Time Series

Peter Lukianchenko

18 September 2021

Markov process

Definition

$\{x(t,\omega)\}_{t\in J}$ is called a **Markov process (MP)**, if the following **Markov property** holds: for any $t_0 \leq \tau \leq t \leq T$ an all $A \in \mathcal{B}^n$

$$\boxed{\mathrm{P}\left\{x(t,\omega) \in A \mid \mathcal{F}_{[t_0,\tau]}\right\} \overset{a.s.}{=} \mathrm{P}\left\{x(t,\omega) \in A \mid x(\tau,\omega)\right\}}$$

Markov process

Let the *phase space* of a Markov process $\{x(t,\omega)\}_{t\in\mathcal{T}}$ be *discrete*, that is,

$$\boxed{\begin{array}{c} x(t,\omega) \in \mathcal{X} := \{(1,2,\dots,N) \text{ or } \mathbb{N}\cup\{0\}\} \\ \mathbb{N} = 1,2,\dots \text{ is a countable set, or finite} \end{array}}$$

Definition

A Markov process $\{x(t,\omega)\}_{t\in\mathcal{T}}$ with a discrete phase space X is said to be a **Markov chain (or Finite Markov Chain** if $\mathbb{N}$ is finite)

a) in continuous time if

$$\mathcal{T} := [t_0, T)\,, \quad T \text{ is admitted to be } \infty$$

b) in discrete time if

$$\mathcal{T} := \{t_0, t_1, \dots, t_T\}\,, \quad T \text{ is admitted to be } \infty$$

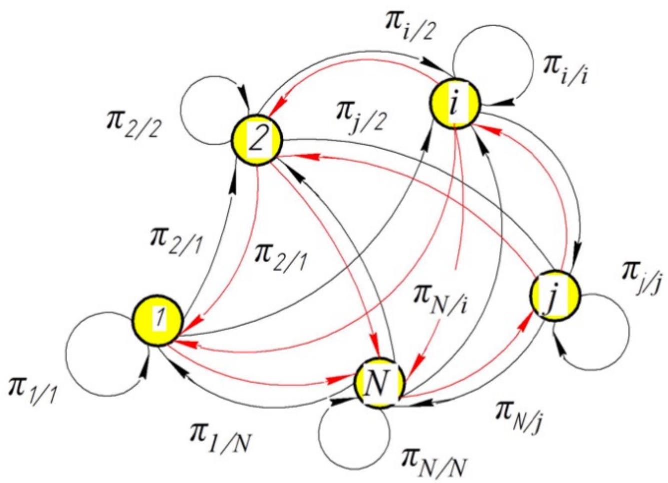
Markov process

Homogeneous

Definition

A Markov Chain is said to be Homogeneous **(Stationary)** if the transition probabilities are constant, that is,

$$\boxed{\pi_{j|i}(n) = \pi_{j|i} = \text{const for all } n = 0, 1, 2, \ldots}$$



Markov process

Definition

A Markov Chain is called **ergodic** if all its states are returnable.

The result below shows that homogeneos ergodic Markov chains posses some additional property:
after a long time such chains "forget" the initial states from which they have started.

Markov process

Show that the Finite Markov Chain with the transition matrix

$$\Pi := \begin{bmatrix} 0 & 0.3 & 0 & 0.7 \\ 1 & 0 & 0 & 0 \\ 0.1 & 0 & 0.9 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

is *ergodic*. Indeed, after 2 steps $(n_0 = 2)$

$$\Pi^2 = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 \\ 0 & 0.3 & 0 & 0.7 \\ 0.09 & 0.03 & 0.81 & 0.07 \\ 1.0 & 0 & 0 & 0 \end{bmatrix}, \Pi^3 = \begin{bmatrix} 0.7 & \mathbf{0.09} & 0 & 0.21 \\ 0.3 & \mathbf{0.7} & 0 & 0 \\ 0.111 & \mathbf{0.097} & 0.729 & 0.063 \\ 0 & \mathbf{0.3} & 0 & 0.7 \end{bmatrix}$$

$\Pi^3 = \Pi^{1+n_0}$ contains the column $j = 2$ with strictly positive elements.

Markov process

Theorem (the ergodic theorem)

Let for some state $j_0 \in X$ of a homogeneous Markov chain and some $n_0 > 0$, $\delta \in (0,1)$ for all $i \in (1, ..., N)$

$$\boxed{\left(\Pi^{n_0}\right)_{j_0|i} \geq \delta > 0}$$

i.e., after n_0-times multiplications Π by itself at least one column of the matrix Π^{n_0} has all nonzero elements. Then for any initial state distribution $P\left\{x\left(t_0, \omega\right) = i\right\}$ and for any $i, j \in (1, ..., N)$ there exists the limit

$$\boxed{p_j^* := \lim_{n \to \infty} \left(\Pi^n\right)_{j|i} > 0}$$

such that for any $t \geq 0$ this limit is reachable with an exponential rate, namely,

$$\left|\left(\Pi^n\right)_{j|i} - p_j^*\right| \leq (1-\delta)^{[t_n/n_0]} = e^{-\alpha[t_n/n_0]}, \alpha := \left|\ln(1-\delta)\right|$$

Markov process

Corollary

For any $j \in (1, 2, ..., N)$ of an ergodic homogeneous finite Markov chain the components p_j^* of the **stationary distribution,** *satisfy the folowing* **ergodicity relations**

$$\boxed{\left.\begin{array}{c} p_j^* = \sum\limits_{i \in \mathcal{X}} \pi_{j|i}\, p_i^* \\ \sum\limits_{i \in \mathcal{X}} p_i^* = 1,\, p_i^* > 0 \;\; (i = 1, 2, ..., N) \end{array}\right\}}$$

or equivalently, in the vector format

$$\boxed{p^* = \Pi^\intercal p^*, \qquad p^* := \left(p_1^*, ..., p_N^*\right), \quad \Pi := \left\| \pi_{j|i} \right\|_{i,j=1,...,N}}$$

that is, the positive vector p^* is the eigenvector of the matrix $\Pi^\intercal(t)$ corresponding to its eigenvalue equal to 1.

Markov process

Let $\Pi_k(n) := \left\| \pi_{j|i,k}(n) \right\|_{i,j=1,\ldots,N}$ be the transition matrix with the elements

$$\boxed{\pi_{j|i,k}(n) := P\left\{x(t_{n+1},\omega) = j \mid x(t_n,\omega) = i,\ a(t_n,\omega) = k\right\},\ k = 1,\ldots,K}$$

where the variable $a(t_n,\omega)$ is associated with a control action (decision making) from the given set of possible controls $(1,\ldots,K)$. Each control action $a(t_n,\omega) = k$ may be selected (realized) in state $x(t_n,\omega) = i$ with the probability

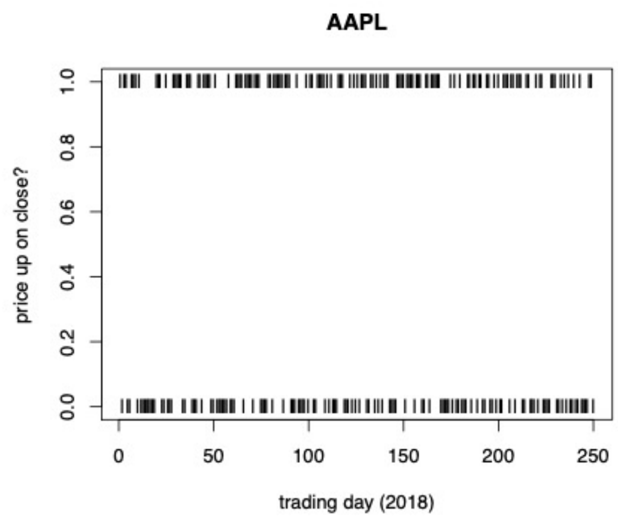
$$\boxed{d_{ki}(n) := P\left\{a(t_n,\omega) = k \mid x(t_n,\omega) = i\right\}}$$

fulfilling the *stochastic constraints*

$$\boxed{d_{ki}(n) \geq 0,\ \sum_{k=1}^{N} d_{ki}(n) = 1 \text{ for all } i = 1,\ldots,N}$$

Markov process

Back to the AAPL data again!



If we model the AAPL price up/down data as a two-state Markov chain, our main job becomes to estimate the transition matrix

$$\Gamma = \begin{bmatrix} \gamma_{00} & \gamma_{01} \\ \gamma_{10} & \gamma_{11} \end{bmatrix} = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}.$$

Markov process

We can summarize observations of a Markov chain in terms of the initial state c_0 and n_{ij}, the number of observed transitions from state i to state j.

- The matrix (n_{ij}) is sometimes called the frequency matrix.
- In the case of the AAPL price up/down data we have

$$\begin{bmatrix} n_{00} & n_{01} \\ n_{10} & n_{11} \end{bmatrix} = \begin{bmatrix} 62 & 59 \\ 60 & 68 \end{bmatrix}$$

- **NB** The total number of transitions is $n_{00} + n_{01} + n_{10} + n_{11} = 249$, i.e. one less than the number of observations.

Markov process

The simplest estimates are

$$\hat{\alpha} = \frac{n_{01}}{n_{00} + n_{01}} = \frac{59}{62 + 59} = 0.488$$

$$\hat{\beta} = \frac{n_{10}}{n_{10} + n_{11}} = \frac{60}{60 + 68} = 0.469$$

- These make intuitive sense:

$$\hat{\alpha} = \frac{\text{number of times chain went from } 0 \text{ to } 1}{\text{number of times chain went from } 0 \text{ to anywhere}}$$

$$\hat{\beta} = \frac{\text{number of times chain went from } 1 \text{ to } 0}{\text{number of times chain went from } 1 \text{ to anywhere}}$$

Markov process

The results of the two-state model generalize in the way you might expect.

If n_{ij} are the observed number of transitions from state i to state j, then the estimates of the transition probabilities for the Markov chain are given by

$$\hat{\gamma}_{ij} = \frac{n_{ij}}{n_i},$$

where $n_i = \sum_j n_{ij}$.

- If $n_i = 0$, then $n_{ij} = 0$, and we set $\hat{\gamma}_{ij} = 0$.
- Importantly, **confidence intervals** can be constructed for the transition probabilities.

Markov process

It is crucially important for statistical modelling that we can also carry out parameter estimation in Markov chain models when $\Gamma = \Gamma(\theta)$. Fortunately, this can be done – though we will omit the details!

- As example, in the mindreader example, we might decide to model the human brain so that $\alpha = \beta = \theta$.
- In this case,
$$\hat{\theta} = \frac{n_{01} + n_{10}}{n},$$
where n is the total number of transitions.

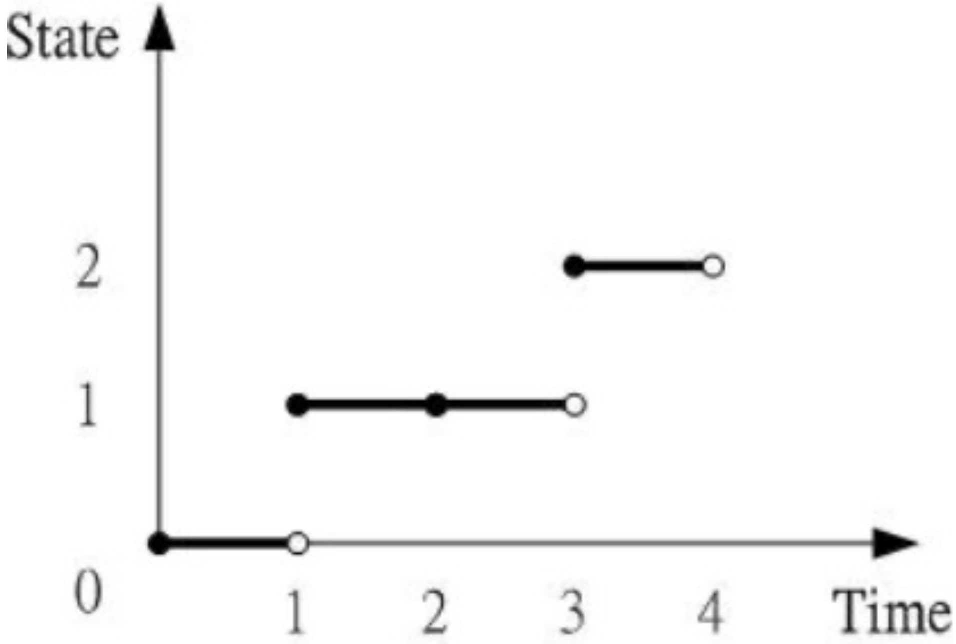
Markov process

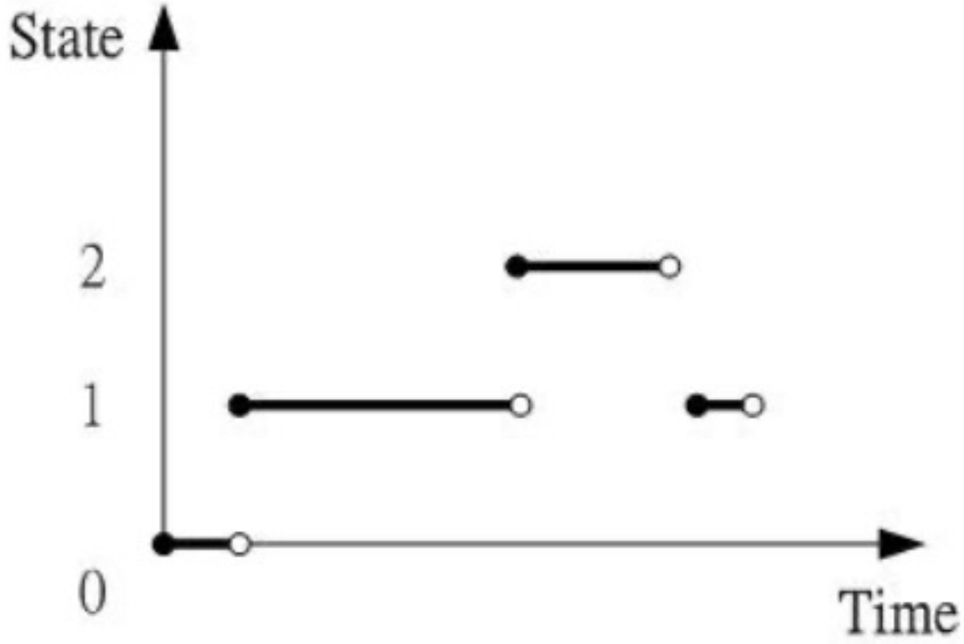
Classical parameter estimation relies on the fact that **the likelihood of the parameters is equal to the probability of the data**. In homogeneous Markov chain models, the probability of the data is

$$
\begin{aligned}
& \mathbb{P}(C_0 = c_0, C_1 = c_1, \ldots, C_n = c_n) \\
\overset{\text{saving ink}}{=} \; & p(c_0, c_1, \ldots, c_n) \\
\overset{\text{conditional probability}}{=} \; & p(c_0)p(c_1|c_0)p(c_2|c_1, c_0) \cdots p(c_n|c_{n-1}, \ldots, c_0) \\
\overset{\text{Markov property}}{=} \; & p(c_0)p(c_1|c_0)p(c_2|c_1) \cdots p(c_n|c_{n-1}) \\
\overset{\text{homogeneous}}{=} \; & \delta_{c_0}\gamma_{c_0c_1}\gamma_{c_1c_2} \cdots \gamma_{c_{n-1}c_n}
\end{aligned}
$$

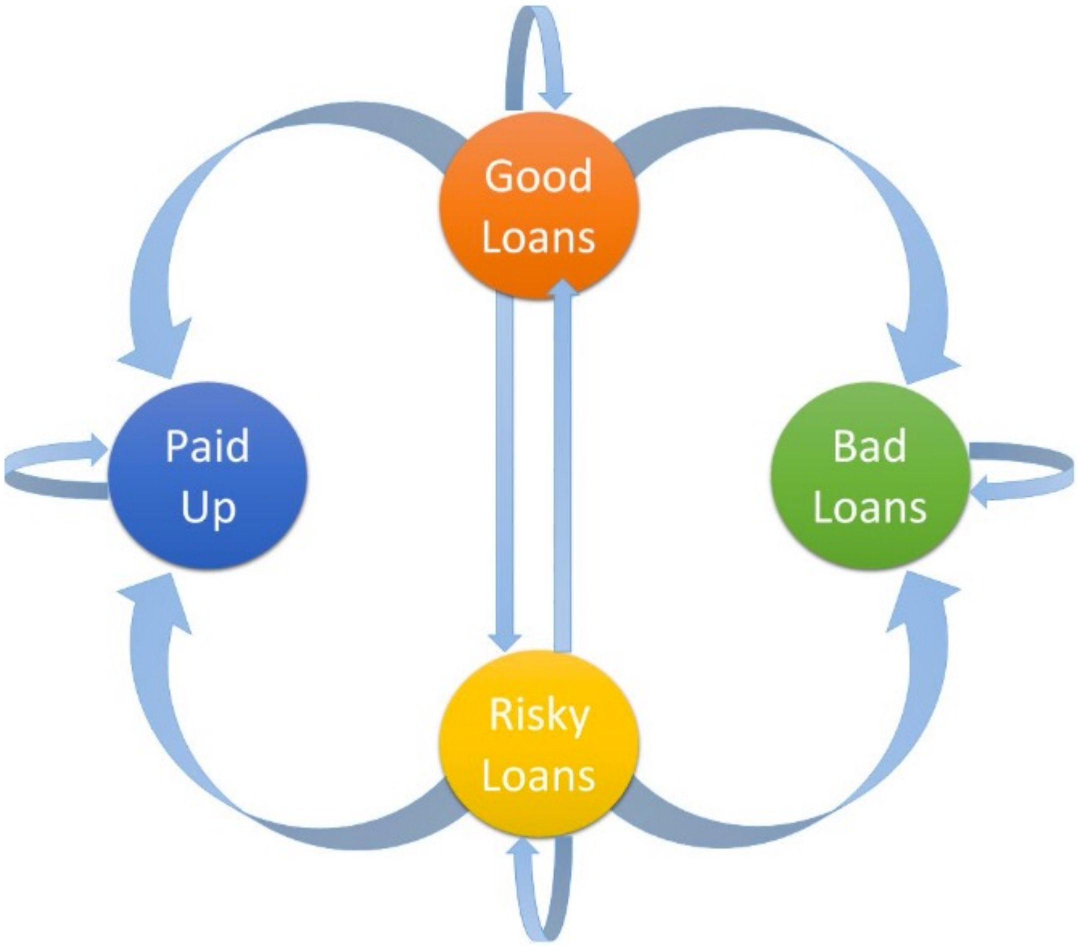
- Parameter estimation is then typically carried out using **maximum likelihood estimation**.

Markov process

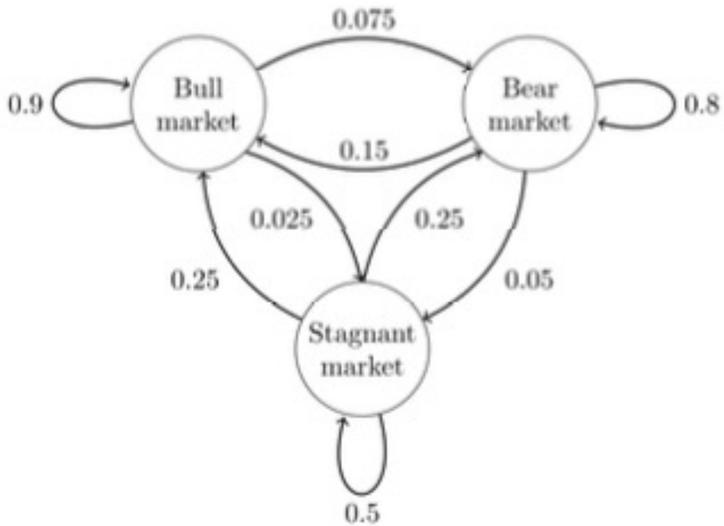




Markov process



Markov process



Markov process

- Continuous-time positive variable $t \in [0, \infty)$
- Time-dependent random state $X(t)$ takes values on a countable set
  - In general denote states as $i = 0, 1, 2, \ldots$, i.e., here the state space is $\mathbb{N}$
  - If $X(t) = i$ we say "the process is in state i at time t"
- **Def:** Process $X(t)$ is a continuous-time Markov chain (CTMC) if

$$\begin{aligned} &\mathrm{P}\left(X(t+s) = j \,\middle|\, X(s) = i, X(u) = x(u), u < s\right) \\ = &\mathrm{P}\left(X(t+s) = j \,\middle|\, X(s) = i\right) \end{aligned}$$

- Markov property $\Rightarrow$ Given the present state $X(s)$

  $\Rightarrow$ Future $X(t+s)$ is independent of the past $X(u) = x(u), u < s$
- In principle need to specify functions $\mathrm{P}\left(X(t+s) = j \,\middle|\, X(s) = i\right)$

  $\Rightarrow$ For all times t and s, for all pairs of states (i, j)

Markov process

- T_i = time until transition out of state i into any other state j
- **Def:** T_i is a random variable called transition time with ccdf

$$\mathrm{P}\left(T_i > t\right) = \mathrm{P}\left(X(0:t] = i \,\middle|\, X(0) = i\right)$$

- Probability of $T_i > t + s$ given that $T_i > s$? Use cdf expression

$$\begin{aligned}
\mathrm{P}\left(T_i > t + s \,\middle|\, T_i > s\right) &= \mathrm{P}\left(X(0:t+s] = i \,\middle|\, X[0:s] = i\right) \\
&= \mathrm{P}\left(X(s:t+s] = i \,\middle|\, X[0:s] = i\right) \\
&= \mathrm{P}\left(X(s:t+s] = i \,\middle|\, X(s) = i\right) \\
&= \mathrm{P}\left(X(0:t] = i \,\middle|\, X(0) = i\right)
\end{aligned}$$

- Used that $X[0:s] = i$ given, Markov property, and homogeneity
- From definition of T_i $\Rightarrow \mathrm{P}\left(T_i > t + s \,\middle|\, T_i > s\right) = \mathrm{P}\left(T_i > t\right)$

  $\Rightarrow$ Transition times are exponential random variables

Markov process

- Exponential transition times is a fundamental property of CTMCs

  $\Rightarrow$ Can be used as "algorithmic" definition of CTMCs

- Continuous-time random process $X(t)$ is a CTMC if

  (a) Transition times T_i are exponential random variables with mean $1/\nu_i$
  (b) When they occur, transition from state i to j with probability P_{ij}

$$\sum_{j=1}^{\infty} P_{ij} = 1, \qquad P_{ii} = 0$$

  (c) Transition times T_i and transitioned state j are independent

- Define matrix $\mathbf{P}$ grouping transition probabilities P_{ij}

- CTMC states evolve as in a discrete-time Markov chain

  $\Rightarrow$ State transitions occur at exponential intervals $T_i \sim \exp(\nu_i)$

  $\Rightarrow$ As opposed to occurring at fixed intervals

Markov process

- Consider a CTMC with transition matrix $\mathbf{P}$ and rates ν_i
- **Def:** CTMC's embedded discrete-time MC has transition matrix $\mathbf{P}$
- Transition probabilities $\mathbf{P}$ describe a discrete-time MC
  - $\Rightarrow$ No self-transitions ($P_{ii} = 0$, $\mathbf{P}$'s diagonal null)
  - $\Rightarrow$ Can use underlying discrete-time MCs to study CTMCs
- **Def:** State j accessible from i if accessible in the embedded MC
- **Def:** States i and j communicate if they do so in the embedded MC
  - $\Rightarrow$ Communication is a class property
- Recurrence, transience, ergodicity. Class properties . . . More later

Markov process

- Expected value of transition time T_i is $\mathbb{E}[T_i] = 1/\nu_i$
  - $\Rightarrow$ Can interpret ν_i as the rate of transition out of state i
  - $\Rightarrow$ Of these transitions, a fraction P_{ij} are into state j
- **Def:** Transition rate from i to j is $q_{ij} := \nu_i P_{ij}$
- Transition rates offer yet another specification of CTMCs
- If q_{ij} are given can recover ν_i as

$$\nu_i = \nu_i \sum_{j=1}^{\infty} P_{ij} = \sum_{j=1}^{\infty} \nu_i P_{ij} = \sum_{j=1}^{\infty} q_{ij}$$

- Can also recover P_{ij} as $\Rightarrow P_{ij} = q_{ij}/\nu_i = q_{ij}\left(\sum_{j=1}^{\infty} q_{ij}\right)^{-1}$

Markov process

- State $X(t) = 0, 1, \ldots$ Interpret as number of individuals
- Birth and deaths occur at state-dependent rates. When $X(t) = i$
- Births $\Rightarrow$ Individuals added at exponential times with mean $1/\lambda_i$
  $\Rightarrow$ Birth or arrival rate $= \lambda_i$ births per unit of time
- Deaths $\Rightarrow$ Individuals removed at exponential times with rate $1/\mu_i$
  $\Rightarrow$ Death or departure rate $= \mu_i$ deaths per unit of time
- Birth and death times are independent
- Birth and death (BD) processes are then CTMCs

Markov process

- Q: Transition times T_i? Leave state $i \neq 0$ when birth or death occur
- If T_B and T_D are times to next birth and death, $T_i = \min(T_B, T_D)$

  $\Rightarrow$ Since T_B and T_D are exponential, so is T_i with rate

$$\nu_i = \lambda_i + \mu_i$$

- When leaving state i can go to $i+1$ (birth first) or $i-1$ (death first)

  $\Rightarrow$ Birth occurs before death with probability $\dfrac{\lambda_i}{\lambda_i + \mu_i} = P_{i,i+1}$

  $\Rightarrow$ Death occurs before birth with probability $\dfrac{\mu_i}{\lambda_i + \mu_i} = P_{i,i-1}$

- Leave state 0 only if a birth occurs, then

$$\nu_0 = \lambda_0, \qquad P_{01} = 1$$

  $\Rightarrow$ If CTMC leaves 0, goes to 1 with probability 1

  $\Rightarrow$ Might not leave 0 if $\lambda_0 = 0$ (e.g., to model extinction)

Markov process

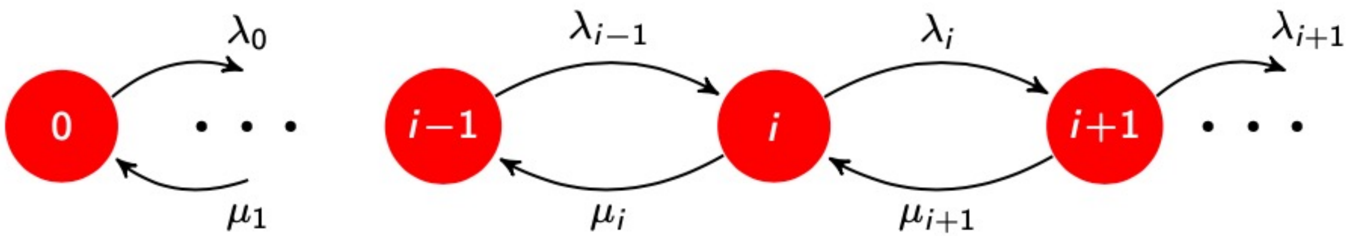
- Rate of transition from i to $i+1$ is (recall definition $q_{ij} = \nu_i P_{ij}$)

$$q_{i,i+1} = \nu_i P_{i,i+1} = (\lambda_i + \mu_i)\frac{\lambda_i}{\lambda_i + \mu_i} = \lambda_i$$

- Likewise, rate of transition from i to $i-1$ is

$$q_{i,i-1} = \nu_i P_{i,i-1} = (\lambda_i + \mu_i)\frac{\mu_i}{\lambda_i + \mu_i} = \mu_i$$

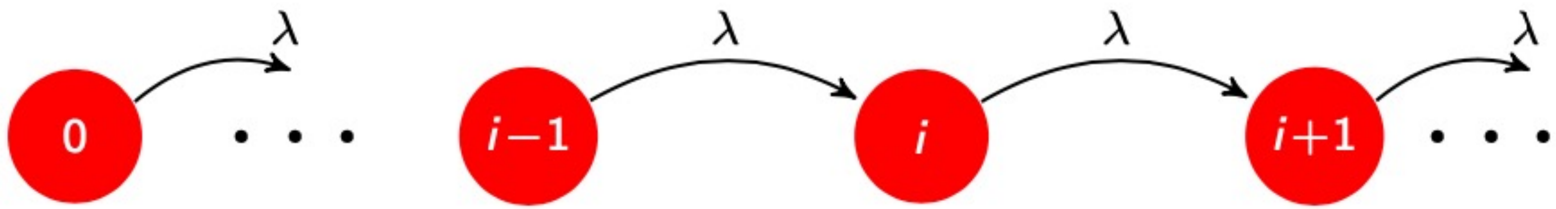
- For $i = 0 \Rightarrow q_{01} = \nu_0 P_{01} = \lambda_0$



- Somewhat more natural representation. Similar to discrete-time MCs

Markov process

- A Poisson process is a BD process with $\lambda_i = \lambda$ and $\mu_i = 0$ constant
- State $N(t)$ counts the total number of events (arrivals) by time t
  - $\Rightarrow$ Arrivals occur a rate of λ per unit time
  - $\Rightarrow$ Transition times are the i.i.d. exponential interarrival times



- The Poisson process is a CTMC