

Project 1 TMA4265

Eivind H. Brataas and Sturla Frydenlund

7 November 2019

Problem 1

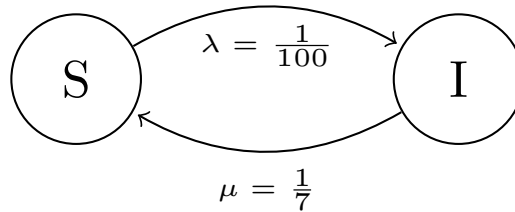
- a) $X(t)$ can be in one of two possible states; either the individual is susceptible to infections or it is infected. The probability of being infected when susceptible (or vice versa) after a given time t , only depends on the current state (at time s), not what happened before:

$$Pr\{X(t+s) = j | X(s) = i, X(u) = x(u), 0 \leq u < s\} = Pr\{X(t+s) = j | X(s) = i\}$$

for all $i, j = 0, 1, 2, \dots$ and for all $t \geq 0, s > 0$. Thus $X(t)$ satisfies the Markov property. Lastly, the probability of transitioning from I to S (or vice versa) does not depend on the current time u , that is:

$$P_{ij}(t) = Pr\{X(t+u) = j | X(u) = i\}$$

is independent of $u \geq 0$ for all $i, j = 1, 2, \dots$. Since $X(t)$ has a countable state space, satisfies the Markov property and has stationary transition probabilities, $X(t)$ is a continuous-time Markov chain. An individual spends 100 days on average being susceptible before transitioning to state I, and an infected person spends on average 7 days being sick before recovering. Therefore the transition rates are $\lambda = \frac{1}{100}$ days and $\mu = \frac{1}{7}$ days. Here is a visualisation of the process:



- b) A person goes from being susceptible to being infected with a frequency of 100 days, and the frequency of recovering from an infection is 7 days. In other words there is a $\frac{100}{107}$ chance of being susceptible and a $\frac{7}{107}$ chance of being infected per unit time. We would therefore expect that the fraction of time spent being sick approaches $\frac{7}{107} = 0.0654$ as the time increases. One year is a good amount of time, so after a year, we would expect the fraction amount of time a person is sick to be 0.654.
- c) Figure 1 shows one realization of $X(t)$ over 5 years. When the state is 2, the person is sick, and when the state is 1, the person is healthy.

According to the simulation, the long-run mean fraction of time per year that an individual has a cold was about 0.06463. This was calculated by simulating the days where the individual was healthy and sick independently, until the sum of days became more than 1000*365.

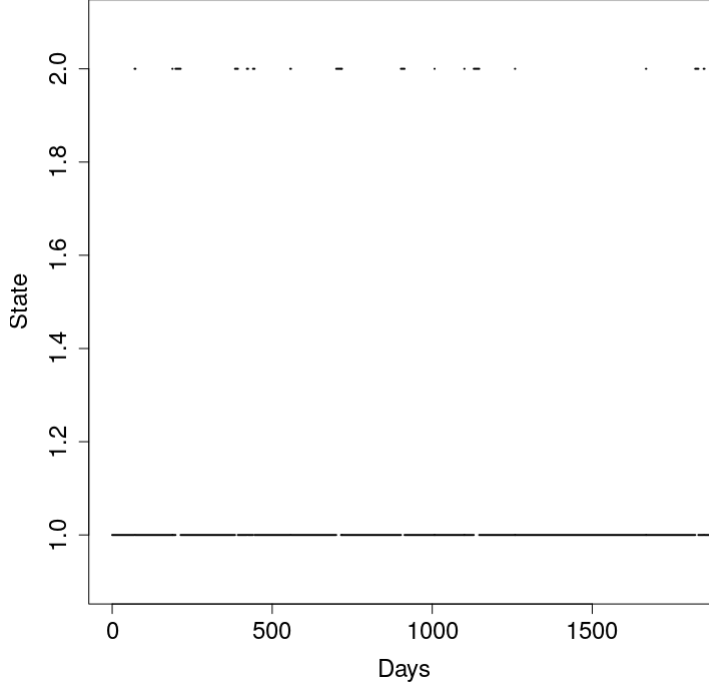
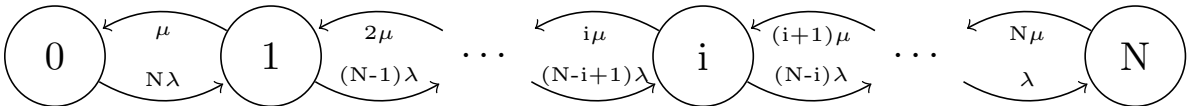


Figure 1: 5 year realization of $X(t)$

- d) $Y(t)$ is the number of infected people in a population at time t . We can view changes in $Y(t)$ as jumps from one state to another, where each state corresponds to a number of infected people in a given population. The jumps can happen at any time, and from each state they happen with certain frequencies. The transitions can only be i to $i - 1$ or i to $i + 1$.

Let λ_i denote the rate of jumping from state i to state $i + 1$, and let μ_i denote the rate of jumping from state i to state $i - 1$, where $i = 0, 1, 2, \dots$. The more susceptible people there are in a given population N , the bigger the chance of someone getting infected will be, so as more and more people get infected, the rate of being infected decreases. We also know that for a given susceptible person, the rate of getting infected is λ . Let us therefore denote $\lambda_i = (N - i)\lambda$ for $i = 0, 1, 2, \dots$.

As more and more people get sick, more and more people can recover, so the rate of recovering increases. For a single individual, the chance of recovering from a disease is μ , so $\mu_i = i\mu$. Initially, at time 0, the disease has just been "developed", so nobody has been infected. We therefore set $Y(0) = 0$. At this point, $Y(t)$ has to increase (if no one is sick, no one can recover). This means that $\mu_0 = 0$. So the birth rates satisfies $\lambda_i > 0$ for $i = 0, 1, \dots$ and the death rates satisfy $\mu_0 = 0$ and $\mu_i > 0$ for $i = 1, 2, \dots$. These are sufficient conditions for modeling $Y(t)$ as a birth and death process.



- e) For a queueing system that has reached its steady state, little's law applies: $L = \lambda^* W$. For our problem, L is the average number of people at a hospital, λ^* is the rate of seriously

infected people arriving at the hospital and W is the average treatment time for each person at the hospital. We now calculate λ^* . Let $\lambda(t) = \lambda_t = (N - Y(t))\lambda$.

$$\lambda^* = \frac{1}{100}E[\lambda(t)] = \frac{1}{100}E[(N - Y(t))\lambda] = \frac{\lambda}{100}(N - E[Y(t)])$$

From b) we have that the the long-run mean fraction of time that an individual has a cold, $E[Y(t)]$ is $\frac{7}{107}$. So

$$\lambda^* = \frac{1}{100 * 100}(5.26 * 10^6 - \frac{7}{107}) = 526$$

We want to find the maximum amount of time a person can be at the hospital without exceeding 2000 people, that is

$$W = \frac{L}{\lambda^*} \leq \frac{2000}{\lambda^*} = \frac{2000}{526} = 3.8$$

So each person have to get treatment within 3.8 days.

Problem 2

Most of the work for this problem has been done in R. Here we are going to present and analyze the different plots.

- a) We first predict the unknown scores based in our 5 evaluation points, then we make the confidence intervals for each prediction and plot this against the θ s with predicted scores.

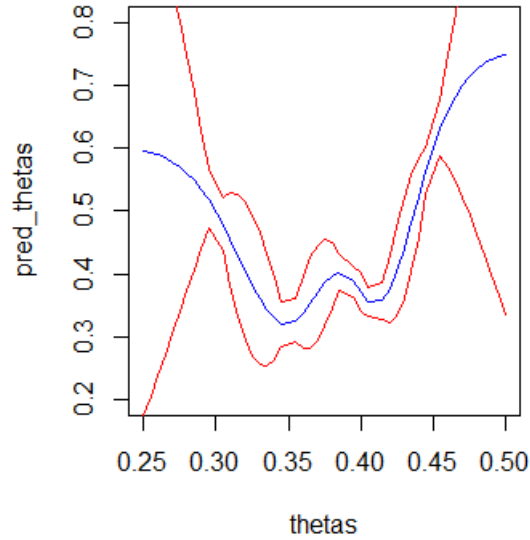


Figure 2: Visualization of 90% confidence intervals for each predicted $Y(\theta)$ s in a)

As figure 2 shows, most of the variation lies at the edges of the set of θ s. This is because that is where we have the least information, or least evaluations.

b) Here we plotted the probability of each predicted $Y(\theta)$ to be less than 0.3.

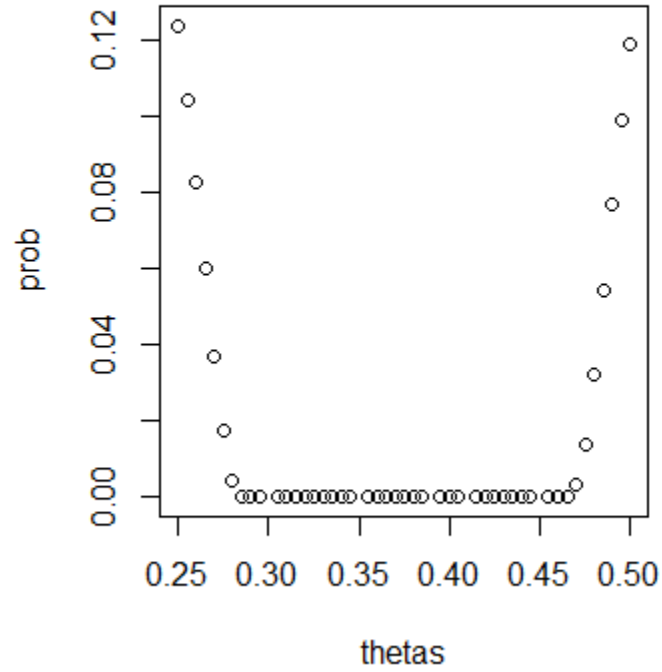


Figure 3: Probability of $Y(\theta)$ being less than 0.30 for each θ in b)

The θ s close to the edges have the highest chance of having scores below 0.3.

- c) We now have one more evaluation point, so our plots of predictions and confidence intervals in figure 4 looks slightly different than figure 2. The biggest difference is that the confidence intervals around our new evaluation point, 0.33, is tighter. However, other than that there is not that much difference. This is because we already have quite a lot of information around these θ s.

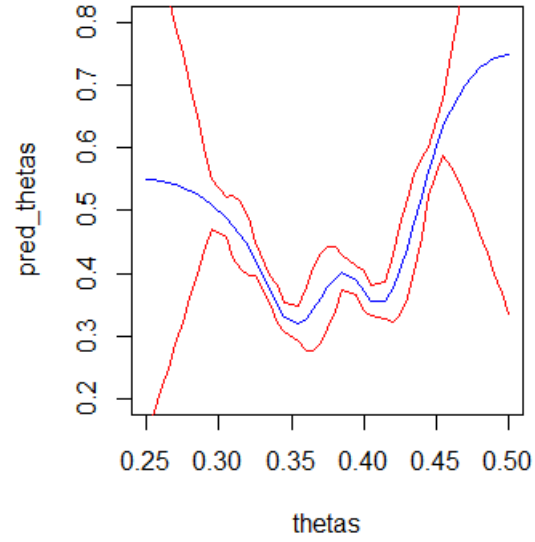


Figure 4: Visualization of 90% confidence intervals for each predicted $Y(\theta)$ s in c)

We now want to find out where we should evaluate to maximize our chances of finding scores below 0.3. We do this by finding the maximum of the plot in figure 5.

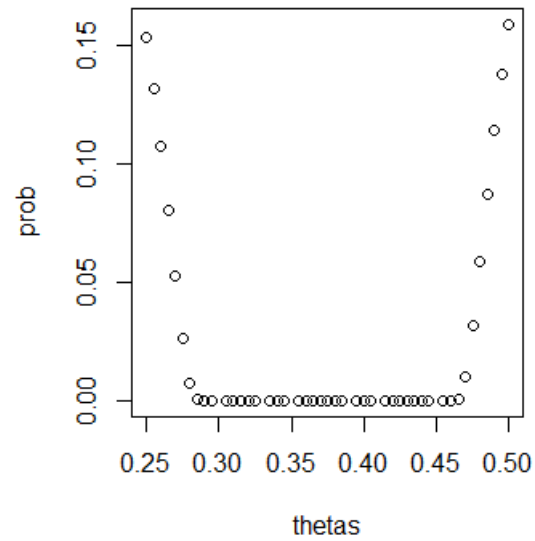


Figure 5: Probability of $Y(\theta)$ being less than 0.30 for each θ in c)

The maximum corresponds to $\theta_0 = 0.50$. This is where we should evaluate.