Stoch. Proc. HW assignment 10. Due Friday, November 17 at start of class

1. There are six machines and six repairmen in a factory. Each machine works for an EXP(2) time interval and then it breaks down. Machines break down independently from one another. If a machine breaks down, a repairmen immediately starts repairing it. The time it takes to repair a machine is EXP(3) distributed. The working times and the repair times are independent. If a machine is fixed, it immediately starts working again, and it works until it breaks down again, etc. Initially all machines work. Denote by X_t the number of working machines at time t.

   (a) Write down the infinitesimal generator of the Markov chain (X_t).

   (b) Find the distribution of X_t. In other words, find $p_t(6, x)$ for all $x = 0, 1, \dots, 6$.

Solution:

(a) If x machines work currently and $6 - x$ machines are being repaired, then x independent EXP(2) „break" random variables compete against $6 - x$ independent EXP(3) „repair" random variables. Thus the rate of jump from x to $x - 1$ is $g(x, x - 1) = 2x$ while the rate of jump from x to $x + 1$ is $g(x, x + 1) = 3(6 - x)$ for any $x \in \{0, 1, \dots, 6\}$. The diagonal elements of the infinitesimal generator matrix are of form $g(x, x) = -(2x + 3(6 - x)) = -(18 - x) = x - 18$. The other entries of the infinitesimal generator matrix are zero. Thus the state space is $S = \{0, 1, \dots, 6\}$ and

$$\underline{\underline{G}} = \begin{pmatrix} -18 & 18 & 0 & 0 & 0 & 0 & 0 \\ 2 & -17 & 15 & 0 & 0 & 0 & 0 \\ 0 & 4 & -16 & 12 & 0 & 0 & 0 \\ 0 & 0 & 6 & -15 & 9 & 0 & 0 \\ 0 & 0 & 0 & 8 & -14 & 6 & 0 \\ 0 & 0 & 0 & 0 & 10 & -13 & 3 \\ 0 & 0 & 0 & 0 & 0 & 12 & -12 \end{pmatrix}$$

(b) First let us choose our favourite machine and find the probability p_t that it works at time t. Here we could use diagonalization (like in the case of the Markov chain on page 142 of the scanned lecture notes), but we will use another method: we solve the Kolmogorov forward equation. We have

$$\frac{\mathrm{d}}{\mathrm{d}t} p_t = -2p_t + 3(1 - p_t) = 3 - 5p_t,$$

since $1 - p_t$ is the probability that our favourite machine is being repaired at time t and a working machine breaks at rate 2, while a broken machine gets repaired at rate 3. This is an inhomogeneous linear ODE, so let's first solve the corresponding homogeneous ODE: $\frac{\mathrm{d}}{\mathrm{d}t} p_t = -5p_t$. The general solution of this is $p_t = Ce^{-5t}$. Now a particular solution of the original inhomogeneous ODE $\frac{\mathrm{d}}{\mathrm{d}t} p_t = 3 - 5p_t$ is the constant solution $p_t \equiv 3/5$. So the general solution of the original inhomogeneous ODE is $p_t = 3/5 + Ce^{-5t}$. We need to find C so that the initial condition $p_0 = 1$ is satisfied (initially the machine works). Thus $p_t = \frac{3}{5} + \frac{2}{5} e^{-5t}$.
Now observe that machines break down and get repaired independently from each other. Thus if $A_{i,t}$ is the event that machine i works at time t, then the events $A_{1,t}, A_{2,t}, \dots, A_{6,t}$ are independent and $\mathbb{P}(A_{i,t}) = p_t$ for $i = 1, \dots, 6$, where $p_t = \frac{3}{5} + \frac{2}{5} e^{-5t}$. Hence $X_t = \mathbb{1}[A_{1,t}] + \cdots + \mathbb{1}[A_{6,t}]$ and thus $X_t \sim \mathrm{BIN}(6, p_t)$. Therefore

$$p_t(6, x) = \binom{6}{x} p_t^x (1 - p_t)^{6-x}, \qquad x = 0, 1, 2, \dots, 6.$$

Remark : The stationary distribution is $\mathrm{BIN}(6, \frac{3}{5})$, since $\lim_{t\to\infty} p_t = \frac{3}{5}$.
Note that (X_t) is a birth and death chain, hence it is reversible, hence the detailed balance condition must be satisfied. Let us check for fun that that indeed we have $\pi(x)g(x, x + 1) = \pi(x + 1)g(x + 1, x)$ for all $x = 0, \dots, 5$:

$$\pi(x)g(x, x + 1) = \left(\binom{6}{x} \left(\frac{3}{5}\right)^x \left(\frac{2}{5}\right)^{6-x} \right) \cdot 3(6 - x) = \frac{6!}{x!(5 - x)!} \frac{3^{x+1} 2^{6-x}}{5^6}$$

$$\pi(x + 1)g(x + 1, x) = \left(\binom{6}{x + 1} \left(\frac{3}{5}\right)^{x+1} \left(\frac{2}{5}\right)^{5-x} \right) \cdot 2(x + 1) = \frac{6!}{x!(5 - x)!} \frac{3^{x+1} 2^{6-x}}{5^6}$$

2. *Forest fire model.* We consider the time evolution of a random graph on three vertices. The edge set changes as time evolves. The dynamics is driven by six independent Poisson point processes:

   - for each unordered pair of vertices there is an „edge" PPP with rate 1/3 and upon each arrival of an edge process, we draw and edge between that pair of vertices (unless there is already an edge between them, in which case we don't do anything).
   - for each vertex there is a „lightning" PPP with rate 1 and upon each arrival of a lightning process, a lightning strikes that vertex and immediately burns all of the edges of the connected component of that vertex.

   Denote by G_t the edge set of the graph at time t and let X_t be number of connected components in G_t.

   (a) Argue that (X_t) is a Markov process and find its infinitesimal generator matrix.

   (b) Find the stationary distribution of (X_t).

Solution:

(a) (X_t) is a Markov process because if we know the value of X_t at time t, then in fact we know the sizes of connected components in G_t, and this is enough to figure out the rates at which X_t changes, and we can forget about the rest of the past of (X_t). The state space of (X_t) is $S = \{1, 2, 3\}$ and the infinitesimal generator matrix is

$$\underline{\underline{G}} = \begin{pmatrix} -3 & 0 & 3 \\ 2/3 & -8/3 & 2 \\ 0 & 1 & -1 \end{pmatrix}.$$

In words:
$g(1,3) = 3$ because if $X_t = 1$ then all three vertices are in one component and each of the vertices is exposed to a lightning PPP of rate 1, so the total rate of lightning is 3, and if a lightning strikes the component then all edges burn and we obtain three singleton connected components.
$g(2,1) = 2/3$ because if $X_t = 2$ then there are two possible edges connecting these two components, and an edge arrives at both of these possible edges at rate 1/3, so the total rate of merger is 2/3.
$g(2,3) = 2$ since if $X_t = 2$ then there is a connected component of size 2 and both of the vertices of this component are exposed to a lightning PPP of rate 1, so the total rate of lightning is 2, and if a lightning strikes the component then it splits into two singleton connected components.
$g(3,2) = 1$ since if $X_t = 3$ then there are three possible edges, each of which arrive at rate 1/3 and reduce the number of connected components to 2.

(b) In order to find the stationary distribution $\underline{\pi} = (\pi_1, \pi_2, \pi_3)$ of (X_t), we need to solve

$$\begin{aligned} -3\pi_1 + \frac{2}{3}\pi_2 + 0\pi_3 &= 0, \\ 0\pi_1 - \frac{8}{3}\pi_2 + \pi_3 &= 0, \\ 3\pi_1 + 2\pi_2 - \pi_3 &= 0, \\ \pi_1 + \pi_2 + \pi_3 &= 1. \end{aligned}$$

Note that these are four equations with only three variables, but this system of linear equations is not overdetermined, as we now explain. The sum of the first three equations is just the trivial identity $0 = 0$, so one of the first three equations can be omitted without changing the set of solutions.
The solution of the above system of linear equations is

$$\pi_1 = \frac{2}{35}, \qquad \pi_2 = \frac{9}{35}, \qquad \pi_3 = \frac{24}{35}.$$

Remark: If we consider the forest fire model on four vertices, then the number of connected components is not a Markov chain, because we cannot recover the component size structure from the number of connected components.

3. There is an amoeba at time 0 in a Petri dish. An amoeba splits after an EXP(1) distributed waiting time into two amoebae which are identical to the original one. The goal of this exercise is to show that at time t the number X_t of amoebae in the Petri dish has optimistic GEO(e^{-t}) distribution.

   (a) Write down the infinitesimal generator matrix of the continuous-time Markov chain (X_t).

   (b) Write down the *Kolmogorov forward equations* for the family of functions $t \mapsto p_t(1,x), x \in \mathbb{N}$.

   (c) Write down the conjectured value of $p_t(1,x)$ for any $t \in \mathbb{R}_+$ and $x \in \mathbb{N}$.

   (d) Verify that the conjectured functions $t \mapsto p_t(1,x), x \in \mathbb{N}$ indeed satisfy the Kolmogorov forward differential equations and the initial condition $p_0(1,x) = \mathbb{1}[x=1]$.

Solution: The Markov chain (X_t) is called the *Yule process*. It is a *pure birth process*, meaning that it is a birth-death process without death. The Yule process is a cousin of supercritical Galton-Watson branching processes, and this exercise is about the exponential growth of the population.

(a) When $X_t = x \in \mathbb{N}$, then each of these x amoebae split independently at rate 1, so the rate of jumping from x to $x+1$ is x. Thus the infinitesimal generator matrix is

$$\underline{\underline{G}} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & \dots \\ 0 & -2 & 2 & 0 & 0 & \dots \\ 0 & 0 & -3 & 3 & 0 & \dots \\ 0 & 0 & 0 & -4 & 4 & \dots \\ 0 & 0 & 0 & 0 & -5 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

(b) $\dot{p}_t(1,x) = (x-1)p_t(1,x-1) - xp_t(1,x)$ for any $x \geq 1$.
In words: the jump rate from state $x-1$ to state x is $x-1$ and the jump rate out of state x is x.

(c) We conjecture $p_t(1,x) = e^{-t}(1-e^{-t})^{x-1}$ for any $x = 1,2,3,\dots$

(d) The initial condition $p_0(1,x) = \mathbb{1}[x=1]$ is satisfied since $\lim_{t\to 0_+} e^{-t}(1-e^{-t})^{x-1} = \mathbb{1}[x=1]$.
The Kolmogorov forward equations are satisfied for $x=1$, since

$$\dot{p}_t(1,1) = \frac{\mathrm{d}}{\mathrm{d}t}e^{-t} = -e^{-t} = -p_1(1,1) \overset{(*)}{=} (1-1)p_t(1,1-1) - 1p_t(1,1),$$

where after $(*)$ the quantity $p_t(1,1-1)$ is just formally defined, it is multiplied with zero anyway.
It remains to show that the Kolmogorov forward equations are satisfied for $x \geq 2$.
First we calculate the left-hand side of the differential equation:

$$\dot{p}_t(1,x) = \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-t}(1-e^{-t})^{x-1}\right) = -e^{-t}(1-e^{-t})^{x-1} + e^{-t}(x-1)(1-e^{-t})^{x-2}e^{-t} =$$
$$\left(-e^{-t}(1-e^{-t}) + (x-1)e^{-2t}\right)(1-e^{-t})^{x-2} = \left(-e^{-t} + xe^{-2t}\right)(1-e^{-t})^{x-2} \quad (1)$$

Next we calculate the right-hand side of the differential equation:

$$(x-1)p_t(1,x-1) - xp_t(1,x) = (x-1)e^{-t}(1-e^{-t})^{x-2} - xe^{-t}(1-e^{-t})^{x-1} =$$
$$(x-1)e^{-t}(1-e^{-t})^{x-2} - xe^{-t}(1-e^{-t})(1-e^{-t})^{x-2} = \left((x-1)e^{-t} - xe^{-t}(1-e^{-t})\right)(1-e^{-t})^{x-2} =$$
$$\left(xe^{-t} - e^{-t} - xe^{-t} + xe^{-2t}\right)(1-e^{-t})^{x-2} = \left(-e^{-t} + xe^{-2t}\right)(1-e^{-t})^{x-2} \quad (2)$$

Now (1) and (2) are equal, so the conjectured functions indeed satisfy the Kolmogorov forward equations for $x \geq 2$.

Remark: So we have found that $X_t \sim \text{OPTGEO}(e^{-t})$. Note that $\mathbb{E}(X_t) = e^t$, the population grows exponentially.

Here is an alternative explanation of the fact that $X_t \sim \text{OPTGEO}(e^{-t})$. Denote by $T(n)$ the n'th jump time of the process (X_t) (with the assumption $T(0) = 0$). Denote by τ_k the time it takes for a new amoeba to be born when there are k amoebae, i.e., $\tau_k = T(k) - T(k-1)$. Now $T(n) = \tau_1 + \dots + \tau_n$, where $\tau_k \sim \text{EXP}(k)$ (since k amoebae are competing to give birth) and $\tau_1, \dots, \tau_n$ are independent. Let us argue that $T(n)$ has the same distribution as $\widehat{T}(n) = \max\{\widehat{\tau}_1, \dots, \widehat{\tau}_n\}$, where $\widehat{\tau}_1, \dots, \widehat{\tau}_n$ are i.i.d. with EXP(1) distribution. Indeed, if you think about $\widehat{\tau}_1, \dots, \widehat{\tau}_n$ as the ringing times of i.i.d. exponential clocks, then $\widehat{T}(n)$ is the time when the last clock rings. The time when the first clock rings has EXP(n) distribution (since n clocks are competing), and the remaining lifetime of the remaining $n-1$ clocks are still i.i.d. EXP(1) by the memoryless property. So the time between the first and second ring has EXP($n-1$) distribution. The time between the second and third ring has EXP($n-2$) distribution, etc. The time between ring $n-1$ and ring n has EXP(1) distribution (since there is only one clock left). This shows that $T(n) \sim \widehat{T}(n)$. Now we can compute

$$\mathbb{P}(X_t > n) = \mathbb{P}(T(n) \le t) = \mathbb{P}(\widetilde{T}(n) \le t) = \mathbb{P}(\widetilde{\tau}_1 \le t, \dots, \widetilde{\tau}_n \le t) \stackrel{(*)}{=} (1 - e^{-t})^n,$$

where in $(*)$ we used that $\widehat{\tau}_1, \dots, \widehat{\tau}_n$ are independent with EXP(1) distribution. Now if $\mathbb{P}(X_t > n) = (1 - e^{-t})^n$ for $n = 0, 1, 2, \dots$ then X_t has optimistic geometric distribution with success parameter e^{-t}.