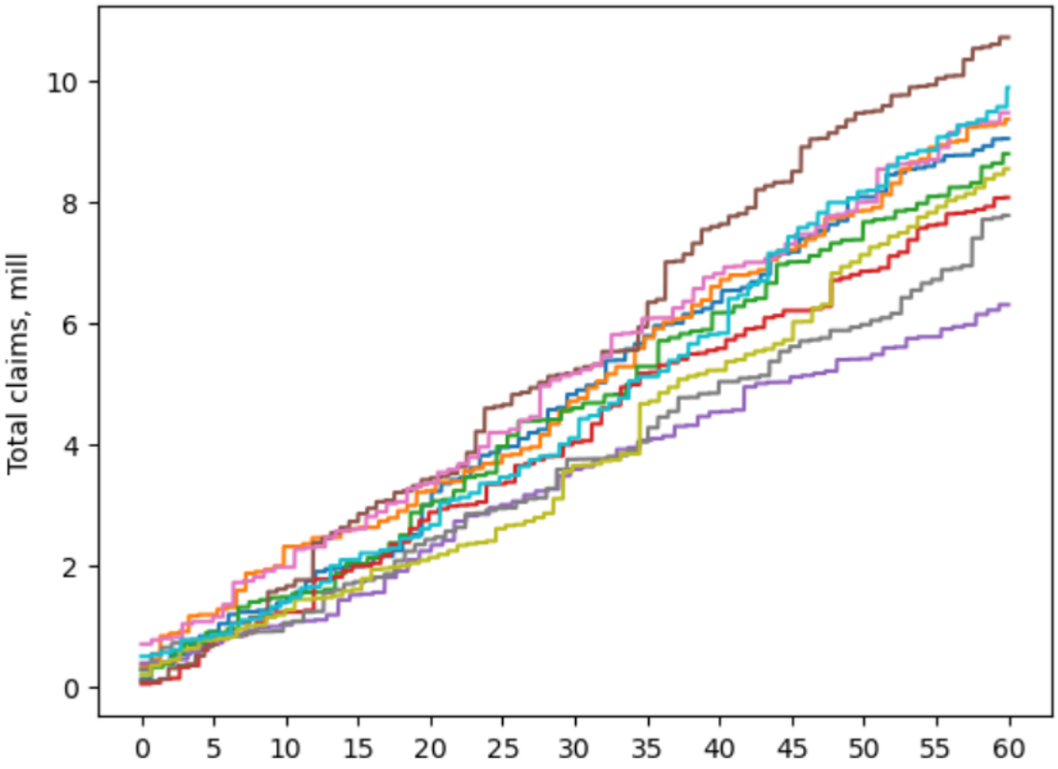


Figure 4: Z(t)

Finally $Y_0 = 0$, since no claims can occur in zero time. Since $\lambda = 10$ and

$$\Pr(C_i > 1/4) = \int_{1/4}^{\infty} 10e^{-10x}\,dx = e^{-5/2}$$

we get that the rate of the Poisson process is $p\lambda = 1.5e^{-5/2}$.