

TMA4265 Stochastic Modelling - Fall 2020

Project 2

Alexander J Ohrt, Jim Totland

October 2020

Problem 1: Modelling the common cold

a) $\{X(t) : t \geq 0\}$ is a continuous-time Markov chain because there is no memory in the process. It is not possible to gain any more information about the process by conditioning on several previous steps in the chain. Hence, $\{X(t) : t \geq 0\}$ satisfies the Markov property. Before going any further we will define the states $\{0,1,2\}$ to represent the states $\{S, I_L, I_H\}$ respectively. The jump probabilities are

$$\begin{aligned}P(S \rightarrow I_L) &= \frac{0.90\lambda}{0.10\lambda + 0.90\lambda} = 0.90, \\P(S \rightarrow I_H) &= \frac{0.10\lambda}{0.10\lambda + 0.90\lambda} = 0.10, \\P(I_L \rightarrow I_H) &= 0, \quad P(I_L \rightarrow S) = \frac{\mu_L}{\mu_L} = 1, \\P(I_H \rightarrow I_L) &= 0, \quad P(I_H \rightarrow S) = \frac{\mu_H}{\mu_H} = 1.\end{aligned}$$

The transition rates are

$$\begin{aligned}q_{S,I_L} &= 0.90\lambda = 0.0090, \quad q_{S,I_H} = 0.10\lambda = 0.0010, \\q_{I_L,I_H} &= 0, \quad q_{I_L,S} = \mu_L = 1/7, \\q_{I_H,I_L} &= 0, \quad q_{I_H,S} = \mu_H = 1/20.\end{aligned}$$

The transition diagram is shown in figure 1.

b) Calculating the stationary distribution, the long-run mean fraction of time in each of the states is

$$\begin{aligned}\pi_0\lambda &= \pi_0/100 = \pi_1\mu_1 + \pi_2\mu_2 = \pi_1/7 + \pi_2/20, \\ \pi_1\mu_1 &= \pi_1/7 = \pi_0 \cdot 0.90\lambda = \pi_0 \cdot 0.0090, \\ \pi_0 + \pi_1 + \pi_2 &= 1.\end{aligned}$$

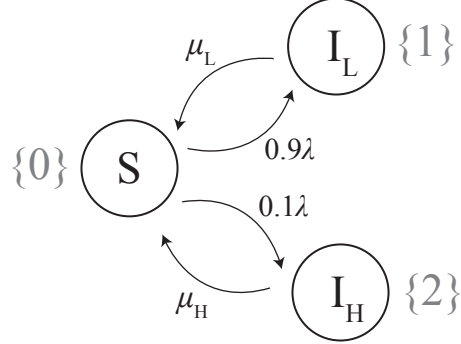


Figure 1: Transition diagram of the continuous-time Markov chain $\{X(t) : t \geq 0\}$.

The solution to this system is

$$\pi_0 = \frac{1000}{1083} \approx 0.923, \quad \pi_1 = \frac{21}{361} \approx 0.058, \quad \pi_2 = \frac{20}{1083} \approx 0.019.$$

which means that the average amount of time per year an individual is infected is

$$\bar{z} = \pi_1 + \pi_2 = \frac{83}{1083} \approx 0.0766,$$

which amounts to $365 \cdot \bar{z} \approx 27.6 \approx 28$ days per year.

c) One realization of the Markov chain over 5 years, after simulating the Markov chain for 1000 years, is shown in figure 2.

d) The estimate, based on one realization of 1000 years, is 0.0749. This estimate is calculated by adding all the sojourn times in state 1 and 2 throughout the 1000 years, before dividing this sum by the total days the simulation was run for. Comparing the estimate to the value calculated in b), it is apparent that the estimation of the infected time in the long run is within the same day as the theoretical result, since 0.0749 amounts to approximately 27.34 days.

e) We begin by defining the expected value

$$v_i = E[\min\{t \geq 0 : X(t) = 2\} | X(0) = i]. \quad (1)$$

We want to determine the value of v_0 , because this is the expected time of returning to state 2 (I_H) if we assume that the last heavy infection ended exactly before $t = 0$. We trivially have that $v_2 = 0$. Furthermore, we have

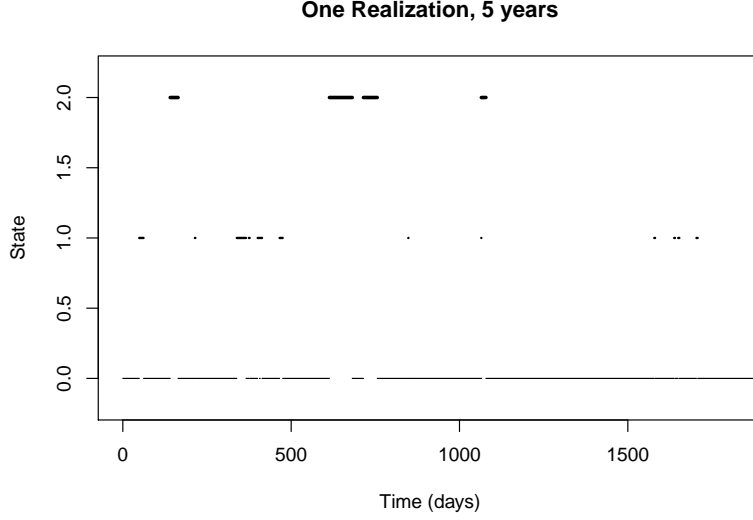


Figure 2: One realization of the Markov chain over 5 years, after simulating the Markov chain for 1000 years $\{X(t) : t \geq 0\}$

$$\begin{aligned} v_0 &= \frac{1}{\lambda} + 0.9v_1 + 0.1v_2 \\ v_1 &= \frac{1}{\mu_L} + v_0 \end{aligned} \tag{2}$$

This results in $v_0 = \frac{1/\lambda + 0.9/\mu_L}{0.1} = 1063$. Thus, the expected time between two heavy infections is 1063 days. One realization of the Markov chain with $0 \leq t \leq 1000 \cdot 365$ gives an approximated expected time between heavy infections $\hat{v}_0 = 1059.84$ days. This result has high variability. Taking a mean of the result over 100 simulations of realizations, yields $\hat{\hat{v}}_0 = 1063.868$, which supports the theoretical result.

f) $\{Y(t) : t \geq 0\}$ is a birth and death process because the number of infected people in the population either increases (birth) or decreases (death). Moreover, all times until next infection (birth) or transition to susceptible (death) are independent and exponentially distributed. Let $n = 5.26 \cdot 10^6$ denote the total population and let the number of infected at a given time $t = \tau$ be $Y(\tau) = i$. Then the birth and death rates are given by

$$\begin{aligned} \lambda_i &= (n - i)\lambda \quad \text{and} \\ \mu_i &= i\mu, \quad \text{for } i = 0, 1, \dots, n, \end{aligned} \tag{3}$$

respectively. The transition diagram of the birth and death process $\{Y(t) : t \geq 0\}$ is shown in figure 3.

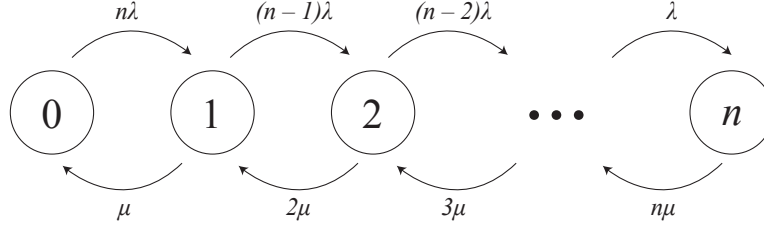


Figure 3: Transition diagram of the birth and death process $\{Y(t) : t \geq 0\}$.

g) We look at one person's state S or I (susceptible or infected). This is a continuous-time Markov chain, and computing the limiting probabilities with the balance equations gives

$$\begin{aligned}\pi_S \lambda &= \pi_I \mu \\ \pi_S + \pi_I &= 1,\end{aligned}\tag{4}$$

which has the solution $\pi_I = \frac{7}{107}$ and $\pi_S = \frac{100}{107}$. We know that the fraction of susceptible people in the long run is π_S . This means that the rate of each person becoming infected in the long run is $\pi_S \cdot \lambda = \frac{1}{107}$. Finally, the total rate of serious complications in the entire population is $\lambda^* = \frac{1}{107} \cdot 0.01 \cdot n$, which in Little's law gives

$$W = \frac{L}{\lambda^*} = \frac{2000}{\lambda^*} \approx 4.07 \text{ days}.\tag{5}$$

Hence the average treatment time required so the average number of individuals in the hospital does not exceed the capacity of 2000 is approximately 4.07 days.

Problem 2: Calibrating climate models

a) Figure 4 shows the prediction as a function of θ , along with 90% prediction intervals.

b) Figure 5 shows the conditional probability that $Y(\theta) < 0.30$ given the 5 evaluation points, computed from the multivariate Gaussian distribution from earlier.

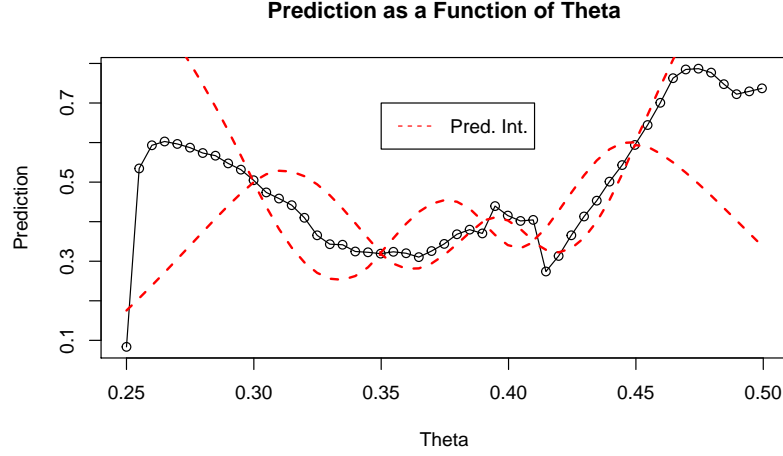


Figure 4: Prediction as a function of θ , along with 90% prediction intervals. The prediction interval is shown with red dotted lines.

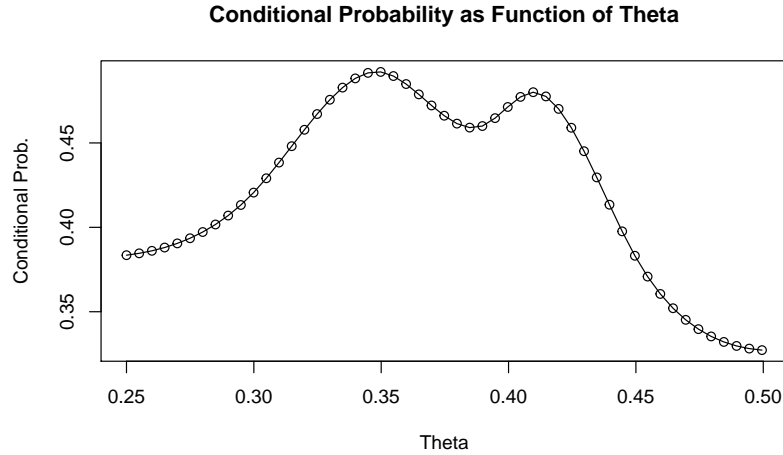


Figure 5: Conditional probability that $Y(\theta) < 0.30$, given the 5 evaluation points, plotted against the grid of parameter values from $\theta = 0.25$ to $\theta = 0.50$.

c) Figures 6 and 7 show the new prediction and the new conditional probability that $Y(\theta) < 0.30$, after adding the new observation to the list of evaluation points. Based on this, the value of θ we would suggest to run the climate model with to have the best chance to achieve $y(\theta) < 0.30$ is $\theta = 0.35$, since it is observed as close to the maximum value in figure 7.

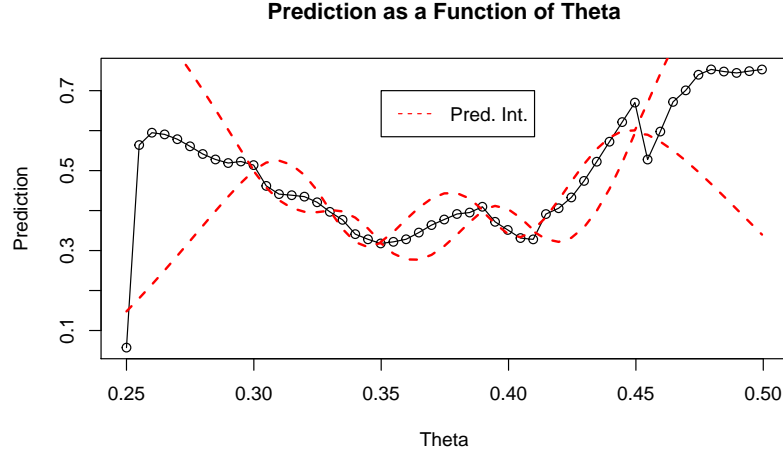


Figure 6: Prediction as a function of θ , along with 90% prediction intervals, after adding the new observation. The prediction interval is shown with red dotted lines.

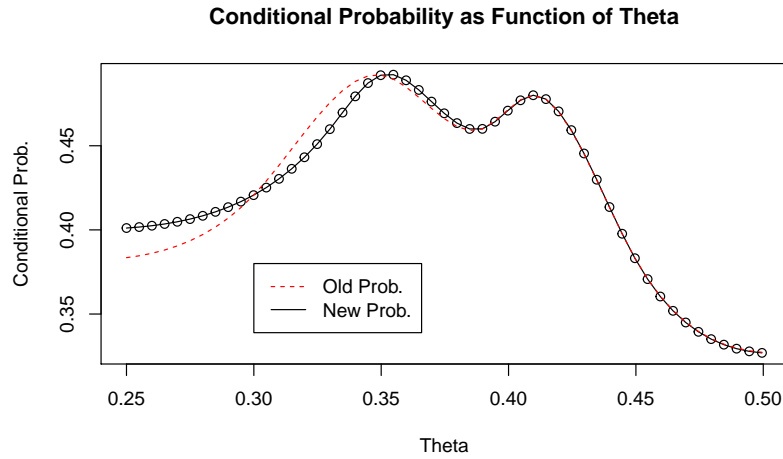


Figure 7: Conditional probability that $Y(\theta) < 0.30$, given the evaluation points after adding the new observation, plotted against the grid of parameter values from $\theta = 0.25$ to $\theta = 0.50$. The red dotted line shows the conditional probability from b), when only using the first 5 observations.