

CS-E5885 Modeling biological networks

Markov processes in discrete and continuous time

Harri Lähdesmäki

Department of Computer Science
Aalto University

January 14, 2022

Contents

- ▶ Markov chains: finite-state, discrete-time
 - ▶ Definitions
 - ▶ Chapman-Kolmogorov equations
 - ▶ Stationary distribution
 - ▶ Simulation
- ▶ Continuous-time Markov chains: finite-state
 - ▶ Definitions
 - ▶ Chapman-Kolmogorov equations
 - ▶ Transition rate matrix
 - ▶ Stationary distribution
 - ▶ Kolmogorov equations
 - ▶ Simulation
 - ▶ Poisson process
 - ▶ Countable state-space
- ▶ Poisson process
- ▶ Reading (see references at the end):
 - ▶ Sections 5.1–5.4 from (Wilkinson, 2011)

Introduction

- ▶ A **random variable** is a function which acts on elements of the sample space (outcomes), e.g.
 - ▶ Discrete random variables
 - ▶ Continuous random variables
- ▶ A **stochastic process** is a random variable which evolves over time
 - ▶ Stochastic processes can have a discrete or continuous state, and be defined in discrete or continuous time
- ▶ A **Markov process** is a stochastic process whose future behavior depends only on the current state, not on the past
 - ▶ Markov processes are relatively easy to handle theoretically and in computer simulations
 - ▶ Dynamic behavior of many biochemical networks can be modeled by Markov processes

Markov chains

- ▶ A discrete time stochastic process is a set random variables

$$\{\theta^{(n)} | n = 1, 2, \dots\}$$

- ▶ Markov chains: finite-state and discrete time

Markov chains

- A discrete time stochastic process is a set random variables

$$\{\theta^{(n)} | n = 1, 2, \dots\}$$

- Markov chains: finite-state and discrete time
- A first-order Markov chain with a (finite) state space $S = \{1, 2, \dots, r\}$ has the property that, given the present state x_n , future does not depend on the past

$$P(\underbrace{\theta^{(n+1)} = x}_{\text{next state}} | \underbrace{\theta^{(n)} = x_n}_{\text{present state}}, \underbrace{\theta^{(n-1)} = x_{n-1}, \dots, \theta^{(0)} = x_0}_{\text{earlier states}})$$
$$= P(\theta^{(n+1)} = x | \theta^{(n)} = x_n), \quad \forall, x_n, \dots, x_0$$

Markov chains

- A discrete time stochastic process is a set random variables

$$\{\theta^{(n)} | n = 1, 2, \dots\}$$

- Markov chains: finite-state and discrete time
- A first-order Markov chain with a (finite) state space $S = \{1, 2, \dots, r\}$ has the property that, given the present state x_n , future does not depend on the past

$$P(\underbrace{\theta^{(n+1)} = x}_{\text{next state}} | \underbrace{\theta^{(n)} = x_n}_{\text{present state}}, \underbrace{\theta^{(n-1)} = x_{n-1}, \dots, \theta^{(0)} = x_0}_{\text{earlier states}})$$
$$= P(\theta^{(n+1)} = x | \theta^{(n)} = x_n), \quad \forall, x_n, \dots, x_0$$

- In general, the **transition probability** depends on n, x, x_n
- If there is no dependency on time n , the chain is said to be **homogeneous** and the **transition kernel** is written

$$P(\theta^{(n+1)} = x | \theta^{(n)} = x_n) = P(x_n, x)$$

Markov chains (2)

- ▶ For finite-state, discrete-time Markov chains, the transition kernel can be represented as a square matrix
 - ▶ Each element defines the probability of moving from state x_i to state x_j : $P(x_i, x_j)$

Markov chains (3)

► An example

$$S = \{x_1, x_2\}$$



$$P(\theta^{(n+1)} = x_1 \mid \theta^{(n)} = x_1) = p_{11} = 0.9$$

$$P(\theta^{(n+1)} = x_2 \mid \theta^{(n)} = x_1) = p_{12} = 0.1$$

⋮

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

Markov chains (4)

- ▶ Definition: a square matrix with non-negative real elements is said to be a **stochastic matrix** if its rows sum to one
- ▶ Transition kernel of a homogeneous finite-state Markov chain (when represented as a square matrix) is obviously a stochastic matrix

Markov chains (4)

- ▶ Definition: a square matrix with non-negative real elements is said to be a **stochastic matrix** if its rows sum to one
- ▶ Transition kernel of a homogeneous finite-state Markov chain (when represented as a square matrix) is obviously a stochastic matrix
- ▶ Propositions:
 - ▶ The product of two stochastic matrices is a stochastic matrix
 - ▶ Every **eigenvalue** of a stochastic matrix satisfy $\lambda \leq 1$
 - ▶ Every stochastic matrix has at least one eigenvalue equal to one

Propagation of the distribution

- ▶ Assume r states, $S = \{x_1, \dots, x_r\}$
- ▶ Assume we know the probabilities of being in any of the state x_i at time n : $P(\theta^{(n)} = x_i)$
- ▶ We can compute the probability that after one time step (from time n to time $n + 1$) the system transitions to any of the states x_1, \dots, x_r
- ▶ This is called the **propagation of the distribution** over one time step

Propagation of the distribution

- ▶ Assume r states, $S = \{x_1, \dots, x_r\}$
- ▶ Assume we know the probabilities of being in any of the state x_i at time n : $P(\theta^{(n)} = x_i)$
- ▶ We can compute the probability that after one time step (from time n to time $n+1$) the system transitions to any of the states x_1, \dots, x_r
- ▶ This is called the **propagation of the distribution** over one time step

$$P(\theta^{(n+1)} = x_1) = P(\theta^{(n)} = x_1)P(x_1, x_1) + P(\theta^{(n)} = x_2)P(x_2, x_1) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_1)$$

Propagation of the distribution

- ▶ Assume r states, $S = \{x_1, \dots, x_r\}$
- ▶ Assume we know the probabilities of being in any of the state x_i at time n : $P(\theta^{(n)} = x_i)$
- ▶ We can compute the probability that after one time step (from time n to time $n+1$) the system transitions to any of the states x_1, \dots, x_r
- ▶ This is called the **propagation of the distribution** over one time step

$$P(\theta^{(n+1)} = x_1) = P(\theta^{(n)} = x_1)P(x_1, x_1) + P(\theta^{(n)} = x_2)P(x_2, x_1) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_1)$$

⋮

$$P(\theta^{(n+1)} = x_i) = P(\theta^{(n)} = x_1)P(x_1, x_i) + P(\theta^{(n)} = x_2)P(x_2, x_i) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_i)$$

Propagation of the distribution

- ▶ Assume r states, $S = \{x_1, \dots, x_r\}$
- ▶ Assume we know the probabilities of being in any of the state x_i at time n : $P(\theta^{(n)} = x_i)$
- ▶ We can compute the probability that after one time step (from time n to time $n+1$) the system transitions to any of the states x_1, \dots, x_r
- ▶ This is called the **propagation of the distribution** over one time step

$$P(\theta^{(n+1)} = x_1) = P(\theta^{(n)} = x_1)P(x_1, x_1) + P(\theta^{(n)} = x_2)P(x_2, x_1) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_1)$$

⋮

$$P(\theta^{(n+1)} = x_i) = P(\theta^{(n)} = x_1)P(x_1, x_i) + P(\theta^{(n)} = x_2)P(x_2, x_i) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_i)$$

⋮

$$P(\theta^{(n+1)} = x_r) = P(\theta^{(n)} = x_1)P(x_1, x_r) + P(\theta^{(n)} = x_2)P(x_2, x_r) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_r)$$

Propagation of the distribution

- ▶ Assume r states, $S = \{x_1, \dots, x_r\}$
- ▶ Assume we know the probabilities of being in any of the state x_i at time n : $P(\theta^{(n)} = x_i)$
- ▶ We can compute the probability that after one time step (from time n to time $n+1$) the system transitions to any of the states x_1, \dots, x_r
- ▶ This is called the **propagation of the distribution** over one time step

$$P(\theta^{(n+1)} = x_1) = P(\theta^{(n)} = x_1)P(x_1, x_1) + P(\theta^{(n)} = x_2)P(x_2, x_1) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_1)$$

⋮

$$P(\theta^{(n+1)} = x_i) = P(\theta^{(n)} = x_1)P(x_1, x_i) + P(\theta^{(n)} = x_2)P(x_2, x_i) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_i)$$

⋮

$$P(\theta^{(n+1)} = x_r) = P(\theta^{(n)} = x_1)P(x_1, x_r) + P(\theta^{(n)} = x_2)P(x_2, x_r) + \dots + P(\theta^{(n)} = x_r)P(x_r, x_r)$$

- ▶ Let $\pi^{(n)} = (P(\theta^{(n)} = x_1), \dots, P(\theta^{(n)} = x_r))$ then the above equations can be written as

$$\pi^{(n+1)} = \pi^{(n)} P$$

Propagation of the distribution (2)

- ▶ Write $\pi^{(n+1)} = \pi^{(n)}P$ for $n = 0$ and $n = 1$:

$$\pi^{(1)} = \pi^{(0)}P \text{ and } \pi^{(2)} = \pi^{(1)}P$$

- ▶ Substitute $\pi^{(1)}$ in $\pi^{(2)} = \pi^{(1)}P$ to get

$$\pi^{(2)} = \pi^{(1)}P = \pi^{(0)}PP = \pi^{(0)}P^2$$

Propagation of the distribution (2)

- ▶ Write $\pi^{(n+1)} = \pi^{(n)}P$ for $n = 0$ and $n = 1$:

$$\pi^{(1)} = \pi^{(0)}P \text{ and } \pi^{(2)} = \pi^{(1)}P$$

- ▶ Substitute $\pi^{(1)}$ in $\pi^{(2)} = \pi^{(1)}P$ to get

$$\pi^{(2)} = \pi^{(1)}P = \pi^{(0)}PP = \pi^{(0)}P^2$$

- ▶ By induction we get

$$\pi^{(n)} = \pi^{(0)}P^n,$$

where $P^n = \underbrace{PP \dots P}_{n \text{ times}}$ defines the transition kernel for n time steps

Chapman-Kolmogorov equations

- ▶ Chapman-Kolmogorov equations: because $P^{m+n} = \underbrace{PP\ldots P}_{m+n \text{ times}} = P^m P^n$ (m and n step probs.)

$$\pi^{(m+n)} = \pi^{(0)} P^{m+n} = \pi^{(0)} P^m P^n$$

Chapman-Kolmogorov equations

- ▶ Chapman-Kolmogorov equations: because $P^{m+n} = \underbrace{PP\ldots P}_{m+n \text{ times}} = P^m P^n$ (m and n step probs.)

$$\pi^{(m+n)} = \pi^{(0)} P^{m+n} = \pi^{(0)} P^m P^n$$

- ▶ The elements of transition kernel P^{m+n} (for $m + n$ time steps) are

$$\begin{aligned} p_{ij}(m+n) &= P(\theta^{(m+n)} = j | \theta^{(0)} = i) \\ &= \sum_{k=1}^r P(\theta^{(m+n)} = j | \theta^{(m)} = k) P(\theta^{(m)} = k | \theta^{(0)} = i) \\ &= \sum_{k=1}^r p_{kj}(n) p_{ik}(m) \end{aligned}$$

- ▶ The probability of moving from state i to state j in $m + n$ steps

Stationary distribution

- ▶ A distribution is said to be a **stationary distribution** of the homogeneous Markov chain if

$$\pi = \pi P$$

- ▶ Loosely speaking, stationary distribution describes the distribution of the state x in a long-run (after a sufficiently long time $n \rightarrow \infty$) when π does not change anymore, i.e., $\pi^{(n+1)} = \pi^{(n)}$
- ▶ Stationary distribution is a **row eigenvector** of P corresponding to eigenvalue one

Stationary distribution

- ▶ A distribution is said to be a **stationary distribution** of the homogeneous Markov chain if

$$\pi = \pi P$$

- ▶ Loosely speaking, stationary distribution describes the distribution of the state x in a long-run (after a sufficiently long time $n \rightarrow \infty$) when π does not change anymore, i.e., $\pi^{(n+1)} = \pi^{(n)}$
- ▶ Stationary distribution is a **row eigenvector** of P corresponding to eigenvalue one
- ▶ Stationary distribution can be found by solving

$$\pi(I - P) = \mathbf{0}$$

- ▶ Always infinitely many solutions
 - ▶ Positive solution together with the unit-sum constraint
- ▶ Unique solution if stochastic matrix has only one eigenvalue equal to one

Simulation of Markov chains

- ▶ Given an initial distribution $\pi^{(0)}$ and a stochastic matrix P
 1. Sample an initial state $\theta^{(0)}$ from $\pi^{(0)}$ using the look-up method
 2. Given $\theta^{(0)}$, sample $\theta^{(1)}$ from the $\theta^{(0)}$:th row of P using the look-up method
 3. Continue for $\theta^{(t)}$, $t = 2, 3, \dots$

An illustration of Markov chains

Figure: Example Figure 5.2 from (Wilkinson, 2011)

Continuous time Markov chains

- ▶ Continuous time Markov chains
 - ▶ Previously we looked at Markov chains in discrete time
 - ▶ Biochemical processes evolve continuously in time

Definitions

- ▶ Consider a finite state space $S = \{1, \dots, r\}$, $X \in S$ and continuous time $t \in \mathbb{R}_+$
- ▶ Markov property in continuous time, where dt is infinitesimal time increment

$$P(\underbrace{X(t+dt) = x}_{\text{next state}} | \underbrace{\{X(t) = x(t) | t \in [0, t]\}}_{\text{history}}) = P(\underbrace{X(t+dt) = x}_{\text{next state}} | \underbrace{X(t) = x(t)}_{\text{current state}})$$

Definitions

- ▶ Consider a finite state space $S = \{1, \dots, r\}$, $X \in S$ and continuous time $t \in \mathbb{R}_+$
- ▶ Markov property in continuous time, where dt is infinitesimal time increment

$$P(\underbrace{X(t+dt) = x}_{\text{next state}} | \underbrace{\{X(t) = x(t) | t \in [0, t]\}}_{\text{history}}) = P(\underbrace{X(t+dt) = x}_{\text{next state}} | \underbrace{X(t) = x(t)}_{\text{current state}})$$

- ▶ Given the current state x at time t , future behavior can be characterized by a transition kernel

$$p(x, t, x', t') = P(X(t+t') = x' | X(t) = x)$$

Definitions

- ▶ Consider a finite state space $S = \{1, \dots, r\}$, $X \in S$ and continuous time $t \in \mathbb{R}_+$
- ▶ Markov property in continuous time, where dt is infinitesimal time increment

$$P(\underbrace{X(t+dt) = x}_{\text{next state}} | \underbrace{\{X(t) = x(t) | t \in [0, t]\}}_{\text{history}}) = P(\underbrace{X(t+dt) = x}_{\text{next state}} | \underbrace{X(t) = x(t)}_{\text{current state}})$$

- ▶ Given the current state x at time t , future behavior can be characterized by a transition kernel

$$p(x, t, x', t') = P(X(t+t') = x' | X(t) = x)$$

- ▶ If no dependency on time t , then the process is said to be homogeneous
- ▶ For each value of t' , the kernel can be written as an $r \times r$ stochastic matrix $P(t')$
 - ▶ Conceptually, that would correspond to a discrete-time Markov chain where the abstract unit time increment corresponds to an absolute time increment of t'
- ▶ Clearly, $P(0) = I$ as no transition can take place within zero time

Transition rate matrix

- ▶ Define transition rate matrix to be the derivative of $P(t)$ at time $t = 0$, i.e.,

$$\begin{aligned} Q &= \frac{d}{dt'} P(t')|_{t'=0} \\ &= \lim_{\delta t \rightarrow 0} \frac{P(\delta t) - P(0)}{\delta t} \\ &= \lim_{\delta t \rightarrow 0} \frac{P(\delta t) - I}{\delta t} \end{aligned}$$

- ▶ Q defines the **hazards** of moving to different states

Transition rate matrix

- ▶ Define transition rate matrix to be the derivative of $P(t)$ at time $t = 0$, i.e.,

$$\begin{aligned} Q &= \frac{d}{dt'} P(t')|_{t'=0} \\ &= \lim_{\delta t \rightarrow 0} \frac{P(\delta t) - P(0)}{\delta t} \\ &= \lim_{\delta t \rightarrow 0} \frac{P(\delta t) - I}{\delta t} \end{aligned}$$

- ▶ Q defines the **hazards** of moving to different states
- ▶ Rearranging gives infinitesimal transition matrix

$$P(dt) = I + Qdt$$

Transition rate matrix (2)

Because $P(dt)$ is a stochastic matrix (non-negative elements and rows sum to 1), any Q matrix must satisfy

- ▶ Off-diagonal elements of $P(dt)$ and Qdt are the same
- ▶ Rows sum to 0
- ▶ Diagonal elements of Qdt are non-positive, non-diagonal elements are positive
- ▶ $-q_{ii}$ gives the combined hazard of moving away from state i

$$-q_{ii} = \sum_{j \neq i} q_{ij}$$

$$P(dt) = I + Qdt$$

$$= \begin{pmatrix} 1 + q_{11}dt & q_{12}dt & \dots & q_{1r}dt \\ q_{21}dt & 1 + q_{22}dt & \dots & q_{2r}dt \\ \vdots & \vdots & \ddots & \vdots \\ q_{r1}dt & q_{r2}dt & \dots & 1 + q_{rr}dt \end{pmatrix}$$

$$\Rightarrow \begin{cases} -1 \leq q_{ii}dt \leq 0 \\ 0 \leq q_{ij}dt \leq 1, \quad i \neq j \\ \sum_{j=1}^r q_{ij} = 0 \end{cases}$$

Stationary distribution

- ▶ Transition rate matrix gives us a way to compute the stationary distribution of the continuous-time Markov chain

$$\begin{aligned}\pi P(dt) &= \pi \\ \pi(I + Qdt) &= \pi \\ \pi Q &= \mathbf{0}\end{aligned}$$

subject to unit-sum constraint

A single prokaryotic gene example

Figure: An example from p. 139 from (Wilkinson, 2011)

A single prokaryotic gene example (2)

Figure: an example from Wilkinson's book, page 124

Chapman-Kolmogorov equations

- ▶ Because $P(t)$ is a stochastic transition matrix (for a fixed t) we can multiply these to get a combined transition, in particular

$$\begin{aligned} P(t + t') &= P(t')P(t) \\ &= P(t)P(t') \end{aligned}$$

- ▶ Kolmogorov's backward equations
- ▶ Kolmogorov's forward equations

Kolmogorov's backward equations

- ▶ How to obtain the stochastic matrix $P(t)$ for any finite t ? Start by writing the time derivative

$$\begin{aligned}\frac{d}{dt}P(t) &= \lim_{dt \rightarrow 0} \frac{P(t + dt) - P(t)}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{P(dt)P(t) - P(t)}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{P(dt) - I}{dt} P(t) \\ &= QP(t)\end{aligned}$$

- ▶ Solve the matrix differential equation

Kolmogorov's backward equations

- ▶ How to obtain the stochastic matrix $P(t)$ for any finite t ? Start by writing the time derivative

$$\begin{aligned}\frac{d}{dt}P(t) &= \lim_{dt \rightarrow 0} \frac{P(t + dt) - P(t)}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{P(dt)P(t) - P(t)}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{P(dt) - I}{dt} P(t) \\ &= QP(t)\end{aligned}$$

- ▶ Solve the matrix differential equation
- ▶ It is more intuitive to look at the component form of the above matrix differential equation

$$\frac{d}{dt}p(i, j, t) = \sum_{k=1}^r q_{ik}p(k, j, t) \quad i, j = 1, 2, \dots, r$$

- ▶ These are known as Kolmogorov's backward equations

Kolmogorov's forward equations

- ▶ Alternatively, one can expand $P(t + dt)$ as $P(t)P(dt)$ (instead of $P(dt)P(t)$) to get

$$\frac{d}{dt}P(t) = P(t)Q$$

which in component form equals

$$\frac{d}{dt}p(i, j, t) = \sum_{k=1}^r p(i, k, t)q_{kj} \quad i, j = 1, 2, \dots, r$$

- ▶ These are known as Kolmogorov's forward equations

Stochastic simulation

- ▶ Two ways to simulate continuous-time discrete-state Markov process
 1. A finite time-discretization of the infinitesimal transition matrix: approximate

$$P(dt) = I + Qdt$$

as (for small Δt)

$$P(\Delta t) \simeq I + Q\Delta t$$

results in discrete-time Markov chain that assumes time points $0, \Delta t, 2\Delta t, 3\Delta t \dots$ that are fixed a priori

2. Exact simulation of each transition event: so called **discrete event simulation**

Discrete event simulation

- ▶ Changes in the system state correspond to a (discrete) sequence of events in time
 - ▶ Each event occurs at a particular instant in time and marks a change of state in the system
 - ▶ No change in the system between consecutive events
- ▶ How it works?
 - ▶ Assume being in state i at time t
 - ▶ Compute the probability that
 - ▶ The next event will be in time interval $(t + t', t + t' + dt]$, and
 - ▶ Consists of a move from state i to state j
 - ▶ Sample the probability to get the event time t' and the next state j
 - ▶ Update time $t := t + t'$ and set state to j

Discrete event simulation (2)

- ▶ Assume being in state i at time t
- ▶ Compute the probability $f(t', j|t, i)dt$ that
 - ▶ The next event will be in time interval $(t + t', t + t' + dt]$, and
 - ▶ Consists of a move from state i to state j

$$\begin{aligned} f(t', j|t, i)dt &= P(\text{next event in } (t + t', t + t' + dt] | t, i) \\ &\quad \times P(j | \text{next event in } (t + t', t + t' + dt], t, i) \end{aligned}$$

Discrete event simulation (2)

- ▶ Assume being in state i at time t
- ▶ Compute the probability $f(t', j|t, i)dt$ that
 - ▶ The next event will be in time interval $(t + t', t + t' + dt]$, and
 - ▶ Consists of a move from state i to state j

$$\begin{aligned} f(t', j|t, i)dt &= P(\text{next event in } (t + t', t + t' + dt] | t, i) \\ &\quad \times P(j | \text{next event in } (t + t', t + t' + dt], t, i) \end{aligned}$$

- ▶ Time to the next transition has exponential distribution (explained a bit later)

$$P(\text{next event in } (t + t', t + t' + dt] | t, i) = \text{Exp}(-q_{ii}) = -q_{ii}e^{-q_{ii}t'}$$

Discrete event simulation (3)

- ▶ Probability of the move from state i to state $j \neq i$

$$\begin{aligned} & P(\underbrace{X(t + t' + dt) = j}_{A} | \underbrace{[X(t + t') = i]}_{B} \cap \underbrace{[X(t + t' + dt) \neq i]}_{C}) \\ &= \frac{P(A, C|B)}{P(C|B)} = \frac{P(X(t + t' + dt) = j | X(t + t') = i)}{P(X(t + t' + dt) \neq i | X(t + t') = i)} \\ &= \frac{q_{ij} dt}{\sum_{k \neq i} q_{ik} dt} = \frac{q_{ij}}{-q_{ii}} \end{aligned}$$

Discrete event simulation (3)

- ▶ Probability of the move from state i to state $j \neq i$

$$\begin{aligned} & P(\underbrace{X(t + t' + dt) = j}_{A} | \underbrace{[X(t + t') = i]}_{B} \cap \underbrace{[X(t + t' + dt) \neq i]}_{C}) \\ &= \frac{P(A, C|B)}{P(C|B)} = \frac{P(X(t + t' + dt) = j | X(t + t') = i)}{P(X(t + t' + dt) \neq i | X(t + t') = i)} \\ &= \frac{q_{ij} dt}{\sum_{k \neq i} q_{ik} dt} = \frac{q_{ij}}{-q_{ii}} \end{aligned}$$

- ▶ Thus

$$\begin{aligned} f(t', j | t, i) dt &= -q_{ii} e^{-q_{ii} t'} \times \frac{q_{ij}}{-q_{ii}} \\ &= q_{ij} e^{-q_{ii} t'} \end{aligned}$$

- ▶ No j dependency in the probability of t'
- ▶ No t' dependency in the probability of j
- ▶ Thus, random variables j and t' are independent of each other

Standard discrete event simulation algorithm

1. Initialize the process at time $t = 0$ with an initial value i
2. Call current state i . Simulate the time to the next event, t' , as an $\text{Exp}(-q_{ii})$ random quantity
3. Update time $t := t + t'$
4. Simulate new state j as a discrete random quantity with probability mass function (PMF)
 $-q_{ik}/q_{ii}$, $k \neq i$, i.e.,

$$(-q_{i1}/q_{ii}, -q_{i2}/q_{ii}, \dots, -q_{i(i-1)}/q_{ii}, -q_{i(i+1)}/q_{ii}, \dots, -q_{ir}/q_{ii})$$

5. Output the time t and state j
 6. Continue with step 2 if $t < T_{\max}$
- ▶ Simulation gives a single realization of the random process
 - ▶ Simulation only needs the transition rate matrix Q

Countable state space

- ▶ The above construction does not work directly in the case of countably infinite state space, i.e., infinite transition matrices
- ▶ An immigration-death process example from the course book

Countable state space example

Figure: An example from p. 143 from (Wilkinson, 2011)

Countable state space example (2)

Figure: An example from p. 144 from (Wilkinson, 2011)

Countable state space example (3)

Figure: Figure 5.7 from (Wilkinson 2011)

Poisson process

- ▶ Consider a sequence of events/observations that take place at random times $0 < T_1 < T_2 < \dots$
- ▶ Let $N(t')$ denote the number of observations up to time t'
 - ▶ The number of observations at time $t' = 0$ is 0, $N(0) = 0$
- ▶ The number of observations between $(t', t' + t]$ is $X_t = N(t' + t) - N(t')$

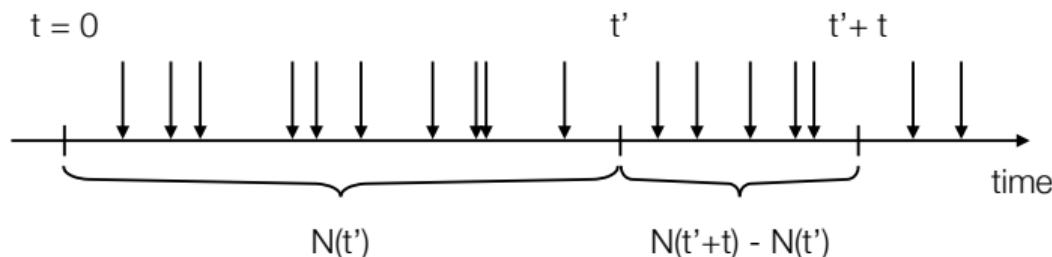


Figure: An illustration: a sequence of events and counting process

Poisson process (2)

- ▶ E.g. the number of radioactive particles/isotopes detected in a time interval
- ▶ Events take place with so-called fixed rate λ
- ▶ In our application, these events correspond to transitions in a continuous time Markov chain or reactions in our coupled chemical reaction model

Poisson process (3)

- ▶ This **counting process** $\{N(t), t \geq 0\}$ is said to follow a homogeneous Poisson process with fixed rate λ if the number of observations, $X_t = N(t' + t) - N(t')$, in any interval of length t is such that

$$X \sim Po(\lambda t)$$

where $Po(\cdot)$ denotes the Poisson distribution

$$P(X = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

and the number of events in disjoint time intervals are independent

Poisson process (4)

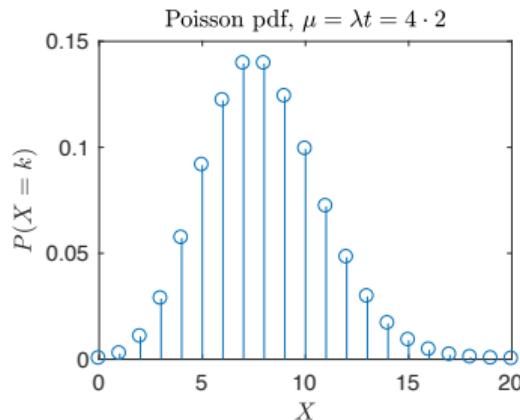
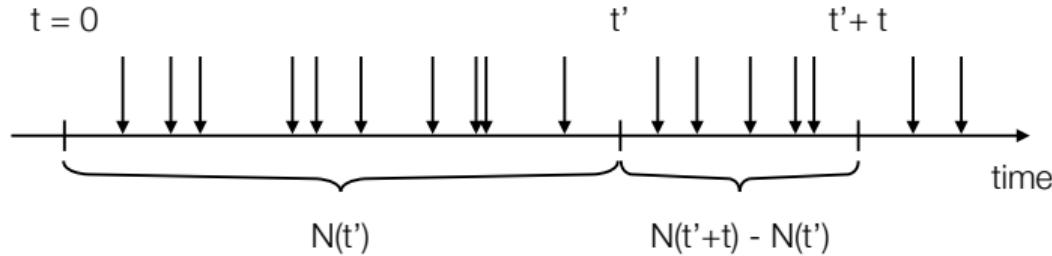


Figure: An illustration of the Poisson process and Poisson PDF

Poisson process and exponential distribution

- ▶ Theorem¹: consider a Poisson process with a fixed rate λ . Let T be the time to the **first** event / observation. Then $T \sim \text{Exp}(\lambda)$

¹Theorem (3.17) in (Wilkinson, 2011)

Poisson process and exponential distribution

- ▶ Theorem¹: consider a Poisson process with a fixed rate λ . Let T be the time to the **first** event / observation. Then $T \sim \text{Exp}(\lambda)$
- ▶ Recall that PDF $f_X(x)$ and CDF $F_X(x)$ of an exponentially distributed random variable X are

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

¹Theorem (3.17) in (Wilkinson, 2011)

Poisson process and exponential distribution

- ▶ Theorem¹: consider a Poisson process with a fixed rate λ . Let T be the time to the first event / observation. Then $T \sim Exp(\lambda)$
- ▶ Recall that PDF $f_X(x)$ and CDF $F_X(x)$ of an exponentially distributed random variable X are

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

- ▶ Let N_t be the number of events in the interval $(0, t]$, then $N_t \sim Po(\lambda t)$. The CDF of T is

$$\begin{aligned} F_T(t) &= P(T \leq t) \\ &= 1 - P(T > t) \\ &= 1 - P(N_t = 0) \\ &= 1 - \frac{(\lambda t)^0}{0!} e^{-\lambda t} \\ &= 1 - e^{-\lambda t} \quad (\text{CDF of } Exp(\lambda)) \end{aligned}$$

¹Theorem (3.17) in (Wilkinson, 2011)

Poisson process and exponential distribution (2)

- ▶ Time to the first event of a Poisson process is exponentially distributed
- Time between any two events has the same exponential distribution (independence property)

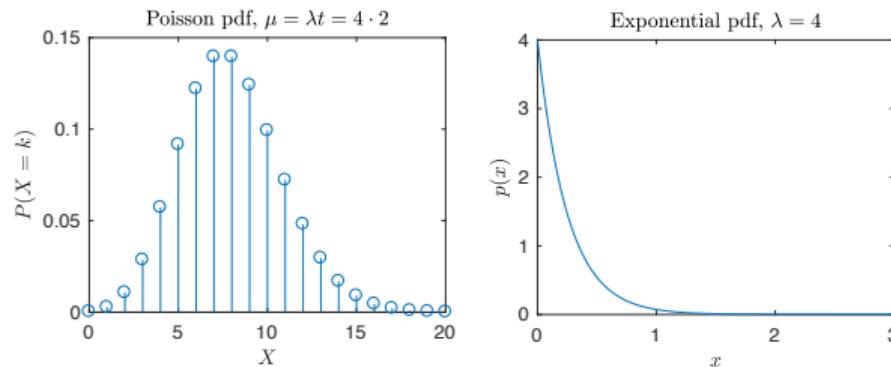
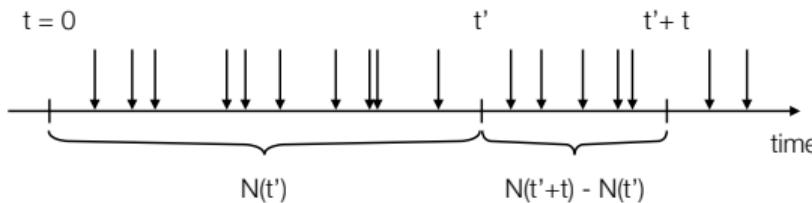


Figure: An illustration of the Poisson process and Poisson and exponential PDFs

Poisson process and exponential distribution (3)

- ▶ Exponential distribution can be reinterpreted as the time to an event of constant hazard λ
- ▶ An alternative (but equivalent) definition of Poisson process states that for infinitesimal time increment dt

$$P(T \leq dt) = \lambda dt$$

- ▶ The hazard defines the “event density” on the time axis

Poisson process and exponential distribution (3)

- ▶ Exponential distribution can be reinterpreted as the time to an event of constant hazard λ
- ▶ An alternative (but equivalent) definition of Poisson process states that for infinitesimal time increment dt

$$P(T \leq dt) = \lambda dt$$

- ▶ The hazard defines the “event density” on the time axis
- ▶ Due to the independence property of the Poisson process the probability $P(T \leq dt) = \lambda dt$ holds for any time interval of length dt

Poisson process and exponential distribution (4)

- Theorem²: if $T_i \sim Exp(\lambda_i)$, $i = 1, \dots, v$, are independent, then

$$T_0 = \min_i \{T_i\} \sim Exp \left(\sum_{i=1}^v \lambda_i \right)$$

²Theorem (3.18) in (Wilkinson, 2011)

Poisson process and exponential distribution (4)

- Theorem²: if $T_i \sim Exp(\lambda_i)$, $i = 1, \dots, v$, are independent, then

$$T_0 = \min_i \{T_i\} \sim Exp\left(\sum_{i=1}^v \lambda_i\right)$$

- Note that for $T \sim Exp(\lambda)$ we have $P(T > t) = e^{-\lambda t}$ and

$$\begin{aligned} P(T_0 > t) &= P(\min_i \{T_i\} > t) \\ &= P(\{T_1 > t\} \cap \dots \cap \{T_v > t\}) \\ &= \prod_{i=1}^v P(T_i > t) = \prod_{i=1}^v e^{-\lambda_i t} \\ &= e^{-(\sum_{i=1}^v \lambda_i)t} \end{aligned}$$

$$\text{so } P(T_0 \leq t) = 1 - P(T_0 > t) = 1 - e^{-(\sum_{i=1}^v \lambda_i)t}$$

²Theorem (3.18) in (Wilkinson, 2011)

Poisson process and exponential distribution (5)

- Related to theorem on the minimum of v independent exponential random variables (previous slide), we can also show³ that the probability that the j th random variable (T_j) is the smallest is

$$\pi_j = \frac{\lambda_j}{\sum_{i=1}^v \lambda_i}$$

³Proposition (3.19) in (Wilkinson, 2011)

References

- ▶ Darren J. Wilkinson, *Stochastic Modelling for Systems Biology*, Chapman & Hall/CRC, 2011