San José State University

Math 263: Stochastic Processes

Continuous-time Markov Chains

Dr. Guangliang Chen

This lecture is based on the following textbook sections:

- Chapter 6 (Sections 6.1 - 6.5)

Outline of the presentation

- Definition of continuous-time Markov chains
- Birth and death processes
- Transition probabilities
- Limiting probabilities

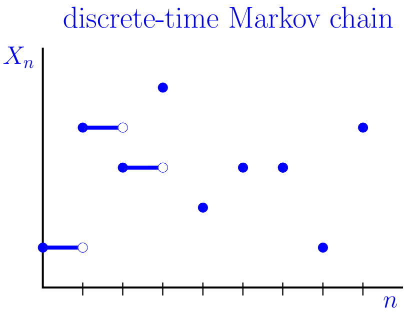
HW7

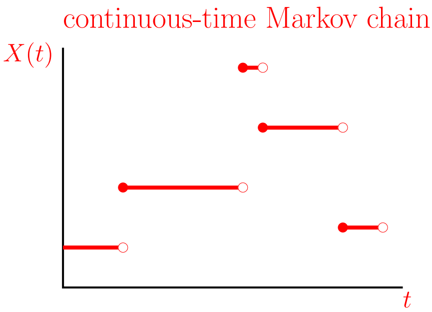
Recall that discrete-time Markov chains $\{X_n, n = 0, 1, 2, \ldots\}$ make transitions only at integer times:

$$P(X_{n+1} = j \mid X_n = i)$$

In other words, the chain can only stay in each state for an integer amount of time before making the next transition.

If we change the integer duration to continuous transition times according to an exponential distribution, then we can obtain a continuous-time Markov chain.


discrete-time Markov chain
X_n
n


continuous-time Markov chain
$X(t)$
t

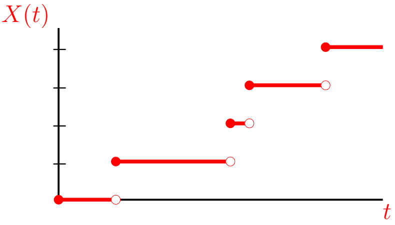
Def 0.1. Let $\{X(t), t \geq 0\}$ be a continuous-time stochastic process taking on values in the set of nonnegative integers. It is called a **continuous-time Markov chain** if each time it enters state i,

- the amount of time it spends in that state before making a transition into a different state is exponentially distributed with mean, say, $1/v_i$, i.e., $T_i \sim \text{Exp}(v_i)$
- when the process leaves state i, it next enters state j with some probability p_{ij}, which collectively satisfies

$$p_{ij} \geq 0, \quad \sum_j p_{ij} = 1$$

Example 0.1. Poisson processes are continuous time Markov chains having states $0,1,2,\ldots$ that always proceed from state i to state $i+1$, i.e., $p_{i,i+1}=1$, where $i\geq 0$. The transition rate vector is $\mathbf{v}=(v_i)$ with $v_i=\lambda$.

Such a process is known as a **pure birth process** since whenever a transition occurs, the state of the system is always increased by one.



The following is an alternative definition of continuous-time Markov chains.

Def 0.2. The process $\{X(t), t \geq 0\}$ is called a **continuous-time Markov chain** if for all $s, t \geq 0$ and for all nonnegative integers $i, j, x(u), 0 \leq u < s$,

$$P(X(t+s) = j \mid X(s) = i, \underline{X(u) = x(u), 0 \leq u < s}) = P(X(t+s) = j \mid X(s) = i)$$

That is, the probability that the chain will be in state j after time t depends only on the current state i (and is independent of the past regarding which states the chain has visited and how long the chain has been in state i).

Remark. If, in addition, $P(X(t+s)=j \mid X(s)=i)$ is independent of s, then the continuous-time Markov chain is said to have **stationary** (or homogeneous) transition probabilities. In that case, we denote

$$p_{ij}(t) = P(X(t+s)=j \mid X(s)=i)$$

and will derive its formula later.

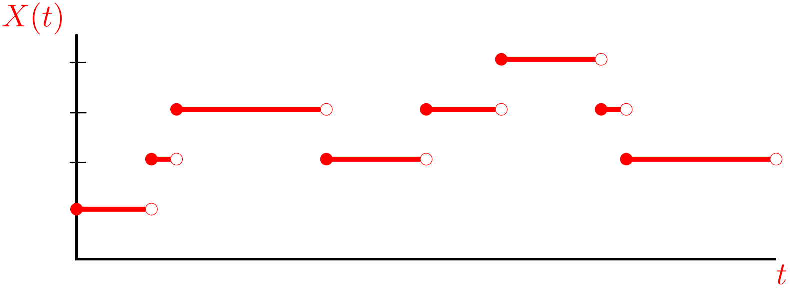
Birth and death processes

Consider a stochastic process $\{X(t), t \geq 0\}$ whose state at any time t is represented by the number of people in the system (e.g., shop, or country) at time t.

Suppose that whenever there are i people in the system, then

(i) new arrivals enter the system at an exponential rate λ_i, and

(ii) people leave the system at an exponential rate μ_i.



Def 0.3. The above continuous-time stochastic process $\{X(t), t \geq 0\}$ having states $0, 1, 2, \ldots$ is called a **birth and death process**, with arrival (or birth) rate $\{\lambda_i\}_{i=0}^{\infty}$ and departure (or death) rates $\{\mu_i\}_{i=0}^{\infty}$ (with $\mu_0 = 0$).

Theorem 0.1. For any birth and death process, the transition rates are

$$v_0 = \lambda_0, \quad v_i = \lambda_i + \mu_i, \ i > 0$$

and the transition probabilities are

$$p_{0,1} = 1, \quad p_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad p_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}, \ i > 0$$

Proof. Suppose the process is in state i at present. Let the waiting time for the next arrival be $T_i^{(a)} \sim \text{Exp}(\lambda_i)$ and the waiting time for the next departure $T_i^{(d)} \sim \text{Exp}(\mu_i)$. Then the waiting time for the next transition is

$$T_i = \min\left(T_i^{(a)}, T_i^{(d)}\right) \sim \text{Exp}(v_i), \quad v_i = \lambda_i + \mu_i.$$

The transition probabilities are

$$p_{i,i+1} = P\left(T_i^{(a)} < T_i^{(d)}\right) = \frac{\lambda_i}{\lambda_i + \mu_i}$$
$$p_{i,i-1} = 1 - p_{i,i+1} = \frac{\mu_i}{\lambda_i + \mu_i}, \quad i \geq 1.$$

In particular, $p_{01} = \frac{\lambda_0}{\lambda_0 + 0} = 1.$ □

Example 0.2. The Poisson process is a birth and death process with $\lambda_i = \lambda$ and $\mu_i = 0$

Example 0.3 (A birth process with linear birth rate, called **Yule process**).

$$\lambda_i = i\lambda, \quad \mu_i = 0$$

Example 0.4 (A linear growth model with immigration).

$$\lambda_i = i\lambda + \theta, \quad \mu_i = i\mu$$

Theorem 0.2. Let $X(t)$ represent the size of the population at time t in the linear growth model with immigration, and $M(t) = \mathrm{E}(X(t))$. If $X(0) = i$, then

$$M(t) = \begin{cases} \left(\frac{\theta}{\lambda-\mu} + i\right) e^{(\lambda-\mu)t} - \frac{\theta}{\lambda-\mu}, & \lambda \neq \mu; \\ \theta t + i, & \lambda = \mu. \end{cases}$$

Proof. (Not rigorous; see a rigorous proof in the textbook)

$$M'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(X(t)) = \mathrm{E}(X'(t)) = \mathrm{E}(\lambda X(t) - \mu X(t) + \theta) = (\lambda - \mu)M(t) + \theta.$$

There is also an initial condition $M(0) = \mathrm{E}(X(0)) = i$.

The differential equation (plus the initial condition) is of the following form:

$$x'(t) = a \cdot x(t) + b, \quad x(0) = x_0$$

The solution is

$$x(t) = \begin{cases} x_0 e^{at} + \frac{b}{a}\left(e^{at} - 1\right), & a \neq 0 \\ bt + x_0, & a = 0 \end{cases}$$

Applying the above formula directly would yield the desired result. □

Remark. If $\theta = 0$ (no immigration), then the formula for $M(t)$ reduces to

$$M(t) = i e^{(\lambda-\mu)t}, \quad t \geq 0.$$

Therefore, the effect of immigration on the population growth is

- when $\lambda \neq \mu$:

$$\frac{\theta}{\lambda-\mu}\left(e^{(\lambda-\mu)t} - 1\right).$$

- when $\lambda = \mu$:

$$\theta t$$

Example 0.5 (The Queuing System M/M/1). Suppose that customers arrive at a single-server service station in accordance with a Poisson process having rate λ.

Upon arrival, each customer goes directly into service if the server is free; if not, then the customer joins the queue (that is to wait in line).

When the server finishes serving a customer, the customer leaves the system and the next customer in line, if any, enters service.

The successive service times are assumed to be independent exponential random variables having mean $1/\mu$.

The preceding example is known as the M/M/1 queuing system:

- The first M refers to the fact that the interarrival process is Markovian (since it is a Poisson process);
- The second M to the fact that the service distribution is exponential (and, hence, Markovian);
- The 1 refers to the fact that there is a single server.

If we let $X(t)$ denote the number in the system (queue + service station) at time t, then $\{X(t), t \geq 0\}$ is a birth and death process with

$$\lambda_i = \lambda,\ i \geq 0, \qquad \text{and} \quad \mu_i = \mu,\ i \geq 1$$

Example 0.6 (The Queuing System M/M/s). Consider an exponential queuing system in which there are s servers available, each serving at rate μ. An entering customer first waits in line and then goes to the first free server. Assuming customers arrive according to a Poisson process with rate λ, this is a birth and death process with parameters:

- arrival rates: $\lambda_i = \lambda$ for each $i \ge 0$, and
- departure rates: $\mu_i = \min(i, s) \cdot \mu$ for each $i \ge 1$

Consider now a general birth and death process with birth rates $\{\lambda_i\}$ and death rates $\{\mu_i\}$, where $\mu_0 = 0$.

Let U_i denote the time, starting from state i, it takes for the process to enter state $i+1$, for any $i \geq 0$. Then we have the following result.

Theorem 0.3.

$$\mathrm{E}(U_0) = \frac{1}{\lambda_0}, \quad \mathrm{E}(U_i) = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i} \cdot \mathrm{E}(U_{i-1}),\ i \geq 1$$

Proof. We will recursively compute $\mathrm{E}(U_i), i \ge 0$, by starting with $i = 0$.

Since $U_0 = T_0$ is exponential with rate λ_0, we have $\mathrm{E}(U_0) = 1/\lambda_0$.

For $i \ge 1$, we condition on the first transition which takes the process into state $i-1$ or $i+1$:

$$\mathrm{E}(U_i) = \frac{1}{\lambda_i + \mu_i} + \frac{\lambda_i}{\lambda_i + \mu_i} \cdot 0 + \frac{\mu_i}{\lambda_i + \mu_i} \cdot [\mathrm{E}(U_{i-1}) + \mathrm{E}(U_i)]$$

Thus,

$$\mathrm{E}(U_i) = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i} \cdot \mathrm{E}(U_{i-1}), \;\; i \ge 1.$$

Remark. A recursive formula for the variance of U_i is the following:

$$\operatorname{Var}(U_0) = \frac{1}{\lambda_0^2}$$

$$\operatorname{Var}(U_i) = \frac{1}{\lambda_i(\lambda_i+\mu_i)} + \frac{\mu_i}{\lambda_i}\operatorname{Var}(U_{i-1}) + \frac{\mu_i}{\lambda_i+\mu_i}\left[\mathrm{E}(U_{i-1})+\mathrm{E}(U_i)\right]^2, \ \ i \geq 1$$

See the textbook for its derivation.

Corollary 0.4. If $\lambda_i = \lambda$ and $\mu_i = \mu$ for all i, then

$$\mathrm{E}(U_i) = \begin{cases} \frac{1-(\mu/\lambda)^{i+1}}{\lambda-\mu}, & \lambda \neq \mu \\ \frac{i+1}{\lambda}, & \lambda = \mu \end{cases}$$

Proof. This is a direct application of the following formula: If

$$a_{i+1} = c + d \cdot a_i, \quad i = 0, 1, 2, \ldots$$

then

$$a_i = \begin{cases} ci + a_0, & d = 1 \\ c\frac{1-d^i}{1-d} + d^i a_0, & d \neq 1 \end{cases}$$

Remark. The expected time for the process to transition from i to $j > i$ is

$$\mathrm{E}(U_i) + \cdots + \mathrm{E}(U_{j-1})$$

and the variance of the overall transition time from i to j is

$$\mathrm{Var}(U_i) + \cdots + \mathrm{Var}(U_{j-1})$$

The transition probability function $p_{ij}(t)$

Consider a continuous-time, homogeneous Markov chain. Let

$$p_{ij}(t) = P(X(t+s) = j \mid X(s) = i) = P(X(t) = j \mid X(0) = i)$$

We consider two different scenarios for the Markov chain and find formulas for $p_{ij}(t)$ separately:

- **Pure birth process with distinct birth rates**
- **General continuous-time Markov chains**

Pure birth process with distinct birth rates

We have the following explicit formula for the transition probability function in the case of a pure birth process ($\mu_i = 0, v_i = \lambda_i$) having distinct birth rates ($\lambda_i \neq \lambda_j$).

Theorem 0.5. For a pure birth process having distinct rates,

$$p_{ii}(t) = P(T_i > t) = e^{-\lambda_i t}$$

$$p_{ij}(t) = \sum_{k=i}^{j} C_{k,i,j} e^{-\lambda_k t} - \sum_{k=i}^{j-1} C_{k,i,j-1} e^{-\lambda_k t}, \quad i < j$$

where

$$C_{k,i,j} = \prod_{i \le \ell \le j, \ell \neq k} \frac{\lambda_\ell}{\lambda_\ell - \lambda_k}$$

Remark. Two special cases:

- $j = i+1$:

$$p_{ij}(t) = \frac{\lambda_{i+1}}{\lambda_{i+1}-\lambda_i} e^{-\lambda_i t} + \frac{\lambda_i}{\lambda_i - \lambda_{i+1}} e^{-\lambda_{i+1} t} - e^{-\lambda_i t}$$
$$= \frac{\lambda_i}{\lambda_{i+1}-\lambda_i} (e^{-\lambda_i t} - e^{-\lambda_{i+1} t})$$

- $\lambda_k = k\lambda$ for all $k \geq 1$ (Yule process): $\lambda_i \neq \lambda_j$ if $i \neq j$.

  Suppose $X_0 = 1$. Then it can be shown that (textbook Example 6.8)

  $$p_{1j}(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{j-1}$$

  implying that

  $$X(t) \mid X(0) = 1 \ \sim \ \text{Geom}(p = e^{-\lambda t}).$$

  That is, starting with a single individual, the population size at time t has a geometric distribution with mean $e^{\lambda t}$. If initially there are i individuals, then the population size at time t has a negative binomial distribution NB$(n = i, p = e^{-\lambda t})$.

Proof of the theorem. First, we write

$$p_{ij}(t) = P(X(t) < j+1 \mid X(0) = i) - P(X(t) < j \mid X(0) = i).$$

Next, letting T_k represent the duration of the chain in state k, we have

$$P(X(t) < j \mid X(0) = i) = P(T_i + \cdots + T_{j-1} > t)$$

and similarly,

$$P(X(t) < j+1 \mid X(0) = i) = P(T_i + \cdots + T_j > t)$$

It remains to determine the distribution of a sum of independent exponential random variables with distinct rates, by using the result on next page.

Proposition 0.6. If $X_i \sim \text{Exp}(\lambda_i), i = 1,2$ are independent random variables having distinct rates ($\lambda_1 \neq \lambda_2$), then

$$f_{X_1+X_2}(t) = \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_2 - \lambda_1} \lambda_1 e^{-\lambda_1 t}, \quad t > 0$$

More generally, for n such random variables (in that case, the sum is called a hyperexponential random variable),

$$f_{X_1+\cdots+X_n}(t) = \sum_{k=1}^{n} C_{k,1,n} \lambda_k e^{-\lambda_k t}, \quad C_{k,1,n} = \prod_{1 \le \ell \le n, \ell \neq k} \frac{\lambda_\ell}{\lambda_\ell - \lambda_k}$$

Proof. We prove only the special case of $n=2$ (the proof of the general case can be found in Section 5.2.4):

$$\begin{aligned}
f_{X_1+X_2}(t) &= \int_0^t f_{X_1}(s) f_{X_2}(t-s)\,\mathrm{d}s \\
&= \int_0^t \lambda_1 e^{-\lambda_1 s} \cdot \lambda_2 e^{-\lambda_2 (t-s)}\,\mathrm{d}s \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 t} \int_0^t e^{-(\lambda_1-\lambda_2)s}\,\mathrm{d}s \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 t} \frac{1}{\lambda_1-\lambda_2}(1-e^{-(\lambda_1-\lambda_2)t}) \\
&= \frac{\lambda_1}{\lambda_1-\lambda_2} \lambda_2 e^{-\lambda_2 t} + \frac{\lambda_2}{\lambda_2-\lambda_1} \lambda_1 e^{-\lambda_1 t}
\end{aligned}$$

Remark. In the case of $n = 2$ with equal rate $\lambda_1 = \lambda_2 = \lambda$,

$$f_{X_1+X_2}(t) = \lambda^2 t e^{-\lambda t}, t > 0$$

This is the Gamma($\alpha = 2, \lambda$) density.

Remark. The survival function of the hyperexponential random variable $S = X_1 + \cdots + X_n$ is

$$P(S > t) = \sum_{k=1}^{n} C_{k,1,n} e^{-\lambda_k t}$$

General continuous-time Markov chains

First, we also have the so-called Chapman–Kolmogorov equations.

Theorem 0.7 (Chapman–Kolmogorov equations). For all $s, t \geq 0$,

$$p_{ij}(t+s) = \sum_k p_{ik}(t) p_{kj}(s)$$

Proof.

$$\begin{aligned}
p_{ij}(t+s) &= P(X(t+s)=j \mid X(0)=i) \\
&= \sum_k P(X(t+s)=j, X(t)=k \mid X(0)=i) \\
&= \sum_k P(X(t+s)=j \mid X(t)=k, X(0)=i) P(X(t)=k \mid X(0)=i) \\
&= \sum_k p_{kj}(s) p_{ik}(t).
\end{aligned}$$

From the Chapman–Kolmogorov equations we can obtain the following differential equations for all $p_{ij}(t)$.

Theorem 0.8 (Kolmogorov's Backward Equations). In any continuous-time Markov chain,

$$p'_{ij}(t) = \sum_{k \neq i} q_{ik} p_{kj}(t) - v_i p_{ij}(t)$$

where

$$q_{ik} = v_i p_{ik}$$

are called the instantaneous transition rates (from state i to state k).

Remark. We have

$$\sum_k q_{ik} = v_i \sum_k p_{ik} = v_i, \qquad \text{and} \quad p_{ik} = \frac{q_{ik}}{v_i} = \frac{q_{ik}}{\sum_k q_{ik}}.$$

Remark. In a birth and death process, the instantaneous transition rates are just birth and death rates:

$$q_{i,i+1} = v_i p_{i,i+1} = (\lambda_i + \mu_i) \cdot \frac{\lambda_i}{\lambda_i + \mu_i} = \lambda_i$$
$$q_{i,i-1} = v_i p_{i,i-1} = (\lambda_i + \mu_i) \cdot \frac{\mu_i}{\lambda_i + \mu_i} = \mu_i$$

Proof. For any $h>0$,

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h} = \frac{1}{h}\left(\sum_k p_{ik}(h)p_{kj}(t) - p_{ij}(t)\right)$$
$$= \sum_{k\neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) - \frac{1-p_{ii}(h)}{h} p_{ij}(t)$$

It remains to show that

$$\lim_{h\to 0} \frac{p_{ik}(h)}{h} = q_{ik}, \quad k\neq i$$
$$\lim_{h\to 0} \frac{1-p_{ii}(h)}{h} = v_i.$$

First, since

$$p_{ii}(h) = P(X(h) = i \mid X(0) = i) = P(T_i > h) = e^{-h v_i},$$

we have

$$\lim_{h\to 0} \frac{1 - p_{ii}(h)}{h} = \lim_{h\to 0} \frac{1 - e^{-h v_i}}{h} = v_i.$$

Next, due to

$$p_{ik}(h) = P(X(h) = k \mid X(0) = i) = P(T_i < h) p_{ik} = (1 - e^{-h v_i}) p_{ik},$$

we can correspondingly obtain that

$$\lim_{h\to 0} \frac{p_{ik}(h)}{h} = \lim_{h\to 0} \frac{(1 - e^{-h v_i}) p_{ik}}{h} = v_i p_{ik}.$$

Example 0.7 (A continuous-time Markov chain consisting of two states). Consider a machine that works for an exponential amount of time having mean $1/\lambda$ before breaking down; and suppose that it takes an exponential amount of time having mean $1/\mu$ to repair the machine. If the machine is in working condition at time 0, then what is the probability that it will be working at time $t = 10$?

Solution. We note that the process is a birth and death process (with state 0 meaning that the machine is working and state 1 that it is being repaired) having parameters

$$\lambda_0 = \lambda,\ \lambda_1 = 0,\ \mu_1 = \mu,\ \mu_0 = 0, \qquad \text{and} \quad p_{01} = 1 = p_{10}$$

We can write down the Chapman–Kolmogorov backward equations:

$$p'_{00}(t) = \lambda(p_{10}(t) - p_{00}(t))$$
$$p'_{10}(t) = \mu(p_{00}(t) - p_{10}(t))$$

or in matrix form

$$\begin{pmatrix} p_{00}(t) \\ p_{10}(t) \end{pmatrix}' = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \begin{pmatrix} p_{00}(t) \\ p_{10}(t) \end{pmatrix}$$

along with the intial conditions:

$$p_{00}(0) = 1, \; p_{10}(0) = 0$$

The matrix has the following decomposition

$$\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 1 & -\mu \end{pmatrix} \begin{pmatrix} 0 & \\ & -\lambda-\mu \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 1 & -\mu \end{pmatrix}^{-1}$$

Therefore, the solution is given by

$$\begin{pmatrix} p_{00}(t) \\ p_{10}(t) \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 1 & -\mu \end{pmatrix} \begin{pmatrix} 1 & \\ & e^{-(\lambda+\mu)t} \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 1 & -\mu \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} \\ \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t} \end{pmatrix}$$

Similarly, we can obtain the Kolmogorov's forward equations.

Theorem 0.9 (Kolmogorov's Forward Equations). Under suitable regularity conditions,

$$p'_{ij}(t) = \sum_{k \neq j} q_{kj} p_{ik}(t) - v_j p_{ij}(t)$$

Remark. For a pure birth process, Kolmogorov's forward equations become

$$p'_{ij}(t) = \lambda_{j-1} p_{i,j-1}(t) - \lambda_j p_{ij}(t)$$

which then yield that

$$\begin{aligned} p_{ij}(t) &= 0, \quad j < i \\ p_{ii}(t) &= e^{-\lambda_i t}, \\ p_{ij}(t) &= \lambda_{j-1} e^{-\lambda_j t} \int_0^t e^{\lambda_j s} p_{i,j-1}(s)\,\mathrm{d}s, \quad j \ge i+1 \end{aligned}$$

Limiting probabilities

For each state j of a continuous-time Markov-chain, let

$$P_j = \lim_{t\to\infty} p_{ij}(t)$$

The limit exists and is independent of the initial state i if all states communicate and the chain is positive recurrent.

The P_j are called stationary probabilities.

Additionally, P_j also have the interpretation of being the long-run proportion of time that the process is in state j.

We have the following result.

Theorem 0.10.

$$v_j P_j = \sum_{k \neq j} q_{kj} P_k \qquad \left(\sum P_j = 1\right)$$

Proof. Let $t \to \infty$ in the forward equations and use $\lim_{t\to\infty} p'_{ij}(t) = 0$ □

Interpretation:

- LHS: rate at which the process leaves state j;
- RHS: rate at which the process enters state j.

Example 0.8 (A continuous-time Markov chain consisting of two states, continued)**.** Find the proportion of time when the machine is in working condition.

Example 0.9. For a birth and death process, a sufficient and necessary condition for the limiting probabilities to exist is

$$\sum_{i=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{i-1}}{\mu_1 \mu_2 \ldots \mu_i} < \infty$$

In this case, it can be shown that

$$P_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{i-1}}{\mu_1 \mu_2 \ldots \mu_i}}, \qquad P_i = \frac{\lambda_0 \lambda_1 \cdots \lambda_{i-1}}{\mu_1 \mu_2 \ldots \mu_i} P_0, \quad n \geq 1$$

In the M/M/1 queue $(\lambda_i = \lambda, \mu_i = \mu)$,

$$P_i = \frac{(\lambda/\mu)^i}{1 + \sum_{i=1}^{\infty} (\lambda/\mu)^i} = (\lambda/\mu)^i (1 - \lambda/\mu)$$

provided that $\lambda/\mu < 1$.

Proof. For the given birth and death process,

$$j=0: \qquad \lambda_0 P_0 = \mu_1 P_1$$

$$j\ge 1: \qquad (\lambda_j+\mu_j)P_j = \lambda_{j-1}P_{j-1} + \mu_{j+1}P_{j+1}$$

and further that (by induction)

$$\lambda_j P_j = \mu_{j+1}P_{j+1}, \quad j \ge 0$$

or equivalently,

$$\frac{P_{j+1}}{P_j} = \frac{\lambda_j}{\mu_{j+1}}, \quad j \ge 0$$

Multiplying such equations from $j=0$ to $j=i-1$ gives that

$$P_i = \frac{\lambda_0\lambda_1\cdots\lambda_{i-1}}{\mu_1\mu_2\ldots\mu_i}P_0, \quad i \ge 1.$$

Using $P_0 + \sum_{i=1}^{\infty} P_i = 1$, we can find

$$P_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{i-1}}{\mu_1 \mu_2 \ldots \mu_i}}$$

assuming

$$\sum_{i=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{i-1}}{\mu_1 \mu_2 \ldots \mu_i} < \infty.$$