

Probabilità e Processi Stocastici (455AA)

Lezione 13

Dario Trevisan – <https://web.dm.unipi.it/trevisan>

03/11/2025

Processi a stati discreti

Presentazione

- ▶ Introduciamo il linguaggio di base della teoria dei processi con alcune definizioni generali.

Presentazione

- ▶ Introduciamo il linguaggio di base della teoria dei processi con alcune definizioni generali.
- ▶ Vediamo la teoria di base delle catene di Markov

Presentazione

- ▶ Introduciamo il linguaggio di base della teoria dei processi con alcune definizioni generali.
- ▶ Vediamo la teoria di base delle catene di Markov
- ▶ e poi dei processi di Markov a salti.

Presentazione

- ▶ Introduciamo il linguaggio di base della teoria dei processi con alcune definizioni generali.
- ▶ Vediamo la teoria di base delle catene di Markov
- ▶ e poi dei processi di Markov a salti.
- ▶ Studiamo esistenza e unicità delle distribuzioni invarianti associate ad un processo di Markov

Presentazione

- ▶ Introduciamo il linguaggio di base della teoria dei processi con alcune definizioni generali.
- ▶ Vediamo la teoria di base delle catene di Markov
- ▶ e poi dei processi di Markov a salti.
- ▶ Studiamo esistenza e unicità delle distribuzioni invarianti associate ad un processo di Markov
- ▶ Accenniamo al problema della stima dei parametri di un processo a partire dalle osservazioni.

Presentazione

- ▶ Introduciamo il linguaggio di base della teoria dei processi con alcune definizioni generali.
- ▶ Vediamo la teoria di base delle catene di Markov
- ▶ e poi dei processi di Markov a salti.
- ▶ Studiamo esistenza e unicità delle distribuzioni invarianti associate ad un processo di Markov
- ▶ Accenniamo al problema della stima dei parametri di un processo a partire dalle osservazioni.
- ▶ Concludiamo con degli esempi fondamentali dalla teoria delle code.

Definizioni generali

Un **processo stocastico** è una collezione di variabili aleatorie $(X_t)_{t \in \mathcal{T}}$,

- ▶ tutte a valori nello stesso insieme E , detto insieme degli **stati** del processo,

Definizioni generali

Un **processo stocastico** è una collezione di variabili aleatorie $(X_t)_{t \in \mathcal{T}}$,

- ▶ tutte a valori nello stesso insieme E , detto insieme degli **stati** del processo,
- ▶ indicizzate da un insieme $\mathcal{T} \subseteq \mathbb{R}$ detto insieme dei **tempi** del processo.

Il calcolo delle probabilità fornisce strumenti utili per affrontare problemi circa

- ▶ il *futuro* di un processo (questo è il problema della *previsione*)

Il calcolo delle probabilità fornisce strumenti utili per affrontare problemi circa

- ▶ il *futuro* di un processo (questo è il problema della *previsione*)
- ▶ il *passato*,

Il calcolo delle probabilità fornisce strumenti utili per affrontare problemi circa

- ▶ il *futuro* di un processo (questo è il problema della *previsione*)
- ▶ il *passato*,
- ▶ oppure anche il *presente* (se non è esattamente osservato, è il *filtraggio*).

Classificazione generale

Analogamente alle singole variabili aleatorie, si classificano i processi stocastici in base all'insieme degli stati:

- ▶ a **stati discreti** se E discreto (quindi finito oppure infinito numerabile, ad esempio $E = \mathbb{Z}$ oppure \mathbb{N})

In base all'insieme \mathcal{T} dei tempi:

Classificazione generale

Analogamente alle singole variabili aleatorie, si classificano i processi stocastici in base all'insieme degli stati:

- ▶ a **stati discreti** se E discreto (quindi finito oppure infinito numerabile, ad esempio $E = \mathbb{Z}$ oppure \mathbb{N})
- ▶ a **stati continui** se E è infinito continuo, $E = \mathbb{R}$, $E = \mathbb{R}^k$ (e di solito ciascuna X_t ammetta densità continua)

In base all'insieme \mathcal{T} dei tempi:

Classificazione generale

Analogamente alle singole variabili aleatorie, si classificano i processi stocastici in base all'insieme degli stati:

- ▶ a **stati discreti** se E discreto (quindi finito oppure infinito numerabile, ad esempio $E = \mathbb{Z}$ oppure \mathbb{N})
- ▶ a **stati continui** se E è infinito continuo, $E = \mathbb{R}$, $E = \mathbb{R}^k$ (e di solito ciascuna X_t ammetta densità continua)

In base all'insieme \mathcal{T} dei tempi:

- ▶ a **tempi discreti** se \mathcal{T} è discreto (ad esempio finito, oppure $\mathcal{T} = \mathbb{N}$),

Classificazione generale

Analogamente alle singole variabili aleatorie, si classificano i processi stocastici in base all'insieme degli stati:

- ▶ a **stati discreti** se E discreto (quindi finito oppure infinito numerabile, ad esempio $E = \mathbb{Z}$ oppure \mathbb{N})
- ▶ a **stati continui** se E è infinito continuo, $E = \mathbb{R}$, $E = \mathbb{R}^k$ (e di solito ciascuna X_t ammetta densità continua)

In base all'insieme \mathcal{T} dei tempi:

- ▶ a **tempi discreti** se \mathcal{T} è discreto (ad esempio finito, oppure $\mathcal{T} = \mathbb{N}$),
- ▶ a **tempi continui** se $\mathcal{T} = [0, T]$ è un intervallo (anche illimitato, ad esempio $\mathcal{T} = [0, \infty)$).

Traiettorie e marginali

È utile pensare a $(X_t)_{t \in \mathcal{T}}$ come ad una variabile aleatoria vettoriale a valori in uno spazio di **traiettorie**,

$$E^{\mathcal{T}} = \{x : \mathcal{T} \rightarrow E\} = \{(x_t)_{t \in \mathcal{T}}\}.$$

- ▶ Ad esempio, se $\mathcal{T} = \{1, \dots, d\}$, allora un processo $(X_i)_{i=1}^d$ può essere pensato come una variabile aleatoria congiunta X , a valori in E^d .

Traiettorie e marginali

È utile pensare a $(X_t)_{t \in \mathcal{T}}$ come ad una variabile aleatoria vettoriale a valori in uno spazio di **traiettorie**,

$$E^{\mathcal{T}} = \{x : \mathcal{T} \rightarrow E\} = \{(x_t)_{t \in \mathcal{T}}\}.$$

- ▶ Ad esempio, se $\mathcal{T} = \{1, \dots, d\}$, allora un processo $(X_i)_{i=1}^d$ può essere pensato come una variabile aleatoria congiunta X , a valori in E^d .
- ▶ Ricordiamo la differenza tra la legge delle marginali

$$P(X_t \in U | I),$$

al variare di $U \subseteq E$ e $t \in \mathcal{T}$, e la legge congiunta, in questo caso detta semplicemente **legge del processo** $(X_t)_{t \in \mathcal{T}}$, che è definita come tutte le probabilità del tipo

$$P(X_{t_1} \in U_1, X_{t_2} \in U_2, \dots, X_{t_k} \in U_k | I),$$

Nel caso di processi a stati discreti, per ogni $t \in \mathcal{T}$ la densità discreta della marginale al tempo t è la collezione delle probabilità

$$P(X_t = x | I).$$

- ▶ La densità discreta del processo è la collezione di tutte le probabilità

$$P(X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_k} = x_k | I),$$

Nel caso di processi a stati discreti, per ogni $t \in \mathcal{T}$ la densità discreta della marginale al tempo t è la collezione delle probabilità

$$P(X_t = x | I).$$

- ▶ La densità discreta del processo è la collezione di tutte le probabilità

$$P(X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_k} = x_k | I),$$

- ▶ Nel caso di processi a stati continui (con densità continua), basta sostituire la “ P ” di probabilità con “ p ” della densità di probabilità.

Come parametrizzare un processo?

Determinare la legge di un processo tramite pochi parametri è un problema difficile, soprattutto se l'insieme dei tempi diventa grande.

- ▶ Se $E = \{0, 1\}$, la densità discreta di un processo su $\mathcal{T} = \{1, \dots, d\}$ è praticamente una qualsiasi funzione da $\{0, 1\}^d$ a valori in $[0, 1] \rightarrow$ circa 2^d “parametri”.

Come parametrizzare un processo?

Determinare la legge di un processo tramite pochi parametri è un problema difficile, soprattutto se l'insieme dei tempi diventa grande.

- ▶ Se $E = \{0, 1\}$, la densità discreta di un processo su $\mathcal{T} = \{1, \dots, d\}$ è praticamente una qualsiasi funzione da $\{0, 1\}^d$ a valori in $[0, 1] \rightarrow$ circa 2^d “parametri”.
- ▶ Le d densità marginali si ottengono descrivendo solo d “parametri” (la probabilità $P(X_t = 1|I)$), anche meno se le leggi sono tutte uguali.

Come parametrizzare un processo?

Determinare la legge di un processo tramite pochi parametri è un problema difficile, soprattutto se l'insieme dei tempi diventa grande.

- ▶ Se $E = \{0, 1\}$, la densità discreta di un processo su $\mathcal{T} = \{1, \dots, d\}$ è praticamente una qualsiasi funzione da $\{0, 1\}^d$ a valori in $[0, 1] \rightarrow$ circa 2^d “parametri”.
- ▶ Le d densità marginali si ottengono descrivendo solo d “parametri” (la probabilità $P(X_t = 1|I)$), anche meno se le leggi sono tutte uguali.
- ▶ Non si può ricostruire la densità del processo a partire dalle densità marginali, senza ulteriori ipotesi.

Proprietà di Markov

La proprietà afferma che **il futuro e il passato sono condizionatamente indipendenti, noto esattamente il presente.**

- ▶ Un processo $(X_t)_{t \in \mathcal{T}}$ è *di Markov* (o markoviano) se, per ogni $x \in E$, $t \in \mathcal{T}$, le due variabili congiunte relative ai tempi “passati” $(X_s)_{s < t}$ e “futuri” $(X_r)_{r > t}$ sono indipendenti, nota l’informazione esatta sul presente, ossia $\{X_t = x\}$.

Proprietà di Markov

La proprietà afferma che **il futuro e il passato sono condizionatamente indipendenti, noto esattamente il presente.**

- ▶ Un processo $(X_t)_{t \in \mathcal{T}}$ è *di Markov* (o markoviano) se, per ogni $x \in E$, $t \in \mathcal{T}$, le due variabili congiunte relative ai tempi “passati” $(X_s)_{s < t}$ e “futuri” $(X_r)_{r > t}$ sono indipendenti, nota l’informazione esatta sul presente, ossia $\{X_t = x\}$.
- ▶ Se A è una affermazione che riguarda solo le variabili $(X_s)_{s < t}$, e B è una riguarda solamente le variabili $(X_r)_{r > t}$, allora A , B sono indipendenti rispetto all’informazione $\{X_t = x\}$:

$$P(A, B | X_t = x) = P(A | X_t = x)P(B | X_t = x),$$

oppure

$$P(A | X_t = x, B) = P(A | X_t = x),$$

o anche

$$P(B | X_t = x, A) = P(B | X_t = x).$$

In termini grafici, la proprietà di Markov si traduce in una rete bayesiana associata al processo $(X_t)_{t \in \mathcal{T}}$ del seguente tipo:

Densità di un processo di Markov

Processi omogenei

Per procedere ulteriormente e sviluppare una teoria semplice:

1. consideriamo come insiemi di tempi \mathcal{T} intervalli discreti $\mathcal{T} = \{0, 1, 2, \dots, n\}$ o continui $\mathcal{T} = [0, T]$.

Processi omogenei

Per procedere ulteriormente e sviluppare una teoria semplice:

1. consideriamo come insiemi di tempi \mathcal{T} intervalli discreti $\mathcal{T} = \{0, 1, 2, \dots, n\}$ o continui $\mathcal{T} = [0, T]$.
2. consideriamo processi di Markov **omogenei**, ossia tali che le probabilità di transizione dal tempo s al tempo t dipendano solamente dalla differenza dei tempi $t - s$, o equivalentemente, per ogni $\Delta t \geq 0$ si abbia

$$P(X_t = y | X_s = x) = P(X_{t+\Delta t} = y | X_{s+\Delta t} = x)$$

per stati $x, y \in E$.

Processi stazionari

Un processo $(X_t)_{t \in \mathcal{T}}$ si dice **stazionario** se, per ogni $\Delta t \geq 0$, la legge (congiunta) del processo coincide con quella del “traslato” $(X_{t+\Delta t})_{t \in \mathcal{T}}$ (purché i tempi $t + \Delta t$ appartengano a \mathcal{T}).

- ▶ Più precisamente, per ogni $k \geq 1$ e $t_1, t_2, \dots, t_k \in \mathcal{T}$ e $\Delta t \geq 0$, la legge congiunta di $(X_{t_1}, \dots, X_{t_k})$ coincide con quella di $(X_{t_1+\Delta t}, \dots, X_{t_k+\Delta t})$, purché i tempi $t_i + \Delta t$ appartengano a \mathcal{T} .

Processi stazionari

Un processo $(X_t)_{t \in \mathcal{T}}$ si dice **stazionario** se, per ogni $\Delta t \geq 0$, la legge (congiunta) del processo coincide con quella del “traslato” $(X_{t+\Delta t})_{t \in \mathcal{T}}$ (purché i tempi $t + \Delta t$ appartengano a \mathcal{T}).

- ▶ Più precisamente, per ogni $k \geq 1$ e $t_1, t_2, \dots, t_k \in \mathcal{T}$ e $\Delta t \geq 0$, la legge congiunta di $(X_{t_1}, \dots, X_{t_k})$ coincide con quella di $(X_{t_1+\Delta t}, \dots, X_{t_k+\Delta t})$, purché i tempi $t_i + \Delta t$ appartengano a \mathcal{T} .
- ▶ In particolare, nel caso di stati discreti, vale
$$P(X_{t_1} = x_1, \dots, X_{t_k} = x_k) = P(X_{t_1+\Delta t} = x_1, \dots, X_{t_k+\Delta t} = x_k),$$
per qualsiasi scelta di stati $x_1, \dots, x_k \in E$.

Processi stazionari

Un processo $(X_t)_{t \in \mathcal{T}}$ si dice **stazionario** se, per ogni $\Delta t \geq 0$, la legge (congiunta) del processo coincide con quella del “traslato” $(X_{t+\Delta t})_{t \in \mathcal{T}}$ (purché i tempi $t + \Delta t$ appartengano a \mathcal{T}).

- ▶ Più precisamente, per ogni $k \geq 1$ e $t_1, t_2, \dots, t_k \in \mathcal{T}$ e $\Delta t \geq 0$, la legge congiunta di $(X_{t_1}, \dots, X_{t_k})$ coincide con quella di $(X_{t_1+\Delta t}, \dots, X_{t_k+\Delta t})$, purché i tempi $t_i + \Delta t$ appartengano a \mathcal{T} .
- ▶ In particolare, nel caso di stati discreti, vale
$$P(X_{t_1} = x_1, \dots, X_{t_k} = x_k) = P(X_{t_1+\Delta t} = x_1, \dots, X_{t_k+\Delta t} = x_k),$$
per qualsiasi scelta di stati $x_1, \dots, x_k \in E$.
- ▶ Nel caso continuo l’identità sopra vale per le densità continue (scrivendo la densità p al posto della probabilità P).

Processi stazionari

Un processo $(X_t)_{t \in \mathcal{T}}$ si dice **stazionario** se, per ogni $\Delta t \geq 0$, la legge (congiunta) del processo coincide con quella del “traslato” $(X_{t+\Delta t})_{t \in \mathcal{T}}$ (purché i tempi $t + \Delta t$ appartengano a \mathcal{T}).

- ▶ Più precisamente, per ogni $k \geq 1$ e $t_1, t_2, \dots, t_k \in \mathcal{T}$ e $\Delta t \geq 0$, la legge congiunta di $(X_{t_1}, \dots, X_{t_k})$ coincide con quella di $(X_{t_1+\Delta t}, \dots, X_{t_k+\Delta t})$, purché i tempi $t_i + \Delta t$ appartengano a \mathcal{T} .
- ▶ In particolare, nel caso di stati discreti, vale
$$P(X_{t_1} = x_1, \dots, X_{t_k} = x_k) = P(X_{t_1+\Delta t} = x_1, \dots, X_{t_k+\Delta t} = x_k),$$
per qualsiasi scelta di stati $x_1, \dots, x_k \in E$.
- ▶ Nel caso continuo l’identità sopra vale per le densità continue (scrivendo la densità p al posto della probabilità P).
- ▶ Se un processo X è stazionario, necessariamente tutte le leggi delle marginali X_t coincidono: basta usare $k = 1$ nella definizione sopra.

Catene di Markov

Catene di Markov (omogenee)

Studiamo processi di Markov $(X_t)_{t \in \mathcal{T}}$ a *tempi discreti*, con

$$\mathcal{T} = \{0, 1, 2, \dots, n\} \subseteq \mathbb{N}$$

e *stati discreti* (spesso finiti).

- ▶ Ci limitiamo a processi *omogenei*. Le *probabilità di transizione* tra k e $k+1 \in \mathcal{T}$ dipendono solo dagli stati $x, y \in E$, e non da k :

$$P(X_{k+1} = y | X_k = x) = P(X_1 = y | X_0 = x).$$

Matrice di transizione

Le probabilità di transizione

$$Q_{x \rightarrow y} = P(X_1 = y | X_0 = x)$$

al variare di $x, y \in E$ sono spesso raccolte in una *matrice quadrata*, con tante righe e colonne quanti gli stati (eventualmente infinite),
 $Q \in \mathbb{R}^{E \times E}$

Un esempio

$$E = \{\text{OFF}, \text{ON}\}.$$

$$\begin{pmatrix} Q_{\text{OFF} \rightarrow \text{OFF}} & Q_{\text{OFF} \rightarrow \text{ON}} \\ Q_{\text{ON} \rightarrow \text{OFF}} & Q_{\text{ON} \rightarrow \text{ON}} \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.7 \end{pmatrix}$$

Matrici stocastiche

- ▶ la somma delle probabilità in ciascuna riga della matrice di transizione vale 1, perché

$$(Q_{x \rightarrow y})_{y \in E} = (P(X_1 = y | X_0 = x))_{y \in E}$$

sono la densità discreta della variabile X_1 (rispetto all'informazione $X_0 = x$).

Matrici stocastiche

- ▶ la somma delle probabilità in ciascuna riga della matrice di transizione vale 1, perché

$$(Q_{x \rightarrow y})_{y \in E} = (P(X_1 = y | X_0 = x))_{y \in E}$$

sono la densità discreta della variabile X_1 (rispetto all'informazione $X_0 = x$).

- ▶ In generale, una qualsiasi matrice Q con entrate a valori in $[0, 1]$ e tale che la somma sulle righe sia costante e uguale ad 1 è detta **matrice stocastica**.

Matrici stocastiche

- ▶ la somma delle probabilità in ciascuna riga della matrice di transizione vale 1, perché

$$(Q_{x \rightarrow y})_{y \in E} = (P(X_1 = y | X_0 = x))_{y \in E}$$

sono la densità discreta della variabile X_1 (rispetto all'informazione $X_0 = x$).

- ▶ In generale, una qualsiasi matrice Q con entrate a valori in $[0, 1]$ e tale che la somma sulle righe sia costante e uguale ad 1 è detta **matrice stocastica**.
- ▶ Se anche la somma sulle colonne è uguale ad 1 è detta **matrice bistocastica**

Legge di una catena di Markov

- Vale

$$P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = P(X_0 = x_0) \prod_{k=1}^n Q_{x_{k-1} \rightarrow x_k},$$

Legge di una catena di Markov

- Vale

$$P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = P(X_0 = x_0) \prod_{k=1}^n Q_{x_{k-1} \rightarrow x_k},$$

- un **cammino** è una sequenza ordinata di stati

$\gamma = (x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n)$:

$$P(X = \gamma) = P(X_0 = x_0) Q_\gamma$$

dove il “peso” di γ è

$$Q_\gamma = \prod_{k=1}^n Q_{x_{k-1} \rightarrow x_k}.$$

Un esempio

Una strategia generale

Per calcolare la probabilità di una qualsiasi affermazione A circa una catena di Markov $(X_n)_{n=0}^{\infty}$, basta

- ▶ rappresentare A in termini di cammini (a partire dal tempo iniziale)

Una strategia generale

Per calcolare la probabilità di una qualsiasi affermazione A circa una catena di Markov $(X_n)_{n=0}^{\infty}$, basta

- ▶ rappresentare A in termini di cammini (a partire dal tempo iniziale)
- ▶ sommare le probabilità corrispondenti ottenute tramite la formula sopra,

Una strategia generale

Per calcolare la probabilità di una qualsiasi affermazione A circa una catena di Markov $(X_n)_{n=0}^{\infty}$, basta

- ▶ rappresentare A in termini di cammini (a partire dal tempo iniziale)
- ▶ sommare le probabilità corrispondenti ottenute tramite la formula sopra,
- ▶ purché gli eventi relativi a cammini diversi siano a due a due incompatibili!.

Una strategia generale

Per calcolare la probabilità di una qualsiasi affermazione A circa una catena di Markov $(X_n)_{n=0}^{\infty}$, basta

- ▶ rappresentare A in termini di cammini (a partire dal tempo iniziale)
- ▶ sommare le probabilità corrispondenti ottenute tramite la formula sopra,
- ▶ purché gli eventi relativi a cammini diversi siano a due a due incompatibili!.
- ▶ Questo avviene anche se i cammini hanno lunghezze diverse, ma *nessun cammino considerato si può ottenere come prolungamento di un altro.*

Un esempio

Calcoliamo

$$P(X_2 = \text{OFF} \text{ oppure } X_3 = \text{ON})$$

tramite cammini (sapendo che $X_0 = \text{OFF}$)

X_0	X_1	X_2	X_3	Q_γ
OFF	$\rightarrow \text{OFF}$	$\rightarrow \text{OFF}$		$0.9 \cdot 0.9$
OFF	$\rightarrow \text{ON}$	$\rightarrow \text{OFF}$		$0.1 \cdot 0.7$
OFF	$\rightarrow \text{OFF}$	$\rightarrow \text{ON}$	$\rightarrow \text{ON}$	$0.9 \cdot 0.1 \cdot 0.3$
OFF	$\rightarrow \text{ON}$	$\rightarrow \text{ON}$	$\rightarrow \text{ON}$	$0.1 \cdot 0.3 \cdot 0.3$

Grafo associato alla catena

- ▶ Il calcolo dei cammini si può effettuare tramite diagrammi ad albero (ma diventano troppo grandi!).

Grafo associato alla catena

- ▶ Il calcolo dei cammini si può effettuare tramite diagrammi ad albero (ma diventano troppo grandi!).
- ▶ Nella pratica si introduce un grafo pesato orientato i cui nodi corrispondono agli *stati* $i \in E$, e l'arco da i ad $j \in E$ è pesato con la probabilità di transizione Q_{ij} (se la probabilità è nulla non viene rappresentato).

Grafo associato alla catena

- ▶ Il calcolo dei cammini si può effettuare tramite diagrammi ad albero (ma diventano troppo grandi!).
- ▶ Nella pratica si introduce un grafo pesato orientato i cui nodi corrispondono agli *stati* $i \in E$, e l'arco da i ad $j \in E$ è pesato con la probabilità di transizione Q_{ij} (se la probabilità è nulla non viene rappresentato).
- ▶ Il grafo associato permette facilmente di calcolare le probabilità

$$Q_\gamma = \prod_{k=1}^n Q_{x_{k-1} x_k}$$

associate ad un cammino $\gamma = (x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n)$,

Grafo associato alla catena

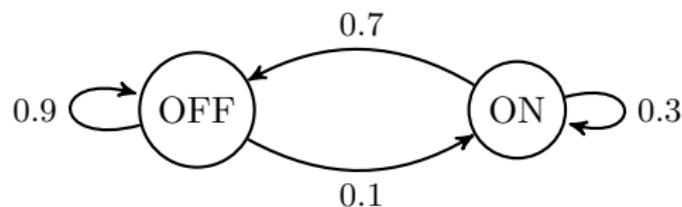
- ▶ Il calcolo dei cammini si può effettuare tramite diagrammi ad albero (ma diventano troppo grandi!).
- ▶ Nella pratica si introduce un grafo pesato orientato i cui nodi corrispondono agli *stati* $i \in E$, e l'arco da i ad $j \in E$ è pesato con la probabilità di transizione Q_{ij} (se la probabilità è nulla non viene rappresentato).
- ▶ Il grafo associato permette facilmente di calcolare le probabilità

$$Q_\gamma = \prod_{k=1}^n Q_{x_{k-1} x_k}$$

associate ad un cammino $\gamma = (x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n)$,

- ▶ Nel grafo non è rappresentata la probabilità marginale al tempo iniziale ($t = 0$).

Un esempio



Tempo di permanenza

Supponiamo di osservare che al tempo 0 la catena si trova nello stato x_0 . Poniamo T_1 il più piccolo tempo $k \in \{1, 2, \dots\}$ tale che $X_k \neq x_0$.

- ▶ $T_1 = k$ significa che osserviamo il cammino $X_0 = x_0, X_1 = x_0, \dots, X_k = y$ dove $y \neq x_0$.

Tempo di permanenza

Supponiamo di osservare che al tempo 0 la catena si trova nello stato x_0 . Poniamo T_1 il più piccolo tempo $k \in \{1, 2, \dots\}$ tale che $X_k \neq x_0$.

- ▶ $T_1 = k$ significa che osserviamo il cammino $X_0 = x_0, X_1 = x_0, \dots, X_k = y$ dove $y \neq x_0$.
- ▶ Si trova $P(T_1 = k | X_0 = x_0) = Q_{x_0 \rightarrow x_0}^{k-1} (1 - Q_{x_0 \rightarrow x_0})$ ossia $T_1 - 1$ ha densità geometrica di parametro $p = 1 - Q_{x_0 \rightarrow x_0}$.

Tempo di permanenza

Supponiamo di osservare che al tempo 0 la catena si trova nello stato x_0 . Poniamo T_1 il più piccolo tempo $k \in \{1, 2, \dots\}$ tale che $X_k \neq x_0$.

- ▶ $T_1 = k$ significa che osserviamo il cammino $X_0 = x_0, X_1 = x_0, \dots, X_k = y$ dove $y \neq x_0$.
- ▶ Si trova $P(T_1 = k | X_0 = x_0) = Q_{x_0 \rightarrow x_0}^{k-1} (1 - Q_{x_0 \rightarrow x_0})$ ossia $T_1 - 1$ ha densità geometrica di parametro $p = 1 - Q_{x_0 \rightarrow x_0}$.
- ▶ *Esercizio:* mostrare che

$$P(X_{T_1} = y | X_0 = x_0) = \frac{Q_{x_0 \rightarrow y}}{1 - Q_{x_0 \rightarrow x_0}} \propto Q_{x_0 \rightarrow y} \quad \text{per } y \neq x_0$$

mentre ovviamente $P(X_{T_1} = x_0 | X_0 = x_0) = 0$.

Tempo di permanenza

Supponiamo di osservare che al tempo 0 la catena si trova nello stato x_0 . Poniamo T_1 il più piccolo tempo $k \in \{1, 2, \dots\}$ tale che $X_k \neq x_0$.

- ▶ $T_1 = k$ significa che osserviamo il cammino $X_0 = x_0, X_1 = x_0, \dots, X_k = y$ dove $y \neq x_0$.
- ▶ Si trova $P(T_1 = k | X_0 = x_0) = Q_{x_0 \rightarrow x_0}^{k-1} (1 - Q_{x_0 \rightarrow x_0})$ ossia $T_1 - 1$ ha densità geometrica di parametro $p = 1 - Q_{x_0 \rightarrow x_0}$.
- ▶ *Esercizio:* mostrare che

$$P(X_{T_1} = y | X_0 = x_0) = \frac{Q_{x_0 \rightarrow y}}{1 - Q_{x_0 \rightarrow x_0}} \propto Q_{x_0 \rightarrow y} \quad \text{per } y \neq x_0$$

mentre ovviamente $P(X_{T_1} = x_0 | X_0 = x_0) = 0$.

- ▶ Possiamo interpretare X come una successione di **permanenze** aventi distribuzione geometrica **salti** con distribuzione ottenuta da Q .

Densità marginali

Le densità marginali di una catena di Markov si ottengono sommando la densità congiunta su tutti i possibili valori delle altre variabili.

$$P(X_n = x_n) = \sum_{x_0, x_1, \dots, x_{n-1} \in E} P(X_0 = x_0) \prod_{k=1}^n Q_{x_{k-1} x_k}.$$

- ▶ Vale un'equazione *ricorsiva*:

$$\begin{aligned} P(X_n = x_n) &= \sum_{x_{n-1} \in E} P(X_n = x_n | X_{n-1} = x_{n-1}) P(X_{n-1} = x_{n-1}) \\ &= \sum_{x_{n-1} \in E} Q_{x_{n-1} x_n} P(X_{n-1} = x_{n-1}). \end{aligned}$$

Se Q è una matrice in $\mathbb{R}^{E \times E}$ e $P(X_n = \cdot)$ è un vettore riga

$$\mu_n(x) = P(X_n = x),$$

allora

$$\mu_n = \mu_{n-1} Q$$

- ▶ Iterando otteniamo

$$\mu_n = \mu_{n-1} Q = \mu_{n-2} Q^2 = \dots = \mu_0 Q^n,$$

dove le potenze Q^k sono intese nel senso del prodotto di matrici.

Un esempio in R

In R il prodotto tra matrici si ottiene tramite il comando `%*%`.

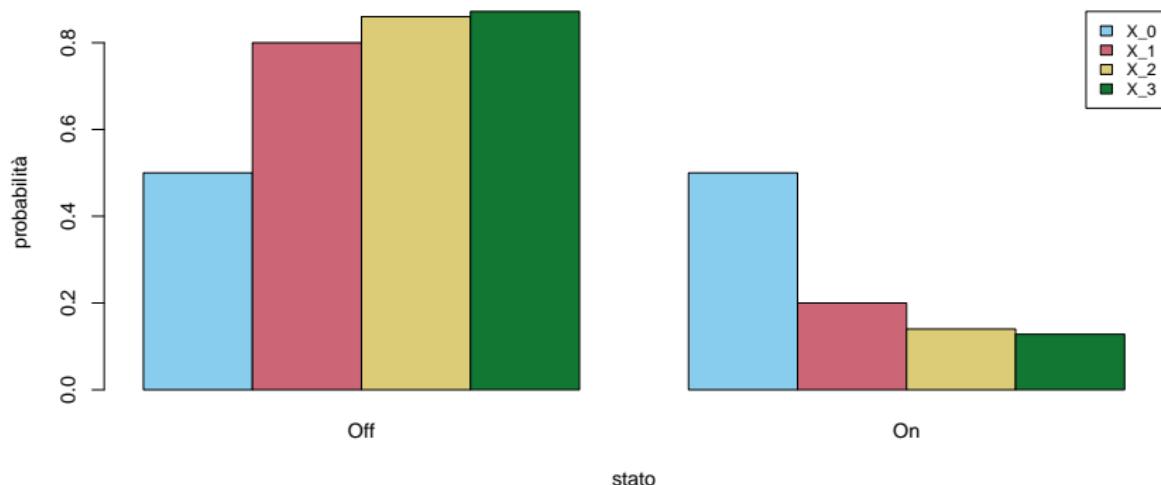


Figure 1: Grafico a barre della densità marginale della catena ai tempi 0 (uniforme), 1, 2 e 3.

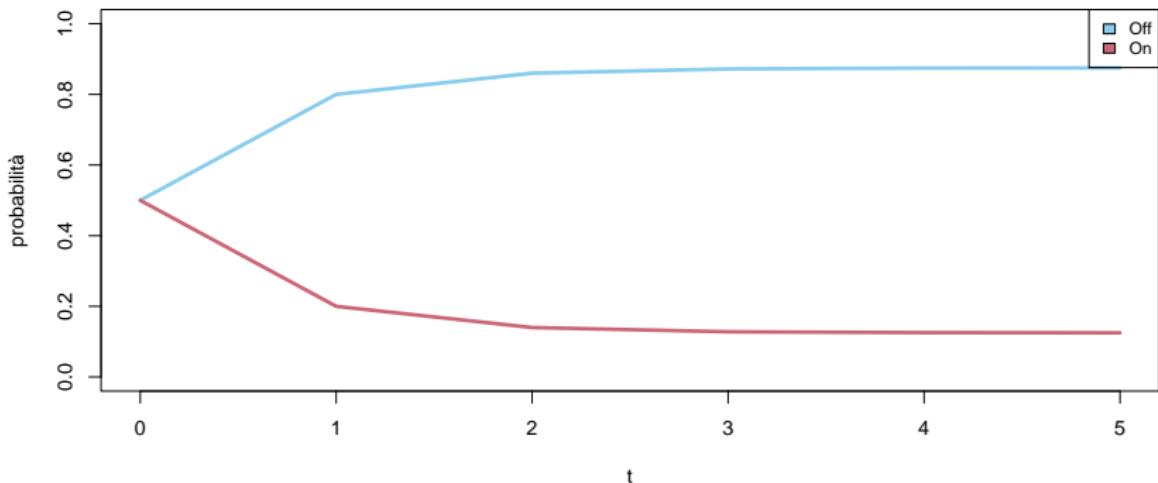


Figure 2: grafico delle densità marginali in funzione del tempo
 $t = 0, 1, \dots, 5$ partendo dalla densità uniforme al tempo 0.

Processi di Markov a salti

Richiami sulle catene di Markov

Abbiamo definito una *catena di Markov* $(X_n)_{n=0}^N$ come un processo stocastico **tempi discreti, stati discreti, di Markov, omogeneo**.

- ▶ Matrice di transizione $Q_{x \rightarrow y} = P(X_1 = y | X_0 = x)$ (stocastica)

Richiami sulle catene di Markov

Abbiamo definito una *catena di Markov* $(X_n)_{n=0}^N$ come un processo stocastico **tempi discreti, stati discreti, di Markov, omogeneo**.

- ▶ Matrice di transizione $Q_{x \rightarrow y} = P(X_1 = y | X_0 = x)$ (stocastica)
- ▶ Rappresentazione grafica

Richiami sulle catene di Markov

Abbiamo definito una *catena di Markov* $(X_n)_{n=0}^N$ come un processo stocastico **tempi discreti, stati discreti, di Markov, omogeneo**.

- ▶ Matrice di transizione $Q_{x \rightarrow y} = P(X_1 = y | X_0 = x)$ (stocastica)
- ▶ Rappresentazione grafica
- ▶ Peso di un cammino $\gamma = x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n$,

$$Q_\gamma = Q_{x_0 x_1} Q_{x_1 x_2} \dots Q_{x_{n-1} x_n}$$

Richiami sulle catene di Markov

Abbiamo definito una *catena di Markov* $(X_n)_{n=0}^N$ come un processo stocastico **tempi discreti, stati discreti, di Markov, omogeneo**.

- ▶ Matrice di transizione $Q_{x \rightarrow y} = P(X_1 = y | X_0 = x)$ (stocastica)
- ▶ Rappresentazione grafica
- ▶ Peso di un cammino $\gamma = x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n$,

$$Q_\gamma = Q_{x_0 x_1} Q_{x_1 x_2} \dots Q_{x_{n-1} x_n}$$

- ▶ Legge del processo

$$P(X = \gamma) = P(X_0 = x_0) Q_\gamma$$

Richiami sulle catene di Markov

Abbiamo definito una *catena di Markov* $(X_n)_{n=0}^N$ come un processo stocastico **tempi discreti, stati discreti, di Markov, omogeneo**.

- ▶ Matrice di transizione $Q_{x \rightarrow y} = P(X_1 = y | X_0 = x)$ (stocastica)
- ▶ Rappresentazione grafica
- ▶ Peso di un cammino $\gamma = x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n$,

$$Q_\gamma = Q_{x_0 x_1} Q_{x_1 x_2} \dots Q_{x_{n-1} x_n}$$

- ▶ Legge del processo

$$P(X = \gamma) = P(X_0 = x_0) Q_\gamma$$

- ▶ Legge marginale $\mu_n = \mu_{n-1} Q$, da cui $\mu_n = \mu_0 Q^n$.

Processi di Markov a salti: definizione

- ▶ I processi di Markov $(X_t)_{t \in \mathcal{T}}$ a tempi continui

$$\mathcal{T} = [0, T]$$

e a stati discreti sono detti *a salti*, perché le traiettorie “saltano” da uno stato all’altro.

Processi di Markov a salti: definizione

- ▶ I processi di Markov $(X_t)_{t \in \mathcal{T}}$ a tempi continui

$$\mathcal{T} = [0, T]$$

e a stati discreti sono detti *a salti*, perché le traiettorie “saltano” da uno stato all’altro.

- ▶ Ci limitiamo a considerare processi omogenei, ossia tali che, per ogni $t \in [0, T]$, $\delta t > 0$ tale che $t + \delta t \in [0, T]$, si abbia

$$P(X_{t+\delta t} = y | X_t = x) = P(X_{\delta t} = y | X_0 = x).$$

Matrice delle intensità di salto

Nel caso dei processi di Markov a salti, invece di Q si usa la **matrice delle intensità di salto**

$$\begin{aligned} L_{x \rightarrow y} &= \frac{d}{dt} \Big|_{t=0} P(X_t = y | X_0 = x) \\ &= \lim_{\delta t \rightarrow 0^+} \frac{P(X_{\delta t} = y | X_0 = x) - P(X_0 = y | X_0 = x)}{\delta t}, \end{aligned}$$

Proprietà della matrice L

La derivata della matrice (stocastica) delle intensità di salto *non* è una matrice stocastica. Dati $x, y \in E$,

- ▶ se $x \neq y$, vale $P(X_0 = y | X_0 = x) = 0$ e quindi

$$L_{xy} = \frac{d}{dt} P(X_t = y | X_0 = x) = \lim_{\delta t \rightarrow 0} \frac{P(X_{\delta t} = y | X_0 = x)}{\delta t} \geq 0$$

Proprietà della matrice L

La derivata della matrice (stocastica) delle intensità di salto *non* è una matrice stocastica. Dati $x, y \in E$,

- ▶ se $x \neq y$, vale $P(X_0 = y | X_0 = x) = 0$ e quindi

$$L_{xy} = \frac{d}{dt} P(X_t = y | X_0 = x) = \lim_{\delta t \rightarrow 0} \frac{P(X_{\delta t} = y | X_0 = x)}{\delta t} \geq 0$$

- ▶ se invece $x = y$, allora $P(X_0 = y | X_0 = x) = 1$ e quindi

$$L_{xy} = \frac{d}{dt} P(X_t = y | X_0 = x) = \lim_{\delta t \rightarrow 0} \frac{P(X_{\delta t} = y | X_0 = x) - 1}{\delta t} \leq 0.$$

Proprietà della matrice L

La derivata della matrice (stocastica) delle intensità di salto *non* è una matrice stocastica. Dati $x, y \in E$,

- ▶ se $x \neq y$, vale $P(X_0 = y | X_0 = x) = 0$ e quindi

$$L_{xy} = \frac{d}{dt} P(X_t = y | X_0 = x) = \lim_{\delta t \rightarrow 0} \frac{P(X_{\delta t} = y | X_0 = x)}{\delta t} \geq 0$$

- ▶ se invece $x = y$, allora $P(X_0 = y | X_0 = x) = 1$ e quindi

$$L_{xy} = \frac{d}{dt} P(X_t = y | X_0 = x) = \lim_{\delta t \rightarrow 0} \frac{P(X_{\delta t} = y | X_0 = x) - 1}{\delta t} \leq 0.$$

- ▶ La condizione che la somma su ciascuna riga valga 1 diventa

$$\sum_{y \in E} L_{xy} = \frac{d}{dt} \sum_{y \in E} P(X_t = y | X_0 = x) = \frac{d}{dt} 1 = 0,$$

equivalentemente $L_{xx} = - \sum_{y \neq x} L_{xy}$.

Un esempio

$E = \{\text{Off}, \text{Standby}, \text{On}\}$, con matrice delle intensità di salto

$$L = \begin{pmatrix} * & 5 & 10 \\ 1 & * & 3 \\ 0 & 4 & * \end{pmatrix}.$$

Approssimazione tramite tempi discreti

Per ricondursi alle catene di Markov, l'idea è di discretizzare i tempi

$$\mathcal{T}^\delta = \{0, \delta, 2\delta, \dots, \lfloor T/\delta \rfloor \delta\},$$

e passare al limite $\delta \rightarrow 0$.

- ▶ Il processo

$$X_k^\delta := X_{k\delta},$$

è una catena di Markov con matrice di transizione

$$P_{xy}^\delta = P(X_\delta = y | X_0 = x) = Id_{xy} + L_{xy}\delta + O(\delta^2) \approx Id + L\delta.$$

Densità marginali

Indicando con $\mu_t(x) = P(X_t = x)$ il vettore (riga) della densità discreta marginale al tempo t , si ha per i tempi $t = h\delta$,

$$\mu_t = \mu_0(P^\delta)^h \approx \mu_0 \left(Id + \frac{tL}{h} \right)^h,$$

avendo scritto $\delta = t/h$.

- ▶ Fissato $t \in [0, T]$ si possono trovare $h \rightarrow \infty$ in modo che $\delta \rightarrow 0$ e quindi, ricordando il limite notevole (che vale anche per le matrici)

$$\lim_{h \rightarrow \infty} \left(Id + \frac{A}{h} \right)^h = \exp(A),$$

si trova che

$$\mu_t = \mu_0 \exp(tL).$$

Un esempio (in R)

L'esponenziale di matrice è la funzione `expm()` nella libreria `Matrix`.

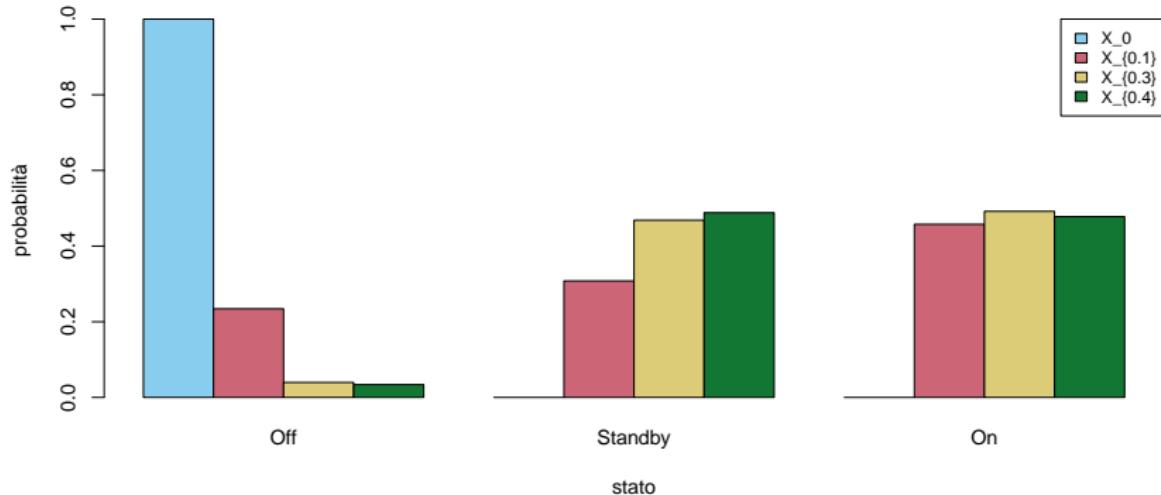


Figure 3: Grafico a barre della densità marginale del processo a salti ai tempi 0 (Off), 0.1, 0.3 e 0.4.

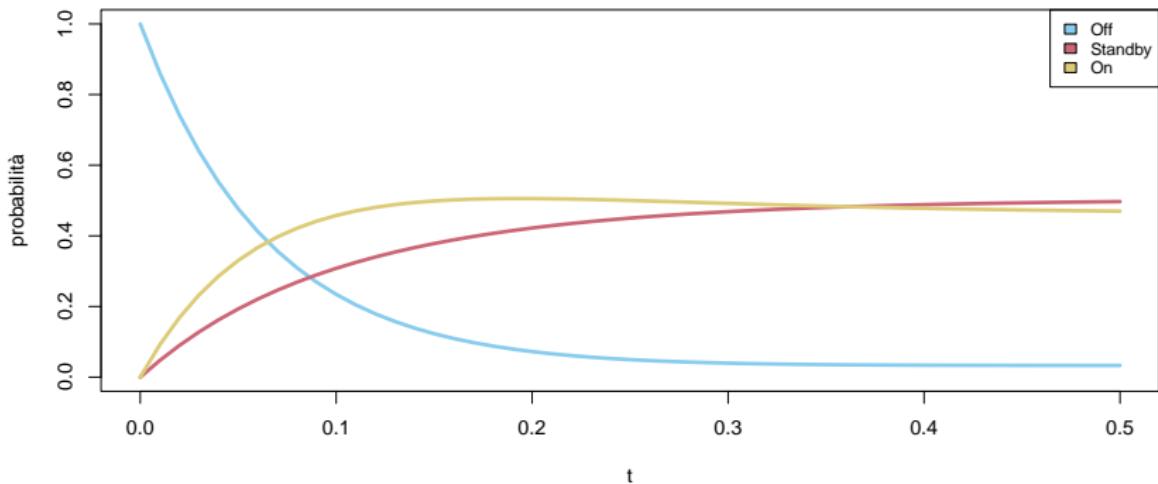


Figure 4: grafico delle densita' marginali in funzione del tempo
 $t \in [0, 0.5]$ partendo dallo stato $X_0 = \text{Off}$.

La master equation

L'analoga dell'equazione ricorsiva è l'equazione differenziale ottenuta derivando la formula sopra:

$$\frac{d}{dt}\mu_t = \mu_t L$$

- ▶ Equazione differenziale *lineare* è detta anche equazione di Kolmogorov in avanti (o anche *master equation*).

Legge del processo di Markov a salti

Consideriamo la probabilità di osservare un cammino che visiti nell'ordine gli stati $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n$.

- ▶ Trattandosi di tempi continui, dobbiamo specificare i *tempi di permanenza* in ciascuno stato t_1, \dots, t_n , in modo che il processo “salti” al tempo t_1 dallo stato x_0 verso x_1 , al tempo $t_1 + t_2$ da x_1 verso x_2 , e così via.

Legge del processo di Markov a salti

Consideriamo la probabilità di osservare un cammino che visiti nell'ordine gli stati $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n$.

- ▶ Trattandosi di tempi continui, dobbiamo specificare i *tempi di permanenza* in ciascuno stato t_1, \dots, t_n , in modo che il processo “salti” al tempo t_1 dallo stato x_0 verso x_1 , al tempo $t_1 + t_2$ da x_1 verso x_2 , e così via.
- ▶ Fissato δ tale che $t_1 = \delta h_1$, $t_2 = \delta h_2$ ecc., si trova l'approssimazione

$$\begin{aligned} P(X^\delta = (x_0 \rightarrow \dots \rightarrow x_n)) &\approx P(X_0 = x_0) \left(Id_{x_0 x_0} + \frac{t_1 L_{x_0 x_0}}{h_1} \right)^{h_1} \delta L_{x_0 x_1} \\ &\quad \cdot \left(Id_{x_1 x_1} + \frac{t_2 L_{x_1 x_1}}{h_2} \right)^{h_2} (\delta L_{x_1 x_2}) \cdot \dots \\ &\quad \cdot \left(Id_{x_{n-1} x_{n-1}} + \frac{t_n L_{x_{n-1} x_{n-1}}}{h_n} \right)^{h_n} (\delta L_{x_{n-1} x_n}). \end{aligned}$$

- Al tendere di δ a zero si trova che

$$\begin{aligned} P(X^\delta = (x_0 \rightarrow \dots \rightarrow x_n)) &\approx P(X_0 = x_0) \left(1 + \frac{t_1 L_{x_0 x_0}}{h_1}\right)^{h_1} (\delta L_{x_0 x_1}) \cdot \\ &\quad \cdot \left(1 + \frac{t_2 L_{x_1 x_1}}{h_2}\right)^{h_2} (\delta L_{x_1 x_2}) \cdot \dots \\ &\quad \cdot \left(1 + \frac{t_n L_{x_{n-1} x_{n-1}}}{h_n}\right)^{h_n} (\delta L_{x_{n-1} x_n}) \rightarrow 0 \end{aligned}$$

- ▶ Al tendere di δ a zero si trova che

$$\begin{aligned}
 P(X^\delta = (x_0 \rightarrow \dots \rightarrow x_n)) &\approx P(X_0 = x_0) \left(1 + \frac{t_1 L_{x_0 x_0}}{h_1}\right)^{h_1} (\delta L_{x_0 x_1}) \cdot \\
 &\quad \cdot \left(1 + \frac{t_2 L_{x_1 x_1}}{h_2}\right)^{h_2} (\delta L_{x_1 x_2}) \cdot \dots \\
 &\quad \cdot \left(1 + \frac{t_n L_{x_{n-1} x_{n-1}}}{h_n}\right)^{h_n} (\delta L_{x_{n-1} x_n}) \rightarrow 0
 \end{aligned}$$

- ▶ Dividendo per δ^n si ottiene una “densità di probabilità” non nulla, data dall’espressione

$$p(X = \gamma) = P(X_0 = x_0) \prod_{k=1}^n \exp(t_k L_{x_{k-1} x_{k-1}}) \prod_{k=1}^n L_{x_{k-1} x_k}.$$

La derivazione rigorosa si ottiene introducendo

- ▶ le variabili aleatorie T_1, T_2, \dots, T_n come tempi di permanenza del processo in ciascuno degli stati visitati x_0, x_1, \dots, x_{n-1}

La derivazione rigorosa si ottiene introducendo

- ▶ le variabili aleatorie T_1, T_2, \dots, T_n come tempi di permanenza del processo in ciascuno degli stati visitati x_0, x_1, \dots, x_{n-1}
- ▶ tempi di salto sono dati dalle somme $S_1 = T_1, S_2 = T_1 + T_2, \dots, S_n = T_1 + T_2 + \dots + T_n.$

La derivazione rigorosa si ottiene introducendo

- ▶ le variabili aleatorie T_1, T_2, \dots, T_n come tempi di permanenza del processo in ciascuno degli stati visitati x_0, x_1, \dots, x_{n-1}
- ▶ tempi di salto sono dati dalle somme $S_1 = T_1, S_2 = T_1 + T_2, \dots, S_n = T_1 + T_2 + \dots + T_n$.
- ▶ La densità è quindi

$$\begin{aligned} p(T_1 = t_1, X_{S_1} = x_1, T_2 = t_2, X_{S_2} = x_2, \dots, T_n = t_n, X_{S_n} = x_n) \\ = P(X_0 = x_0) \prod_{k=1}^n \exp(t_k L_{x_{k-1} x_k}) L_{x_{k-1} x_k}, \end{aligned}$$

La derivazione rigorosa si ottiene introducendo

- ▶ le variabili aleatorie T_1, T_2, \dots, T_n come tempi di permanenza del processo in ciascuno degli stati visitati x_0, x_1, \dots, x_{n-1}
- ▶ tempi di salto sono dati dalle somme $S_1 = T_1, S_2 = T_1 + T_2, \dots, S_n = T_1 + T_2 + \dots + T_n$.
- ▶ La densità è quindi

$$\begin{aligned} p(T_1 = t_1, X_{S_1} = x_1, T_2 = t_2, X_{S_2} = x_2, \dots, T_n = t_n, X_{S_n} = x_n) \\ = P(X_0 = x_0) \prod_{k=1}^n \exp(t_k L_{x_{k-1} x_k}) L_{x_{k-1} x_k}, \end{aligned}$$

- ▶ le variabili T_1, T_2, \dots, T_n sono continue, mentre le variabili $X_{S_1}, X_{S_2}, \dots, X_{S_n}$ sono discrete.

Una interpretazione alternativa dei processi a salti

La formula si può riscrivere separando le variabili continue da quelle discrete:

$$P(X_{S_1} = x_1, X_{S_2} = x_2 \dots, X_{S_n} = x_n) = P(X_0 = x_0) \prod_{k=1}^n \frac{L_{x_{k-1} x_k}}{-L_{x_{k-1} x_{k-1}}},$$

e

$$\begin{aligned} & p(T_1 = t_1, T_2 = t_2, \dots, T_n = t_n | X_{S_1} = x_1, \dots, X_{S_n} = x_n) \\ &= P(X_0 = x_0) \prod_{k=1}^n (-L_{x_{k-1} x_{k-1}}) \exp(t_k L_{x_{k-1} x_{k-1}}). \end{aligned}$$

$$P(X_{S_1} = x_1, X_{S_2} = x_2 \dots, X_{S_n} = x_n) = P(X_0 = x_0) \prod_{k=1}^n \frac{L_{x_{k-1}x_k}}{-L_{x_{k-1}x_{k-1}}},$$

- ▶ La prima equazione mostra che le variabili $X_0, X_{S_1}, \dots, X_{S_n}$ che indicano gli stati visitati dal processo sono una catena di Markov (a tempi discreti) con probabilità di transizione (per $x \neq y$)

$$Q_{xy} = \frac{L_{xy}}{-L_{xx}},$$

$$\begin{aligned} p(T_1 = t_1, T_2 = t_2, \dots, T_n = t_n | X_{S_1} = x_1, \dots, X_{S_n} = x_n) \\ = P(X_0 = x_0) \prod_{k=1}^n (-L_{x_{k-1}x_k}) \exp(t_k L_{x_{k-1}x_k}). \end{aligned}$$

- ▶ la seconda mostra che, noti gli stati visitati, i *tempi di permanenza* T_1, T_2, \dots, T_n sono variabili aleatorie indipendenti tra loro, e ciascuna T_k ha densità continua esponenziale di parametro $-L_{x_{k-1}x_k}$.

Simulazione di processi di Markov

- ▶ Catene: campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:

Simulazione di processi di Markov

- ▶ Catene: campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:
 - ▶ tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \rightarrow x_i}$

Simulazione di processi di Markov

- ▶ Catene: campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:
 - ▶ tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \rightarrow x_i}$
 - ▶ stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \rightarrow y}$, $y \neq x_i$

Simulazione di processi di Markov

- ▶ *Catene:* campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:
 - ▶ tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \rightarrow x_i}$
 - ▶ stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \rightarrow y}$, $y \neq x_i$
- ▶ *Processi a salti:* campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:

Simulazione di processi di Markov

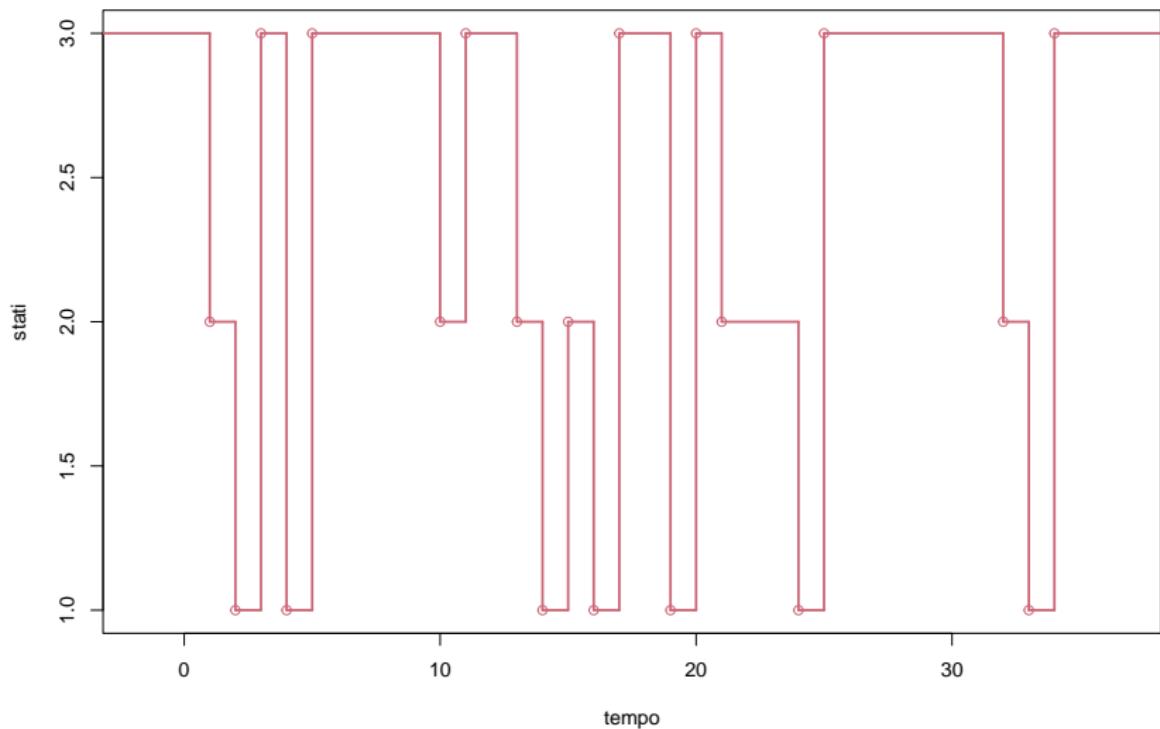
- ▶ *Catene:* campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:
 - ▶ tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \rightarrow x_i}$
 - ▶ stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \rightarrow y}$, $y \neq x_i$
- ▶ *Processi a salti:* campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:
 - ▶ tempo di permanenza T_i in X_i è v.a. esponenziale di parametro $-L_{x_i \rightarrow x_i}$

Simulazione di processi di Markov

- ▶ *Catene:* campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:
 - ▶ tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \rightarrow x_i}$
 - ▶ stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \rightarrow y}$, $y \neq x_i$
- ▶ *Processi a salti:* campionare X_0 con probabilità μ_0 , poi per $i \geq 0$:
 - ▶ tempo di permanenza T_i in X_i è v.a. esponenziale di parametro $-L_{x_i \rightarrow x_i}$
 - ▶ stato al tempo X_{T_i} è scelta con probabilità $\propto L_{x_i \rightarrow y}$, $y \neq x_i$

Esempi

traiettoria simulata (Catena di Markov)



traiettoria simulata (processo di Markov a salti)

