Transition probabilities $P_{ij}(t)$

Recall: We characterize a CTMC by defining P_{ij} = prob. next jump (when it occurs) is to j, given current state is i, and the time T_i spent at state i is $\mathcal{E}xp(\nu_i)$.

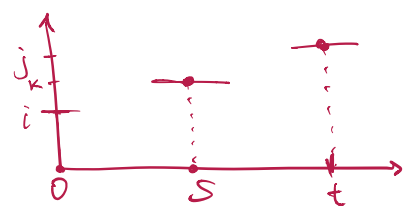
- Our general goal is to study

$$P_{ij}(t) = P(X(s+t)=j \mid X(s)=i) \underset{\text{stationarity}}{=} P(X(t)=j \mid X(0)=i).$$

$P_{ij}(t)$ is the continuous-time analogue of $P_{ij}^{(n)}$ that we studied in discrete time. Like in discrete time, $P_{ij}(t)$ satisfies Chapman-Kolmogorov equations:

Proposition (Chapman-Kolmogorov equations): $\forall s,t \geq 0$,

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



Proof: $\forall s,t \geq 0$

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

- How can we compute $P_{ij}(t)$?

→ In 6.4, see explicit computation of $P_{ij}(t)$ for pure birth processes, e.g., the Yule process: $P_{ij}(t) = \binom{j-1}{i-1} e^{-i\lambda t}(1-e^{-\lambda t})^{j-i}$ for $j \geq i$.

→ More generally, $P_{ij}(t)$ obeys differential equations (Kolmogorov backward and forward equations). To derive these, we will need to describe $P_{ij}(t+h)$ as $h \to 0$.

<u>Recall</u>: For $\{X(t), t \geq 0\}$ a general CTMC (not necessarily a birth-death process), we characterized $X(t)$ by

- ν_i : rate of transitions from state i, $T_i \sim Exp(\nu_i)$.
- p_{ij} : probability of transition from state i to state $j \neq i$. $\sum_{j \neq i} p_{ij} = 1$.

Let $q_{ij} = \nu_i p_{ij}$

<u>Claim</u>: If we know q_{ij}, we also know ν_i and p_{ij} $\forall j \neq i$.

<u>Proof</u>: $\nu_i = \nu_i \cdot \sum_{j \neq i} p_{ij} = \sum_{j \neq i} \nu_i p_{ij} = \sum_{j \neq i} q_{ij}$

Then $p_{ij} = \frac{q_{ij}}{\nu_i} = \frac{q_{ij}}{\sum_{k \neq i} q_{ik}}$.

We will assume for the rest of this chapter that the ν_i's are "small enough" (and this is always the case when the state space is finite!) so that the probability an interval of time contains ∞-many transitions is 0. Otherwise, the process is explosive, and may not be defined for all $t \geq 0$.

Lemma: As $h \to 0$, $P_{ii}(h) = 1 - \nu_i h + o(h)$, i.e.

$$\lim_{h\to 0} \frac{P_{ii}(h) - \overset{=P_{ii}(0)}{1}}{h} = -\nu_i \quad \text{meaning} \quad \frac{d}{dh} P_{ii}(h)\Big|_{h=0} = -\nu_i$$

And for $i \neq j$, $P_{ij}(h) = q_{ij} h + o(h)$, i.e.,

$$\lim_{h\to 0} \frac{P_{ij}(h) - \overset{=0}{P_{ij}(0)}}{h} = q_{ij} \quad \text{meaning} \quad \frac{d}{dh} P_{ij}(h)\Big|_{h=0} = q_{ij}.$$

Proof: • $P(\geq 2 \text{ events on } [0,h] \mid P(0)=i)$

$$= \sum_{j\neq i} p_{ij} P(T_i < h, T_i + T_j \leq h)$$

$$= \sum_{j\neq i} p_{ij} \int_0^h p_{T_i}(t) P(T_j + T_i \leq h \mid T_i = t)\, dt$$

$$= \sum_{j\neq i} p_{ij} \int_0^h \nu_i e^{-t\nu_i} P(T_j \leq h - t)\, dt$$

$$= \sum_{j\neq i} p_{ij} \int_0^h \nu_i e^{-t\nu_i} \left(1 - e^{-(h-t)\nu_j}\right) dt$$

$$= \sum_{j\neq i} p_{ij} \left(\int_0^h \nu_i e^{-t\nu_i} dt - \int_0^h \nu_i e^{-t\nu_i} e^{-(h-t)\nu_j} dt \right)$$

$$= \sum_{j\neq i} p_{ij} \left(1 - e^{-h\nu_i} - \nu_i \int_0^h (1 - t\nu_i + o(h))(1 - (h-t)\nu_j + o(h))\, dt \right)$$

$$= \sum_{j\neq i} p_{ij} \left(\underbrace{1 - e^{-h\nu_i}}_{h\nu_i + o(h)} - \nu_i \int_0^h (1 - t\nu_i - (h-t)\nu_j + o(h))\, dt \right)$$

$$= \sum_{j\neq i} p_{ij} \left(h\nu_i + o(h) - h\nu_i + o(h) \right) = \sum_{j\neq i} p_{ij}\, o(h) = o(h).$$

- $P(1 \text{ jump to } j \text{ in } [0,h] \mid X(0)=i) =$

$= p_{ij} \; P(T_i < h, \; T_i + T_j > h) =$ same exact proof as in Monday's lecture ...

$= p_{ij} \, \nu_i h + o(h).$

- $P(\text{no jumps in } [0,h] \mid X(0)=i) = 1 - \sum_{j \neq i} p_{ij} \nu_i h + o(h)$

$= 1 - \nu_i h + o(h)$

Alternatively,

$P(\text{no jumps in } [0,h] \mid X(0)=i) = P(T_i > h) = e^{-\nu_i h}$

$= 1 - \nu_i h + o(h).$

Therefore, $\rightarrow P_{ii}(h) = o(h) + P(\text{no jumps in } [0,h] \mid X(0)=i)$

$= 1 - \nu_i h + o(h)$

$\rightarrow P_{ij}(h) = o(h) + P(1 \text{ jump to } j \text{ in } [0,h] \mid X(0)=i)$

$= \underbrace{p_{ij} \nu_i}_{q_{ij}} h + o(h) = q_{ij} h + o(h).$

This lemma leads to the <u>Kolmogorov-Backward equations</u>:

<u>Theorem</u>: $P'_{ij}(t) = \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t)$

we look "backward" from $k \rightarrow j$

<u>Proof</u>:

From the Chapman-Kolmogorov equations, we have that

$P_{ij}(t+h) - P_{ij}(t) = \sum_k P_{ik}(h) P_{kj}(t) - P_{ij}(t)$

$$= \sum_{k \neq i} \underbrace{p_{ik}(h)}_{q_{ik}h + o(h)} P_{kj}(t) + p_{ii}(h) P_{ij}(t) - P_{ij}(t)$$

$$= \sum_{k \neq i} (q_{ik}h + o(h)) P_{kj}(t) - (\underbrace{1 - p_{ii}(h)}_{\nu_i h + o(h)}) P_{ij}(t)$$

$$\Rightarrow \lim_{h \to 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h} = \lim_{h \to 0} \frac{\sum_{k \neq i} (q_{ik}h + o(h)) P_{kj}(t) - \nu_i h P_{ij}(t) + o(h)}{h}$$

$$= \lim_{h \to 0} \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t) + \frac{o(h)}{h}$$

$$= \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t).$$

<u>Remark</u>: Here we assume that

$\lim_{h \to 0} \sum_{k \neq i} \frac{p_{ik}(h) P_{kj}(t)}{h} < \infty$. More precisely, we assume a stronger condition: $\lim_{h \to 0} \sum_{k \neq i} \frac{p_{ik}(h)}{h} < \infty$.

$$\sum_{k \neq i} \frac{p_{ik}(h)}{h} = \sum_{k \neq i} \frac{q_{ik}h + o(h)}{h} = \sum_{k \neq i} q_{ik} + \underbrace{\sum_{k \neq i} \frac{o(h)}{h}}_{\substack{\text{we assume} \\ \text{this is } < \infty \\ \text{as } h \to 0.}}$$

Such a process is called conservative and this is always the case when the state space is finite.

<u>Example</u>: Birth and death processes:

$p_{ij} = 0$, if $j \neq i \pm 1$,

$p_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}$ $\quad p_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}$ for $i \geq 1$. $\quad \mu_0 = 0$.

$\nu_i = \lambda_i + \mu_i$

$$q_{ij} = \nu_i \, p_{ij} = \begin{cases} \lambda_i, & \text{if } j = i+1 \\ \mu_i, & \text{if } j = i-1 \\ 0, & \text{if } j \neq i \pm 1. \end{cases}$$

So, the backward equation is:

if $i \neq 0$, $P'_{i,j}(t) = q_{i,i+1} P_{i+1,j}(t) + q_{i,i-1} P_{i-1,j}(t) - (\lambda_i + \mu_i) P_{i,j}(t)$

$$= \lambda_i P_{i+1,j}(t) + \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i) P_{ij}(t).$$

if $i = 0$, $P'_{0j}(t) = \lambda_0 P_{1,j}(t) - \lambda_0 P_{0j}(t)$.