9. Continuous-time Markov chain (CTMC)

We now discuss an important class of stochastic process, called continuous-time Markov chain (CTMC).

Definition 9.1. Let $X = \{X_t : t \ge 0\}$ be a stochastic process with countable state space S.

i) We say that the process is a continuous-time Markov chain (CTMC) if
$$\mathbb{P}\Big(X_{t_n} = j \Big| X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}\Big) = \mathbb{P}\Big(X_{t_n} = j \Big| X_{t_{n-1}} = i_{n-1}\Big) \tag{9.1}$$
for all $j, i_1, i_2, \dots, i_{n-1} \in S$ and for all $0 \le t_1 < t_2 < \dots < t_n$.

ii) We say that the CTMC is homogeneous if and only if the probabilities
$$\mathbb{P}\big(X_{t+s} = j \big| X_s = i\big)$$
is independent of s for all t.

iii) The probability
$$p_{ij}(t) = \mathbb{P}\big(X_{t+s} = j \big| X_s = i\big) \quad s, t \ge 0$$
is called transition probability for the CTMC. Denote $P(t) = (p_{ij}(t))$ for all $i, j \in S$. We say that $P(t)$ is a transition probability matrix.

Let $\{X_t : t \ge 0\}$ be a homogeneous CTMC. Then the following properties hold:

a) $p_{ij}(t) = \mathbb{P}\big(X_t = j \big| X_0 = i\big)$ and hence $0 \le p_{ij}(t) \le 1$ for any $t \ge 0$ and $i, j \in S$.

b) $p_{ij}(0) = \mathbb{P}\big(X_0 = j \big| X_0 = i\big) = \delta_{ij}$.

c) $\displaystyle\sum_{j \in S} p_{ij}(t) = \sum_{j \in S} \mathbb{P}\big(X_t = j \big| X_0 = i\big) = 1.$

Example 9.1. *At time $t = 0$, exactly N systems start operating. Their lifetimes are independent, identically distributed exponential random variables with parameter λ . If $X(t)$ denotes the number of systems still operating at time t, then $\{X(t) : t \ge 0\}$ is a CTMC with state space $S = \{0, 1, 2, \dots, N\}$.*

9.1. **Kolmogorov forward and backward equations:** Observe that
$$\begin{aligned}
p_{ij}(t+T) &= \sum_k \mathbb{P}\Big(X_{t+T} = j, X_t = k \Big| X_0 = i\Big) \\
&= \sum_k \frac{\mathbb{P}\Big(X_{t+T} = j, X_0 = i, X_t = k\Big)}{\mathbb{P}(X_0 = i)} \\
&= \sum_k \frac{\mathbb{P}\big(X_t = k, X_0 = i\big)}{\mathbb{P}(X_0 = i)} \cdot \frac{\mathbb{P}\Big(X_{t+T} = j, X_0 = i, X_t = k\Big)}{\mathbb{P}\big(X_t = k, X_0 = i\big)} \\
&= \sum_k \mathbb{P}\Big(X_t = k \Big| X_0 = i\Big) \cdot \mathbb{P}\Big(X_{t+T} = j \Big| X_0 = i, X_t = k\Big) \\
&= \sum_k p_{ik}(t) \mathbb{P}\Big(X_{t+T} = j \Big| X_0 = i, X_t = k\Big).
\end{aligned}$$

Since the process $\{X_t : t \ge 0\}$ is time homogeneous Markov, we have
$$\mathbb{P}\Big(X_{t+T} = j \Big| X_0 = i, X_t = k\Big) = \mathbb{P}\Big(X_{t+T} = j \Big| X_t = k\Big) = p_{kj}(T).$$

Thus, we get
$$p_{ij}(t+T) = \sum_k p_{ik}(t) p_{kj}(T)$$

which holds for all states i, j and $t, T \geq 0$. It is called **Chapman-Kolmogorov equation**. In matrix notation, it can be written as

$$P(t+T) = P(t)P(T).$$

Remark 9.1. Let $p_j(t) = \mathbb{P}(X_t = j)$. Then $\sum_{j \in S} p_j(t) = 1$ for each $t \geq 0$. Indeed,

$$p_j(t) = \sum_i \mathbb{P}\big(X_t = j, X_0 = i\big) = \sum_{i \in S} \mathbb{P}(X_0 = i)\mathbb{P}\big(X_t = j \big| X_0 = i\big) = \sum_i p_{ij}(t)\mathbb{P}(X_0 = i).$$

Since $\sum_j p_{ij}(t) = 1$, we get that $\sum_j p_j(t) = 1$.

Transition density matrix. Assume that $q_{ij} = \frac{d}{dt} p_{ij}(t)\Big|_{t=0}$ exists, i.e.,

$$q_{ij} = \lim_{h \to 0} \frac{p_{ij}(h) - p_{ij}(0)}{h}.$$

Then, it is easily seen that

$$p_{ij}(h) = \begin{cases} 1 + hq_{ii} + o(h) & i = j \\ hq_{ij} + o(h) & i \neq j \end{cases}$$

where $f(h) = o(h)$ as $h \to 0$ if $\frac{f(h)}{h} \to 0$ as $h \to 0$. Let $Q = (q_{ij})$. Then one can easily check that

$$q_{ii} \leq 0, \quad q_{ij} \geq 0 \text{ for } i \neq j, \quad \sum_j q_{ij} = 0.$$

The matrix Q is known as *transition density matrix* or *rate matrix* of the process.

Example 9.2. *Let $N(t)$ be a Poisson process with intensity λ. Then it is a time-continuous Markov chain. Observe that, for small Δt,*

$$\begin{aligned} &p_{i,i+1}(\Delta t) = \mathbb{P}(N(\Delta t) = 1) = \lambda \Delta t + o(\Delta t), \\ &p_{i,j}(\Delta t) = o(\Delta t) \quad \textit{for } j \neq i, i+1\,, \\ &p_{ii}(\Delta t) = \mathbb{P}\big(N(t+\Delta t) = i \big| N(t) = i\big) = \mathbb{P}(N(\Delta t) = 0) = 1 - \lambda \Delta t + o(\Delta t). \end{aligned}$$

So, Q, the transition density matrix, is given by $Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \dots \\ 0 & -\lambda & \lambda & 0 & 0 & \dots \\ 0 & 0 & -\lambda & \lambda & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}.$

Kolmogorov forward and backward equations. For $h > 0$ and $t \geq 0$, we have from Chapman-Kolmogorov equation

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \frac{1}{h}\Big(\sum_{k \in S} p_{ik}(h)p_{kj}(t) - p_{ij}(t)\Big) = \frac{1}{h}\big(p_{ii}(h) - 1\big)p_{ij}(t) + \frac{1}{h}\sum_{k \neq i} p_{ik}(h)p_{kj}(t)$$

$$\implies p'_{ij}(t) = \lim_{h \to 0} \frac{p_{ij}(t+h) - p_{ij}(h)}{h} = q_{ii}p_{ij}(t) + \sum_{k \neq i} q_{ik}p_{kj}(t).$$

Hence, in matrix notation

$$P'(t) = QP(t).$$

This is known as *Kolmogorov backward equation.* Similarly, by using the fact that

$$p_{ij}(t+h) = \sum_{k \in S} p_{ik}(t)p_{kj}(h),$$

j

we have $p'_{ij}(t) = q_{jj}p_{ij}(t) + \sum_{k\neq j} q_{kj}p_{ik}(t)$. In matrix notation,

$$P'(t) = P(t)Q, \quad P(0) = I.$$

This is called *Kolmogorov forward equation.* Formally the solution of the above equation can be cast in the form

$$P(t) = I + \sum_{n=1}^{\infty} \frac{t^n Q^n}{n!}.$$

If Q is a finite dimensional matrix, the above series is convergent and has a unique solution for the system of equations. If Q is infinite, we cannot say anything. Suppose that Q is finite dimensional matrix and diagonalizable , and let $\beta_0, \beta_1, \ldots, \beta_n$ be the distinct eigenvalues of the matrix Q. Then there exists a matrix A such that' $Q = ADA^{-1}$, where D is a diagonal matrix with diagonal entries $\beta_0, \beta_1, \ldots \beta_n$. In this case, the matrix P takes the form $P(t) = AD_1(t)A^{-1}$, where $D_1(t)$ is a diagonal matrix with diagonal entries $e^{t\beta_0}, e^{\beta_1 t}, \ldots e^{\beta_n t}$.

Remark 9.2. Let $p(t)$ be the state probability vector i.e., $p(t) = (p_j(t))_{j\in S}$ given by $p_j(t) = \mathbb{P}(X(t) = j)$. Then

$$\begin{cases} p'(t) = p(t)Q \\ p'(t) = Qp(t). \end{cases}$$

These can be solved to get $p(t)$ at time t given initial state probability distribution $p(0)$ and rate matrix Q for a CTMC.

Example 9.3 (The two-state chain). *Consider a two-state continuous-time Markov chain that spends an exponential time with rate λ in state 0 before going to state 1, where it spends an exponential time with rate μ before returning to state 0, i.e., the rate matrix Q is given by $Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$. Find its probability transition matrix.*

Solution: *We first use* **matrix method**. *Observe that Q has two distinct eigen-values namely 0 and $-(\lambda+\mu)$ and corresponding eigen-vectors are $(1,1)$ and $(\lambda, -\mu)$. Thus, Q can be written as $Q = ADA^{-1}$ where*

$$A = \begin{pmatrix} 1 & \lambda \\ 1 & -\mu \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 \\ 0 & -(\lambda+\mu) \end{pmatrix}, \quad A^{-1} = \frac{1}{\lambda+\mu}\begin{pmatrix} \mu & \lambda \\ 1 & -1 \end{pmatrix}.$$

Therefore the probability transition matrix $P(t)$ is given by $P(t) = AD_1(t)A^{-1}$ where $D_1(t)$ is a diagonal matrix with diagonal entries 1 and $e^{-(\lambda+\mu)t}$. Thus, we have

$$\begin{aligned} P(t) &= \begin{pmatrix} 1 & \lambda \\ 1 & -\mu \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & e^{-(\lambda+\mu)t} \end{pmatrix} \cdot \frac{1}{\lambda+\mu}\begin{pmatrix} \mu & \lambda \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{\lambda+\mu}\begin{pmatrix} \mu + \lambda e^{-(\lambda+\mu)t} & \lambda - \lambda e^{-(\lambda+\mu)t} \\ \mu - \mu e^{-(\lambda+\mu)t} & \lambda + \mu e^{-(\lambda+\mu)t} \end{pmatrix}. \end{aligned}$$

Direct approach: *The Kolmogorov forward equation yields*

$$\begin{aligned} & p'_{00}(t) = \mu p_{01}(t) - \lambda p_{00}(t) = -(\lambda+\mu)p_{00}(t) + \mu \quad (\because\ p_{01}(t) = 1 - p_{00}(t)) \\ & \implies \frac{d}{dt}\big(e^{(\lambda+\mu)t}p_{00}(t)\big) = \mu e^{(\lambda+\mu)t} \\ & \implies e^{(\lambda+\mu)t}p_{00}(t) = \frac{\mu}{\lambda+\mu}e^{(\lambda+\mu)t} + C \\ & \implies p_{00}(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \quad (\because p_{00}(0) = 1 \implies C = \frac{\lambda}{\lambda+\mu}). \end{aligned}$$

Similarly, one can show that

$$p_{11}(t) = \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t}.$$

Example 9.4. *Calculate $p_{0j}(t)$ $(j > 0)$ for a Poisson process with rate λ. To do so, we solve the Kolmogorov forward equation. We have*

$$p_{00}'(t) = -\lambda p_{00}(t), \quad p_{0j}'(t) = -\lambda p_{0j}(t) + \lambda p_{0,j-1}(t) \quad j \geq 1.$$

Let $p_{0j}(t) = p_j(t)$. Define $\Pi(s,t) = \sum_{j=0}^{\infty} s^j p_j(t)$. Observe that

$$\begin{aligned}
\frac{\partial}{\partial t}\Pi(s,t) &= \frac{\partial}{\partial t}\Big\{\sum_{j=0}^{\infty} s^j p_j(t)\Big\} = -\lambda \sum_{j=0}^{\infty} p_j(t) s^j + \lambda \sum_{j=1}^{\infty} p_{j-1}(t) s^j \\
&= -\lambda \sum_{j=0}^{\infty} p_j(t) s^j + \lambda s \sum_{j=1}^{\infty} p_{j-1}(t) s^{j-1} \\
&= (\lambda s - \lambda) \sum_{j=0}^{\infty} p_j(t) s^j = \lambda(s-1)\Pi(s,t) \\
&\implies \Pi(s,t) = C \exp(\lambda(s-1)t) \quad \textit{for some constant } C > 0.
\end{aligned}$$

Note that

$$\Pi(s,0) = \sum_{j=0}^{\infty} s^j p_j(0) = \sum_{j=0}^{\infty} s^j p_{0j}(0) = 1.$$

Hence

$$\Pi(s,t) = \exp(\lambda(s-1)t)$$

which is the probability generating function of the Poisson(λt). Thus,

$$p_{0j}(t) = p_j(t) = e^{-\lambda t} \frac{(\lambda t)^j}{j!}.$$

9.2. **Waiting time and probability of jumps:** Let τ_i be the spent time in the state i before moving to another state. Then

$$\begin{aligned}
\mathbb{P}(\tau_i > t) &= \mathbb{P}\big(X_s = i,\ 0 \leq s \leq t \big| X_0 = i\big) \\
&= \mathbb{P}\big(X_s = i,\ 0 \leq s \leq \frac{t}{n} \big| X_0 = i\big)\mathbb{P}\big(X_s = i,\ \frac{t}{n} \leq s \leq \frac{2t}{n} \big| X_{\frac{t}{n}} = i\big) \\
&\quad \ldots \mathbb{P}\big(X_s = i,\ \frac{(n-1)t}{n} \leq s \leq t \big| X_{\frac{(n-1)t}{n}} = i\big) \\
&= \Big\{\mathbb{P}\big(X_s = i,\ 0 \leq s \leq \frac{t}{n} \big| X_0 = i\big)\Big\}^n \quad \forall\, n
\end{aligned}$$

where the last equality follows from time homogeneity. Thus, we have

$$\begin{aligned}
\mathbb{P}(\tau_i > t) &= \lim_{n\to\infty} \Big\{\mathbb{P}\big(X_s = i,\ 0 \leq s \leq \frac{t}{n} \big| X_0 = i\big)\Big\}^n \\
&= \lim_{n\to\infty} \Big\{1 + q_{ii}\frac{t}{n} + o(\frac{t}{n})\Big\}^n = e^{tq_{ii}} \quad \big(e^x = \lim_{n\to\infty}(1 + x/n)^n\big) \\
&\implies \mathbb{P}(\tau_i \leq t) = 1 - e^{tq_{ii}}.
\end{aligned}$$

Thus $\{\tau_i\}$ is exponentially distributed with mean $-\frac{1}{q_{ii}}$.

Example 9.5. *Consider a CTMC with rate matrix* $Q = \begin{pmatrix} -5 & 3 & 2 \\ 1 & -3 & 2 \\ 2 & 4 & -6 \end{pmatrix}$ *and initial distribution* $\lambda = (0, 0, 1)$*. Find* $\mathbb{P}(\tau > t)$ *where* τ *denotes the first transition time of the chain.*
Solution: *Given chain has the state space* $S = \{1, 2, 3\}$*. From the initial distribution, the chain starts at state* 3*. Hence*

$$\mathbb{P}(\tau > t) = e^{tq_{33}} = e^{-6t}.$$

Probability of jump: Next we discuss the probability of jumps to another state j starting from the state i. Observe that

$-q_{ii}$ $\quad \int_0^\infty \lambda x e^{-\lambda x}\, dx \quad \frac{1}{\lambda}$

$$\begin{aligned}
&\lim_{h \to 0} \mathbb{P}\big(X_{t+h} = j \big| X_t = i, X_{t+h} \neq i\big) \\
&= \lim_{h \to 0} \frac{\mathbb{P}\big(X_{t+h} = j \big| X_t = i\big)}{\mathbb{P}\big(X_{t+h} \neq i \big| X_t = i\big)} = \lim_{h \to 0} \frac{p_{ij}(h)}{\sum_{k \neq i} p_{ik}(h)} = \lim_{h \to 0} \frac{hq_{ij} + o(h)}{\sum_{k \neq i} hq_{ik} + o(h)} \\
&= \frac{q_{ij}}{\sum_{k \neq i} q_{ik}} = -\frac{q_{ij}}{q_{ii}}.
\end{aligned}$$

$\lambda x = t \quad \frac{\int_0^\infty t e^{-t} dt}{\lambda}$

$1 - \lambda \quad [-te^{-t} + -e^{-t} dt]_0^\infty$

Thus, the probability of jumps to another state j, starting from state i, is $-\frac{q_{ij}}{q_{ii}}$ and then it stays in this state for a period exponentially distributed with mean $-\frac{1}{q_{jj}}$. Thus, the process only depends on the rate matrix Q.

$\frac{1}{\lambda} =$

9.3. **Limiting distribution of CTMC.**

Definition 9.2. Let $\{X_t : t \geq 0\}$ be a continuous-time Markov-chain with transition probability matrix $(P(t))_{t \geq 0}$. A measure μ defined over the state space S is called an **invariant measure** for $\{X_t : t \geq 0\}$ if and only if for all $t \geq 0$, μ satisfies the relation

$$\mu = \mu P(t)$$

i.e., for each $j \in S$, μ satisfies

$$\mu(j) = \sum_{i \in S} \mu(i) p_{ij}(t).$$

If $\sum_{j \in S} \mu(j) = 1$, then μ is called stationary distribution.

Example 9.6. *Let* $\{X_t : t \geq 0\}$ *be a CTMC with state space* $S = \{0, 1\}$ *and transition matrix*

$$P(t) = \begin{pmatrix} \frac{2}{3} + \frac{1}{3}e^{-3t} & \frac{1}{3} - \frac{1}{3}e^{-3t} \\ \frac{2}{3} - \frac{2}{3}e^{-3t} & \frac{1}{3} + \frac{2}{3}e^{-3t} \end{pmatrix}.$$

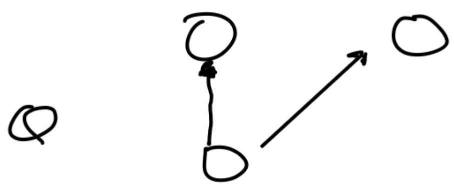
Then $\mu = (\frac{2}{3}, \frac{1}{3})$ *is a stationary distribution for* $\{X_t : t \geq 0\}$*. Indeed,*

$$\begin{aligned}
\mu(0)p_{01}(t) + \mu(1)p_{11}(t) &= \frac{2}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{3} = \frac{1}{3} = \mu(1), \\
\mu(0)p_{00}(t) + \mu(1)p_{10}(t) &= \frac{2}{3} \times \frac{2}{3} + \frac{1}{3} \times \frac{2}{3} = \frac{2}{3} = \mu(0).
\end{aligned}$$

Embedded Markov Chain: For a given rate matrix Q, we can define a stochastic matrix P as follows:

$$p_{ij} = \begin{cases} -\frac{q_{ij}}{q_{ii}}, & i \neq j \text{ and } q_{ii} \neq 0 \\ 0, & i = j. \end{cases}$$

If $q_{ii} = 0$, then $p_{ij} = \delta_{ij}$. The discrete-time Markov chain with initial distribution λ and transition probability matrix P is called **embedded Markov chain**.



$-q_{ij} =$

Example 9.7. *Consider a CTMC $\{X(t) : t \geq 0\}$ with finite state space $\mathcal{S} = \{1, 2, 3\}$ and rate matrix $Q = \begin{pmatrix} -5 & 3 & 2 \\ 1 & -3 & 2 \\ 2 & 4 & -6 \end{pmatrix}$. Find the transition probability matrix of embedded Markov chain. Classify the states of embedded Markov chain.*

Solution: *Let $\{Y_n : n \geq 0\}$ be the embedded Markov chain of the given CTMC. Then the transition probability matrix of $\{Y_n\}$ is given by $P = \begin{pmatrix} 0 & \frac{3}{5} & \frac{2}{5} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}$. From the transition diagram, we can easily see that chain is irreducible. Since it is finite Markov chain, all the states are positive recurrent.*

We make use of the embedded Markov chain to give conditions that will guarantee the existence and uniqueness of a stationary distribution.

Theorem 9.1. *Let $\{X_t : t \geq 0\}$ be a CTMC with rate matrix Q and let $\{Y_n : n \in \mathbb{N}\}$ be the corresponding Markov chain. If $\{Y_n\}$ is irreducible, and positive recurrent with*

$$\widetilde{\lambda} = \inf\{\lambda_i : i \in S\} > 0,$$

then there exists a unique stationary distribution $\Pi = (\Pi_i)$ for $\{X_t : t \geq 0\}$ i.e., for given any state j

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

exists and is the same for all initial states $i \in S$. Moreover, Π_j can be determined as solutions of the system of linear equations

$$0 = \sum_i \Pi_j q_{ij}, \quad and \quad \sum_j \Pi_j = 1.$$

Example 9.8. *Consider a CTMC $\{X(t) : t \geq 0\}$ with finite state space $S = \{1, 2, 3\}$, rate matrix $Q = \begin{pmatrix} -5 & 3 & 2 \\ 1 & -2 & 1 \\ 1 & 3 & -4 \end{pmatrix}$, and the initial distribution $(1/3, 1/3, 1/3)$. Explain whether the stationary distribution for the given CTMC exists or not. If so, then find it.*

Solution: *The transition probability matrix of the embedded Markov chain is given by $P = \begin{pmatrix} 0 & \frac{3}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} & 0 \end{pmatrix}$. Observe that, from the transition diagram, embedded chain is irreducible and since state space is finite, it is positive recurrent. Since $\tilde{\lambda} = \inf\{\lambda_i : i \in \mathcal{S}\} = \frac{1}{3} > 0$, there exists unique limiting distribution i.e., there exists invariant measure $\Pi = \big(\Pi_i\big)$ such that*

$$\lim_{t \to \infty} p_{ij}(t) = \Pi_j \quad \forall\, j \in \mathcal{S}.$$

To find Π, we need to solve

$$\Pi Q = \vec{0}, \quad \Pi_1 + \Pi_2 + \Pi_3 = 1, \quad \Pi_i \geq 0$$

i.e., we need to solve the following system of equations:

$$-5\Pi_1 + \Pi_2 + \Pi_3 = 0, \quad 3\Pi_1 - 2\Pi_2 + 3\Pi_3 = 0$$
$$2\Pi_1 + \Pi_2 - 4\Pi_3 = 0, \quad \Pi_1 + \Pi_2 + \Pi_3 = 1\,.$$

Solving above system, we get

$$\Pi_1 = \frac{1}{6}, \quad \Pi_2 = \frac{3}{5}, \quad \Pi_3 = \frac{7}{30}.$$

Thus, we obtain

$$\lim_{t\to\infty} p_{i1}(t) = \frac{1}{6}, \quad \lim_{t\to\infty} p_{i2}(t) = \frac{3}{5}, \quad \lim_{t\to\infty} p_{i3}(t) = \frac{7}{30}.$$

9.4. **Birth and Death process:** A continuous-time Markov chain with state space $S = \{0, 1, 2, \ldots\}$ for which $q_{ij} = 0$ whenever $|i - j| > 1$ is called a **birth and death process**. Thus, the transitions from state i can only go to either state $i-1$ or state $i+1$. The state of the process is usually thought of as representing the size of some population, and when the state increases by 1, we say that a *birth* occurs, and when it decreases by 1, we say that a *death* occurs. Let $\lambda_i = q_{i,i+1}$ and $\mu_i = q_{i,i-1}$ be the birth and death rates. Then the density /rate matrix Q is given by

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & 0\ldots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & 0\ldots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & 0\ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ldots \end{pmatrix}$$

Observe that $\lambda_i h + o(h)$ represents the probability of a birth in the interval of infinitesimal length $(t, t+h)$ given that $X_t = i$ and $\mu_i h + o(h)$ represents the probability of death in the interval of infinitesimal length $(t, t+h)$ given that $X_t = i$. From the Kolmogorov's backward equation $P'(t) = QP(t)$, we have

$$p'_{0j}(t) = -\lambda_0 p_{0j}(t) + \lambda_0 p_{1j}(t), \quad j = 0, 1, 2, \ldots$$
$$p'_{ij}(t) = -(\lambda_i + \mu_i)p_{ij}(t) + \lambda_i p_{i+1,j}(t) + \mu_i p_{i-1,j}(t) \quad i \geq 1.$$

Similarly, from the forward Kolmogorov equation, we obtain

$$p'_{i0}(t) = -\lambda_0 p_{i0}(t) + \mu_1 p_{i1}(t), \quad i \in S$$
$$p'_{ij}(t) = -(\lambda_j + \mu_j)p_{ij}(t) + \lambda_{i-1} p_{i,j-1}(t) + \mu_{j+1} p_{i,,j+1}(t) \quad j \geq 1.$$

9.4.1. **Stationary distribution for birth and death process.** . For that we need to solve

$$\Pi Q = \vec{0}, \quad \sum_{i\in S} \Pi_i = 1.$$

Q

$\pi P(t) = \pi$

We have

$$-\lambda_0 \Pi_0 + \mu_1 \Pi_1 = 0$$
$$\lambda_{i-1}\Pi_{i-1} - (\lambda_i + \mu_i)\Pi_i + \mu_{i+1}\Pi_{i+1} = 0, \quad i \geq 1.$$

$\pi(P(t) - I) = 0 \ \forall t$

$t, t+\delta t$ h

Thus, we have

$\pi(P(t+\delta t) - P(t))$

$\pi P'(0) = \pi Q = \vec{0}$

$$\begin{aligned} \mu_{i+1}\Pi_{i+1} - \lambda_i \Pi_i &= \mu_i \Pi_i - \lambda_{i-1}\Pi_{i-1} \\ &= \mu_{i-1}\Pi_{i-1} - \lambda_{i-2}\Pi_{i-2} \\ &\vdots \\ &= \mu_1 \Pi_1 - \lambda_0 \Pi_0 = 0 \end{aligned}$$

Therefore

$$\Pi_{i+1} = \frac{\lambda_i}{\mu_{i+1}}\Pi_i \quad \forall\, i \geq 0$$
$$\implies \Pi_n = \frac{\lambda_{n-1}}{\mu_n}\Pi_{n-1} = \ldots = \frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i}\Pi_0$$

$$\implies \sum_{n=0}^{\infty}\Pi_n = \Pi_0 + \sum_{n=1}^{\infty}\Pi_n = \Pi_0\Big(1+\sum_{n=1}^{\infty}\frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i}\Big).$$

The stationary distribution exists if and only if

$$\sum_{n=1}^{\infty}\frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i} < \infty.$$

If this sum converges, then

$$\Pi_0 = \left(1+\sum_{n=1}^{\infty}\frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i}\right)^{-1},$$

$$\Pi_n = \frac{\prod_{i=0}^{n-1}\lambda_i}{\prod_{i=1}^{n}\mu_i}\Pi_0, \quad n \ge 1.$$

Remark 9.3. If the state space $\mathcal{S}$ if finite and that $\lambda_i > 0$, $\mu_i > 0$ for $i \in \mathcal{S}$, then the embedded Markov chain is irreducible and positive recurrent and hence there exists a stationary distribution for $\{X(t) : t \ge 0\}$, say $\Pi = (\Pi_1, \Pi_2, \ldots, \Pi_m)$. Then one has

$$\Pi_i = \frac{\lambda_0\lambda_1\cdots\lambda_{i-1}}{\mu_1\mu_2\cdots\mu_i}\Pi_0, \quad i = 1, 2, \ldots, m$$

$$\Pi_0 = \frac{1}{1+\sum_{i=1}^{m}\frac{\lambda_0\lambda_1\cdots\lambda_{i-1}}{\mu_1\mu_2\cdots\mu_i}}.$$

9.4.2. **Linear Growth with immigration:** A birth and death process is called a linear growth process if

$$\begin{cases}\lambda_n = \lambda n + a\\ \mu_n = n\mu\end{cases}$$

with $\lambda, \mu, a > 0$. Such processes occur naturally in the study of biological reproduction and population growth. From the forward Kolmogorov equation, we obtain

$$p'_{i0}(t) = -ap_{i0}(t) + \mu p_{i1}(t),$$
$$p'_{ij}(t) = \big(\lambda(j-1)+a\big)p_{i,j-1}(t) - \big((\lambda+\mu)j+a\big)p_{ij}(t) + \mu(j+1)p_{i,j+1}(t), \quad j \ge 1.$$

We will determine the mean population size for large t. Let $m(t)$ be the mean population size at time t. Then

$$m(t) = \sum_{j=1}^{\infty} jp_{ij}(t) \implies m'(t) = \sum_{j=1}^{\infty} jp'_{ij}(t)$$

$$\implies m'(t) = \sum_{j=1}^{\infty} j\Big\{\big(\lambda(j-1)+a\big)p_{i,j-1}(t) - \big((\lambda+\mu)j+a\big)p_{ij}(t) + \mu(j+1)p_{i,j+1}(t)\Big\}$$
$$= \lambda\sum_{j=1}^{\infty} j(j-1)p_{i,j-1}(t) - (\lambda+\mu)\sum_{j=1}^{\infty} j^2p_{ij}(t) + \mu\sum_{j=1}^{\infty} j(j+1)p_{i,j+1}(t)$$
$$+ a\sum_{j=1}^{\infty} jp_{i,j-1}(t) - a\sum_{j=1}^{\infty} jp_{ij}(t).$$

Observe that

$$a\sum_{j=1}^{\infty} jp_{i,j-1}(t) - a\sum_{j=1}^{\infty} jp_{ij}(t) = a\sum_{j=0}^{\infty}(j+1)p_{ij}(t) - a\sum_{j=0}^{\infty} jp_{ij}(t) = a\sum_{j=0}^{\infty} p_{ij}(t) = a,$$

and

$$\begin{aligned}
&\lambda \sum_{j=1}^{\infty} j(j-1)p_{i,j-1}(t) - (\lambda+\mu)\sum_{j=1}^{\infty} j^2 p_{ij}(t) + \mu \sum_{j=1}^{\infty} j(j+1)p_{i,j+1}(t) \\
&= \lambda \sum_{j=0}^{\infty} j(j+1)p_{ij}(t) - (\lambda+\mu)\sum_{j=0}^{\infty} j^2 p_{ij}(t) + \mu \sum_{j=2}^{\infty} j(j-1)p_{ij}(t) \\
&= \lambda \sum_{j=0}^{\infty} j(j+1)p_{ij}(t) - (\lambda+\mu)\sum_{j=0}^{\infty} j^2 p_{ij}(t) + \mu \sum_{j=0}^{\infty} j(j-1)p_{ij}(t) \\
&= \lambda \sum_{j=0}^{\infty} jp_{ij}(t) - \mu \sum_{j=0}^{\infty} jp_{ij}(t) = (\lambda-\mu)m(t).
\end{aligned}$$

Thus, $m(t)$ satisfies the differential equation

$$m'(t) = a + (\lambda-\mu)m(t)$$

with initial condition $m(0) = i$ if $X(0) = i$. Ths solution of this equation is

$$m(t) = \begin{cases} at+i, & \lambda = \mu, \\ \frac{a}{\lambda-\mu}\big(e^{(\lambda-\mu)t} - 1\big) + ie^{(\lambda-\mu)t}, & \lambda \neq \mu\,. \end{cases}$$

$$\implies m(t) \to \infty \quad (t\to\infty) \text{ if } \lambda \ge \mu; \quad m(t) \sim \frac{a}{\mu-\lambda}, \quad \text{if } \lambda < \mu.$$

9.4.3. **Pure Birth process:**

Definition 9.3 (Pure birth process)**.** A birth and death process is said to be a **pure birth process** if $\mu_n = 0$ for all n.

Example 9.9. *The simplest example of a pure birth process is the Poisson process, which has a constant birth rate* $\lambda_n = \lambda,\ n \ge 0$.

Example 9.10. *Suppose $X(t)$ represents the population size at time t. Each member in the population acts independently and gives birth at an exponential rate λ. Suppose that no one ever dies. Then $\{X(t) : t \ge 0\}$ is a pure birth process with $\lambda_n = n\lambda,\ n \ge 0$. This pure birth process is called a* **Yule process***.*

Distribution of Yule process: Consider a Yule process starting with a single individual at time 0. Let $\{T_i : i \ge 1\}$ be the time between the $(i-1)$st and ith birth, i.e., T_i is the time it takes from the population size to go from i to $i+1$. $\{T_i\}$ are independent and exponentially distributed with rate $i\lambda$. We wish to find the distribution of $X(t)$. Observe that

$$\begin{aligned}
&\mathbb{P}(T_1 \le t) = 1 - e^{-\lambda t}, \\
&\mathbb{P}(T_1 + T_2 \le t) = \int_0^t \mathbb{P}(T_1+T_2 \le t \big| T_1 = x)\lambda e^{-\lambda x} dx = \int_0^t (1 - e^{-2\lambda(t-x)})\lambda e^{-\alpha x}\, dx \\
&\qquad\qquad = (1-e^{-\lambda t})^2, \\
&\mathbb{P}(T_1+T_2+T_3 \le t) = \int_0^t \mathbb{P}(T_1+T_2+T_3 \le t \big| T_1+T_2 = x)\, dF_{T_1+T_2}(x) \\
&\qquad\qquad = \int_0^t (1-e^{-3\lambda(t-x)})2\lambda e^{-\lambda x}(1-e^{-\lambda x})\, dx = (1-e^{-\lambda t})^3.
\end{aligned}$$

Hence by induction, one can show that

$$\mathbb{P}(\sum_{i=1}^{j} T_i \le t) = (1-e^{-\lambda t})^j.$$

Now

$$\begin{aligned} p_{1j}(t) &= \mathbb{P}\big(X(t) \geq j \big| X(0) = 1\big) - \mathbb{P}\big(X(t) \geq j+1 \big| X(0) = 1\big) \\ &= \mathbb{P}(T_1 + \ldots + T_{j-1} \leq t) - \mathbb{P}(T_1 + \ldots + T_j \leq t) \\ &= (1 - e^{-\lambda t})^{j-1} - (1 - e^{-\lambda t})^j = (1 - e^{-\lambda t})^j e^{-\lambda t}. \end{aligned}$$

Hence, starting with a single individual, the population size at time t will have a *geometric distribution* with mean $e^{\lambda t}$.

Remark 9.4. If the population starts with i individuals, it follows that its size at t will be the sum of i independent and identically distributed geometric random variables, and will thus have a negative binomial distribution i.e.,

$$p_{ij}(t) = \binom{j-1}{i-1} e^{-i\lambda t}(1 - e^{-\lambda t})^{j-i}, \quad 1 \leq i \leq j.$$

Expected sum of the ages: Let $A(t)$ be the sum of the ages at time t of a Yule process starting with single individual i.e., $X(0) = 1$. Then

$$A(t) = a_0 + \int_0^t X(s)\,\mathrm{d}s$$

where a_0 is the age at $t = 0$ of the initial individual. Hence

$$\mathbb{E}[A(t)] = a_0 + \mathbb{E}[\int_0^t X(s)\,\mathrm{d}s] = a_0 + \int_0^t \mathbb{E}[X(s)]\,\mathrm{d}s = a_0 + \int_0^t e^{\lambda s}\,\mathrm{d}s = a_0 + \frac{e^{\lambda t} - 1}{\lambda}.$$