TMA4265 - Stochastic Modeling
Project 1

Group 39
Anders T. Fagerli
Odin A. Severinsen

Problem 1: Modelling the outbreak of measles

a)

The process X_n is defined as follows: It has three states in its state space: $\Omega = \{S, I, R\}$, with S denoting susceptible to measles, I denoting infected by measles and R denoting recovered from measles.

It is clear that given an individual is susceptible to measles at a given day n, information that the individual was susceptible any previous day will not alter our prediction for whether the individual remains susceptible or transitions into infected, as it is assumed constant from one day to the next. Additionally, once we know that the individual was susceptible at day n, it follows that the individual *must have been susceptible all previous days*, making information about this redundant.

A similar argument holds for state I: Knowledge that the individual is infected makes knowledge that the individual has previously been susceptible or infected redundant, as it does not affect our prediction for whether the individual remains infected or transitions into recovered.

A recovered individual can only remain recovered, again making information that the individual previously has been susceptible or infected redundant. Hence, X_n satisfies the definition of a Markov chain.

The transition matrix $\mathbf{P}$ describes the probabilities $\Pr\{X_n = j \mid X_{n-1} = i\}$, where j and i denotes the column and row of $\mathbf{P}$, respectively. Additionally, i and j denotes states of the state space $\Omega = \{\omega : \omega = 0, \ldots, M-1\}$, with M states. Each state can propagate to another with a given probability and are assumed constant for $\mathbf{P}$. From the problem text it is given that $\Pr\{X_n = I \mid X_{n-1} = S\} = \beta$ and $\Pr\{X_n = R \mid X_{n-1} = I\} = \gamma$. As these states cannot back-propagate, it follows that $\Pr\{X_n = S \mid X_{n-1} = S\} = 1 - \Pr\{X_n = I \mid X_{n-1} = S\} = 1 - \beta$ and $\Pr\{X_n = I \mid X_{n-1} = I\} = 1 - \Pr\{X_n = R \mid X_{n-1} = I\} = 1 - \gamma$. Lastly, as an individual that has recovered from measles cannot become infected again and thus also is not susceptible to the disease, R is an *absorbing state*, and hence $\Pr\{X_n = R \mid X_{n-1} = R\} = 1$ while $\Pr\{X_n = S \mid X_{n-1} = R\} = \Pr\{X_n = I \mid X_{n-1} = R\} = 0$. Combining all this, the $\mathbf{P}$ matrix can be constructed as follows:

$$\mathbf{P} = \begin{array}{c} \\ S_{n-1} \\ I_{n-1} \\ R_{n-1} \end{array} \begin{array}{c} \begin{array}{ccc} S_n & I_n & R_n \end{array} \\ \begin{bmatrix} 1-\beta & \beta & 0 \\ 0 & 1-\gamma & \gamma \\ 0 & 0 & 1 \end{bmatrix} \end{array} \tag{1}$$

b)

The transition diagram for X_n is shown in Figure 1. From this, it can be seen that no states communicate, as it is not possible to transition back to a previous state, but only remain in the current or transition to the next. Thus, X_n is a reducible Markov chain. As no state communicates with another state, each state must form their own equivalence class. Since the probability of returning to the states S and I is less than 1 in finite time, they must both be *transient* states (or equivalently, there is a positive probability of never returning). Correspondingly, since the probability of returning to R is 1 in finite time, it is a *recurrent* state. For each state, there can only be one transition before the system returns to starting state (this is the loops in Figure 1). Because of this, the period of all states are the same and equal to $d = \gcd\{1\} = 1$.

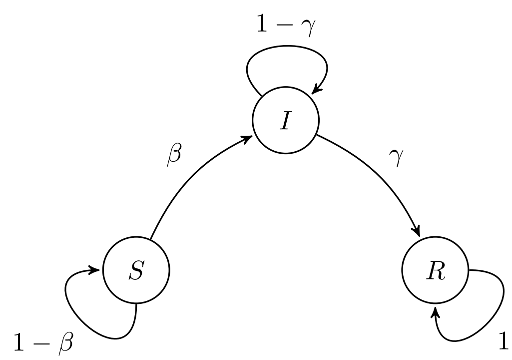

Figure 1: *Transition diagram for X_n.*

c)

As the system has discrete transitions where the system transitions to a different state only after $t-1$ failures, ending with a success, the probability mass function for the transition time follows a geometric distribution, given by

$$\Pr\{T = t\} = \beta\,(1-\beta)^{t-1}\,. \tag{2}$$

For a geometric distribution, we have $E\,[T] = 1/\beta$. By inserting $\beta = 0.05$, we get the expected time from state S to I to be 20 days. The expected time from I to S is computed identically with β exchanged with γ. This leads to the expected time until the system transitions from state I to R to be 5 days.

d)

1 000 000 simulations was run in order to get a narrow, consistent result. A normalized histogram of the transition time realizations for the transition from susceptible to infected is shown in Figure 2. It is evident from the shape of the histogram that the probability distribution converges to an exponential distribution with $\lambda = 1/20$. This is consistent with how the exponential distribution can be considered the continuous counterpart to the geometric distribution.

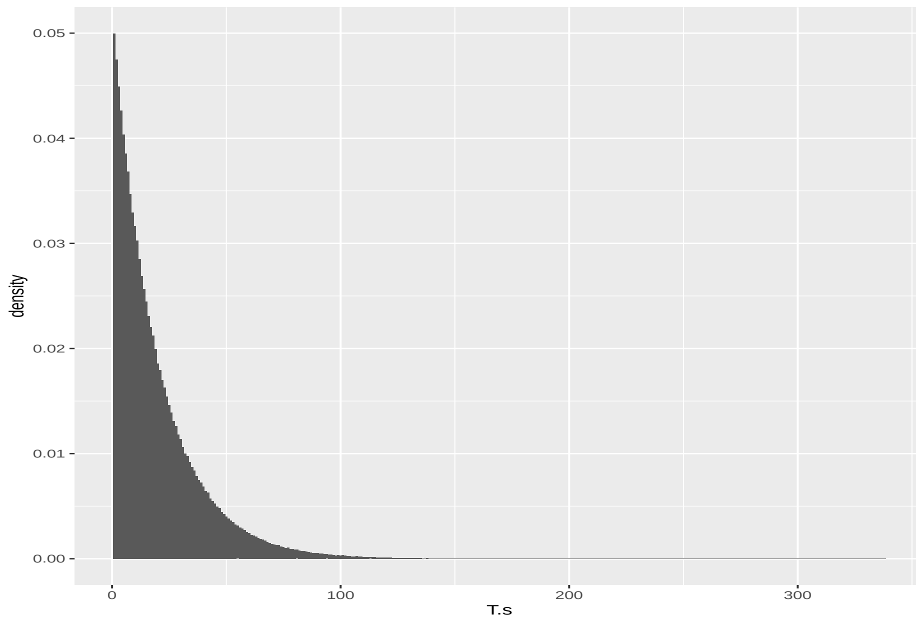

Figure 2: *Normalized histogram of transition time realizations with 1 000 000 simulations. The histogram shown is for the transition from susceptible to infected.*

e)

A realization is plotted in Figure 3. The number of infected individuals converges to zero, which results in the number of susceptible and recovered individuals to stabilise.

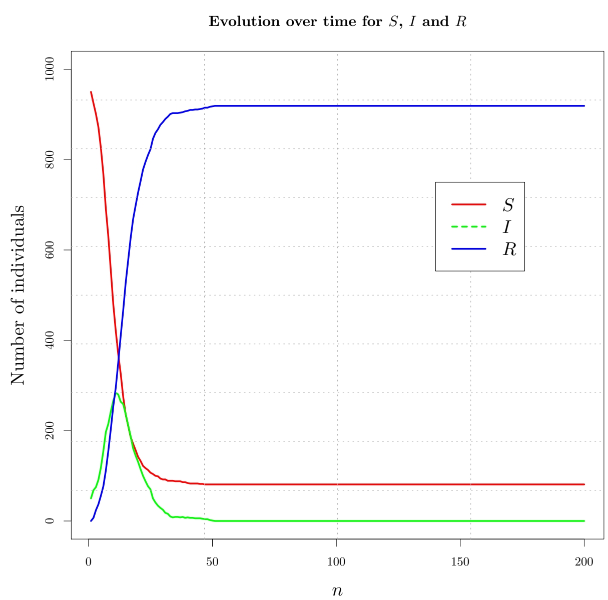

Figure 3: *Evolution of the states S, I and R over time.*

f)

After 1000 simulations, the rounded off expected max of I, $\mathrm{E}[\max\{I_0, \dots, I_{200}\}] = 277$ with the expected time step of max $\mathrm{E}\left[\min\left\{\arg \max_{n \leq 200}\{I_n\}\right\}\right] = 12$, rounded off. A realization can be seen in Figure 3. The results from the simulation is summarised in Table 1.

	$I_{\max}$		$n_{\max}$	
	Average	Standard deviation	Average	Standard deviation
Exact	277.246000	22.368456	12.469000	1.237243
Rounded off	277	22	12	1

Table 1: *Results from simulation for* $\mathrm{E}[I_{\max}]$ *and* $\mathrm{E}[n_{\max}]$ *with* $N = 1000$ *simulations.*

Problem 2: Insurance claims

a)

Let $X(t) \sim \text{Poisson}(\lambda t)$ be a Poisson process denoting the number of insurance claims, with t measured in days from January 1 and stationary intensity $\lambda(t) = 1.5$ for $t \geq 0$. The probability that there are more than 100 claims before March 1$^{\text{st}}$ is then,

$$
\begin{aligned}
P(X(59) > 100) &= 1 - P(X(59) \leq 100) \\
&= 1 - \sum_{x=0}^{100} p(x) \\
&= 1 - \sum_{x=0}^{100} \frac{(1.5 \cdot 59)^x}{x!} e^{-1.5 \cdot 59} \\
&= 0.102822
\end{aligned}
$$

This result can also be shown by simulating realizations from a Poisson process with same intensity λ and taking the mean of all realizations. The attached R code simulates 1000 Poisson realizations, repeated $N = 1000$ times, and estimates the above probability. The estimated probability can be seen from Table 2 to verify the theoretical result.

Variable	Average	Standard deviation
$P(x)$	0.103021	0.009503

Table 2: *Results from simulation with* $N = 1000$ *simulations.*

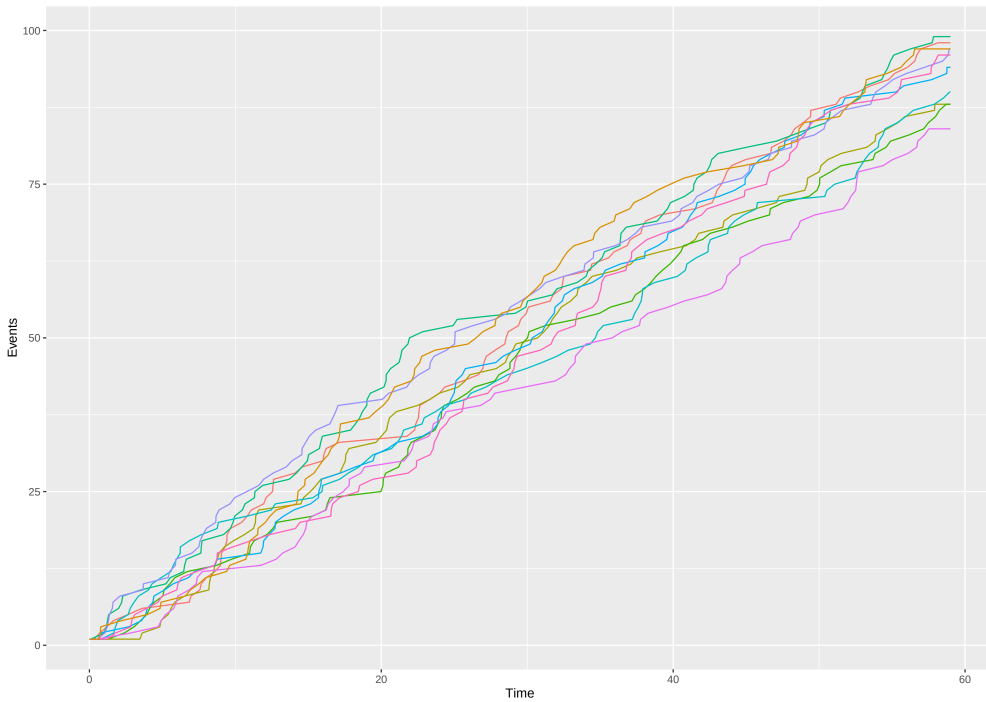

Figure 4: *10 realizations of $X(t)$ for $0 \leq t \leq 59$ with different colors for each realization*

b)

Denoting a claim amount $C_i \sim \text{Exp}(\beta)$ to be exponentially distributed with $\beta = 10$, and the total claim amount at time t to be $Z(t) = \sum_{i=1}^{X(t)} C_i$, the expectation and variance of the total claim amount can be calculated by the law of total expectation and the law of total variance.

$$\text{Law of total expectation:} \quad E(Z(t)) = E(E(Z(t)|X(t))) \tag{3}$$

$$\text{Law of total variance:} \quad Var(Z(t)) = E(Var(Z(t)|X(t))) + Var(E(Z(t)|X(t))) \tag{4}$$

Using the law of total expectation, the expected total claim amount for $t = 59$ days is calculated to,

$$\begin{aligned}
E(Z(t)) &= E\left(E(Z(t)|X(t))\right) \\
&= E\left(E\left(\sum_{i=1}^{X(t)} C_i \middle| X(t)\right)\right) \\
&= E\left(\sum_{i=1}^{X(t)} E(C_i)\right) \\
&= E\left(\frac{X(t)}{\beta}\right) \\
&= \frac{\lambda t}{\beta} \\
&= 8.85
\end{aligned}$$

Using the law of total variance, the variance of the total claim amount for $t = 59$ days is calculated to,

$$
\begin{aligned}
Var(Z(t)) &= E\left(Var(Z(t) \mid X(t))\right) + Var\left(E(Z(t) \mid X(t))\right) \\
&= E\left(Var\left(\sum_{i=1}^{X(t)} C_i \,\middle|\, X(t)\right)\right) + Var\left(E\left(\sum_{i=1}^{X(t)} C_i \,\middle|\, X(t)\right)\right) \\
&= E\left(\sum_{i=1}^{X(t)} Var(C_i)\right) + Var\left(\frac{X(t)}{\beta}\right) \\
&= E\left(\frac{X(t)}{\beta^2}\right) + \frac{1}{\beta^2} Var\left(X(t)\right) \\
&= \frac{\lambda t}{\beta^2} + \frac{\lambda t}{\beta^2} \\
&= \frac{2\lambda t}{\beta^2} \\
&= 1.77
\end{aligned}
$$

These results can also be shown by simulation, given in the attached R code. The estimated expectation and variance can be seen from Table 3, repeating the simulations $N = 1000$ times. The estimated expectation and variance from the simulations can be seen to correspond with the theoretical values.

Variable	Average	Standard deviation
$E(Z(t))$	8.8825	0.9193
$Var(Z(t))$	1.7765	0.1839

Table 3: *Results from simulation with $N = 1000$ simulations.*