Suggested solution, exam TMA4265, Stochastic Modeling, Aug 8, 2018

Task 1

a)

State 1 is an absorbing state (recurrent). The other states; $\{2, 3, 4\}$, are transient - the patient will not return infinitely many times to these states (in the long run). The absorbing state means that a patient will remain healthy, once s/he reaches this state. This is true for some child diseases, and some common virus diseases (like HIV), when a patient gets medicine.

$$P(X_2 = 3|X_0 = 4) = \sum_{k=1}^{4} P(X_2 = 2|X_1 = k)P(X_1 = k|X_0 = 4)$$

$$= 0 + 0 + 0.1 \cdot 0.9 + 0.9 \cdot 0.4 = 0.45$$

b)

The time until leaving state 4 is geometrically distributed with parameter $p = 0.9$. This means that the expected number of time steps is then $1/0.9 = 1.11$.

Define $u_i = E(\text{additional time steps to get healthy}|X_t = i)$.

By a first-step analysis:

$$\begin{aligned}
u_1 &= 0 \\
u_2 &= 1 + 0.3u_1 + 0.6u_2 + 0.1u_3 \\
u_3 &= 1 + 0.5u_2 + 0.4u_3 + 0.1u_4 \\
u_4 &= 1 + 0.9u_3 + 0.1u_4
\end{aligned}$$

This means that

$$\begin{aligned}
u_4 &= \frac{1 + 0.9u_3}{1 - 0.1} = 1.11 + u_3 \\
u_3 &= 2(1 + 0.1 \cdot 1.11) + u_2 \\
u_2 &= \frac{1 + 0.1 \cdot 2(1 + 0.1 \cdot 1.11)}{1 - 0.6 - 0.1} = 4.1
\end{aligned}$$

Then

$$u_3 = 2.2 + 4.1 = 6.3$$

$$u_4 = 1.1 + 6.3 = 7.4$$

c)

The fastest time to get healthy is three time stages.
The probability of this is

$$P(X_3 = 1, X_2 = 2, X_1 = 2|X_0 = 4) = P(X_3 = 1|X_2 = 2)P(X_2 = 2|X_1 = 3)P(X_1 = 3|X_0 = 4)$$

$$= 0.3 \cdot 0.5 \cdot 0.9 = 0.135$$

The patient can only get healthy from state 2, so to fulfil the condition of getting healthy at time step 4, we must have $(X_4 = 1, X_3 = 2)$. The last transition then has probability $P(X_4 = 1|X_3 = 2) = 0.3$. To get to state 2 in 4 steps, the patient must stay in any one of the states for $\{2, 3, 4\}$ for one time step, and otherwise move down in the classes at every time step. This means three ways of getting healthy at step $t = 4$.

The probability is

$$\begin{aligned} P(X_4 = 1, X_3 = 2|X_0 = 4) &= P(X_4 = 1, X_3 = 2, X_2 = 3, X_1 = 4|X_0 = 4) \\ &+ P(X_4 = 1, X_3 = 2, X_2 = 3, X_1 = 3|X_0 = 4) \\ &+ P(X_4 = 1, X_3 = 2, X_2 = 2, X_1 = 3|X_0 = 4) \\ &= 0.1 \cdot 0.9 \cdot 0.5 \cdot 0.3 + 0.9 \cdot 0.4 \cdot 0.5 \cdot 0.3 + 0.9 \cdot 0.5 \cdot 0.6 \cdot 0.3 \\ &= 0.148 \end{aligned}$$

Task 2

a)

The first arrival must then occur after 15. It time T to the first arrival is exponential distributed

$$P(T > 15) = \exp(-0.2 \cdot 15) = 0.05$$

The number of persons X arriving in the time interval is Poisson distributed with parameter $0.2 \cdot 15 = 3$.

$$P(X = 2) = \frac{3^2}{2} \exp(-3) = 0.22$$

b)

The expected number of arrivals is $t\lambda = 30 \cdot 0.2 = 6$.

Given the 7 arrivals, their time points are uniformly distributed over the 30 min time interval. So the probability of an arrival before 12:15 is $15/30 = 0.5$. The people can have any order. This means

$$P(5\text{arrivals } 15|7\text{arrivals } 30) = P(5\text{arrivals } 15, 2\text{arrivals last } 15) = \frac{7!}{5!2!}0.5^5 0.5^2 = 0.16$$

c)

Given that this is at NTNU, which has 45 min teaching sessions, there is a break 12:00-12:15. There will be many more using the restrooms during breaks. It is more likely to have an inhomogeneous rate than a fixed rate λ.

There are also likely inhomogeneous arrival rates during the day because of more people present, and because of lunch time, etc.

d)

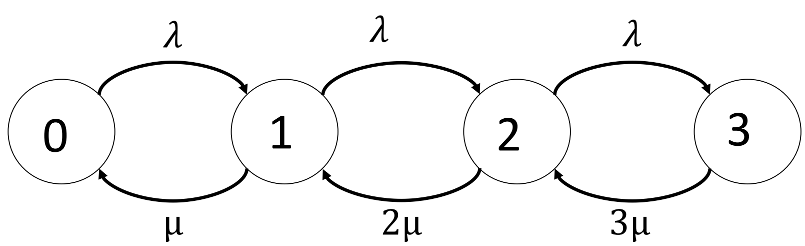

Figure 1: Transition diagram for the number of people in the restroom.

The long-term probabilities $\lim_{t\to\infty} P_{ij}(t) = P_j$, for $j = 0, 1, 2, 3$ states.

The probabilities can be determined by long-term moves in and out of states

$$\begin{aligned} P_0\lambda &= P_1\mu \\ P_1(\mu+\lambda) &= P_0\lambda + P_2 2\mu \\ P_2(2\mu+\lambda) &= P_1\lambda + P_3 3\mu \\ 1 &= P_0 + P_1 + P_2 + P_3 \end{aligned}$$

This gives

$$P_0(1+\mu/\lambda+\lambda^3/2\mu+\lambda^3/6\mu^3) = 1, \quad P_0 = \frac{1}{1+1+0.5+0.17} = 0.375$$

$$P_1 = 0.375, \quad P_2 = 0.19, \quad P_3 = 0.0625$$

e)

With $N = 4$ toilets, the long-term probabiliites become

$$P_0 = \frac{1}{1+\mu/\lambda+\mu^2/(2\lambda^2)+\lambda^3/6\mu^3+\lambda^4/24\mu^4} = 1/(1{+}1{+}0.5{+}0.17{+}0.04) = 0.3692$$

And the probability of having all toilets occupied is then

$$P_4 = P_0/24 = 0.0154$$

In the long run this probability is decreased from 0.0625 to 0.0154, with the additional cost of 1 kr / min.

The long-run expected cost of irritation is $0.0625 \cdot \lambda \cdot 100 = 1.25$ kr /min for the case with $N = 3$ toilets.

With $N = 4$ toilets it it $0.0154 \cdot \lambda \cdot 100 = 0.31$ kr/min.

It is optimal to use $N = 3$ toilets, because the long term additional gain is $1.25 - 0.31 < 1$. The investment of 1 kr /min is not worth it in the long run, if the decision is based on expected costs alone.

Task 3

a)

The conditional distribution is Gaussian with mean and variance

$$\begin{aligned} m(r) = E(x(r)|x(0.4) = 11.5) &= 10 + \frac{(1+5|r-0.4|)e^{-5|r-0.4|}}{1}(11.5-10), \\ \sigma^2(r) = \mathrm{Var}(x(r)|x(0.4) = 11.5) &= 1 - \frac{(1+5|r-0.4|)^2 e^{-2\cdot 5|r-0.4|}}{1}. \end{aligned}$$

$$
\begin{aligned}
P(x(0.5) > 11.5 | x(0.4) = 11.5) &= P\left(Z > \frac{11.5 - m(0.5)}{\sigma(0.5)}\right) \\
&= P\left(Z > \frac{11.5 - 11.36}{0.41}\right) = 0.37
\end{aligned}
$$

$$
\begin{aligned}
P(x(0.6) > 11.5 | x(0.4) = 11.5) &= P\left(Z > \frac{11.5 - m(0.6)}{\sigma(0.6)}\right) \\
&= P\left(Z > \frac{11.5 - 11.10}{0.67}\right) = 0.28
\end{aligned}
$$

b)

For $x = x(r)$, set $p(x) = N(m(r), \sigma^2(r))$. We use a transformation $x = \sigma(r)z + m(r)$, for $z \sim N(0,1)$:

$$
\begin{aligned}
EI &= \int \max\{x - 11.5, 0\} p(x)dx = \int_{11.5}^{\infty} (x - 11.5)p(x)dx \\
&= \int_{\frac{11.5 - m(r)}{\sigma(r)}}^{\infty} (\sigma(r)z + m(r) - 11.5)p(z)dz, \\
&= (m(r) - 11.5)\int_{-\infty}^{v} p(z)dz + \sigma(r)\int_{-v}^{\infty} zp(z)dz \\
&= (m(r) - 11.5)\Phi(v) + \sigma(r)\left[\phi(\infty) - \phi(-v)\right] = (m(r) - 11.5)\Phi(v) + \sigma(r)\phi(v).
\end{aligned}
$$

Assuming first that $\sigma(r)$ is constant. For $m(r) << 11.5$, the EI is small because $\phi(v) \approx 0$ and $\Phi(v) \approx 0$. For $m(r) >> 1.5$, the EI is large beacuse $\Phi(v) = 1$, while $\phi(v) \approx 0$ still. The EI increases with $m(r)$, and EI leads to evaluation points with large $m(r)$.

Assuming next that $m(r)$ is constant, say $m(r) = 11.5$, we get $\Phi(v) = 0.5$, and $\phi(v) = 1/\sqrt{2\pi}$, and the EI increases with $\sigma(r)$. This means that EI encourages evaluation at points with large uncertainty.

$$
EI(0.5) = [11.36 - 11.5]\Phi(\frac{11.37 - 11.5}{0.41}) + 0.41\phi(\frac{11.37 - 11.5}{0.41}) = 0.1067
$$

$$
EI(0.6) = [11.10 - 11.5]\Phi(\frac{11.10 - 11.5}{0.67}) + 0.67\phi(\frac{11.10 - 11.5}{0.67}) = 0.1170
$$

The EI is larger at 0.6, even though the probability of exceeding 11.5 is smaller there (a). This occurs because of the larger uncertainty at $r = 0.6$.