June 3, 2024

Project 2: Aksel Stenvold, Casper Lindeman and Gonchigsuren Bor

Problem 1: Urgent care centre

a)

The patients at the UCC arrival follow a Poisson process with the rate $\lambda > 0$. The probability of arrival in any interval of time is proportional to the length of the interval and does not depend on the history of the process. Thus, we have a Markovian Arrival Process. Furthermore, the time of treatment follows an exponential distribution with an expected value of $\frac{1}{\mu} > 0$. The probability for the treatment to be completed does not depend on how long the treatment has been going on. The treatment time does not depend on the history of the process which makes the process a Markovian Service Process. Lastly, the UCC has a capacity of 1 patient at a time. Therefore, the UCC is by definition an M/M/1 queue.

For the stochastic process $\{X(t) : t \geq 0\}$ to be viewed as a birth-and-death process, we need $X(t)$ to have only two transitions at a given t. Now since the only possible outcomes are being done with the treatments of a patient, such that $X(t+h) = X(t)-1$, or getting an extra patient, that is $X(t+h) = X(t)+1$, we have a birth-and-death process. The arrival of new patients is given as a Poisson process with a rate of λ, meaning that the "birth" rate is λ. Conversely, the "death" of patients represents the patients leaving after treatment with a rate of μ.

The average time a patient spends in the UCC includes both the treatment time and the waiting time. The total waiting time, W, is determined by the average waiting time and the waiting time W_q combined. This results in

$$W = W_q + \frac{1}{q} = \frac{\lambda}{\mu(\mu-\lambda)} + \frac{1}{q} = \frac{1}{\mu-\lambda}$$

for the average waiting time.

b)

Assume that $\lambda = 5$ patients per hour and $\frac{1}{\mu} = 10$ minutes. We can simulate the UCC in R.

From the simulation, we can estimate the expected time a patient spends in the UCC. We take the total time spent in the UCC and the number of patients served. We find the average time a patient spends in the UCC, W, using Little's law. Stating that

$$W = \frac{L}{\lambda},$$

L is the average number of patients in the queueing system. We find L from the simulation using

$$L = \frac{\text{totalt time spent in the UCC}}{\text{the number of patients served}}.$$

We get $W \approx 1.06$ hours.

Repeating the same simulation 30 times we can approximate a 95% confidence interval. We get $CI \approx [0.93, 1.54]$. Comparing the simulation with the calculated expression, that is,

$$W = \frac{1}{\mu - \lambda} = \frac{1}{6-5} = 1 \text{ hour per patient},$$

we can see that the calculated time is within our confidence interval. Below, one can find a plot of a simulation of the UCC for 12 hours.

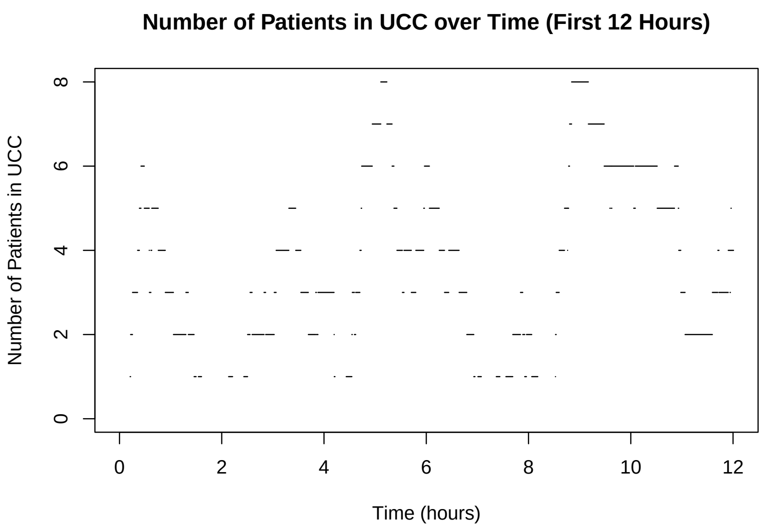

Figure 1: Illustration of the number of patients over a 12-hour period.

c)

The process for urgent patients is unaltered, with one exception. The arrival rate will be changed. They arrive with a Poisson distribution. They still get treated with an exponential distribution. And the UCC still only treats one patient at a time. Therefore $U(T) : T < 0$ still satisfies the conditions of an M/M/1 queue.

We determine the arrival rate by using the law of iterated expectations, that is,

$$\mathrm{E}[g(X)] = \mathrm{E}\Big[E[g(x) \mid Y]\Big]. \tag{1}$$

Given that the number of patients waiting for treatment is $X(t) = U(t)$, and the probability of the treatment being urgent is p, we get

$$\lambda_U = \mathrm{E}[U] = \mathrm{E}\Big[\mathrm{E}[U(t) \mid X]\Big] = \mathrm{E}[Xp] = p\mathrm{E}[X] = \lambda p \tag{2}$$

for the arrival rate.

To find the long-run mean number of urgent patients in the UCC, we use the previous results and insert the new arrival rate. This gives us

$$L_U = \frac{\lambda p}{\mu - \lambda p}. \tag{3}$$

d)

The reason why $N(t)$ does not behave as an M/M/1 queue is that whenever there is an urgent patient that is being treated, treatment of other, "normal" patients is not possible. This is not the case in a standard M/M/1 queue where service is continuously available. Thus, the system does not follow the M/M/1 model's rules.

We determine the long-run mean number using the following expression for the total number of patients $X(t)$

$$X(t) = N(t) + U(t). \tag{4}$$

Now recall that the number of patients waiting for treatment is proportional with its mean value, we get

$$L = L_N + L_U, \tag{5}$$

which results in

$$L_N = L - L_U = \frac{\lambda\mu(1-p)}{(\mu-\lambda)(\mu-\lambda p)}, \tag{6}$$

which is the long-run mean number of normal patients in the UCC.

e)

Using Little's law to determine the expected time in the UCC for both urgent and normal patients, we get

$$W_U = \frac{L_U}{\lambda_U} = \frac{\frac{p\lambda}{\mu - p\lambda}}{p\lambda} = \frac{1}{\mu - p\lambda},$$

for urgent patients. To find the expected time for a normal patient in the UCC, we first need to find λ_N. To do so, we use the same approach as in **1c)**, that is

$$\lambda_N = \mathrm{E}[N] = \mathrm{E}\Big[\mathrm{E}[N|X]\Big] = (1-p)\mathrm{E}[X] = \lambda(1-p). \tag{7}$$

This gives us

$$W_N = \frac{L_N}{\lambda_N} = \frac{\frac{\lambda\mu(1-p)}{(\mu-\lambda)(\mu-\lambda p)}}{\lambda(1-p)} = \frac{\mu}{(\mu-\lambda)(\mu-p\lambda)},$$

for the waiting time for normal patients.

f)

Using the results from the previous sub-problem we can see W_N and W_U as functions of p i Figure 2.

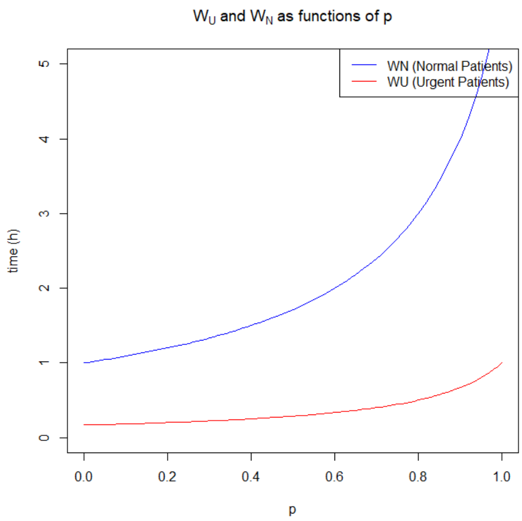

Figure 2: W_N and W_U as functions of p.

When $p \approx 0$ almost all the patients that arrive will be normal patients, this means that no patients will be prioritized and we have the same M/M/1 queue as described in problem a. Making $W_N = \frac{\mu}{(\mu-\lambda)(\mu-p\lambda)} = \frac{1}{\mu-\lambda} = 1$ hour. When $p \approx 1$ almost all the patients that arrive will be urgent patients, this means that none of the patients can be prioritized over any of the other patients and we have the same M/M/1 queue as described in problem a. Making $W_U = \frac{1}{\mu-p\lambda} = \frac{1}{\mu-\lambda} = 1$ hour. If a normal patient arrives when $p \approx 1$ the patient will be down-prioritized by all other patients and can expect to wait for $W_N = \frac{\mu}{(\mu-\lambda)^2} = 5$ hours.

The p for which the expected time spent at the UCC for a normal patient corresponding to 2 hours is:

$$W_U = \frac{L_U}{p\lambda} = \frac{\frac{p\lambda}{\mu-p\lambda}}{p\lambda} = \frac{1}{\mu-p\lambda}. \implies p = \frac{\mu}{\lambda} - \frac{\mu}{(\mu-\lambda)W_N\lambda} = \frac{3}{5}.$$

g)

After running the simulation 30 times we get a confidence interval for urgent patients $CI \approx [2.74, 3.48]$. And for normal patients, we get $CI \approx [3.16, 6.80]$. Comparing the simulation with the calculated expression

$$W_U = \frac{1}{\mu-p\lambda} = \frac{1}{2} \text{ hour per patient.} W_N = \frac{\mu}{(\mu-\lambda)(\mu-p\lambda)} = 3 \text{ hour per patient.}$$

The times are not found within our confidence interval, which means that we probably have done something wrong in our simulation when finding the CI.

Lastly, in Figure 3 we can see a simulation over 12 hours at the UCC, with the total amount of patients and the number of urgent patients.

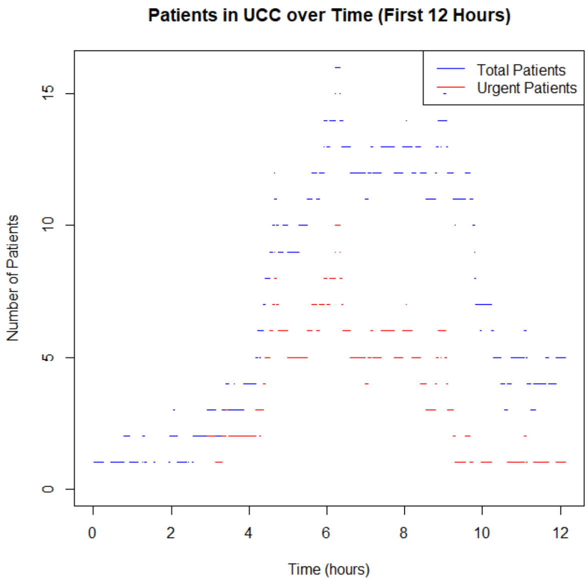

Figure 3: Illustration of total number of patients with the number of urgent patients over a 12-hour period.

Problem 2: Calibrating climate models

In this task, we will use a Gaussian process model $\{Y(\theta) : \theta \in [0,1]\}$ to model the unknown relationship between the parameter value and the score. We will use $\mathrm{E}[Y(\theta)] \equiv 0.5$, $\mathrm{Var}[Y(\theta)] \equiv 0.5^2$ and $\mathrm{Corr}[Y(\theta_1), Y(\theta_2)] = (1 + 15|\theta_1 - \theta_2|)\exp(-15|\theta_1 - \theta_2|)$ for $\theta_1, \theta_2 \in [0,1]$.

a)

In this exercise, we denote by $\boldsymbol{\theta}_A$ a regular grid of parameter values from $\theta = 0.25$ to $\theta = 0.5$, and by $\mathbf{Y}(\boldsymbol{\theta}_B) = (Y(\theta_{B,1}), \ldots, Y(\theta_{B,5}))$ we denote the measurements for the Gaussian process at the points $\boldsymbol{\theta}_B = (\theta_{B,1}, \ldots, \theta_{B,5})$. Furthermore, we denote by $\boldsymbol{\Sigma}_{A,A}$, $\boldsymbol{\Sigma}_{A,B}$, $\boldsymbol{\Sigma}_{B,A}$ and $\boldsymbol{\Sigma}_{B,B}$ the covariance between the stochastic variables $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_A$, $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$, $\boldsymbol{\theta}_B$ and $\boldsymbol{\theta}_A$, and $\boldsymbol{\theta}_B$ and $\boldsymbol{\theta}_B$ respectively. Now since we already know the outcome of the Gaussian process at some locations, we will have to be sure that the resulting curve for the mean of the entire process goes through those points. Therefore, we will compute both conditional mean and variance for the process.

To do so, we first find the covariances $\boldsymbol{\Sigma}_{A,A}$, $\boldsymbol{\Sigma}_{A,B}$, $\boldsymbol{\Sigma}_{B,A}$ and $\boldsymbol{\Sigma}_{B,B}$ using the given formula for the correlation

$$
\begin{aligned}
\boldsymbol{\Sigma}_{A,A} &= \sigma^2(1 + 15|(\theta_{A,i} - \theta_{A,j}|)\exp\{-15|\theta_{A,i} - \theta_{A,j}|\} \\
\boldsymbol{\Sigma}_{A,B} &= \sigma^2(1 + 15|(\theta_{A,i} - \theta_{B,j}|)\exp\{-15|\theta_{A,i} - \theta_{B,j}|\} \\
\boldsymbol{\Sigma}_{B,A} &= \sigma^2(1 + 15|(\theta_{B,i} - \theta_{A,j}|)\exp\{-15|\theta_{B,i} - \theta_{A,j}|\} \\
\boldsymbol{\Sigma}_{B,B} &= \sigma^2(1 + 15|(\theta_{B,i} - \theta_{B,j}|)\exp\{-15|\theta_{B,i} - \theta_{B,j}|\}\,.
\end{aligned}
$$

Using those results we then compute the conditional expected value and the conditional variance

$$\mathrm{E}(\boldsymbol{\theta}_A|\boldsymbol{\theta}_B) = \boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{A,B}\boldsymbol{\Sigma}_{B,B}^{-1}(\mathbf{x}_B - \boldsymbol{\mu}_B)$$
$$\mathrm{Var}(\boldsymbol{\theta}_A|\boldsymbol{\theta}_B) = \boldsymbol{\Sigma}_{A,A} - \boldsymbol{\Sigma}_{A,B}\boldsymbol{\Sigma}_{B,B}^{-1}\boldsymbol{\Sigma}_{B,A}.$$

The 90% prediction interval at a given point $\theta_{A,i}$ is thus

$$\left[\mathrm{E}(\boldsymbol{\theta}_A|\boldsymbol{\theta}_B) - z_{0.05}\sqrt{\mathrm{Var}(\boldsymbol{\theta}_A|\boldsymbol{\theta}_B)},\ \mathrm{E}(\boldsymbol{\theta}_A|\boldsymbol{\theta}_B) + z_{0.05}\sqrt{\mathrm{Var}(\boldsymbol{\theta}_A|\boldsymbol{\theta}_B)}\right], \tag{8}$$

where $z_{0.05} = 1.645$.

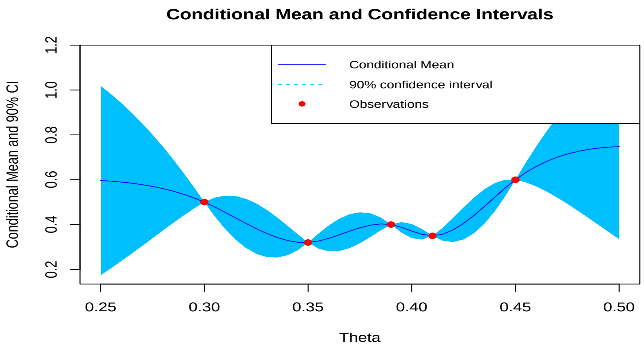

Figure 4: Illustration of the conditional mean and its 90% prediction interval as function of θ, as well as the given observation points.

In figure 4, one can find a plot for the conditional mean as a function of θ (solid blue line), the 90% prediction interval for the conditional mean (sky blue surface), as well as the given observations (red dots). As we can see, the conditional mean goes through our observations which makes sense as some of the covariances used to find the conditional mean are dependent on the observations. In addition, one can also see that the prediction interval gets narrower the closer we are to the observations, as there is no insecurity about the correctness of the observations. Conversely, the prediction interval gets broader the further away we are from the observations, as the insecurity about the correctness of the calculations reflects on points further away from the observations. Note that the prediction interval is at its largest on the endpoints as we do not have any further observations to narrow down the prediction interval at the endpoints.

b)

Here, we compute the conditional probability that $Y(\theta) < 0.3$ given the 5 evaluation points by standardising $Y(\boldsymbol{\theta}) \sim \mathrm{N}(\mathrm{E}(\mathbf{x}_A|\mathbf{x}_B), \mathrm{Var}(\mathbf{x}_A|\mathbf{x}_B))$ to $Z(\boldsymbol{\theta}) \sim \mathrm{N}(0,1)$. This results in

$$\Pr\Big(Y(\boldsymbol{\theta}) < 0.3 \mid Y_B(\boldsymbol{\theta}_B)\Big) = \Pr\left(Z < \frac{0.3 - \mathrm{E}(\mathbf{x}_A|\mathbf{x}_B)}{\sqrt{\mathrm{Var}(\mathbf{x}_A|\mathbf{x}_B)}}\right) \approx \Pr\left(Z \leq \frac{0.3 - \mathrm{E}(\mathbf{x}_A|\mathbf{x}_B)}{\sqrt{\mathrm{Var}(\mathbf{x}_A|\mathbf{x}_B)}}\right).$$

We then compute the probability using the function *pnorm* in R. This yields us figure 5 in which we can see that the conditional probability for $\Pr(Y(\boldsymbol{\theta}) < 0.3 \mid Y_B(\boldsymbol{\theta}_B))$ is zero for all θ observed, since $Y(\boldsymbol{\theta}) > 0.3$ for all observations. However, we can see that the probability for $Y(\boldsymbol{\theta}) < 0.3$ is largest in-between the observation points, with its greatest value being at around 0.24 for $0.3 < \theta < 0.35$.

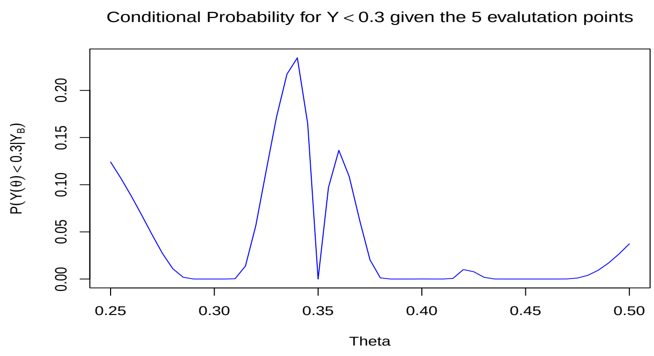

Figure 5: Illustration of the conditional probability $\Pr\left(Y(\boldsymbol{\theta}) < 0.3 \mid Y_B(\boldsymbol{\theta}_B)\right)$ as a function of θ

c)

Consider figure 7, that is, the plot for the conditional probability for $Y(\theta) < 0.3$ given the new observations. We can see that the probabilities have changed from the ones in **2b)**, with now only 2 big spikes at the start of the graph and for $0.35 < \theta < 0.39$ - which is the biggest spike at around 0.18. Since the conditional probability for $Y(\theta) < 0.3$ is highest when $0.35 < \theta < 0.39$, we would suggest to run the model at around $\theta = 0.37$, that is, a value in-between the closed interval $[0.35, 0.39]$.

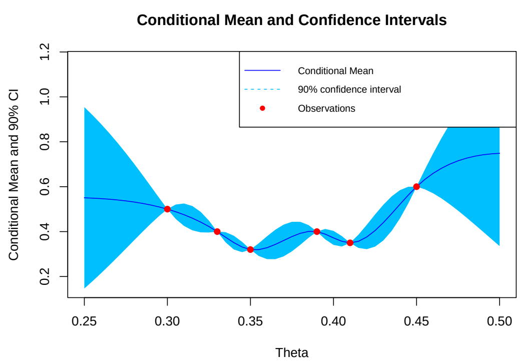

Figure 6: Illustration of the conditional mean and confidence interval given the new observation.

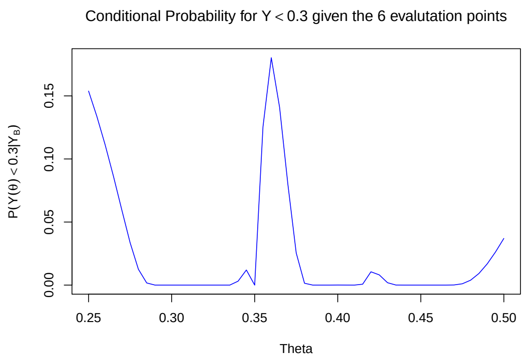

Figure 7: Illustration of the conditional probability for $Y(\theta) < 0.3$ given the new observation.