6. Introduction to stochastic processes

Contents

Stochastic processes (1)

- Consider a teletraffic (or any) system
- It typically **evolves** in time **randomly**
  - Example 1: the number of occupied channels in a telephone link at time t or at the arrival time of the n^{th} customer
  - Example 2: the number of packets in the buffer of a statistical multiplexer at time t or at the arrival time of the n^{th} customer
- This kind of evolution is described by a stochastic process
  - At any individual time t (or n) the system can be described by a random variable
  - Thus, the stochastic process is a collection of random variables

Stochastic processes (2)

- **Definition**: A (real-valued) **stochastic process** $X = (X_t \mid t \in I)$ is a collection of random variables X_t
  - taking values in some (real-valued) set S, $X_t(\omega) \in S$, and
  - indexed by a real-valued (time) parameter $t \in I$.
  - Stochastic processes are also called **random processes** (or just **processes)**
- The index set $I \subset \Re$ is called the **parameter space** of the process
- The value set $S \subset \Re$ is called the **state space** of the process
  - Note: Sometimes notation X_t is used to refer to the whole stochastic process (instead of a single random variable)

Stochastic processes (3)

- Each (individual) random variable X_t is a mapping from the sample space Ω into the real values $\Re$:

$$X_t : \Omega \to \Re, \quad \omega \mapsto X_t(\omega)$$

- Thus, a stochastic process X can be seen as a mapping from the sample space Ω into the set of real-valued functions $\Re^I$ (with $t \in I$ as an argument):

$$X : \Omega \to \Re^I, \quad \omega \mapsto X(\omega)$$

- Each sample point $\omega \in \Omega$ is associated with a real-valued function $X(\omega)$. Function $X(\omega)$ is called a **realization** (or a **path** or a **trajectory**) of the process.

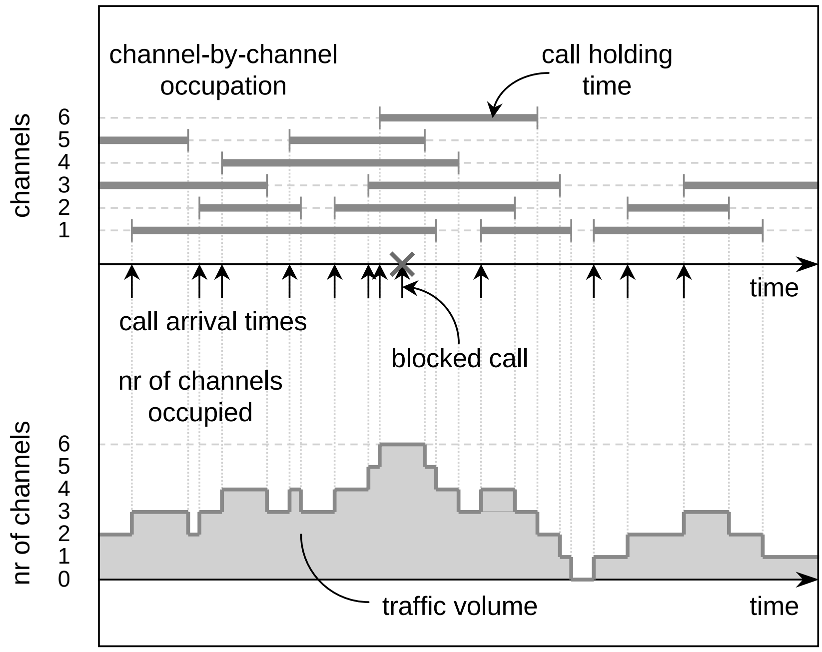
Summary

- Given the sample point $\omega \in \Omega$
  - $X(\omega) = (X_t(\omega) \mid t \in I)$ is a real-valued function (of $t \in I$)
- Given the time index $t \in I$,
  - $X_t = (X_t(\omega) \mid \omega \in \Omega)$ is a random variable (as $\omega \in \Omega$)
- Given the sample point $\omega \in \Omega$ and the time index $t \in I$,
  - $X_t(\omega)$ is a real value

Example

- Consider traffic process $X = (X_t \mid t \in [0,T])$ in a link between two telephone exchanges during some time interval $[0,T]$
  - X_t denotes the number of occupied channels at time t
- Sample point $\omega \in \Omega$ tells us
  - what is the number X_0 of occupied channels at time 0,
  - what are the remaining holding times of the calls going on at time 0,
  - at what times new calls arrive, and
  - what are the holding times of these new calls.
- From this information, it is possible to construct the realization $X(\omega)$ of the traffic process X
- Note that all the randomness is included in the sample point omega
  - Given the sample point, the realization of the process is just a (deterministic) function of time

Traffic process



Categories of stochastic processes

- Reminder:
  - Parameter space: set I of indices $t \in I$
  - State space: set S of values $X_t(\omega) \in S$
- Categories:
  - Based on the parameter space:
    - **Discrete-time processes**: parameter space discrete
    - **Continuous-time processes**: parameter space continuous
  - Based on the state space:
    - **Discrete-state processes**: state space discrete
    - **Continuous-state processes**: state space continuous
- In this course we will concentrate on the discrete-state processes (with either a discrete or a continuous parameter space)
  - Typical processes describe the number of customers in a queueing system (the state space being thus $S = \{0,1,2,...\}$)

Examples

- Discrete-time, discrete-state processes
  - Example 1: the number of occupied channels in a telephone link at the arrival time of the n^{th} customer, $n = 1,2,...$
  - Example 2: the number of packets in the buffer of a statistical multiplexer at the arrival time of the n^{th} customer, $n = 1,2,...$
- Continuous-time, discrete-state processes
  - Example 3: the number of occupied channels in a telephone link at time $t > 0$
  - Example 4: the number of packets in the buffer of a statistical multiplexer at time $t > 0$

Notation

- For a **discrete-time process**,
  - the parameter space is typically the set of positive integers, $I = \{1,2,\ldots\}$
  - Index t is then (often) replaced by n: X_n, $X_n(\omega)$

- For a **continuous-time process**,
  - the parameter space is typically either a finite interval, $I = [0, T]$, or all non-negative real values, $I = [0, \infty)$
  - In this case, index t is (often) written not as a subscript but in parentheses: $X(t)$, $X(t;\omega)$

Distribution

- The **stochastic characterization** of a stochastic process X is made by giving **all** possible **finite-dimensional distributions**

$$P\{X_{t_1} \le x_1, \ldots, X_{t_n} \le x_n\}$$

  where $t_1,\ldots, t_n \in I$, $x_1,\ldots, x_n \in S$ and $n = 1,2,\ldots$
- In general, this is not an easy task because of **dependencies** between the random variables X_t (with different values of time t)

Dependence

- The most simple (but not so interesting) example of a stochastic process is such that all the random variables X_t are **independent** of each other. In this case

$$P\{X_{t_1} \le x_1, ..., X_{t_n} \le x_n\} = P\{X_{t_1} \le x_1\} \cdots P\{X_{t_n} \le x_n\}$$

- The most simple non-trivial example is a **Markov process**. In this case

$$P\{X_{t_1} \le x_1, ..., X_{t_n} \le x_n\} =$$
$$P\{X_{t_1} \le x_1\} \cdot P\{X_{t_2} \le x_2 \mid X_{t_1} \le x_1\} \cdots P\{X_{t_n} \le x_n \mid X_{t_{n-1}} \le x_{n-1}\}$$

- This is related to the so called **Markov property**:
  - Given the current state (of the process), the future (of the process) does not depend on the past (of the process)

Stationarity

- **Definition**: Stochastic process X is **stationary** if all finite-dimensional distributions are invariant to time shifts, that is:

$$P\{X_{t_1+\Delta} \le x_1, \ldots, X_{t_n+\Delta} \le x_n\} = P\{X_{t_1} \le x_1, \ldots, X_{t_n} \le x_n\}$$

  for all Δ, n, $t_1, \ldots, t_n$ and $x_1, \ldots, x_n$
- **Consequence**: By choosing n = 1, we see that all (individual) random variables X_t of a stationary process are identically distributed:

$$P\{X_t \le x\} = F(x)$$

  for all $t \in I$. This is called the **stationary distribution** of the process.

Stochastic processes in teletraffic theory

- In this course (and, more generally, in teletraffic theory) various stochastic processes are needed to describe
  - the arrivals of customers to the system (**arrival process**)
  - the state of the system (**state process**, **traffic process**)

Arrival process

- An arrival process can be described as
  - a **point process** $(\tau_n \mid n = 1,2,...)$ where τ_n tells the arrival time of the n^{th} customer (discrete-time, continuous-state)
    - typically it is assumed that the interarrival times $\tau_n - \tau_{n-1}$ are independent and identically distributed (IID) ⇒ renewal process
    - then it is sufficient to specify the interarrival time distribution
    - exponential IID interarrival times ⇒ Poisson process
  - a **counter process** $(A(t) \mid t \geq 0)$ where $A(t)$ tells the number of arrivals up to time t (continuous-time, discrete-state)
    - non-decreasing: $A(t+\Delta) \geq A(t)$ for all $t,\Delta \geq 0$
    - thus non-stationary!
    - independent and identically distributed (IID) increments $A(t+\Delta) - A(t)$ with Poisson distribution ⇒ Poisson process

State process

- In simple cases
  - the state of the system is described by just an integer
  - e.g. the number $X(t)$ of calls or packets at time t
  - This yields a state process that is continuous-time and discrete-state
- In more complicated cases,
  - the state process is e.g. a vector of integers (cf. loss and queueing network models)
- Now it is reasonable to ask whether the state process is stationary
  - Although the state of the system did not follow the stationary distribution at time 0, in many cases state distribution approaches the stationary distribution as t tends to ∞

Contents

- Basic concepts
- Poisson process
- Markov processes
- Birth-death processes

Bernoulli process

- **Definition**: **Bernoulli process** with success probability p is an infinite series $(X_n \mid n = 1,2,...)$ of independent and identical random experiments of Bernoulli type with success probability p
- Bernoulli process is clearly discrete-time and discrete-state
  - Parameter space: $I = \{1,2,\dots\}$
  - State space: $S = \{0,1\}$
- Finite dimensional distributions (note: X_n's are IID):

$$P\{X_1 = x_1,..., X_n = x_n\} = P\{X_1 = x_1\} \cdots P\{X_n = x_n\}$$
$$= \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} = p^{\sum_i x_i} (1-p)^{n-\sum_i x_i}$$

- Bernoulli process is stationary (stationary distribution: Bernoulli(p))

6. Introduction to stochastic processes

Poisson process (1)

- **Definition 1**: A point process $(\tau_n \mid n = 1,2,...)$ is a **Poisson process** with **intensity** λ if the probability that there is an event during a short time interval $(t, t+h]$ is $\lambda h + o(h)$ independently of the other time intervals
  - τ_n tells the occurrence time of the n^{th} event
  - $o(h)$ refers to any function such that $o(h)/h \to 0$ as $h \to 0$
  - new events happen with a constant intensity λ: $(\lambda h + o(h))/h \to \lambda$
  - Poisson process can be seen as the continuous-time counter-part of a Bernoulli process
- Defined as a point process,
  Poisson process is discrete-time and continuous-state
  - Parameter space: $I = \{1,2,\dots\}$
  - State space: $S = (0, \infty)$

Poisson process (2)

- Consider the interarrival time τ_n - τ_{n-1} between two events ($\tau_0 = 0$)
  - Since the intensity that something happens remains constant λ, the interarrival time distribution is clearly memoryless. On the other hand, we know that this is a property of an exponential distribution.
  - Due to the same reason, different interarrival times are also independent
  - This leads to the following (second) characterization of a Poisson process
- **Definition 2**: A point process ($\tau_n \mid n = 1,2,...$) is a **Poisson process** with **intensity** λ if the interarrival times τ_n - τ_{n-1} are independent and identically distributed (IID) with joint distribution $\mathrm{Exp}(\lambda)$
  - τ_n tells (again) the occurrence time of the n^{th} event

6. Introduction to stochastic processes

Poisson process (3)

- Consider finally the number of events $A(t)$ during time interval $[0,t]$
  - In a Bernoulli process, the number of successes in a fixed interval would follow a binomial distribution. As the “time slice” tends to 0, this approaches a Poisson distribution.
  - On the other hand, since the intensity that something happens remains constant λ, the number of events occurring in disjoint time intervals are clearly independent.
  - This leads to the following (third) characterization of a Poisson process
- **Definition 3**: A counter process ($A(t) \mid t \geq 0$) is a **Poisson process** with **intensity** λ if its increments in disjoint intervals are independent and follow a Poisson distribution as follows:

$$A(t+\Delta) - A(t) \sim \mathrm{Poisson}(\lambda\Delta)$$

Poisson process (4)

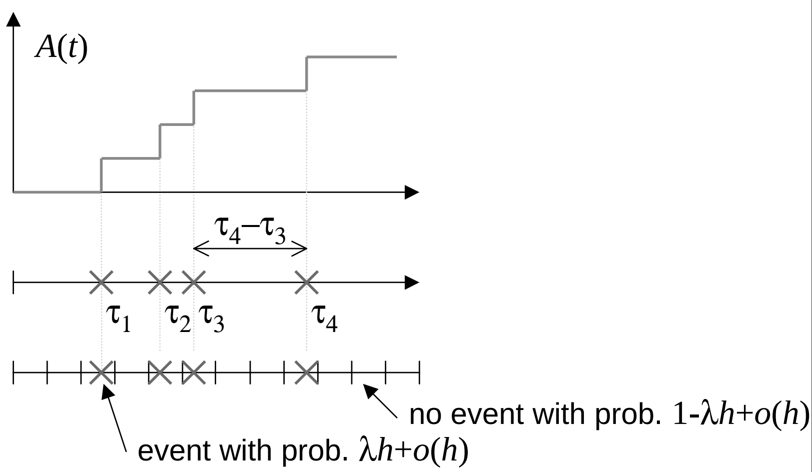
- Defined as a counter process, Poisson process is continuous-time and discrete-state
  - Parameter space: $I = [0, \infty)$
  - State space: $S = \{0,1,2,\dots\}$
- One dimensional distribution: $A(t) \sim \text{Poisson}(\lambda t)$
  - $E[A(t)] = \lambda t$, $D^2[A(t)] = \lambda t$
- Finite dimensional distributions (due to indep. of disjoint intervals):

$$
\begin{aligned}
&P\{A(t_1) = x_1, \dots, A(t_n) = x_n\} = \\
&\quad P\{A(t_1) = x_1\} P\{A(t_2) - A(t_1) = x_2 - x_1\} \cdots \\
&\quad P\{A(t_{n-1}) - A(t_n) = x_n - x_{n-1}\}
\end{aligned}
$$

- No stationary distribution (but identically distributed increments)

Three ways to characterize the Poisson process

- It is possible to show that all three definitions for a Poisson process are, indeed, equivalent



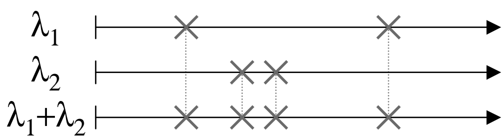
Properties (1)

- **Property 1** (**Sum**): Let $A_1(t)$ and $A_2(t)$ be two independent Poisson processes with intensities λ_1 and λ_2. Then the sum (superposition) process $A_1(t) + A_2(t)$ is a Poisson process with intensity $\lambda_1 + \lambda_2$.
- Proof: Consider a short time interval $(t, t{+}h]$
  - Probability that there are no events in the superposition is

$$(1-\lambda_1 h + o(h))(1-\lambda_2 h + o(h)) = 1-(\lambda_1+\lambda_2)h + o(h)$$

  - On the other hand, the probability that there is exactly one event is

$$(\lambda_1 h + o(h))(1-\lambda_2 h + o(h)) + (1-\lambda_1 h + o(h))(\lambda_2 h + o(h))$$
$$= (\lambda_1+\lambda_2)h + o(h)$$



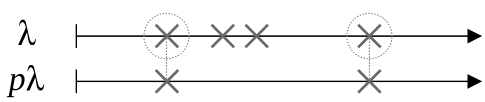
Properties (2)

- **Property 2** (**Random sampling**): Let τ_n be a Poisson process with intensity λ. Denote by σ_n the point process resulting from a random and independent sampling (with probability p) of the points of τ_n. Then σ_n is a Poisson process with intensity $p\lambda$.
- Proof: Consider a short time interval $(t, t{+}h]$
  - Probability that there are no events after the random sampling is

$$(1-\lambda h + o(h)) + (1-p)(\lambda h + o(h)) = 1-p\lambda h + o(h)$$

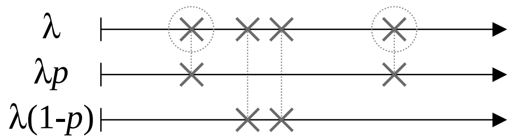
  - On the other hand, the probability that there is exactly one event is

$$p(\lambda h + o(h)) = p\lambda h + o(h)$$



Properties (3)

- **Property 3** (**Random sorting**): Let τ_n be a Poisson process with intensity λ. Denote by $\sigma_n^{(1)}$ the point process resulting from a random and independent sampling (with probability p) of the points of τ_n. Denote by $\sigma_n^{(2)}$ the point process resulting from the remaining points. Then $\sigma_n^{(1)}$ and $\sigma_n^{(2)}$ are independent Poisson processes with intensities λp and $\lambda(1 - p)$.
- Proof: Due to property 2, it is enough to prove that the resulting two processes are independent.



Properties (4)

- **Property 4** (**PASTA**): Consider any simple (and stable) teletraffic model with Poisson arrivals. Let $X(t)$ denote the state of system at time t (continuous-time process) and Y_n denote the state of the system seen by the nth arriving customer (discrete-time process). Then the stationary distribution of $X(t)$ is the same as the stationary distribution of Y_n.
- Thus, we can say that
  - arriving customers see the system in the stationary state
- PASTA property is only valid for Poisson arrivals
  - Consider e.g. your own PC. Whenever you start a new session, the system is idle. In the continuous time, however, the system is not only idle but also busy (when you use it).

Contents

- Basic concepts
- Poisson process
- Markov processes
- Birth-death processes

Markov process

- Consider a continuous-time and discrete-state stochastic process $X(t)$
  - with state space $S = \{0,1,\ldots,N\}$ or $S = \{0,1,...\}$
- **Definition**: The process $X(t)$ is a **Markov process** if

$$P\{X(t_{n+1}) = x_{n+1} \mid X(t_1) = x_1, \ldots, X(t_n) = x_n\} =$$
$$P\{X(t_{n+1}) = x_{n+1} \mid X(t_n) = x_n\}$$

for all n, $t_1 < \ldots < t_{n+1}$ and $x_1, \ldots, x_{n+1}$
  - This is called the Markov property
  - Given the current state, the future of the process does not depend on its past
  - As regards the future of the process, it is important to know the current state (not how the process has evolved to this state)

Example

- Process $X(t)$ with independent increments is always a Markov process:

$$X(t_n) = X(t_{n-1}) + (X(t_n) - X(t_{n-1}))$$

- It follows that Poisson process is a Markov process:
  - according to Definition 3, the increments of a Poisson process are independent

Time-homogeneity

- **Definition**: Markov process $X(t)$ is **time-homogeneous** if

$$P\{X(t+\Delta) = y \mid X(t) = x\} = P\{X(\Delta) = y \mid X(0) = x\}$$

  for all $t, \Delta \geq 0$ and $x, y \in S$
- In other words,
  probabilities $P\{X(t+\Delta) = y \mid X(t) = x\}$ are independent of t

State transition rates

- Consider a time-homogeneous Markov process $X(t)$
- The state transition rates q_{ij}, where $i, j \in S$, are defined as follows:

$$q_{ij} := \lim_{h \downarrow 0} \frac{1}{h} P\{X(h) = j \mid X(0) = i\}$$

- The initial distribution $P\{X(0) = i\}$, $i \in S$, and the state transition rates q_{ij} together determine the state probabilities $P\{X(t) = i\}$, $i \in S$, by the Kolmogorov (backwards/forwards) equations

Exponential holding times

- When in state i, the conditional probability that there is a transition from state i to state j during a short time interval $(t, t+h]$ is $q_{ij}h + o(h)$ independently of the other time intervals
- Let q_i denote the total transition rate out of state i, that is:

$$q_i := \sum_{j \neq i} q_{ij}$$

- Then, the conditional probability that there is a transition from state i to any other state during a short time interval $(t, t+h]$ is $q_i h + o(h)$ independently of the other time intervals
- Thus, the holding time in (any) state i is exponentially distributed with intensity q_i

State transition probabilities

- Let T_i denote the holding time in state i
- It can be seen as the minimum of independent (potential) holding times T_{ij} corresponding to (potential) transitions from state i to state j:

$$T_i = \min_{j \neq i} T_{ij}$$

- Let then p_{ij} denote the conditional probability that, when in state i, there is a transition from state i to state j
- Since potential holding times T_{ij} are exponentially distributed with intensity q_{ij}, we have (by slide 5.45)

$$T_i \sim \mathrm{Exp}(q_i), \qquad p_{ij} = P\{T_i = T_{ij}\} = \frac{q_{ij}}{q_i}$$

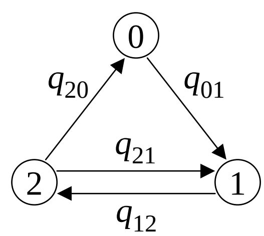
State transition diagram

- A time-homogeneous Markov process can be represented by a **state transition diagram**, which is a directed graph where
  - nodes correspond to states and
  - one-way links correspond to potential state transitions

$$\text{link from state } i \text{ to state } j \quad \Leftrightarrow \quad q_{ij} > 0$$

- Example: Markov process with three states, $S = \{0,1,2\}$

$$Q = \begin{pmatrix} - & + & 0 \\ 0 & - & + \\ + & + & - \end{pmatrix}$$



Irreducibility

- **Definition**: There is a **path** from state i to state j ($i \to j$) if there is a directed path from state i to state j in the state transition diagram.
- In this case, starting from state i, the process visits state j with positive probability
- **Definition**: States i and j **communicate** ($i \leftrightarrow j$) if $i \to j$ and $j \to i$.
- **Definition**: Markov process is **irreducible** if all states $i \in S$ communicate with each other
- Example: The Markov process presented in the previous slide is irreducible

Global balance equations, equilibrium distribution

- Consider an irreducible Markov process $X(t)$
- **Definition**: Let $\pi = (\pi_i \mid \pi_i \geq 0, i \in S)$ be a distribution defined on the state space S, that is:

$$\sum_{i \in S} \pi_i = 1 \qquad \text{(N)}$$

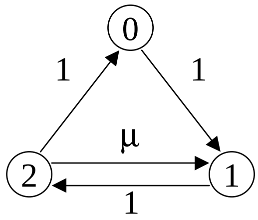
It is the **equilibrium distribution** of the process if the following **global balance equations** (GBE) are satisfied for each $i \in S$:

$$\sum_{j \neq i} \pi_i q_{ij} = \sum_{j \neq i} \pi_j q_{ji} \qquad \text{(GBE)}$$

  - It is possible that no equilibrium distribution exists
  - However, if the state space is finite, a unique equilibrium distribution exists
  - By choosing the equilibrium distribution (if it exists) as the initial distribution, the Markov process $X(t)$ becomes stationary (with stationary distribution π)

Example

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



$$\pi_0 + \pi_1 + \pi_2 = 1 \qquad \text{(N)}$$

$$\begin{aligned} \pi_0 \cdot 1 &= \pi_2 \cdot 1 \\ \pi_1 \cdot 1 &= \pi_0 \cdot 1 + \pi_2 \cdot \mu \qquad \text{(GBE)} \\ \pi_2 \cdot (1+\mu) &= \pi_1 \cdot 1 \end{aligned}$$

$$\Rightarrow \quad \pi_0 = \frac{1}{3+\mu}, \quad \pi_1 = \frac{1+\mu}{3+\mu}, \quad \pi_2 = \frac{1}{3+\mu}$$

Local balance equations

- Consider still an irreducible Markov process $X(t)$. Next we will give sufficient (but not necessary) conditions for the equilibrium distribution.
- **Proposition**: Let $\pi = (\pi_i \mid \pi_i \geq 0, i \in S)$ be a distribution defined on the state space S, that is:

$$\sum_{i \in S} \pi_i = 1 \qquad \text{(N)}$$

If the following **local balance equations** (LBE) are satisfied for each $i,j \in S$:

$$\pi_i q_{ij} = \pi_j q_{ji} \qquad \text{(LBE)}$$

then π is the equilibrium distribution of the process.

  - Proof: (GBE) follows from (LBE) by summing over all $j \neq i$
  - In this case the Markov process $X(t)$ is called **reversible** (looking stochastically the same in either direction of time)

Contents

Birth-death process

- Consider a continuous-time and discrete-state Markov process $X(t)$
  – with state space $S = \{0,1,\dots,N\}$ or $S = \{0,1,...\}$
- **Definition**: The process $X(t)$ is a **birth-death process** (BD) if state transitions are possible only between neighbouring states, that is:

$$|i-j|>1 \quad \Rightarrow \quad q_{ij}=0$$

- In this case, we denote

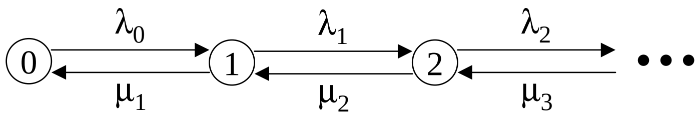
$$\mu_i := q_{i,i-1} \geq 0$$
$$\lambda_i := q_{i,i+1} \geq 0$$

  – The former is called the **death rate** and the latter the **birth rate**.
  – In partcular, we define $\mu_0 = 0$ and $\lambda_N = 0$ (if $N < \infty$)

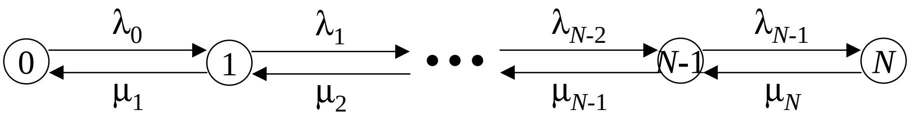
Irreducibility

- **Proposition**: A birth-death process is irreducible if and only if $\lambda_i > 0$ for all $i \in S \backslash \{N\}$ and $\mu_i > 0$ for all $i \in S \backslash \{0\}$

- State transition diagram of an infinite-state irreducible BD process:



- State transition diagram of a finite-state irreducible BD process:



Equilibrium distribution (1)

- Consider an irreducible birth-death process $X(t)$
- Let $\pi = (\pi_i \mid i \in S)$ denote the equilibrium distribution (if it exists)
- Local balance equations (LBE):

$$\pi_i \lambda_i = \pi_{i+1} \mu_{i+1} \qquad \text{(LBE)}$$

- Thus we get the following recursive formula:

$$\pi_{i+1} = \frac{\lambda_i}{\mu_{i+1}} \pi_i \quad \Rightarrow \quad \pi_i = \pi_0 \prod_{j=1}^{i} \frac{\lambda_{j-1}}{\mu_j}$$

- Normalizing condition (N):

$$\sum_{i \in S} \pi_i = \pi_0 \sum_{i \in S} \prod_{j=1}^{i} \frac{\lambda_{j-1}}{\mu_j} = 1 \qquad \text{(N)}$$

Equilibrium distribution (2)

- Thus, the equilibrium distribution exists if and only if

$$\sum_{i\in S}\prod_{j=1}^{i}\frac{\lambda_{j-1}}{\mu_j}<\infty$$

- **Finite state space**:
  The sum above is always finite, and the equilibrium distribution is

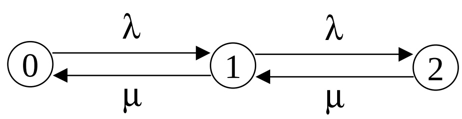
$$\pi_i=\pi_0\prod_{j=1}^{i}\frac{\lambda_{j-1}}{\mu_j},\quad \pi_0=\left(1+\sum_{i=1}^{N}\prod_{j=1}^{i}\frac{\lambda_{j-1}}{\mu_j}\right)^{-1}$$

- **Infinite state space**:
  If the sum above is finite, the equilibrium distribution is

$$\pi_i=\pi_0\prod_{j=1}^{i}\frac{\lambda_{j-1}}{\mu_j},\quad \pi_0=\left(1+\sum_{i=1}^{\infty}\prod_{j=1}^{i}\frac{\lambda_{j-1}}{\mu_j}\right)^{-1}$$

Example

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



$$\pi_i\lambda=\pi_{i+1}\mu$$

$$\Rightarrow\quad \pi_{i+1}=\rho\pi_i\quad(\rho:=\lambda/\mu)\qquad \text{(LBE)}$$

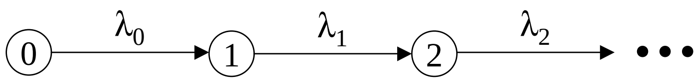
$$\Rightarrow\quad \pi_i=\pi_0\rho^i$$

$$\pi_0+\pi_1+\pi_2=\pi_0(1+\rho+\rho^2)=1\qquad \text{(N)}$$

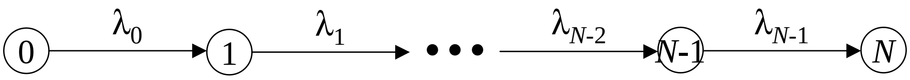
$$\Rightarrow\quad \pi_i=\frac{\rho^i}{1+\rho+\rho^2}$$

Pure birth process

- **Definition**: A birth-death process is a **pure birth process** if $\mu_i = 0$ for all $i \in S$
- State transition diagram of an infinite-state pure birth process:



- State transition diagram of a finite-state pure birth BD process:



- Example: Poisson process is a pure birth process (with constant birth rate $\lambda_i = \lambda$ for all $i \in S = \{0,1,\ldots\}$)
- Note: Pure birth process is never irreducible (nor stationary)!

THE END