

# Appunti di Modelli Probabilistici

Simone Fabbri

Aggiornato al November 14, 2020

## Disclaimer

Non mi assumo responsabilita' per inesattezze o errori dei contenuti. Se anzi ne trovaste, sarei infinitamente grato se li segnalaste per mail a `simone.fabbri8 at studio.unibo.it` oppure tramite la pagina del repository [https://github.com/fabbrus97/mod\\_prob](https://github.com/fabbrus97/mod_prob).

# Contents

<b>1</b>	<b>Introduzione</b>	<b>1</b>
1.1	Notazione . . . . .	1
<b>2</b>	<b>Passeggiate aleatorie</b>	<b>3</b>
2.1	Schema di Bernoulli . . . . .	3
2.2	Passeggiata aleatoria . . . . .	3
2.3	Problema del ballottaggio . . . . .	5
2.4	Problema della rovina del giocatore (gambler's ruin) . . . . .	18
2.5	Processi di diramazione di Galton-Watson (branching processes)	21
2.6	Simulatori . . . . .	26
<b>3</b>	<b>Catene di Markov</b>	<b>29</b>
3.1	Catene di Markov omogenee . . . . .	29
3.1.1	Proprietà di Markov . . . . .	30
3.1.2	Modello di Ebembat . . . . .	33
3.1.3	Modello di Bernoulli-Laplace . . . . .	33
3.2	Classificazione degli stati . . . . .	34
3.2.1	Ricorrenza . . . . .	37
3.2.2	Verificare se uno stato è transiente o ricorrente . . . . .	38
3.3	Teorema ergodico per catene di Markov . . . . .	41
3.4	Modello di coda . . . . .	47
3.5	Reversibilità . . . . .	50
3.6	Metodo Montecarlo basato sulle catene di Markov (MCMC) . . . . .	54
3.6.1	Algoritmo di Metropolis-Hastings . . . . .	57
3.6.2	Campionatore di Gibbs (Gibbs sampler) . . . . .	58
<b>4</b>	<b>Catene di Markov con tempo continuo</b>	<b>59</b>
4.1	Distribuzione esponenziale . . . . .	59
4.1.1	Assenza di memoria . . . . .	61
4.1.2	Tasso di rischio . . . . .	62
4.1.3	Distribuzione iperesponenziale . . . . .	63

4.1.4	Distribuzione di Erlanger . . . . .	66
4.2	Processo di Poisson . . . . .	68
4.3	Decomposizione di un processo di Poisson . . . . .	71
4.4	Definizione equivalente di processo di Poisson . . . . .	73

# Chapter 1

## Introduzione

### 1.1 Notazione

**Evento** quantità che vale 1 se si verifica, 0 se non si verifica.

Data questa premessa, possiamo far corrispondere le operazioni insiemistiche sugli eventi a operazioni sui numeri  $E$  ed  $F$ ; possiamo associare a questi eventi una variabile aleatoria (che può assumere valori compresi tra 0 e 1):

<b>Unione</b> $E \cup F$	massimo tra E ed F, vale 1 se almeno un evento di verifica; somma logica
<b>Intersezione</b> $E \cap F$	è il minimo tra i due eventi; è detto anche prodotto logico
<b>Complementazione</b> $\tilde{E} = 1 - E$	si verifica se E non si verifica

**Attesa (expectation)** concetto associato ai ??? aleatori. Indicheremo con lo stesso simbolo sia l'attesa, sia la probabilità:

$$P(E) \qquad P(X)$$

con  $E$  un evento,  $X$  indica una variabile aleatoria.

Possiamo considerare un evento come una variabile aleatoria che può assumere solo i valori 0 e 1.



# Chapter 2

## Passeggiate aleatorie

### 2.1 Schema di Bernoulli

Effettuiamo prove indipendenti che possono dare due risultati sempre nelle stesse condizioni e che non si influenzano tra loro.

Supponiamo di avere una successione di eventi infiniti (**nota bene** negli esempi useremo un numero finito). Tutti gli eventi hanno la stessa probabilità:

$$P(E_i) = p$$

Quando si parla di schema di Bernoulli, bisogna specificare  $p$ .

Come formalizziamo l'altra ipotesi, ovvero che gli eventi siano indipendenti?

$$P(F_1, \dots, F_i) = p^k (1 - p)^{n-k}$$

con  $F_i = E_i$  oppure  $F_i = \tilde{E}_i$ ,  $p$  la costante associata allo schema di Bernoulli,  $k$  numero di volte in cui appare l'evento: in alcuni  $F_i$  infatti apparirà l'evento, in altri il complementare;  $n - k$  quindi è il numero di volte in cui appare l'evento complementare.

### 2.2 Passeggiata aleatoria

È il primo esempio che incontriamo di un **processo stocastico**.

Un processo stocastico è una successione di variabili aleatorie che dipende da un parametro, tipicamente il tempo. Il tempo nello specifico viene considerato a valori discreti.

Esempio: una particella che a  $t=0$  si trova nel punto 0, ad ogni istante fa un passo a destra con probabilità  $p$  oppure a sinistra con probabilità  $1 - p$ .

La probabilità di andare a destra (sinistra) è costante, e le probabilità di girare in una direzione sono indipendenti.

Dati  $E_1, E_2, E_3 \dots$  definiamo

$$x_i = 2E_i - 1$$

$$\text{se } E_i = 1 \Rightarrow x = 1,$$

$$\text{se } E_i = 0 \Rightarrow x = -1$$

1, -1 rappresentano lo spostamento nella nostra passeggiata aleatoria.

Chiamiamo  $S_n$  lo spostamento, che rappresenta la posizione al tempo  $n$ :

$$\begin{aligned} S_n &= \sum_{i=1}^n x_i \\ &= 2\left(\sum_{i=1}^n E_i\right) - n && \text{detto } z_n = \sum_{i=1}^n E_i \\ &= 2z_n - n \end{aligned}$$

Qual è la probabilità che ci può interessare? Dato un istante  $n$ , possiamo ad esempio chiederci la probabilità che la particella si trovi in una certa posizione (ovvero  $S_n = x$ ):

$$P(S_n = x)$$

Che possiamo riscrivere come:

$$P(2z_n - n = x) = P\left(z_n = \frac{n+x}{2}\right)$$

**Distribuzione binomiale** distribuzione del numero di successi, quando facciamo delle prove indipendenti.

Distribuzione binomiale di parametri  $n$  e  $p$ :

$$P(u = k) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$$

con  $n$  il numero delle prove e  $p$  la probabilità di successo.



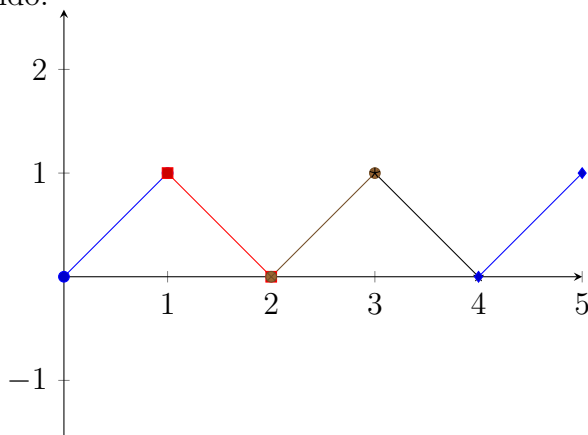
$\frac{n+x}{2}$  deve essere un numero intero, quindi  $n+x$  deve essere pari; quindi se  $n$  è pari (il tempo è pari)  $x$  deve essere pari, la particella cioè deve essere in una posizione pari:

$$\Rightarrow P(z_n = \frac{n+x}{2}) = ???$$

La passeggiata aleatoria è definita su numeri interi, sarebbe graficamente cioè una successione di punti, ma si può visualizzare come una linea spezzata unendo i punti.

**Passeggiata aleatoria simmetrica** è un caso interessante di passeggiata aleatoria, dove  $p = \frac{1}{2}$ : la probabilità di andare a sinistra o a destra è uguale.

Nel grafico, quando aumento di 1 la linea sale di 1; quando  $x$  vale -1, scendo.



Vogliamo calcolare la probabilità di un evento che si riferisce al cammino. Dobbiamo fare la somma di tutti i casi elementari dell'evento. Sia  $A$  un evento:

$$A = \frac{\text{numero dei casi elementari}}{2^n}$$

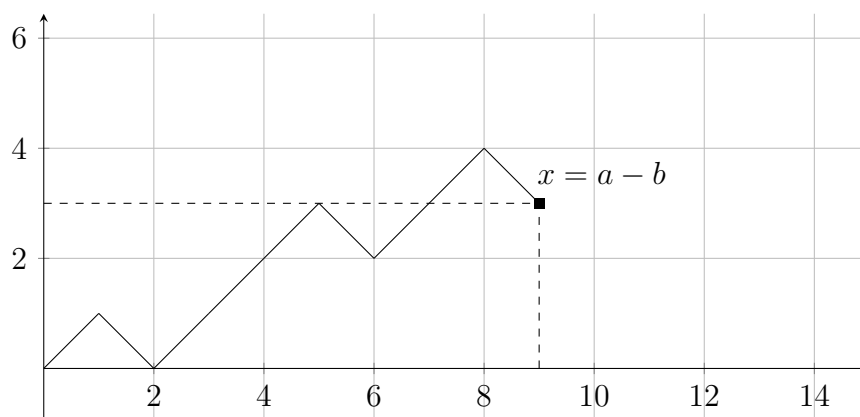
In questo caso quindi ci riduciamo ad un *problema combinatorio* (contare i casi elementari). Possiamo quindi calcolare la probabilità in modo combinatorio.

## 2.3 Problema del ballottaggio

Supponiamo che ci sia un'elezione con 2 candidati. L'urna dei voti conterrà voti a favore del candidato A oppure del candidato B; Sia  $n$  il numero dei

voti, e sia  $a$  il numero dei voti per  $A$  e  $b$  il numero dei voti per  $B$ . Supponiamo ancora  $a > b$ .

Mentre estraiamo le varie schede, possiamo disegnare una passeggiata aleatoria: estraendo un voto per  $A$  facciamo un passo a destra, per  $B$  a sinistra:



Dove sarà la linea spezzata quando avremo estratto tutte le schede? Quando avremo fatto  $a$  passi a destra e  $b$  a sinistra, la posizione sarà  $a - b$ .

Ci chiediamo ora: qual è la probabilità che **A** sia sempre in vantaggio?

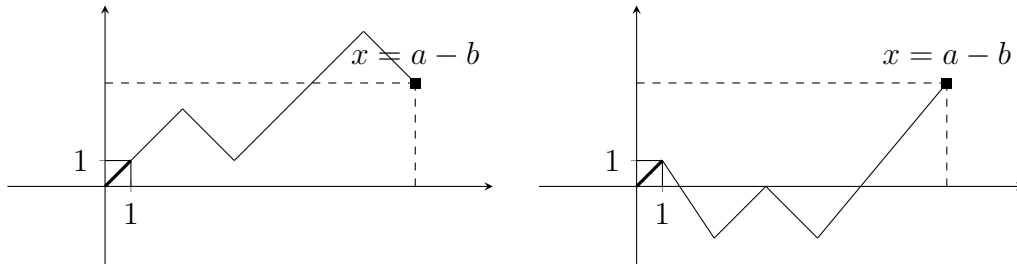
In questo caso non abbiamo tutti i cammini con probabilità uguale, perché alla fine andremo nel punto  $(n, x)$ . Quanti sono i cammini che finiscono nel punto  $x$ ?

Abbiamo  $n$  estrazioni e dobbiamo scegliere un sottoinsieme di  $a$  elementi. Questo è quello che si chiama una combinazione. Le possibili combinazioni sono:

$$\begin{aligned} \binom{n}{a} &= \frac{n!}{a!(n-a)!} \\ &= \binom{n}{b} = \frac{n!}{b!(n-b)!} \quad (n-b=a) \end{aligned}$$

Dobbiamo contare i cammini che sono sempre sopra l'asse delle  $x$ , che non toccano mai quest'asse.

**Principio di riflessione** com'è fatto un cammino che non tocca mai l'asse delle  $x$ ? Il primo passo sarà positivo. Il principio di riflessione dice che il numero dei cammini che partono da  $(1,1)$  è uguale al numero dei cammini che partono da  $(1,-1)$  [e arrivano a  $x$ ]; voglio allora contare, anziché i cammini che non passano per  $x$ , quelli che toccano l'asse:



L'idea è di riflettere specularmente ogni cammino che parte da  $(1,1)$  e tocca l'asse delle  $x$ .

Questa è una corrispondenza biunivoca.

Ci sono  $n - 1$  passi per partire da  $(1,1)$  e arrivare a  $x$ . Ci sono  $a - 1$  passi a destra negli  $n - 1$  cammini da  $(1,1)$ ; ci sono  $a$  passi a destra partendo invece da  $(-1,1)$ .

Il numero dei cammini è  $\binom{n-1}{a}$

Quanti sono i cammini che non toccano? Togliamo da tutti i cammini quelli che toccano:

$$\binom{n-1}{a-1} - \binom{n-1}{a}$$

Per calcolare la probabilità, dividiamo la quantità appena calcolata per il numero totale dei cammini:

$$\begin{aligned} & \frac{\binom{n-1}{a-1} - \binom{n-1}{a}}{\binom{n}{a}} = \\ & = \frac{\frac{(n-1)!}{(a-1)!(n-a)!} - \frac{(n-1)!}{a!(n-a-1)!}}{\frac{(n)!}{a!(n-a)!}} = \\ & = \frac{1}{n}(a-b) = \\ & = \frac{a-b}{a+b} \end{aligned}$$

E questa è la probabilità che il candidato A sia sempre in vantaggio su B.

Applichiamo questo risultato alla passeggiata aleatoria simmetrica. Supponiamo  $S_0, S_1, \dots, S_{2n}$ :

$$u_{2n} = P(S_{2n} = 0) = \binom{2n}{n} \frac{1}{2^{2n}}$$

è la probabilità di ritrovarsi all'origine (ovvero, di fare lo stesso numero di passi sia a destra che a sinistra).  $2^{2n}$  è il numero totale delle passeggiate possibili.

Cosa succede quando  $n$  cresce?

### Formula di Stirling

$$n! = \sqrt{2\pi} \cdot n^{n+\frac{1}{2}} \cdot e^{-n} \cdot e^{\theta_n} \quad |\Theta_n| < \frac{1}{12n}$$

Fornisce un'approssimazione per il fattoriale;  $\Theta_n$  è un errore che tende a 0.

Se  $n$  tende a  $\infty$ ,  $\Theta_n \rightarrow 0$  e quindi  $e^{\theta_n} \rightarrow 1$ .

Usando la formula di Stirling:

$$\begin{aligned} \binom{2n}{n} \frac{1}{2^{2n}} &= \frac{2n!}{(n!)^2} \frac{1}{2^{2n}} = \\ &= \frac{\sqrt{2\pi}(2n)^{2n+\frac{1}{2}} e^{-2n} e^{\theta_{2n}}}{(2\pi)n^{2n+1} e^{-2n} e^{2\theta_n}} \cdot \frac{1}{2^{2n}} = \\ &= \frac{\sqrt{2} \cdot e^{\theta_{2n}-2\theta_n}}{\sqrt{2\pi}\sqrt{n}} = \frac{1}{\sqrt{\pi n}} e^{\theta_{2n}-2\theta_n} \end{aligned}$$

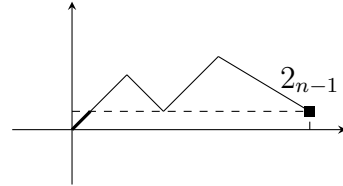
$e^{\theta_{2n}-2\theta_n}$  tende a 1. Il tutto tende a 0, ma abbastanza lentamente.

Chiediamoci ora qual è la probabilità che la passeggiata aleatoria torni a 0 nel tempo  $2n$  e sia sempre positiva prima di  $2n$ :

$$P(S_1 > 0, \dots, S_{2n-1} > 0, S_{2n} = 0)$$

Perché si verifichi occorre che  $S_{2n-1} = 1$ .

Possiamo contare i cammini che partono da 0 e al tempo  $2n-1$  sono a 1, rimanendo sempre positivi. Lo abbiamo visto per il teorema del ballottaggio.



È come avere  $2n-1$  schede e  $n$  voti per A, e  $n-1$  voti per B. Quindi, sapendo che la formula è  $\frac{a-b}{a+b}$  e che  $a$  corrisponde a  $n$  e  $b$  corrisponde a  $n-1$ , per ottenere il numero dei cammini che partono dall'origine e ritornano a 0 in  $2n$  passi e sono sempre positivi possiamo calcolare:

$$\begin{aligned} \frac{1}{2n-1} \cdot \binom{2n-1}{n} &= \frac{(2n-1)!}{n!(n-1)!(2n-1)} = \\ &= \frac{(2n-2)!}{n(n-1)!^2} = \frac{1}{n} \binom{2n-2}{n-1} \end{aligned}$$

Per calcolare la probabilità dobbiamo dividere questa quantità per il numero dei cammini totali:

$$\begin{aligned} P(S_1 > 0, \dots, S_{n-1} > 0, S_n = 0) &= \\ = \frac{1}{n} \binom{2n-2}{n-1} \frac{1}{2^{2n}} &= \frac{1}{n} \binom{2n-2}{n-1} \frac{1}{2^{2n-2}} \frac{1}{4} = \\ &= \frac{1}{4n} u_{2n-2} \end{aligned}$$

Quindi nella passeggiata aleatoria per calcolare la probabilità di un evento che si riferisce ai primi  $2n$  passi, siccome i cammini hanno la stessa probabilità  $p$ , si riduce a contare quanti sono i cammini. Abbiamo visto che

$$P(S_1 > 0, \dots, S_{n-1} > 0, S_n = 0) = \frac{1}{n} u_{2n-2} = \frac{1}{n} \binom{2n-2}{n-1} \frac{1}{n^{n-2}}$$

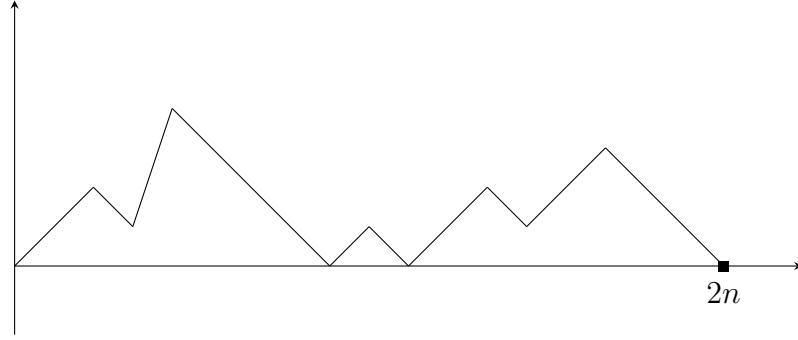
Definiamo  $L_{2n} = \frac{1}{n+1} \binom{2n}{n}$ ; allora possiamo riscrivere l'equazione sopra come

$$P(\dots) = L_{2n-2} \frac{1}{2^{2n-2}}$$

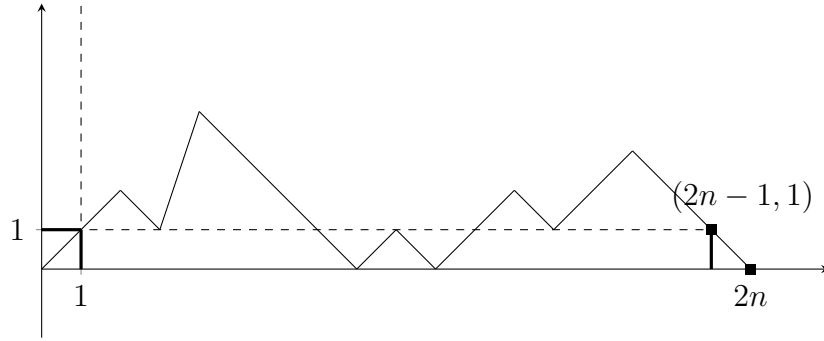
Vogliamo ora calcolare la probabilità di questo evento:

$$P(S_1 \geq 0, S_2 \geq 0, \dots, S_{2n-1} \geq 0, S_{2n} = 0)$$

Dobbiamo calcolare i cammini che rimangono maggiori o uguali a 0; la nostra passeggiata sarà di questo tipo:



Ma possiamo ridurre al problema precedente, spostando l'origine nel punto  $(1, 1)$ :



Il cammino spostato sarà  $\geq 0$ . Possiamo fare una corrispondenza tra i cammini compresi tra 0 e  $2n$  di lunghezza  $2n$ , e i cammini compresi tra  $(1, 1)$  e  $(2n-1, 1)$  di lunghezza  $2n-2$ .

Il numero di cammini tra 0 e  $2n$  sempre positivi era dato da  $\frac{1}{n} \binom{2n-2}{n-1} = L_{2n-2}$ ; quindi la probabilità sarà:

$$P(S_1 \geq 0, S_{2n} = 0) = L_{2n} \frac{1}{2^{2n}} = \frac{u_{2n}}{n+1}$$

**Teorema** Dato  $u_{2n} = \binom{2n}{n} \frac{1}{2^{2n}}$ , allora:

1.  $u_{2n} = P(S_{2n} = 0)$
2.  $u_{2n} = P(S_1 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} \neq 0)$
3.  $u_{2n} = P(S_1 \geq 0, \dots, S_{2n} \geq 0)$

Siccome abbiamo già visto 1, vogliamo dimostrare 2 e 3.

Dimostriamo **2**; definiamo

$$f_{2n} = P(S_1 \neq 0, S_2 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} = 0)$$

Avevamo precedentemente trovato la probabilità che il cammino fosse sempre positivo; per simmetria, è uguale alla probabilità che sia sempre negativo. Allora

$$f_{2n} = 2P(S_1 > 0, \dots, S_{2n-1} < 0, S_{2n} = 0) = 2 \frac{u_{2n} - 2}{n}$$

Consideriamo l'evento complementare di  $S_1 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} \neq 0$ ; è l'evento che qualche volta tra 0 e  $2n$ ,  $S_i$  valga 0:

$$P(S_1 \neq 0, \dots) = 1 - f_2 - f_4 - \dots - f_{2n}$$

con  $f_n$  la probabilità che  $S$  sia tornato a 0 nel tempo  $n$ .

**Osservazione:**

$$\overbrace{\binom{2n-2}{n-1} \frac{1}{2^{2n-2}} - \binom{2n}{n} \frac{1}{2^{2n}}}^{u_{2n-2} - u_{2n}} = f_{2n}$$

$$\begin{aligned} \Rightarrow f_{2n} &= P(S_1 \neq 0, S_2 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} = 0) = \\ &= u_{2n-2} \left( 1 - \frac{1}{4} \frac{(2n)(2n-1)}{n^2} \right) = \\ &= u_{2n-2} \left( \frac{2n-2n+1}{2n} \right) = \\ &= \frac{u_{2n-2}}{2n} = f_{2n} \end{aligned}$$

$$\Rightarrow 1 - f_2 - f_4 - \dots - f_{2n} = 1 - (u_0 - u_2) - (u_2 - u_4) - \dots - (u_{2n-2} - u_{2n}) = u_{2n}. \quad \square \text{ QED}$$

$u_0$  per definizione è 1;  $u_2$  si cancella con  $u_2$  (gli  $u_i$  si cancellano due alla volta).

Interpretazione: sappiamo, per Stirling, che possiamo approssimare  $u_{2n} \approx \frac{1}{\sqrt{\pi n}}$ . Se prendiamo  $n$  grande, la probabilità che non si torni mai a 0 tende a 0. Ad esempio, se prendiamo 2 giocatori A, B che tirano una moneta, e se esce testa A vince 1 euro (e B perde), mentre se esce croce A perde un euro (e B vince), la probabilità che il gioco non termini mai e che i giocatori vanno avanti all'infinito tende a 0.

Dimostriamo **3**; vogliamo dimostrare che  $u_{2n} = P(S_1 \geq 0, \dots, S_{2n} \geq 0)$ , che rappresenta l'evento che il primo giocatore (sempre nell'esempio dei due giocatori che tirano una moneta) sia sempre in vantaggio o in parità.

Definiamo

$$f_{2n} = P(S_1 \geq 0, \dots, S_{2n-2} \geq 0, S_{2n-1} < 0)$$

in base al conteggio fatto precedentemente sul numero dei cammini  $\geq 0$  fino a  $2n - 2$ , questa quantità è pari a:

$$f_{2n} = \frac{1}{2^{2n} u_{2n-2} \frac{1}{2n}}$$

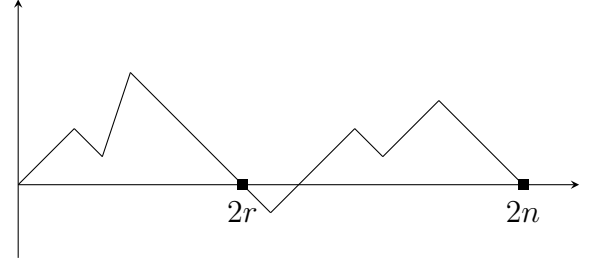
Consideriamo l'evento complementare di  $f_{2n}$ , cioè che ci sia stato un punto negativo. I punti negativi possono essere solo nei punti dispari:

$$\begin{aligned} P(S_1 \geq 0, \dots, S_{2n-2} \geq 0, S_{2n-1} < 0) &= \\ &= 1 - f_2 - \dots - f_{2n} = 1 - (u_0 - u_2) - \dots - (u_{2n-2} - u_{2n}) \\ &= u_{2n} \quad \square \text{ QED} \end{aligned}$$

**Equazione del rinnovamento** c'è una relazione che collega  $u_{2n}$  e  $f_{2n}$ , si chiama *equazione del rinnovamento* (renewal equation):

$$u_{2n} = \sum_{r=1}^n f_{2r} u_{2n-2r}$$

Se abbiamo un cammino che va da 0 a  $2n$ , si può decomporre in due parti: quella che va da 0 a 0 per la prima volta, e da lì a  $2n$ . Chiamiamo il primo tempo  $2r$ .



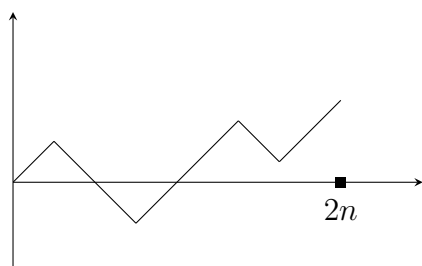
Il numero totale dei cammini è  $\binom{2n}{n}$

$$u_{2n} 2^{2n} \binom{2n}{n} = \sum_{r=1}^n f_{2r} 2^{2r} u_{2n-2r} 2^{2n-2r}$$

se dividiamo a sinistra e a destra per  $2^{2n}$  otteniamo l'equazione di rinnovamento.



Abbiamo due giocatori, A e B, che tirano una moneta: se esce testa, A vince un euro (e B perde), viceversa A perde un euro (e B vince); ci chiediamo qual è la percentuale di tempo in cui uno dei due è in vantaggio. Si potrebbe pensare che sia A che B rimangano in vantaggio per circa il 50% del tempo. In realtà, quando  $N$  (numero di lanci) è grande, è più probabile che un solo giocatore rimanga in vantaggio tutto il tempo.



Come definiamo il tempo in cui un giocatore è in vantaggio sull'altro? Se rappresentiamo nel grafico la passeggiata aleatoria, essa sarà una linea spezzata; ma invece di considerare i tempi discreti, immaginiamo che la linea sia il grafico di una funzione.

Consideriamo la lunghezza dell'intervallo in cui la funzione è positiva. Definiamo

$P_{2k,2n} =$  probabilità che il primo giocatore sia in vantaggio per un tempo  $2k$

È stato dimostrato (Fenner) che  $P_{2k,2n} = u_{2k}u_{2n-2k}$ . Si dimostra per induzione su  $n$ :

**Caso base:**  $p_{0,2n}$  = probabilità che il primo giocatore sia sempre in vantaggio o pari ( $S \geq 0$ )

$$p_{0,2n} = P(S_1 \geq 0, \dots, S_{2n} \geq 0) = u_{2n} \quad \square$$

**Caso induttivo:** supponiamo che l'uguaglianza valga per  $n-1$ ; dimostriamo che vale per  $n$ .

Sia  $1 \leq k \leq n-1$ . Ci sarà un intervallo in cui il grafico è sopra e uno in cui il grafico della passeggiata aleatoria è sotto l'asse x; ad un certo punto quindi passerà per l'asse x ( $y=0$ ). Quindi, il primo intervallo può essere positivo o negativo; il secondo sarà l'opposto.

$$p_{2k,2n} = \underbrace{\sum_{r=1}^k \frac{1}{2} f_{2r} p_{2k-2r, 2n-2r}}_{\substack{\text{probabilità di tornare a 0 per} \\ \text{la prima volta al tempo } 2r; \text{ il} \\ \text{cammino può essere positivo} \\ \text{o negativo, per questo} \\ \text{moltiplichiamo per } \frac{1}{2}}} + \sum_{r=1}^{n-k} \frac{1}{2} f_{2r} p_{2k, 2n-2r} =$$

$$= \frac{1}{2} \sum_{r=1}^k f_{2r} u_{2k-2r} u_{2n-2k} + \frac{1}{2} \sum_{r=1}^{n-k} f_{2r} u_{2k} u(2r-2k) - 2r$$

Applichiamo la formula del rinnovamento:

$$\begin{aligned} &= \frac{1}{2} u_{2n-2} + \sum_{r=1}^k f_{2n} u_{2k-2r} + \frac{1}{2} u_{2k} \sum_{r=1}^{n-k} f_{2r} u(2(n-k) - 2r) \\ &= \frac{1}{2} u_{2n-2k} u_{2k} + \frac{1}{2} u_{2k} u_{2n-2k} \\ &= u_{2k} u_{2n-2k} \quad \square \end{aligned}$$

**Esempio** consideriamo i soliti 2 giocatori che scommettono su testa o croce; supponiamo 20 lanci di moneta.

$p_{2k,20}$  è la probabilità che il primo giocatore sia in vantaggio esattamente un tempo  $2k$ .

$P_{2k,20}$  è la probabilità che il primo giocatore sia in vantaggio almeno un tempo  $2k$ .

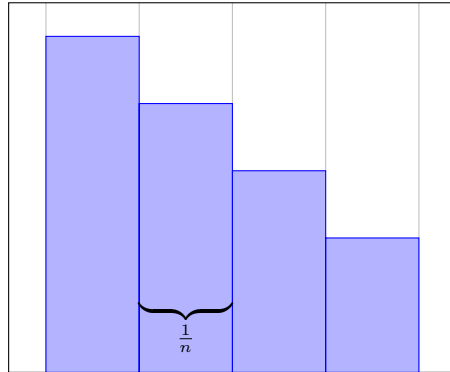
	$k = 0, k = 10$	$k = 1, k = 9$	$k = 2, k = 8$	$k = 3, k = 7$	$k = 4, k = 6$	$k = 5$
p	0,1702	0,0927	0.0736	0.0655	0.0617	0.064
P	0.3504	0.5379	0.6851	0.8160	0.9394	1

Possiamo ricavare una legge generale, facendo tendere  $n$  all'infinito e osservando come si comporta la quantità  $p_{2k,2n}$ .

Anziché  $k$  consideriamo  $\frac{k}{n} = x$ , cioè consideriamo la percentuale del tempo in cui il primo giocatore è in vantaggio.  $x$  assume valori discreti:

$$x = 0, \frac{1}{n}, \dots, 1$$

Un modo per rappresentare la distribuzione di probabilità è attraverso l'istogramma:



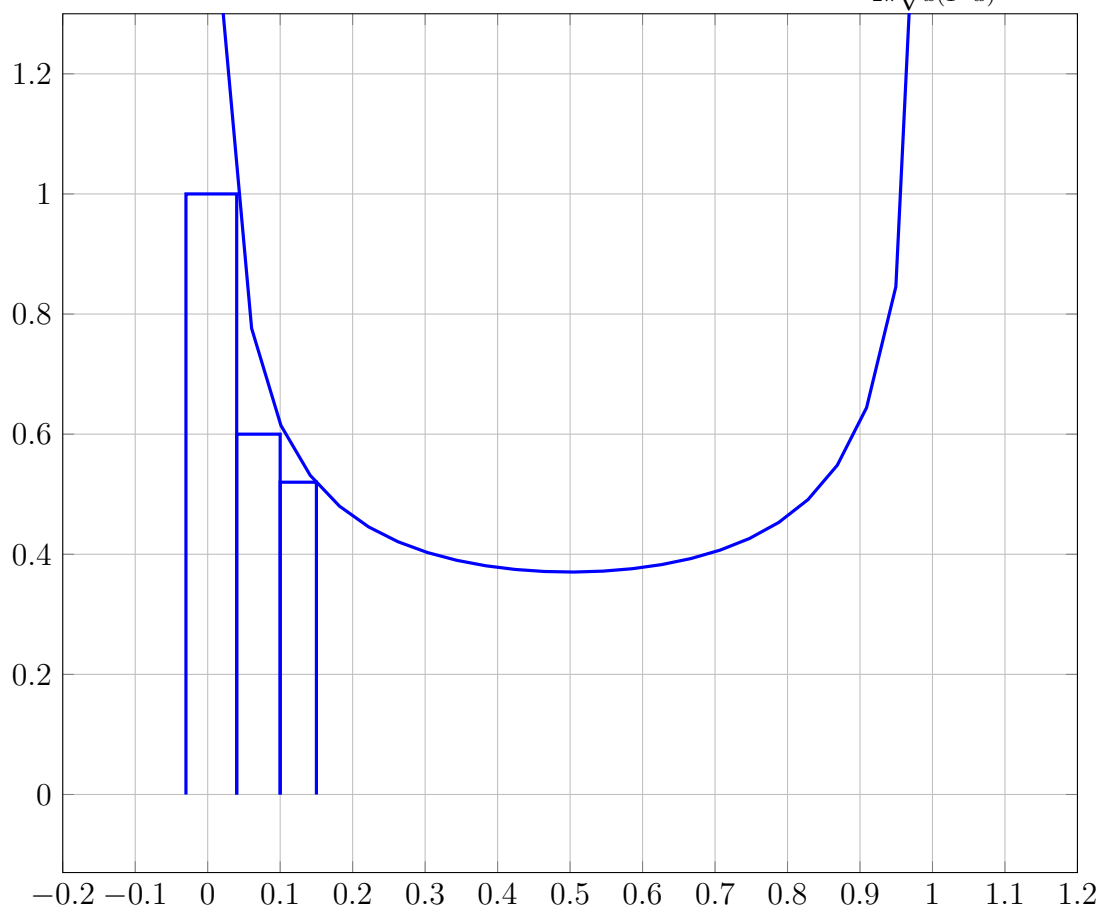
L'area del rettangolo è uguale alla probabilità che assuma un certo valore. Abbiamo visto per Stirling che  $u_{2n} \approx \frac{1}{\sqrt{\pi n}}$ . Supponiamo  $n$  grande e  $0 \leq k \leq n$ ; allora possiamo scrivere:

$$u_{2k} u_{2n-2k} \approx \frac{1}{2\pi} \frac{1}{\sqrt{k(n-k)}}$$

Riscriviamo la funzione in termini di  $x$  e chiamiamola  $g$ :

$$g(x) \approx \frac{1}{n2\pi} \frac{1}{\sqrt{x(1-x)}}$$

Da questa formula si vede che l'istogramma è vicino alla funzione  $\frac{1}{2\pi\sqrt{x(1-x)}}$ :



Dobbiamo fare la somma di tutti i rettangoli sotto la funzione  $\frac{1}{2\pi\sqrt{x(1-x)}}$ . Questa è la somma di Riemann, un'approssimazione dell'integrale:

$$\begin{aligned} &\approx \sum_{\frac{1}{2} < x < z} \frac{1}{n2\pi} \frac{1}{\sqrt{x(1-x)}} \\ &\rightarrow_{n \rightarrow \infty} \int_{\frac{1}{2}}^z \frac{1}{\pi\sqrt{x(1-x)}} dx \end{aligned}$$

Con riferimento al grafico precedente, si vede che quando  $x$  tende a 0 oppure a 1, la probabilità tende all'infinito, mentre a  $\frac{1}{2}$  è nel punto più basso.

Questa è detta **prima legge dell'arcoseno** perché la primitiva dell'integrale è uguale a:

$$\int \frac{1}{\pi \sqrt{x(1-x)}} dx = 2\pi^{-1} \arcsin \sqrt{x} + c$$

La prima legge dell'arcoseno è un'approssimazione della probabilità che uno dei due giocatori sia in vantaggio una certa percentuale di volte, per un  $n$  grande.

Questo vale per la passeggiata aleatoria asimmetrica libera.

Vediamo un caso simile: consideriamo la probabilità che il primo giocatore sia in vantaggio per un tempo  $2k$  e che a  $2n$  la passeggiata sia 0, ovvero i due giocatori siano in pareggio; definiamo

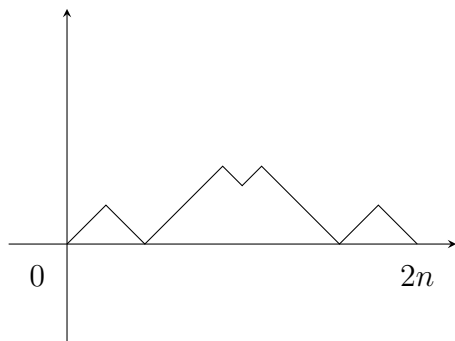
$$\left(\frac{1}{2^{2n}}\right)(L_{2k,2n}) = \begin{array}{l} \text{probabilità che il primo giocatore sia in} \\ \text{vantaggio per un tempo } 2k \text{ e che } s_{2n} = 0 \end{array}$$

$L_{2k,2n}$  è il numero dei cammini; dividendolo per il numero totale dei cammini  $2^{2n}$  dà la probabilità.

Allora  $L_{2k,2n} = L_{2n}$ , e avevamo definito  $L_{2n} = \frac{1}{n+1} \binom{2n}{n}$ .

È possibile dimostrare che il numero dei cammini non dipende da  $k$ , ma è lo stesso per tutti i valori di  $k$ ; qualsiasi percentuale ha la stessa probabilità.

**Dimostrazione** per induzione su  $n$ :



Supponiamo sia vero per  $n-1$ ; vogliamo vedere che la proprietà vale anche per  $n$ . Supponiamo  $k=n$ . Dobbiamo contare i cammini che tornano a 0 e sono sempre  $\geq 0$ ; ma lo abbiamo già visto: i cammini sono  $L_{2n,2n} = \frac{1}{n+1} \binom{2n}{n}$ .

Anche il caso  $k=0$  lo abbiamo già visto.

Supponiamo  $1 \leq k \leq n-1$ . Questo cammino deve passare per forza per 0, perché una parte del tempo è in vantaggio il primo giocatore, l'altra il

secondo. Sia  $2r$  il primo istante in cui la passeggiata passa per lo 0. Fino al tempo  $2r$ , il cammino può essere positivo o negativo.

$$L_{2k,2n} = \sum_{r=1}^k L_{2r-2} L_{2n-2r} + \sum_{r=1}^{n-k} L_{2r-2} L_{2n-2r}$$

Modifichiamo la variabile  $r$  al secondo membro, con  $\rho = n - r + 1$ . Quindi

$$\begin{array}{ll} \text{per } r = n - k, & \rho = k + 1 \\ \text{per } r = 1, & \rho = n \end{array}$$

$$\begin{aligned} &= \sum_{r=1}^k L_{2n-2} L_{2n-2r} + \sum_{\rho=k+1}^n L_{2\rho-2} L_{2n-2\rho} \\ &= \sum_{r=1}^n L_{2r-2} L_{2n-2r} \quad \square \end{aligned}$$

Abbiamo dimostrato che  $L_{2k,2n}$  non dipende da  $k$ , ma è uguale per tutti i valori. Ma quanto vale effettivamente?

$$L_{0,2n} = L_{2n,2n} = L_{2n}$$

$L_{k,2n}$  per  $k = 1, n = 1$  sono uguali

Supponiamo che:

$$\sum_{k=0}^n L_{2k,2n} = \binom{2n}{n}$$

facendo variare  $k$  da 0 a  $n$  otteniamo tutti i cammini di lunghezza  $2n$

Supponiamo  $1 \leq k \leq n - 1$ :

$$\begin{aligned} L_{2k,2n} &= \frac{\binom{2n}{n} - 2L_{2n}}{n - 1} = \\ &= \frac{\binom{2n}{n} - \frac{2\binom{2n}{n}}{n+1}}{n - 1} = \binom{2n}{n} \frac{1 - \frac{2}{n+1}}{n - 1} \\ &= \binom{2n}{n} \frac{\frac{n+1-2}{n+1}}{n - 1} = \frac{\frac{2n}{n}}{n + 1} \\ &= L_{2n} \end{aligned}$$

## 2.4 Problema della rovina del giocatore (gambler's ruin)

Consideriamo un altro problema, la rovina del giocatore. È sempre basato su una passeggiata aleatoria. Consideriamo 2 giocatori A e B che tirano una moneta; se esce testa, A vince 1 euro; se esce croce, perde 1 euro (e lo vince B). Ci chiediamo quale sia il capitale del primo giocatore dopo un certo numero di giocate.

Supponiamo che al tempo 0 il capitale iniziale di A sia  $x$ , mentre il capitale di B ammonti a  $N - x$ . Il capitale totale (capitale di A + capitale di B) ammonta quindi a  $N$ . Se il capitale di uno dei due giocatori arriva a 0, il gioco si ferma (e si ha la rovina del giocatore).

Considereremo anche il caso della **moneta non simmetrica**, dove cioè la probabilità  $p$  che esca testa è  $0 < p < 1$ . I lanci sono indipendenti tra loro.

Ci chiediamo qual è la probabilità che dopo un certo numero di giocate A oppure B sia rovinato. Non supponiamo ci sia un numero finito di lanci, il gioco potrebbe teoricamente continuare all'infinito.

Definiamo le seguenti quantità:

$E_i =$  i-esimo lancio è testa (evento)

$$P(E_1^*, \dots, E_n^*) = p^k (1 - p)^{n-k}$$

$$E_i^* = \begin{cases} E_i, & \text{se l'evento è verificato} \\ \tilde{E}_i, & \text{altrimenti} \end{cases} \quad (2.1)$$

Con  $k$  numero degli eventi  $E_i$  e  $n - k$  numero degli eventi  $\tilde{E}_i$ .

Questo schema di Bernoulli formalizza il fatto che i lanci siano indipendenti.

Definiamo

$u_x =$  probabilità di rovina del giocatore A con capitale iniziale  $x$ , contro B con capitale iniziale  $N - x$

**Caso simmetrico**  $p = \frac{1}{2}$

Supponiamo  $0 < x < N$ . Osserviamo in generale che  $u_0 = 1$  e  $u_N = 0$ .

## 2.4. PROBLEMA DELLA ROVINA DEL GIOCATORE (GAMBLER'S RUIN) 19

Possiamo scrivere  $u_x$  come la probabilità di rovina del giocatore sapendo che il primo lancio ha dato testa sommata alla probabilità di rovina del giocatore sapendo che il primo lancio ha dato croce:

$$\begin{aligned} u_x &= p \cdot u_{x+1} + (1-p) \cdot u_{x-1} \\ &= \frac{1}{2}u_{x-1} + \frac{1}{2}u_{x+1} = \frac{1}{2}(u_{x+1} - u_x) = \frac{1}{2}(u_x - u_{x-1}) \\ u_{x+1} - u_x &= c \quad (\text{la differenza tra due valori successivi è sempre uguale, la chiamiamo } c) \end{aligned}$$

$$\begin{aligned} 0 &= u_n = u_0 + \sum_{x=0}^{N-1} (u_{x+1} - u_x) \\ 0 &= u_n = 1 + N_c \Rightarrow c = -\frac{1}{N} \\ u_x &= u_0 + \sum_{y=0}^{x-1} (u_{y+1} - u_y) = 1 - \frac{x}{N} \end{aligned}$$

Quindi se  $x \rightarrow N$ , la probabilità che A si rovini tende a 0. Cosa succede se  $N$  tende all'infinito? A parte da  $x$ , e B da  $N - x$ . La probabilità di rovina in questo caso tende a 1.

La probabilità che il gioco continui all'infinito è 0.

**Caso non simmetrico**  $0 < p < 1$  Nel caso  $p = \frac{1}{2}$  le differenze erano uguali (al valore  $c$ ). Nel caso non simmetrico questo non vale. Occorre introdurre la **somma geometrica**:

supponiamo  $\alpha \neq 1$ :

$$\begin{aligned} S &= \sum_{k=0}^n \alpha^k = \frac{\alpha^{n+1} - 1}{\alpha - 1} \\ \alpha S &= S - 1 + \alpha^{n+1} - 1 \\ (\alpha - 1)S &= \alpha^{n+1} - 1 \\ S &= \frac{\alpha^{n+1} - 1}{\alpha - 1} \quad \begin{cases} \alpha < 1, & \text{la serie è convergente} \\ \alpha > 1, & \text{la serie va avanti all'infinito} \end{cases} \end{aligned}$$

Tornando a noi, abbiamo detto che nel caso simmetrico  $0 < p < 1$ ; se  $p > \frac{1}{2}$  A è favorito, altrimenti è favorito B. Sia ancora  $u_x$  la probabilità di rovina di A, e scriviamola ancora come la somma delle probabilità della rovina del giocatore se il primo lancio è testa oppure è croce:

$$0 < x < N \quad u_x = p \cdot u_{x+1} + q \cdot u_{x-1}$$

Con  $q = 1 - p$ .

$$(p + q)u_x = pu_x + qu_x = pu_{x+1} + qu_{x-1}$$

$$q(u_x - u_{x-1}) = p(u_{x+1} - u_x)$$

$$u_{x+1} - u_x = \frac{q}{p}(u_x - u_{x-1})$$

$$u_{k+1} - u_k = \left(\frac{q}{p}\right)^k(u_1 - u_0)$$

$$u_0 = 1 \text{ (rovina all'istante iniziale), } u_N = 0$$

$$\begin{aligned} u_k &= u_0 + \sum_{j=0}^{k-1} (u_{j+1} - u_j) \\ &= 1 + \sum_{j=0}^{k-1} \left(\frac{q}{p}\right)^j (u_1 - u_0) \\ &= 1 + (u_1 - u_0) \sum_{j=0}^{k-1} \left(\frac{q}{p}\right)^j \end{aligned}$$

possiamo applicare la formula della somma geometrica ( $p \neq \frac{1}{2} \Rightarrow \frac{q}{p} \neq 1$ ):

$$= 1 + (u_1 - u_0) \frac{\left(\frac{q}{p}\right)^k - 1}{\frac{q}{p} - 1}$$

Applichiamo la formula per  $k = N$ , sapendo che  $u_N = 0$ :

$$\begin{aligned} 0 &= 1 + (u_1 - u_0) \frac{\left(\frac{q}{p}\right)^N - 1}{\frac{q}{p} - 1} \\ u_1 - u_0 &= - \frac{\frac{q}{p} - 1}{\left(\frac{q}{p}\right)^N - 1} \end{aligned}$$

Possiamo calcolare  $u_k$  in generale:

$$u_k = 1 - \frac{\left(\frac{q}{p}\right)^k - 1}{\left(\frac{q}{p}\right)^N - 1}$$



che è la probabilità di rovina di A che parte con capitale iniziale pari a  $k$ .

Consideriamo ora  $v_k$  la probabilità di rovina di B; in questo caso B vince con probabilità  $q$ .

$$v_k = 1 - \frac{\left(\frac{p}{q}\right)^{N-k} - 1}{\left(\frac{p}{q}\right)^N - 1}$$

$v_k + u_k = 1$ . La probabilità che il gioco continui all'infinito è 0 (uno dei due giocatori va in rovina prima o poi). Cosa succede se  $N \rightarrow \infty$  (con  $p \neq \frac{1}{2}$ )? Supponiamo  $p > \frac{1}{2}$ .

$$\frac{q}{p} < 1 \quad u_k \xrightarrow{N \rightarrow \infty} \left(\frac{q}{p}\right)^k \quad \left(\text{perché } \left(\frac{q}{p}\right)^N \rightarrow 0, \text{ perché } \frac{q}{p} < 1\right)$$

Quindi se A è favorito, la sua probabilità di rovina è minore di 1, e più è grande  $k$  e più si abbassa.

Se  $p < \frac{1}{2}$ :  $\frac{q}{p} > 1$  e  $u_k \xrightarrow{N \rightarrow \infty} 1$ .

## 2.5 Processi di diramazione di Galton-Watson (branching processes)

Sono esempi particolari delle catene di Markov, un modello più generale. Servono per studiare l'evoluzione di una popolazione. È un modello molto semplificato; partiamo supponendo di avere un solo individuo che avrà un certo numero di figli.

Supponiamo che ogni individuo abbia le stesse probabilità di avere figli, indipendentemente dagli altri individui, e gli intervalli (numero di figli) siano regolari. Consideriamo una successione di variabili aleatorie, e vogliamo vedere quanti figli ci sono nelle varie generazioni.

Definiamo:

$x_i =$  numero di figli della generazione  $i$ -esima

$p_k =$  probabilità di avere  $k$  figli per ogni individuo;  $\sum_{k=0}^{\infty} p_k = 1$  (sono eventi incompatibili, è una partizione).

Supponiamo che l'attesa del numero di figli sia finita; la indichiamo con  $\mu$ :

$$\mu = \sum_{k=0}^{\infty} k \cdot p_k < +\infty$$

Quindi abbiamo supposto  $x_0 = 1$ . Qual è l'attesa della  $k$ -esima generazione?

**Formula delle probabilità totali** date due variabili aleatorie  $A$  e  $B$ , possiamo chiederci quanto vale l'*attesa condizionata*, ovvero la probabilità che si verifichi  $A$  sapendo che  $B$  assume un certo valore  $x$ :

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

Possiamo conoscere l'attesa non condizionata con la formula delle probabilità totali:

$$P(X) = \sum_{x=0}^{\infty} P(A|B = x)P(B = x)$$

Siano  $x$  e  $y$  due variabili aleatorie a valori interi. Ci possiamo chiedere l'attesa di  $x$  quando  $y$  assume certi valori:  $P(x|y = j)$ ; usando la formula delle probabilità totali:

$$P(x) = \sum_{j=0}^{\infty} P(x|y = j)P(y = j)$$

Vogliamo calcolare l'attesa di  $k$ :

$$\begin{aligned} P(x_k) &= \sum_{j=0}^{\infty} P(x_k|x_1 = j)P(x_1 = j) \\ &= \sum_{j=0}^{\infty} j p_j P(x_{k-1}) \\ &= \mu P(x_{k-1}) \end{aligned}$$

Possiamo dividere la  $j$ -esima generazione nei figli dell'individuo 0; possiamo suddividerla in  $j$  sottoinsiemi e calcolare l'attesa dei sottoinsiemi:

$$P(x_k) = \mu P(x_{k-1}); \quad P(x_0) = 1 \text{ (per ipotesi);} \quad P(x_k) = \mu^k$$

Se  $\mu < 1$ , avremo l'estinzione:

$$E_k(x_k = 0) \quad \text{evento che al tempo } k, x_k = 0$$

$E_k \subset E_{k+1}$        $A$  è contenuto in  $B$  se quando si verifica il primo, allora si verifica anche il secondo

$E = \bigvee_{k=1}^{\infty} E_k$       si verifica almeno uno degli eventi

(Nota:  $\bigvee$  indica la somma logica, si può usare anche  $\bigcup$ ).

Se abbiamo una successione di eventi crescenti ( $E_k \subset E_{k+1}$ ), allora

$$P(E) = \sup_k P(E_k)$$

Vogliamo mostrare che se  $\mu < 1$ , allora  $P(E) = 1$ :

$$\begin{aligned} P(E_k) &= 1 - P(\tilde{E}_k) \\ &= 1 - \sum_{j=1}^{\infty} P(x_k = j) \geq 1 - \sum_{j=1}^{\infty} j P(x_k = j) \\ &= 1 - P(x_k) \\ &= 1 - \mu^k \rightarrow 1 \quad \text{perché abbiamo supposto } \mu < 1 \Rightarrow \mu^k \xrightarrow{k \rightarrow \infty} 0 \end{aligned}$$

$$P(E) = 1$$

(

Se invece  $\mu = 1$ , l'attesa di  $x_k(P(x_k))$  rimane 1.

Vediamo infine il caso generale: non facciamo ipotesi su  $\mu$ , ma può assumere qualsiasi valore.

Sia  $\Pi^k = P(x_k = 0)$ .

Possiamo usare la formula delle probabilità condizionate:

$$\begin{aligned} \Pi^k &= P(x_k = 0) = \sum_{j=0}^{\infty} P(x_k = 0 | x_1 = j) P(x_1 = j) \\ &= \sum_{j=0}^{\infty} p_j \Pi^{(k-1)j} \end{aligned}$$

Affinché nella  $k$ -esima generazione non ci siano più individui, in tutti i rami ci deve essere l'estinzione.

Definiamo una funzione  $f$ :

$$\begin{aligned} f(s) &= \sum_{j=0}^{\infty} p_j s^j \\ \pi^k &= f(\pi^{k-1}) \\ \pi^k &\geq \pi^{k-1} \qquad \pi^k \leq 1 \end{aligned}$$

Abbiamo una successione non decrescente:

$$\pi = P(E) = \sup_k \pi^k$$

Ipotizziamo che  $p_0 + p_1 < 1$  e  $p_0 > 0$ .

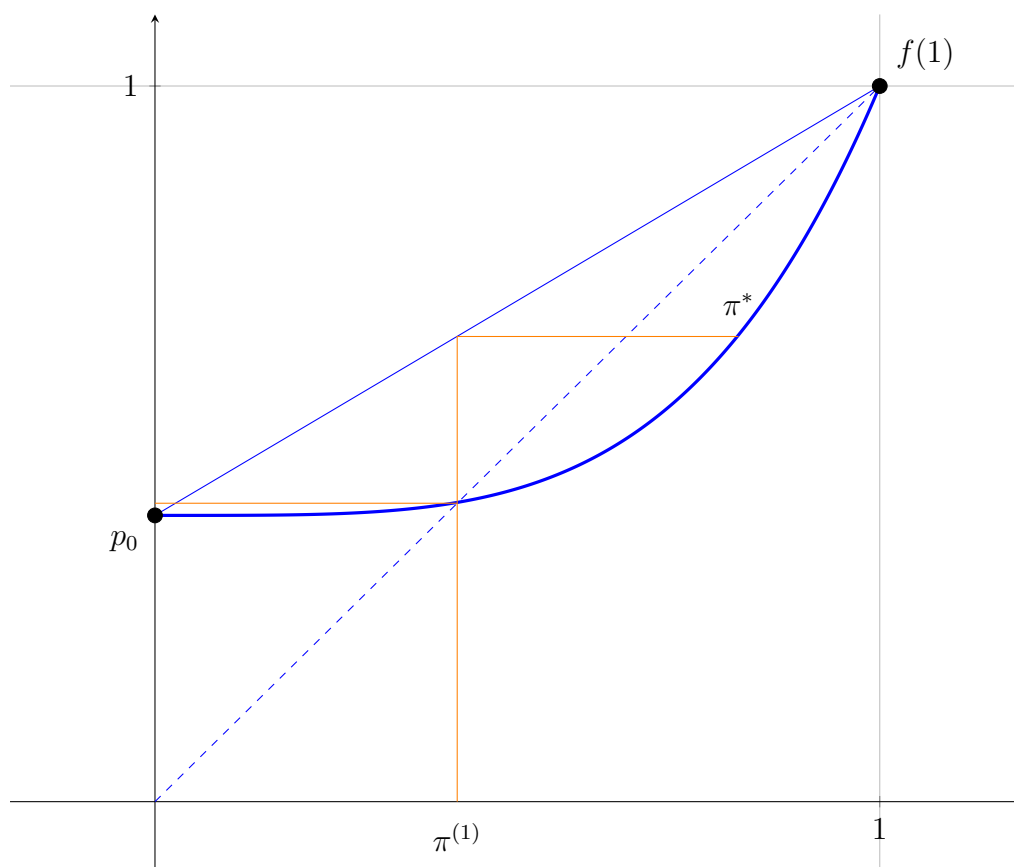
Se  $p_0 = 0$ , vorrebbe dire che la  $p$  che un individuo non abbia figli sia 0. Allora la  $P(E)$  (probabilità di estinzione) sarebbe 0.

Supponiamo ora  $p_0 + p_1 = 1$  e  $p_0 > 0$ , ovvero ogni individuo non ha figli, oppure ha un solo figlio: allora  $P(E) = 1$  (si ha estinzione).

Studiamo com'è fatta  $f(s)$ :

$$\begin{aligned} f(1) &= 1 & f'(s) &= \sum_{j=1}^{\infty} j \cdot p_j \cdot s^{j-1} & f''(s) &> 0 \\ f(0) &= p_0 > 0 & f'(1) &= \sum_{j=1}^{\infty} j \cdot p_j = \mu & f''(s) &= \sum_{j=2}^{\infty} p_j \cdot j \cdot (j-1) s^{j-2} > 0, \quad 0 < s < 1 \end{aligned}$$

La derivata seconda è positiva, quindi la funzione è strettamente convessa.



Qual è la probabilità di estinzione?  $\pi = f(\pi)$ ; la probabilità di estinzione è una soluzione di una delle funzioni sottostanti:

$$\pi^{(1)} = p_0 \text{ (prima generazione)}$$

$$\pi^{(2)} = f(\pi^{(1)})$$

$$\pi^{(3)} = f(\pi^{(2)})$$

$$\pi^{(k+1)} = f(\pi^{(k)})$$

Siccome  $f$  è crescente, se applichiamo da entrambe le parti:

$$p_0 < \pi^*$$

$$\pi^{(1)} < \pi^*$$

...

$$\pi^t < \pi^*$$

Questa successione rimane sempre più piccola di  $\pi^*$ , quindi il limite tenderà a  $\pi^*$ , che è minore di 1.

Quindi se  $\mu > 1$ , c'è una probabilità positiva che non ci sia estinzione.

**Esempi** vediamo alcuni esempi:

- $p_0 = \frac{1}{2}, p_1 = \frac{1}{4}, p_2 = 0, p_3 = \frac{1}{4}$

$$\mu = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 2 \cdot 0 + 3 \cdot \frac{1}{4} = 1 \Rightarrow \pi = 1$$

- $p_0 = \frac{1}{3}, p_1 = \frac{1}{3}, p_2 = 0, p_3 = \frac{1}{3}$

$$\mu = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 2 \cdot 0 + 3 \cdot \frac{1}{3} = \frac{4}{3} > 1$$

Per calcolare  $\pi$  dobbiamo risolvere:

$$f(s) - s = 0 \Rightarrow f(s) = \frac{1}{3} + \frac{s}{3} + \frac{s^3}{3}$$

$$\frac{s^3}{3} + \frac{s}{3} + \frac{1}{3} - s = 0$$

$$\frac{s^3}{3} - \frac{2}{3}s + \frac{1}{3} = 0$$

$$s^3 - 2s + 1 = 0 \Rightarrow (s - 1)(s^2 - As + C) = 0$$

Quindi  $s = 1$  è una soluzione; risolvendo il secondo elemento dell'equazione otterremo 2 radici, una di queste comprese tra 0 e 1, che è quella che ci interessa.

- $p_0 = \frac{1}{4}, p_1 = \frac{1}{4}, p_2 = \frac{1}{4}, p_3 = \frac{1}{4}$

$$f(s) = \frac{1}{4}(s^3 + s^2 + s + 1)$$

$$\frac{1}{4}s^3 + \frac{1}{4}s^2 + \frac{1}{4}s + \frac{1}{4} = 0$$

$$s^3 + s^2 - 3s + 1 = 0$$

$$(s - 1)(s^2 + As + B) = 0$$

Soluzione come nel caso precedente.

## 2.6 Simulatori

Supponiamo di voler generare una passeggiata aleatoria simmetrica tra 0 e  $2n$ . Prendiamo una successione  $x_1, \dots, x_{2n}$ . Definiamo una nuova variabile  $z_i$ :

$$z_i = \begin{cases} 1, & \text{per } x_i > \frac{1}{2} \\ -1, & \text{per } x_i \leq \frac{1}{2} \end{cases}$$

Definiamo  $S_k = z_1 + \dots + z_k$ .

Facciamo tante prove e vediamo quante passeggiate sono positive: la percentuale del numero di passeggiate aleatorie sempre positive dovrebbe

tendere valore delle probabilità che una passeggiata sia sempre positiva per la legge dei grandi numeri.

Si può mostrare un altro esempio per la rovina del giocatore: questo problema si può vedere una passeggiata aleatoria in generale non simmetrica. In questo caso, a differenza di prima, teoricamente potremmo continuare all'infinito. Prendiamo la successione  $x_1, x_2, \dots$  e definiamo:

$$z_i = \begin{cases} 1, & \text{per } 0 < x_i \leq p \\ -1, & \text{per } p < x_i < 1 \end{cases} \quad s_n = x + z_1 + \dots + z_n$$

Con  $0 < x < n$ . Quindi la passeggiata parte da un certo  $0 < x < N$  e ogni volta che genero  $z_n$  vado avanti. Se  $z_n = 0$  ho la rovina del giocatore A, se  $z_n = N$  ho la rovina di B, altrimenti genero un altro numero. Sappiamo che con  $p = 1$  prima o poi ci fermeremo.

Se facciamo il rapporto tra il numero di volte che A si è rovinato e le passeggiate totali, facendo infinite prove, il valore dovrebbe tendere alla probabilità di rovina di A.

Si può vedere anche per i processi di diramazione. Supponiamo di essere nell'intervallo  $[0, 1]$  e di dividerlo in tanti intervalli più piccoli, supponendo che un individuo possa avere  $k$  figli:  $p_0, p_1, \dots, p_k \in [0, 1]$ . I  $k + 1$  intervalli hanno lunghezza pari alla loro probabilità.

Scegliamo, per il primo individuo, un intervallo a caso; diciamo che esca  $p_i$ : il primo individuo avrà  $i$  figli. Ripetiamo per ogni individuo della seconda generazione, e così via. Tipicamente la probabilità di estinzione è minore di 1; l'estinzione avviene nelle prime generazioni (perché la prole cresce esponenzialmente ed è più difficile che avvenga).

Questi tre esempi appena illustrati sono un caso particolare di un modello più generale detto **catene di Markov**.





# Chapter 3

## Catene di Markov

### 3.1 Catene di Markov omogenee

Se ripensiamo ai problemi e ai modelli visti precedentemente, la caratteristica comune è che se sappiamo la posizione in un certo istante, possiamo calcolare la probabilità di andare nello stato successivo, poiché questa probabilità è appunto indipendente dalle misurazioni precedenti. Questa proprietà è chiamata **proprietà di Markov**.

Studieremo le **catene di Markov omogenee con numero finito o numerabile di stati**.

L'attributo "omogeneo" indica che la probabilità di passare da uno stato all'altro. Esistono anche fenomeni dipendenti dal tempo (es. se studiamo la temperatura, bisogna tenere conto di in quale giorno dell'anno si stanno effettuando le misurazioni, e non è quindi un fenomeno omogeneo nel tempo).

Lo "stato", nel caso della probabilità aleatoria, è il punto in cui ci troviamo. Possiamo rappresentare lo stato con un numero intero, che quindi può assumere infiniti valori numerabili. Nel caso dei processi di diramazione, ad esempio, lo stato può essere il numero degli individui (un intero non negativo).

Per **definire** una catena di Markov occorre quindi indicare:

1.  $S$ : spazio degli stati, finito o numerabile;
2. **distribuzione iniziale**  $u_s$ : valori compresi tra 0 e 1 tali che la loro somma è 1. La distribuzione iniziale ci dice al tempo 0, supponendo che il tempo sia discreto, qual è la probabilità di trovarci in un certo stato. Possiamo prendere come stato iniziale qualsiasi stato oltre a 0;

$$0 \leq u_s \leq 1 \qquad \sum u_s = 1 \qquad s \in S$$

3. **probabilità di transizione:** è una matrice; siccome  $S$  può essere infinito, anche questa matrice può avere infinite righe e colonne. Indica la probabilità di passare da uno stato  $s$  di partenza ad uno stato  $s'$  d'arrivo:

$$s, s' \in S \quad \forall s \sum_{s' \in S} p_{s,s'} = 1$$

Possiamo definire una successione  $x_0, x_1, x_2, \dots$  e, dopo aver definito anche i 3 parametri sopra descritti, possiamo dare una distribuzione delle variabili aleatorie. Dato  $n$  finito, e  $p$ , possiamo definire la distribuzione di  $x_0, x_1, \dots$ :

$$\begin{aligned} P(x_0 = s_0, x_1 = s_1, \dots, x_n = s_n) \\ = u_{s_0} \cdot p_{s_0, s_1} \cdot u_{s_1} \cdot p_{s_1, s_2} \cdot \dots \cdot p_{s_{n-1}, s_n} \end{aligned}$$

C'è un problema di compatibilità: se assegno la distribuzione dei primi  $n + 1$ , posso sapere la distribuzione dei primi  $n$ .

**Distribuzione marginale** (dei primi  $n$ ): si calcola facendo la somma dei primi  $n + 1$  valori.

Devo verificare quindi che l'equazione sopra sia valida. Queste assegnazioni delle probabilità sono compatibili, è coerente.

### 3.1.1 Proprietà di Markov

Supponiamo di avere una successione  $s_0, \dots, s_{n-1}, s$  e che  $u_{s_0} \cdot p_{s_0, s_1} \cdot p_{s_1, s_2} \cdot \dots \cdot p_{s_{n-1}, s} > 0$  (se il prodotto fosse 0, la probabilità che si verifichi la successione sarebbe 0). Ci chiediamo quanto vale la seguente probabilità:

$$P(x_{n-1} | x_0 = s_0, x_1 = s_1, \dots, x_{n-1} = s_{n-1}, x_n = s)$$

Usiamo la formula della probabilità condizionata (richiamo: dati due eventi  $A, B$ :  $P(A|B) = \frac{P(AB)}{P(B)}$ , con  $P(B) > 0$ ):

$$\begin{aligned} &= \frac{P(x_0 = s_0, \dots, x_{n-1} = s_{n-1}, x_n = s, x_{n-1} = s')}{P(x_0 = s_0, \dots, x_{n-1} = s_{n-1}, x_n = s)} \\ &= \frac{u_{s_0} \cdot p_{s_0, s_1} \cdot \dots \cdot p_{s_{n-1}, s} \cdot p_{s, s'}}{u_{s_0} \cdot p_{s_0, s_1} \cdot \dots \cdot p_{s_{n-1}, s} \cdot p_{s_{n-1}, s}} \\ &= p_{s, s'} \end{aligned}$$

Quindi, se voglio calcolare la probabilità di andare da  $s$  a  $s'$  mi basta conoscere l'ultimo stato, sapere tutto quello che è successo prima è irrilevante.

Qual è la probabilità di transizione in più passi, ad esempio in due passi?

$$\begin{aligned}
 & P(s_2 = s' | x_0 = s) \\
 &= \frac{P(x_0 = s, x_2 = s')}{P(\underbrace{x_0 = s}_{\substack{=u_s \\ \text{distribuzione} \\ \text{iniziale}}})} \\
 &= \frac{\sum_{s'' \in S} P(x_0 = s, x_1 = s'', x_2 = s')}{P(x_0 = s)} \\
 &= \frac{\sum_{s'' \in S} u_s p_{s,s''} p_{s'',s'}}{u_s} \\
 &= \sum_{s'' \in S} p_{s,s''} p_{s'',s'} \\
 &= p_{s,s'}^{(2)} \quad (\text{probabilità di transizione in 2 passi})
 \end{aligned}$$

Supponendo di partire in un tempo arbitrario  $\neq 0$ , la notazione diventa:

$$p_{s,s'}^{(2)} = P(x_{k+2} = s' | x_k = s)$$

Se immaginiamo di rappresentare  $p_{s,s'}$  come una matrice,  $p^{(2)}$  è il prodotto riga colonna della matrice per se stessa:

$$p_{s,s'}^{(2)} = (\Pi^2)_{s,s'}$$

Generalizzando:

$$P(x_{k+n} = s' | x_k = s) = (\Pi^n)_{s,s'}$$

Quindi mi basta moltiplicare la matrice  $n$  volte e il risultato corrisponderà all'elemento corrispondente alla riga  $s$  e alla colonna  $s'$ .

**Esempio 1** Vediamo un esempio di matrice. Supponiamo di avere due stati; la matrice avrà dimensione  $2 \times 2$ ; la somma degli elementi sulla matrice deve dare 1:

$$\begin{aligned}
 & S = 0, 1 \\
 & 0 < p < 1
 \end{aligned}
 \quad
 \Pi = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix}$$

$p$  è la probabilità di passare dallo stato 0 allo stato 0;  $1-p$  è la probabilità di passare dallo stato 0 allo stato 1.

Possiamo finalmente calcolare le probabilità di passare da uno stato all'altro in  $n$  passi:

$$\begin{aligned}
 p_{0,0}^{(n)} \quad \quad \quad \Pi^{n+1} &= \Pi^n \Pi \\
 p_{0,0}^{(n+1)} &= p_{0,0}^{(n)} p + p_{0,1}^{(1-p)} \\
 &= p_{0,0}^{(n)} p + (1 - p_{0,0}^{(n)}) (1 - p) \\
 p_{0,0}^{(n+1)} &= p_{0,0}^{(n)} (p - 1 - q) + (1 - q) \\
 p_{0,0}^{(2)} &= p(p + q - 1) + (1 - q)
 \end{aligned}$$

Il calcolo è lasciato per esercizio (è una somma geometrica nota,  $p_{0,1}^{(n)} = 1 - p^n$ )

**Esempio 2** Consideriamo una passeggiata aleatoria nel caso generale (non simmetrico). Sia  $0 < p < 1$  la probabilità di fare un passo a destra.

$$\begin{aligned}
 S = Z \quad \quad u_0 &= 1, u_s = z \\
 &\text{per } s \neq 0
 \end{aligned}$$

$$\begin{aligned}
 p_{s,s_0} &= p & p_{s,s-1} &= 1 - p \\
 p_{s,s'} &= 0 & s' + s + 1 &= s - 1
 \end{aligned}$$

Esempio:

$$p_{s,s}^{(2)} = p(1 - p) + (1 - p)p = 2p(1 - p)$$

Ci sono solo 2 possibilità: faccio prima un passo a sinistra e poi a destra, oppure il contrario.

$$p_{s,s+2}^{(2)} = p^2 \cdot p_{s,s-2}^{(2)} = (1 - p)^2$$

$$p_{s,s'}^{(2)} = 0 \quad \text{per } s \neq s', s + 2, s - 2$$

Anziché considerare la passeggiata aleatoria su  $\mathbb{Z}$ , posso considerare un intervallo finito  $[a, b] \subset \mathbb{Z}$ ; devo però definire cosa succede agli estremi dell'intervallo.

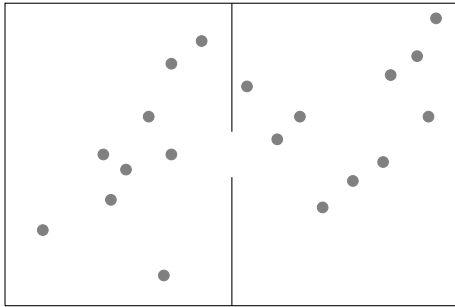
**Condizioni periodiche:** danno un senso orario o antiorario. Se sono in  $a$  e faccio un passo a sinistra, vado in  $b$ :

$$a < s < b, \quad p_{s,s+1} = p, \quad p_{s,s-1} = 1 - p$$

$$\begin{aligned}
 p_{a,b} &= 1 - p & p_{a,a+1} &= p \\
 p_{b,a} &= p & p_{b,b-1} &= 1 - p
 \end{aligned}$$

### 3.1.2 Modello di Ebembat

Descrive la componente di un gas. Supponiamo di avere  $n$  particelle; gli stati saranno  $S = \{0, 1, \dots, N\}$ .



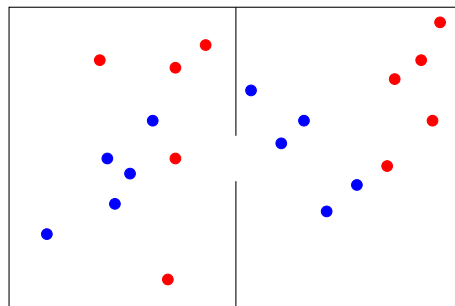
Le probabilità che una particella vada a destra o a sinistra sono rispettivamente:

$$p_{s,s+1} = \frac{N-s}{N}$$

$$p_{s,s-1} ???$$

### 3.1.3 Modello di Bernoulli-Laplace

Descrive l'evoluzione di un gas; in questo caso abbiamo 2 gas diversi, A e B, ed  $N$  particelle del gas A e altrettante del gas B.



Supponiamo che in ogni istante sia scelta a caso una particella nel contenitore sinistro, una a caso nel contenitore destro e vengano scambiate.

Lo stato  $S = \{0, 1, \dots, N\}$  rappresenta il numero di palline di gas A a sinistra.

$p_{s,s}$  è la probabilità che ci siano  $n$  particelle A a sinistra e che dopo lo scambio questo valore rimanga invariato; questo può capitare se scambiamo una particella di A a sinistra con un'altra particella, sempre di A, a destra (oppure se scegliamo due particelle di tipo B).

$$p_{s,s} = ???$$

$$p_{s,s-1} = \frac{s}{N} \cdot \frac{s}{N} = \frac{s^2}{N^2} ???$$

## 3.2 Classificazione degli stati

Due stati  $s$  e  $s'$  si dicono **comunicanti** quando  $s$  comunica con  $s'$ , ovvero la probabilità di andare da  $s$  a  $s'$  è positiva:

$$p_{s,s'}^n = 0$$

$$p_{s,s'}^{(??)} = 0, \text{ con } s' \neq s$$

Uno stato comunica con sè stesso.

**Definizione** (più precisa) Uno stato  $s$  comunica con uno stato  $s'$  se esiste  $n$  tale che la successione da  $s$  a  $s'$  in  $n$  passi è sempre positiva ???

**Proprietà** Siano  $s$  e  $s'$  due stati comunicanti:  $s \prec s'$

Se  $us < s'$  e  $s' < s''$ , allora  $s \prec s''$

In generale, se ho due cammini:

$$s, s_1, \dots, s_{n-1}, s''$$

$$s', \bar{s}_1, \dots, \bar{s}_{m-1}, s''$$

Allora possiamo metterli insieme e fare un cammino che colleghi  $s, s''$ .

Partendo dalla nozione di comunicazione possiamo definire una **relazione di equivalenza**. Nota: la comunicazione non è simmetrica.

$$s \simeq s' \text{ se } s \prec s' \text{ e } s' \prec s$$

$s \prec s$  perché lo stato comunica sempre con sè stesso.

Proprietà transitiva: se  $s \simeq s'$  e  $s' \simeq s''$ , allora  $s \simeq s''$

Proprietà simmetrica:  $s \simeq s'$  e  $s' \simeq s$

Abbiamo una relazione di equivalenza, possiamo definire una classe di equivalenza:

$$[s] = \{s' | s' \simeq s\}$$

La nozione di comunicazione si può estendere alle classi di equivalenza:

$$[s] \prec [s'] \text{ se } s \prec s'$$

Una classe equivalente ad  $s$  si dice **massimale** se non comunica con un'altra classe di equivalenza.

Quando la catena di Markov arriva in uno stato che appartiene ad una classe massimale, rimane "bloccata" sempre dentro stati di quella classe.

Una catena di Markov si dice **irriducibile** se da una sola classe di equivalenza ???

cioè lo spazio degli eventi  $S$  è una sola classe di equivalenza.

Consideriamo la passeggiata aleatoria su  $[a, b] \subset \mathbb{Z}$  con condizioni assorbenti:

$$a < s < b \quad \begin{aligned} p_{s,s+1} &= p \\ p_{s,s-1} &= 1 - p \end{aligned}$$

$$\begin{aligned} p_{a,a} &= 1 & p_{a,s} &= 0 \\ p_{b,b} &= 1 & p_{b,s} &= 0 \end{aligned}$$

Ovvero, se partiamo da uno stato interno lo spostamento è "normale", mentre se siamo in  $a$  o  $b$  rimaniamo sempre in quello stato (come ad esempio nella rovina dei giocatori).

**Esempio** Definiamo ora 3 classi di equivalenza:

$$c_1 = \{a\} \quad c_2 = \{a+1, \dots, b-1\} \quad c_3 = \{b\}, \quad c_2 \prec c_1, c_2 \prec c_3$$

$c_1$  e  $c_3$  sono massimali. In questo caso la catena di Markov non è irriducibile (il modello di Bernoulli-Laplace è un esempio di catena irriducibile).

**Periodo** Dato uno stato  $s \in S$ :

$$A_s^+ = \{? | p_{s,s}^{(n)} > 0\}$$

$$A_s^+ \neq 0$$

Se  $n_1 \in A_s^+$  e  $n_2 \in A_s^+$ , allora  $n_1 + n_2 \in A_s^+$  (ovvero la probabilità di andare dallo stato  $s$  allo stato  $s$  in  $n_1 + n_2$  passi è positiva).

Il **periodo** di  $s$  (con  $A_s^+ \neq 0$ ) è:

$$q = \text{MCD}(A_s^+)$$

In una **passeggiata aleatoria con condizioni periodiche** possiamo fare un movimento in senso orario o antiorario (dallo stato  $b$  torniamo ad  $a$ ). Non è detto che tutti gli  $A^+$  siano divisibili per 2: il numero degli stati può essere dispari.

Se  $q = 1$ ,  $s$  si dice aperiodico.

**Osservazione** due stati equivalenti hanno lo stesso periodo.

Se  $s \simeq s'$ , allora  $s \prec s'$ ; infatti,

$$\exists n_1 : p_{s,s'}^{(n_1)} > 0$$

$$\exists n_2 : p_{s',s}^{(n_2)} > 0$$

Supponiamo  $q$  periodo di  $s$  e  $q'$  periodo di  $s'$ . Vogliamo mostrare che  $q = q'$

Osserviamo che  $n_1 + n_2$  è divisibile sia per  $q$ , sia per  $q'$ :

$$p_{s,s}^{(n_1+n_2)} > 0 \qquad p_{s',s'}^{(n_1+n_2)} > 0$$

Prendiamo un  $n \in A_s^+$ :

$$n + n_1 + n_2 \in A_s^+$$

$n + n_1 + n_2$  è divisibile per  $q' \Rightarrow n$  è divisibile per  $q'$ .

Viceversa, preso  $n' \in A_{s'}^+$ ,  $n'$  è divisibile per  $q$ .

???

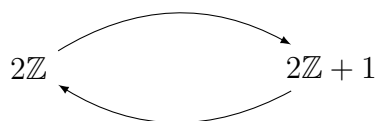
Il periodo è una caratteristica delle classi di equivalenza.

Supponiamo di avere una passeggiata aleatoria su  $\mathbb{Z}$  (che è una catena di Markov irriducibile). Il periodo di ogni stato è 2:

$$\mathbb{Z} = 2\mathbb{Z} \dot{\cup} 2\mathbb{Z} + 1$$

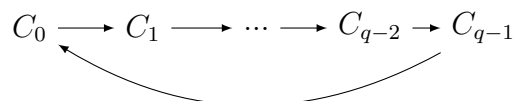
(dove  $\dot{\cup}$  è l'unione disgiunta).

La passeggiata aleatoria passa da un elemento di  $2\mathbb{Z}$  ad un elemento di  $2\mathbb{Z} + 1$ :



In generale, una classe di equivalenza  $C$  con periodo  $q > 1$  può essere suddivisa in  $q$  sottoinsiemi:

$$C = C_0 \dot{\cup} C_1 \dot{\cup} \dots \dot{\cup} C_{q-1}$$



E la catena di Markov percorre in maniera ciclica le classi.



### 3.2.1 Ricorrenza

Vediamo ora un altro tipo di classificazione, detto appunto **ricorrenza**: questa è la proprietà di uno stato di ritornare (infinite) volte nella catena di Markov.

Definiamo, dati due stati  $s$  e  $s'$  e dato  $n > 0$ :

$$f_{s,s'}^{(n)} = P(x_n = s', x_{n-1} \neq s', \dots, x_1 \neq s' | x_0 = s)$$

Questa è la probabilità di passare da  $s$  a  $s'$  per la prima volta in un tempo  $n$ .

Osserviamo che se consideriamo questi eventi per un dato  $n$  e per un  $n' \neq n$ , questi eventi sono incompatibili: se partendo da  $s$  arriviamo per la prima volta a  $s'$  in un tempo  $n$ , ovviamente è impossibile che succeda la stessa cosa per un tempo  $n' \neq n$ . In particolare, se facciamo la somma di queste probabilità, troveremo che questa è pari alla probabilità di uno di questi eventi:

$$\sum_n f_{s,s'}^{(n)} = f_{s,s'}$$

è appunto la probabilità che si verifichi uno di quegli eventi, o in altre parole che arriveremo in uno stato  $s'$  partendo da uno stato  $s$  in un certo tempo.

Consideriamo il caso  $s' = s$ ; in questo caso scriveremo per semplicità  $f_{s,s} = f_s$  e indica la probabilità di tornare, nel futuro, a  $s$ .

$s$  è uno stato **ricorrente** se  $f_s = 1$  (torniamo almeno una volta - in realtà infinite);

$s$  è uno stato **transiente** se  $f_s < 1$ .

La probabilità di tornare infinite volte in uno stato ricorrente è 1. La probabilità di tornare almeno 2 volte in uno transiente è  $f_s^2$ ; la probabilità di tornare almeno  $n$  volte è  $f_s^n$ .

Negli stati transienti, la probabilità di tornare infinite volte invece è 0. La catena di Markov passa per gli stati transienti con probabilità 1 un numero finito di volte.

**Esempio: passeggiata aleatoria simmetrica** La probabilità di tornare a 0 dopo  $2n$  passi, oppure la probabilità che la passeggiata sia diversa da 0 per  $2n$  passi si indica con la stessa quantità:

$$P(x_{2n} \neq 0, x_{2n-1} \neq 0, \dots, x_1 \neq 0 | x_0 = 0) = u_{2n} = \binom{2n}{n} \frac{1}{2^{2n}}$$

$$u_{2n} \stackrel{\text{per Stirling}}{\simeq} \frac{1}{\sqrt{\pi n}} \xrightarrow{n \rightarrow \infty} 0$$

0 quindi è uno stato ricorrente per la passeggiata aleatoria simmetrica.

Ad esempio, se consideriamo i soliti due giocatori che usano una moneta simmetrica, la probabilità che il guadagno non arrivi mai al pareggio è pari a 0.

### 3.2.2 Verificare se uno stato è transiente o ricorrente

C'è una proprietà che semplifica il verificare se uno stato è transiente o ricorrente:

$s$  è ricorrente se e solo se:

$$\sum_{n=0}^{\infty} p_{s,s}^n = \infty$$

$s$  è transiente se e solo se:

$$\sum_{n=0}^{\infty} p_{s,s}^n < \infty \quad (\text{la serie è convergente})$$

**Esempio** Applichiamo il metodo appena esposto considerando per esempio una passeggiata aleatoria su  $\mathbb{Z}$ .

$$p = \frac{1}{2} \quad p_{0,0}^{(2n)} = u_{2n} = \binom{2n}{n} \frac{1}{2^{2n}} \simeq \frac{1}{\sqrt{\pi n}}$$

$$\sum_{n=0}^{\infty} p_{0,0}^{(n)} = +\infty$$

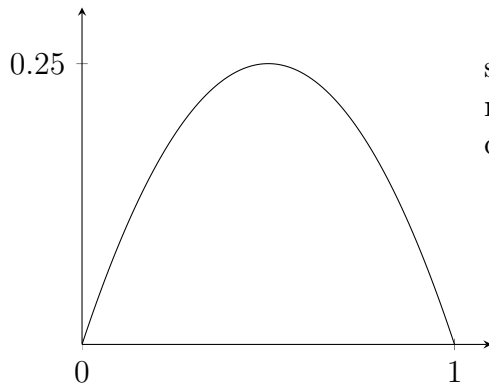
Dobbiamo considerare la serie  $\frac{1}{\sqrt{n}}$  che è divergente per  $n \geq \frac{1}{2}$  e converge per  $n < \frac{1}{2}$ . In questo caso è divergente.

$$\begin{aligned} p \neq \frac{1}{2} \quad p_{0,0}^{(2n)} &= \binom{2n}{n} p^n (1-p)^n = \underbrace{\binom{2n}{n} \frac{1}{2^{2n}}}_1 \cdot \underbrace{2^{2n} p^n (1-p)^n}_{= 4p(1-p)^n} \\ &= u_{2n} \simeq \frac{1}{\sqrt{\pi n}} \end{aligned}$$

Consideriamo la funzione  $p(1-p)$ :

$$\frac{\partial}{\partial p} p(1-p) = 1-p-p = 1-2p$$

È positiva per  $p < \frac{1}{2}$  e negativa per  $p > \frac{1}{2}$ .



Questa quantità:  $(4p(1-p))^n$  è sempre minore di 1. È una serie geometrica convergente; allora, in questo caso lo stato 0 è transiente.

Quindi nel caso di passeggiata aleatoria non simmetrica i due giocatori andranno in pareggio solo un numero finito di volte.

Verifichiamo ora la seconda proprietà, quella che ci permette di determinare se uno stato è transiente. Introduciamo la **funzione generatrice**. È uno strumento che si usa in molti casi, quando abbiamo una successione.

$$\begin{aligned}
 & f_{s,s}^{(0)} \\
 U(z) &= \sum_{n=0}^{\infty} p_{s,s}^{(n)} z^n \\
 & |z| < 1 \\
 F(z) &= \sum_{n=1}^{\infty} f_{s,s}^{(n)} z^n
 \end{aligned}$$

Osserviamo che  $U(z) = 1 + F(z) \cdot U(z)$ . Si vede dall'equazione del rinnovamento.

La probabilità di tornare a  $s$  dopo  $n$  passi è:

$$p_{s,s}^{(n)} = \sum_{k=1}^n f_{s,s}^{(k)} p_{s,s}^{(n-k)} \quad \text{con } n \geq 1$$

C'è quindi una prima volta dove arriviamo a  $s$  ( $f_{s,s}^{(k)}$ ) e poi ci torniamo una seconda volta (al tempo  $n-k$ ).

Per Markov è come se la catena ripartisse dall'inizio.

$$U(z) = \sum_{n=0}^{\infty} p_{s,s}^{(n)} z^n$$

Per  $n=0$ ,  $p_{s,s}^0 = 1$  quindi

$$U(z) = 1 + \sum_{n=1}^n \left( \sum_{k=1}^n f_{s,s}^{(k)} p_{s,s}^{(n-k)} \right) z^n$$

Introduciamo  $m = n - k \Rightarrow n = m + k$

$$\begin{aligned} U(z) &= 1 + \sum_{m=0}^{\infty} \sum_{k=1}^{\infty} f_{s,s}^{(k)} p_{s,s}^{(m)} z^{m+k} \\ &\quad \parallel \\ &\quad z^m \cdot z^k \\ &= 1 + \left( \sum_{m=0}^{\infty} p_{s,s}^{(m)} z^m \right) \left( \sum_{k=1}^{\infty} f_{s,s}^{(k)} z^k \right) \\ &= 1 + F(z) \cdot U(z) \\ \lim_{z \uparrow 1} \sum_{n=1}^{\infty} f_{s,s}^{(n)} z^n &= \sum_{n=1}^{\infty} f_{s,s}^{(n)} = f_s \\ \lim_{z \uparrow 1} U(z) &= \sum_{n=0}^{\infty} p_{s,s}^{(n)} \\ U(z) &= \frac{1}{1 - F(z)} \end{aligned}$$

Nel caso del periodo abbiamo che due stati con la stessa classe di equivalenza hanno lo stesso periodo. Anche la ricorrenza/transienza hanno la stessa proprietà (ovvero, due classi di equivalenza hanno la stessa ricorrenza/transienza).

**Lemma** Dato uno stato  $s$  ricorrente, se  $s \prec s'$  allora  $s' \prec s$ .

Supponiamo per assurdo  $s' \not\prec s$ .

$$\exists n_1 : p_{s,s'}^{(n_1)} > 0.$$

Quindi posso tornare un numero finito di volte su  $s$ . Ma se uno stato è ricorrente, posso tornare infinite volte in uno stato. ASSURDO Quindi  $s' \approx s$

**Lemma** Se  $s \approx s'$  e  $s$  è ricorrente, allora  $s'$  è ricorrente.

Questo lemma ci dice che una classe di equivalenza ha tutti stati ricorrenti, oppure tutti stati transienti.

Supponiamo  $s$  ricorrente; supponiamo  $s \approx s'$ . Quindi  $s \prec s'$  e  $s' \prec s$ .

$$\Rightarrow \exists n_1 : P_{s,s'}^{(n_1)} > 0 \wedge \exists n_2 : p_{s,s'}^{(n_2)} > 0$$

$$\sum_{n=0}^{\infty} p_{s',s'}^{(n)} \geq \sum_{n=n_1+n_2}^{\infty} p_{s',s'}^{(n)}$$

Se invece di partire da  $n = 0$  partiamo da  $n = n_1 + n_2$ , come facciamo ad andare da  $s'$  a  $s'$ ? Andiamo in  $n_2$  da  $s'$  a  $s$ , e in  $n - n_1 - n_2$  possiamo andare a  $s$ , e infine in  $n_1$  da  $s$  a  $s'$ :

$$\geq \sum_{n=n_1+n_2}^{\infty} p_{s',s}^{(n_2)} p_{s,s}^{(n-n_1-n_2)} p_{s,s'}^{(n_1)}$$

Sia  $k = n - n_1 - n_2$ :

$$= p_{s',s}^{n_2} p_{s,s'}^{(n_1)} \sum_{k=0}^{\infty} p_{s,s}^{(k)} = +\infty$$

$\Rightarrow s'$  è ricorrente.  $\square$

### 3.3 Teorema ergodico per catene di Markov

Questo risultato riguarda le catene di Markov con una sola classe di equivalenza con stati ricorrenti e aperiodica. Possiamo dimostrare il teorema ergodico per catene di Markov.

Formalmente,

$$\begin{array}{ll} s \in S & v_s = \sum_{n=1}^{\infty} n f_{s,s}^{(n)} \text{ (attesa del tempo di ritorno)} \\ s \text{ ricorrente} & f_s = \sum_{n=1}^{\infty} f_{s,s}^{(n)} = 1 \end{array}$$

$$\begin{array}{l} p_{s,s'}^{(n)} \xrightarrow{n \rightarrow \infty} \frac{1}{v_s} \\ \text{se } v_s = +\infty, \quad \frac{1}{v_s} = 0 \end{array}$$

Se  $v_s = +\infty$  allora non è necessario che sia aperiodica.

Questo teorema (che non dimostreremo) è interessante perché ci dice che quando abbiamo una sola classe di equivalenza aperiodica con gli stati ricorrenti, è come se la catena di Markov si dimenticasse del punto di partenza: quando faccio tendere  $n$  all'infinito, la probabilità di andare da uno stato qualunque a  $s$  tende ad un limite che non dipende dal punto di partenza.

Si può anche vedere che facendo andare avanti la catena di Markov e contando quante volte passa per  $s$ , la frequenza/percentuale di tempo tende a questo limite.

È come nel caso dello schema di Bernoulli; per la legge dei grandi numeri, se contiamo quante volte otteniamo il successo, questa quantità tende alla probabilità. La probabilità ha il significato di frequenza.

Anche nelle catene di Markov che sono dipendenti, nel caso ergodico la probabilità di andare da uno stato ad un altro tende ad un certo limite che è anche la frequenza con cui la catena di Markov passa per quello stato.

Questo valore è l'attesa del tempo di ritorno; c'è però un metodo più semplice per calcolarlo.

Supponiamo nel caso finito che uno stato sia ricorrente, ci sia una sola classe di equivalenza (e quindi tutti gli stati sono ricorrenti) e aperiodici.

Sappiamo che  $p_{s',s}^{(n)} \rightarrow v_s^{-1} = u_s$ .

Osservazione:  $p_{s',s}^{(n+1)} = \sum_{s''} p_{s',s''}^{(n)} p_{s'',s}$

$$\begin{aligned} p_{s,s'}^{(n+1)} & \xrightarrow{n \rightarrow \infty} u_s & \sum_{s''} p_{s',s''}^{(n)} p_{s'',s} \\ u_s &= \sum_{s''} u_{s''} & \sum_{s \in S} u_s = 1 \end{aligned}$$

$u_s$  è una distribuzione di probabilità sugli stati;

$$\sum_{s \in S} p_{s'',s}^{(n)} = 1 \quad \forall s \in S$$

$$u_s = \sum_{s''} u_{s''} p_{s'',s}$$

Ho un sistema di equazioni lineari: abbiamo  $k$  incognite (gli stati) e  $k + 1$  equazioni. Ma le equazioni  $\sum_{s''} p_{s'',s} \forall s \in S$  non sono linearmente indipendenti, quindi ne posso prendere solo una. Il mio sistema diventa:

$$\begin{cases} \sum u_s = 1 \\ u_s = \sum_{s''} u_{s''} p_{s'',s} \forall s \in S \end{cases}$$

Supponiamo ora:

$$\sum_{s \in S} v_s = 1 \quad v_s = \sum_{s''} v_{s''} p_{s'',s}$$

Con  $S$  finito. C'è un'unica soluzione:

$$v_s = \sum_{s''} v_{s''} p_{s'',s}^{(n)}$$

Faccio tendere  $n$  all'infinito:

$$= \sum_{s''} v_{s''} u_s = u_s$$

Nell'equazione  $\sum_{s''} v_{s''} u_s$  posso portare fuori  $u_s$  perché non dipende dalla sommatoria; ma la sommatoria è uguale a 1, quindi ottengo proprio  $u_s$ .

Definiamo ora  $\mu_s = \sum_{n=1}^{\infty} n \cdot f_s^{(n)}$

Allora  $p_{s,s'}^{(n)} \rightarrow \frac{1}{\mu_s}$

Quindi è uguale a 0 se  $\mu = +\infty$ . Se  $\mu = +\infty$ , non è necessaria aperiodicità.

$$\begin{aligned} \text{Nel caso } S \text{ finito, } u_s &= \frac{1}{\mu_s} & \sum_{s \in S} \mu_s &= 1 \\ \forall s' & \sum_s u_s p_{s,s'} &= u_{s'} \end{aligned}$$

Queste relazioni ci dicono che  $u_s$  è una distribuzione invariante o stazionaria, ovvero se mettiamo  $u_s$  come stato iniziale in una catena di Markov, vediamo che nei tempi successivi ???

Vogliamo verificare se la proprietà sia vera anche nel caso  $S$  infinito.

Dato  $\bar{s} \in S$ ,

$$p_{\bar{s},s'}^{(n+1)} = \sum_{s \in S} p_{\bar{s},s}^{(n)} p_{s,s'}$$

Questa è una serie, dobbiamo passare al limite. Nel caso infinito però non è sempre vero che il limite della serie è uguale alla serie dei limiti. Consideriamo  $\bar{S} \subset S$  finito anziché  $S$ :

$$\begin{aligned} p_{\bar{s},s'}^{(n+1)} &\geq \sum_{s \in \bar{S}} p_{\bar{s},s}^{(n)} p_{s,s'} && \text{per il teorema ergodico} \\ u_{s'} &\geq \sum_{s \in \bar{S}} u_s p_{s,s'} \\ u_{s'} &\geq \sum_{s \in S} u_s p_{s,s'} && \text{(I)} \end{aligned}$$

Abbiamo dimostrato una disuguaglianza. Definiamo

$$0 < \alpha = \sum_{s \in S} u_s \leq 1 \quad (\text{II})$$

$\alpha > 0$  perché gli  $u_s > 0$ , ma  $\alpha \leq 1$  perché  $\sum_{s' \in S} p_{s,s'}^{(n)} = 1$

Dobbiamo dimostrare che in (I) c'è uguaglianza, e in (II) che la somma è pari a 1.

Supponiamo per assurdo che non ci sia uguaglianza in (I) per un particolare stato  $\bar{s}'$ :

$$u_{\bar{s}'} > \sum_{s \in S} u_s p_{s, \bar{s}'}$$

Allora quando sommiamo tutti gli stati, avremo la disuguaglianza stretta:

$$\begin{aligned} \sum_{s'} u_{s'} &> \sum_{s'} \sum_s u_s p_{s,s'} \\ \alpha &= \sum_{s'} u_{s'} > \sum_{s'} \sum_s s u_s p_{s,s'} \\ \alpha &> \alpha \quad \text{ASSURDO} \end{aligned}$$

Quindi anche nel caso infinito,  $u_s$  è una distribuzione di probabilità, cioè la somma è pari a 1:

$$\begin{aligned} u_{s'} &= \sum_s u_s p_{s,s'}^{(2)} \\ u_{s'} &= \sum_s u_s p_{s,s'}^{(n)} \end{aligned}$$

Per  $n \rightarrow \infty$ ,  $u_{s'} = \sum_s u_s u_{s'}$

$$\begin{aligned} u_{s'} &= \alpha u_{s'} & u_{s'} &> 0 \\ \Rightarrow \alpha &= 1 \quad \square \end{aligned}$$

Abbiamo ipotizzato finora l'attesa di ritorno  $\mu_s < +\infty$  (cioè  $\mu_s$  finito).

Nel caso  $\mu_s = +\infty$ , allora  $p_{s,s'} \rightarrow 0$ . Lo abbiamo visto nel caso aperiodico, ma questa proprietà è vera anche nel caso generale. Questo ci porta a fare una classificazione degli stati ricorrenti in 2 categorie:

$$\text{supponiamo } s \text{ ricorrente: } \begin{cases} \mu_s < +\infty, & s \text{ ricorrente positivo} \\ \mu_s = +\infty, & s \text{ ricorrente nullo} \end{cases}$$



Nel primo caso quindi l'attesa del tempo di ritorno è finita, nel secondo infinita.

Supponiamo di avere una classe di equivalenza  $S$  irriducibile e ricorrente. Vogliamo mostrare che non è possibile che ci sia uno stato ricorrente positivo e uno stato ricorrente nullo. Supponendo quindi che  $s_1$  sia positivo, vogliamo mostrare che  $s_2$  non può essere ricorrente nullo.

Dato un  $s$ ,

$$p_{s,s_1}^{(n)} \xrightarrow{n \rightarrow \infty} u_{s_1} > 0$$

$$\exists n_1 : p_{s_1,s_2}^{n_1} > 0 \quad \text{perché } s_1, s_2 \in S$$

$$\begin{aligned} &\text{Ma } p_{s,s_2}^{(n)} > p_{s,s_1}^{(n-n_1)} p_{s,s_2}^{(n_1)} \\ &p_{s,s_2}^{(n)} \text{ non tende a } 0 \text{ (per } n \rightarrow \infty) \end{aligned}$$

$$\Rightarrow s_2 \text{ è ricorrente positivo} \quad \square$$

Una conseguenza è che se  $S$  è finito, non ci sono stati ricorrenti nulli, perché

$$\sum_{s' \in S} p_{s,s'}^{(n)} = 1 \quad \begin{array}{l} \text{essendo } S \text{ finito, possiamo passare al lim-} \\ \text{ite della somma. Se gli stati fossero ri-} \\ \text{correnti nulli, il limite tenderebbe a } 0, \\ \text{ma non è possibile.} \end{array}$$

???

Vediamo il caso più semplice di catena di Markov con spazio degli stati finiti, che è una catena di Markov con due stati:  $S = \{0, 1\}$ . La matrice di transizione è:

$$\Pi = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \quad \begin{array}{l} 0 < p < 1 \\ 0 < q < 1 \end{array}$$

$$1. \quad u_0 + u_1 = 1 \quad (\text{perché è una distribuzione di probabilità})$$

$$2. \quad u_0 \cdot p + u_1 \cdot (1-q) = u_0$$

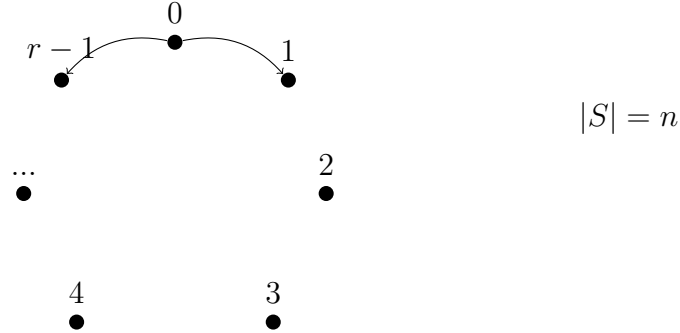
$$3. \quad u_1 \cdot q + u_0 \cdot (1-p) = u_1$$

2. e 3. sono uguali, possiamo usare 1. e 2.

$$\begin{cases} u_0 = u_1 \left( \frac{1-q}{1-p} \right) \\ u_1 \frac{1-q}{1-p} + u_1 = 1 \end{cases}$$

$$\begin{cases} u_1 \frac{(1-q) + (1+p)}{1-p} = 1 \\ u_0 = u_1 \left( \frac{1-q}{1-p} \right) \end{cases} \quad \begin{cases} u_0 = \frac{1-q}{2-p-q} \\ u_1 = \frac{1-p}{2-p-q} \end{cases}$$

Possiamo rappresentare la passeggiata aleatoria con un cerchio:



Se  $r$  è dispari, è aperiodica.

$$\begin{aligned} u_0 + u_1 + \dots + u_{r-1} &= 1 \\ 1 \leq s &\leq r-2 \\ u_s &= u_{s-1}p + u_{s+1}(1-p) \\ u_0 &= u_{r-1}p + u_1(1-p) \\ u_{r-1} &= u_{r-2}p + u_0(1-p) \end{aligned}$$

Abbiamo  $r+1$  equazioni e  $r$  incognite. Ipotizziamo che  $u_s = \frac{1}{r}$ .

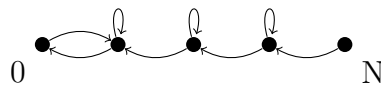
Questa è in realtà una distribuzione invariante anche nel caso pari, ma nel caso pari non è periodico.

???

In entrambi i modelli di **Laplace-Bernoulli** e **Ehrenfest** abbiamo due urne. Nel primo scegliamo a caso una pallina e la spostiamo da un'urna all'altra; nel secondo, abbiamo  $n$  palline bianche e  $n$  palline nere. Di queste  $2n$  palline, metà stanno nell'urna A e metà nell'urna B. Possiamo prendere come stato il numero delle palline bianche in A. Quello che osserviamo è che il modello di Ehrenfest non è aperiodico.

???

Nel modello di Laplace-Bernoulli il numero di palline bianche può anche rimanere invariato:



se scambiamo 2 palline bianche tra le urne A e B, il numero delle palline bianche in A è invariato. Se la probabilità di rimanere nello stato è 0, allora il periodo deve essere 1 e possiamo usare il teorema ergodico. Intuitivamente, comunque, possiamo affermare che

$$u_s < \frac{\binom{N}{s} \binom{N}{N-s}}{\binom{2N}{N}}$$

e questa disuguaglianza dovrà soddisfare le equazioni.

**Teorema** Supponiamo  $S$  infinito, una sola classe di equivalenza aperiodica irriducibile. Supponiamo ci sia una distribuzione invariante  $u_s$ , con  $0 \leq u_s \leq 1$ :

$$\sum_s u_s = 1 \quad \forall s', u_{s'} = \sum_s u_s p_{s,s'}$$

Allora gli stati sono ricorrenti positivi.

Dobbiamo far vedere che non sono ricorrenti nulli o transienti. Supponiamo per assurdo siano transienti:

$$??? \sum_n p_{s,s^*}^{(n)} < +\infty$$

Se prendiamo  $s' \neq s$ ,

$$p_{s',s^*}^{(n)} \xrightarrow{n \rightarrow \infty} 0$$

Possiamo scrivere l'equazione del rinnovamento:

$$p_{s',s}^{(n)} \sum ???$$

???

## 3.4 Modello di coda

In questo modello abbiamo dei clienti, che arrivano e vanno dagli sportelli dove vengono serviti. Il tempo che passano agli sportelli è aleatorio. Supponiamo che il tempo sia discreto.

Uno sportello può contenere un cliente, gli altri clienti sono in coda.

Definiamo  $p$  la **probabilità di arrivo di un cliente**, mentre  $q$  la probabilità di **fine del servizio**,  $0 \leq p \leq 1$  e  $0 \leq q \leq 1$ . Gli arrivi dei clienti e la fine del servizio sono indipendenti.

Questo è descritto da una catena di Markov in cui lo spazio degli stati è  $\mathbb{N}$ :  $S = \mathbb{N} = \{0, 1, \dots\}$ .

Ci chiediamo ora qual è la probabilità, partendo da 0, di rimanere nello stato 0 (ovvero che non ci sia nessun cliente e che non arrivi nessun cliente), e altre probabilità, con  $S = \mathbb{N}$ :

$$\begin{aligned} p_{0,0} &= 1 - p & p_{1,1} &= p \cdot q + (1 - p)1 - q & p_{1,2} &= p(1 - q) \\ p_{0,1} &= p & p_{1,0} &= (1 - p)q \end{aligned}$$

$p_{1,1}$  indica la probabilità che, dopo aver finito di servire un cliente ne arrivi uno nuovo; oppure di non finire di servire un cliente, e che non arrivi nessun nuovo cliente.

In generale, per  $s \geq 1$ :

$$\begin{aligned} p_{s,s} &= p \cdot q + (1 - p)(1 - q) & &= r \\ p_{s,s+1} &= p(1 - q) & &= l \\ p_{s,s-1} &= q(1 - p) & &= d \end{aligned}$$

Questa passeggiata è ricorrente positiva? Se esiste una distribuzione invariante sì, se non esiste o gli stati sono ricorrenti nulli oppure transienti.

Definiamo  $u_i$ , con  $i \in \mathbb{N}$ :

$$\begin{aligned} \sum_{i=0}^{\infty} u_i &= i \\ u_0 &= u_0(1 - p) + u_1 l \\ u_1 &= u_0 p + u_1 r + u_2 l \\ u_s &= u_{s-1} d + u_s r + u_{s+1} l, & s \geq 2 \\ u_1 l &= u_0 p \\ u_1(1 - r - l) &= u_2 l \\ u_1 d &= u_2 l \\ u_s d &= u_{s+1} l \end{aligned}$$

Mettiamo insieme le equazioni:

$$u_1 = u_0 \frac{p}{l} \qquad u_2 = u_1 \frac{d}{l} \qquad u_{s+1} = u_s \frac{d}{l}$$

Riscritte in termini di  $u_0$ :

$$u_2 = u_0 \frac{p}{l} \frac{d}{l} \qquad u_s = u_0 \frac{p}{l} \left( \frac{d}{l} \right)^{s-1}$$

Sapendo che  $\sum_{i=0}^{\infty} u_i = 1$ :

$$u_0 \left( 1 + \frac{p}{l} + \frac{p}{l} \frac{d}{l} + \frac{p}{l} \left( \frac{d}{l} \right)^2 + \dots \right) = 1$$

L'equazione è possibile solo se questa serie:

$$u_0 \left( 1 + \frac{p}{l} \left( \sum_{j=0}^{\infty} \left( \frac{d}{l} \right)^j \right) \right)$$

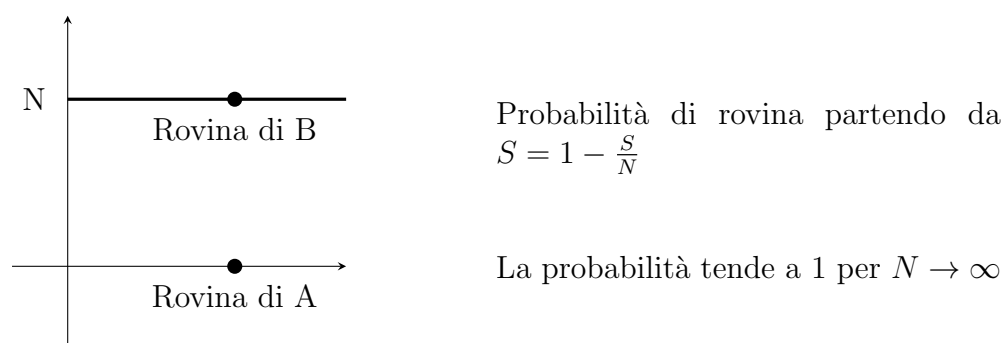
è convergente; ma questa è una serie geometrica, ed è convergente per  $\frac{d}{l} = 1$ .

$$\begin{array}{ll} d = p(1 - q) & l = q(1 - p) \\ p(1 - q) < q(1 - p) & ??? \\ \frac{p}{1-p} < \frac{q}{1-q} & p < q \end{array}$$

Intuitivamente,  $p$  è la probabilità di arrivo di un cliente,  $q$  la probabilità che un cliente vada via. ????

Questo fa sì che non si formi una coda lunga. Se  $p < q$ , allora abbiamo che gli stati sono **ricorrenti positivi**. Se  $p \geq q$ , gli stati sono **ricorrenti nulli** oppure **transienti**.

Ci possiamo ricondurre al problema della rovina del giocatore: in quel caso è una passeggiata aleatoria in cui possiamo aumentare o diminuire di 1. Qui però c'è anche la possibilità di rimanere nello stesso stato all'istante successivo. Ma se trascuriamo questi tempi, nel caso  $p = q$ ,  $\Rightarrow \frac{p}{1-p} = \frac{q}{1-q}$ , siccome le probabilità di andare a destra e a sinistra sono uguali, il modello corrisponde alla passeggiata simmetrica.



Quindi, se  $S = 1$ , siccome andremo a 0, dovremo per forza ripassar e per 1; gli stati sono ricorrenti. Inoltre sono ricorrenti nulli perché non c'è distribuzione di probabilità invariante.

L'interpretazione è che la coda è lunga, ma ogni tanto si svuota.

Se  $p > q$ , è come la passeggiata asimmetrica:  $d = \frac{p}{1-p} > l = \frac{q}{1-q}$ : la probabilità di andare a destra è maggiore della probabilità di andare a sinistra.

C'è una probabilità positiva che il primo giocatore non si rovini, quindi che non torni a 0. Quindi in questo caso gli stati sono transienti. L'interpretazione quindi è che lo sportello si può vuotare un numero finito di volte, prima di allungarsi e non riuscire più a svuotarsi.

### 3.5 Reversibilità

È una proprietà che possono possedere le catene di Markov.

Supponiamo di avere una catena irriducibile ergodica aperiodica. Sia  $S$  il suo spazio degli stati; abbiamo una distribuzione invariante  $u_s$ . Nella catena di Markov, la distribuzione di probabilità è  $x_0, x_1, x_2, \dots$ .

Se come distribuzione iniziale diamo  $u_s$ , ai tempi successivi la distribuzione di  $x_1, x_2, \dots$  è sempre uguale a  $u_s$ .

Possiamo pensare la catena di Markov, anziché su  $\mathbb{N}$ , sui numeri interi:

$$\dots, x_{-n}, x_{-n+1}, \dots, x_{-1}, x_0, x_1, x_2, \dots$$

Possiamo poi considerare un processo stocastico in cui combiniamo la direzione del tempo:  $y_n = x_{-n}$

Quello che si vede è che anche  $y_n$  è una catena di Markov.

Dobbiamo dimostrare che soddisfa le proprietà di Markov:

$$P(y_{n+1} = s' | y_n = ???)$$

Usiamo la formula delle probabilità subordinate:

$$\begin{aligned} &= \frac{P(y_{n+1} = s' | y_n = ???)}{P(???) } \\ &= \frac{P(x_0 = s_0, x_{-1} = s_1, \dots, s_n = s, x_{-n-1} = s')}{P_{??} = s_0, x_{-1} = s_1, \dots, x_{-n} = s} \\ &= \frac{s_{s'} p_{s',s} p_{s,s_n} \dots p_{s_1,s_0}}{u_s p_{s,s_n} \dots p_{s_1,s_0}} \\ &= \frac{s_{s'} p_{s',s}}{u_s} \quad \begin{array}{l} \text{Questa probabilità non dipende da tutti gli stati} \\ \text{precedenti, ma solo da quello immediatamente precedente} \end{array} \\ &= q_{s,s'} \end{aligned}$$

La catena di Markov si dice reversibile se  $p_{s,s'} = q_{s,s'}$ .

Quindi se guardiamo la catena all'indietro è la stessa in senso contrario. Equivalentemente:

$$u_{s'}p_{s',s} = u_s p_{s,s'}$$

Supponiamo  $s, s' \in S$ ; supponiamo, per ogni stato  $s$ :

$$\alpha_s > 0 : \sum_{s \in S} \alpha_s < +\infty, \quad \alpha_s p_{s,s'} = \alpha_{s'} p_{s',s}$$

allora  $u_s = \frac{\alpha_s}{\sum_{s'} \alpha_{s'}}$  è una distribuzione invariante e la catena è reversibile. Dobbiamo far vedere anzitutto che è una distribuzione invariante:

$$\sum_s u_s p_{s,s'} = u_{s'}$$

Sia  $k = \frac{1}{\sum_{s'} \alpha_{s'}}$

$$\sum_s k \alpha_s p_{s,s'} = \sum_s k \alpha_s^{=k \alpha_{s'} = u_{s'}} p_{s',s} ???$$

$$u_s p_{s,s'} = k \alpha_s p_{s,s'} = k \alpha_{s'} p_{s',s} = u_{s'} p_{s',s}$$

**Esempio** Passeggiata aleatoria su intervallo  $[a, b] \subset \mathbb{Z}$  con condizioni simmetriche

$$a + 1 \leq s \leq b - 1$$

$$\begin{aligned} p_{s,s+1} &= p & p_{s,s-1} &= 1 - p \\ p_{a,a+1} &= p & p_{a,b} &= 1 - p \\ p_{b,a} &= p & p_{b,b-1} &= 1 - p \end{aligned}$$

Vogliamo cercare dei numeri tali che:

$$\alpha_s p_{s,s'} = \alpha_{s'} p_{s',s}$$

Supponiamo  $s' = s + 1$

$$\begin{aligned} \alpha_s p &= \alpha_{s+1} (1 - p) \\ \alpha_{s+1} &= \frac{p}{1 - p} \alpha_s \end{aligned}$$

Il numero degli stati è  $b - a + 1$ ; se itero queste relazioni:

$$\alpha_a = \left( \frac{p}{1 - p} \right)^{b-a+1} \alpha_b$$

Quindi  $\left( \frac{p}{1-p} \right)^{b-a+1}$  deve essere uguale a 1; ciò è vero solo se  $p = \frac{1}{2}$ .

Se  $p = \frac{1}{2}$ , allora  $\alpha_s = \alpha_a \forall s$

$$u_s = \frac{1}{b+a+1}$$

La catena è reversibile e ho trovato la distribuzione invariante.

Se  $p \neq \frac{1}{2}$ , la passeggiata andrebbe più in senso orario o antiorario.

Vediamo un criterio generale per la reversibilità. Supponiamo di avere una catena di Markov ergodica e aperiodica:

$$p_{s,s'}^{(n)} u_{s'} \quad \text{distribuzione invariante}$$

Allora la catena di Markov è reversibile se e solo se  $\forall s_0, \dots, s_n$ :

$$p_{s_0,s_1} \cdot p_{s_1,s_2} \cdot \dots \cdot p_{s_{n-1},s_n} \cdot p_{s_n,s_0} = p_{s_0,s_n} \cdot p_{s_n,s_{n-1}} \cdot \dots \cdot p_{s_1,s_0}$$

Supponiamo sia reversibile.

$$\begin{aligned} p_{s,s'} &= \frac{p_{s',s}}{u_s} \cdot u_{s'} \\ &= p_{s_0,s_n} \cdot p_{s_n,s_{n-1}} \cdot \dots \cdot p_{s_1,s_0} \\ &= \left( \frac{p_{s_0,s_1}}{u_{s_1}} \cdot u_{s_0} \right) \cdot \left( \frac{p_{s_1,s_2}}{u_{s_2}} \cdot u_{s_1} \right) \cdot \dots \end{aligned}$$

Gli  $u_{s_i}$  si semplificano. Questa è una condizione necessaria. Dobbiamo ora far vedere che è sufficiente.

Supponiamo che la relazione sia soddisfatta.

$$\begin{aligned} u_s p_{s,s'} &= \lim_{n \rightarrow \infty} p_{s,s'} p_{s',s}^{(n)} \\ &= \lim_{n \rightarrow \infty} \lim_{n \rightarrow \infty} p_{s,s'} \sum_{s_1, s_2, \dots, s_{n-1}} p_{s',s_1} p_{s_1,s_2} \dots p_{s_{n-1},s} \\ &= \lim_{n \rightarrow \infty} \sum_{s_1, s_2, \dots, s_{n-1}} p_{s,s_{n-1}} \dots p_{s_1,s'} p_{s',s} \\ &= \lim_{n \rightarrow \infty} p_{s,s'}^{(n)} p_{s',s} \stackrel{\text{per ergodicità}}{=} u_{s'} p_{s',s} \end{aligned}$$



**corollario? esempio?** Supponiamo di avere una catena di Markov irriducibile su un intervallo tale che:

$$p_{s,s'} = 0 \quad \text{se } |s' - s| \geq 2$$

Allora è reversibile.

Questa catena quindi non può fare dei salti maggiori o uguali a 2; può rimanere ferma, andare a destra oppure a sinistra.

Sappiamo che ci deve essere almeno uno stato ricorrente positivo. Siccome la catena è irriducibile, allora tutti gli stati sono così.

Mostriamo che è irreversibile:

$$s_1, \dots, s_{n-1}, s'$$

$$p_{s_1,s} \cdot p_{s_1,s_2} \cdot \dots \cdot ??? \quad (\text{I})$$

$$p_{s,s'} \cdot p_{s',s_{n-1}} \cdot p_{s_2,s_1} \cdot p_{s_1,s} \quad (\text{II})$$

Devo vedere se i due prodotti sono uguali.

In (I) ci saranno dei passaggi a destra, a sinistra e dei momenti in cui la catena rimane ferma. Siccome devo tornare allo stato di partenza, il numero dei passi a destra deve essere uguale al numero dei passi a sinistra. Per ogni coppia di stati vicini, alcuni andranno in un senso e altri nel senso contrario.

**Distribuzione invariante per il processo di coda** Dobbiamo cercare dei numeri  $\alpha_s$  tali che

$$\alpha_s p_{s,s'} = \alpha_{s'} p_{s',s}$$

$$s \geq 1 : \begin{cases} \alpha_s d = \alpha_{s+1} l \\ \alpha_{s+1} = \frac{\alpha_s d}{l} \end{cases} \quad s = 0 : \begin{cases} \alpha_0 p = \alpha_1 l \\ \alpha_1 = \alpha_0 \frac{p}{l} \end{cases}$$

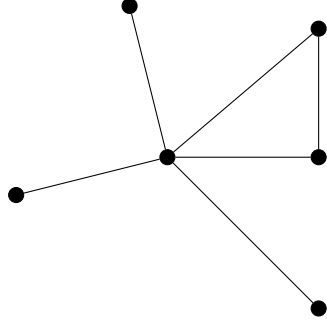
Applichiamo lo stesso criterio per la passeggiata aleatoria (consideriamo solo gli spostamenti):

$$s \geq 1 : \bar{\alpha}_{s-1} = \bar{\alpha}_s \frac{\bar{p}}{1 - \bar{p}} = \frac{d}{l} \quad s = 0 : \bar{\alpha}_0 = \bar{\alpha}_1 \frac{l}{1 - r}$$

Le equazioni sono uguali a parte per il primo stato ( $s = 0$ ), quindi la distribuzione invariante non è uguale.

C'è un altro modo con il quale si può ricercare la distribuzione invariante

Supponiamo sia reversibile. Abbiamo un grafo:



$S = \{1, \dots, n\}$  (lo spazio degli stati sono i vertici del grafo)

$\forall i, j$  abbiamo una costante positiva  $v_{i,j} = v_{j,i} > 0$ .

$\{i, j\}$  è un arco del grafo.

Definiamo la catena di Markov in questo modo:

$$p_{i,j} = \frac{v_{i,j}}{\sum_{j'} v_{i,j'}}$$

Ci sono degli archi che attirano più di altri il passaggio. Cerchiamo una distribuzione invariante in modo che la catena di Markov sia reversibile:

$$\begin{aligned} \alpha_i p_{i,j} &= \alpha_j p_{j,i} \\ \alpha_i \frac{v_{i,j}}{\sum_{j'} v_{i,j'}} &= \alpha_j \frac{v_{j,i}}{\sum_{i'} v_{j,i'}} \\ \frac{\alpha_j}{\alpha_i} &= \frac{\sum_{i'} v_{j,i'}}{\sum_{j'} v_{i,j}} \end{aligned}$$

Devo scegliere  $k$  in modo che la somma faccia 1:

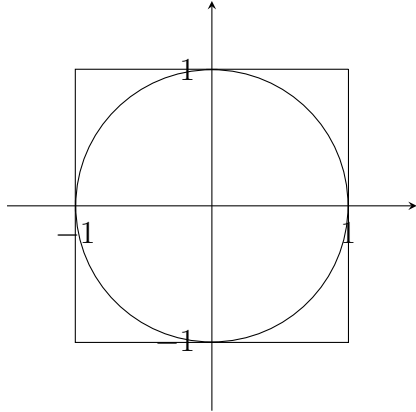
$$\begin{aligned} \sum_i \alpha_i &= k \sum_i \sum_{j'} v_{i,j'} = 1 \\ k &= \frac{1}{\sum_i \sum_{j'} v_{i,j'}} \end{aligned}$$

In questo modo ho trovato la distribuzione invariante e ho dimostrato che è una catena di Markov reversibile. Se non riesco a risolvere l'equazione non è detto che non ci sia la distribuzione inversa, ma la catena di Markov non è reversibile.

### 3.6 Metodo Montecarlo basato sulle catene di Markov (MCMC)

Il metodo Montecarlo è un metodo numerico usato per approssimare usando generatori di numeri casuali. Vediamo un esempio del metodo standard dove

si vuole approssimare  $\pi$ :



L'area del cerchio è data da:

$$A = r^2 \cdot \pi = \pi$$

L'area del quadrato è data da:

$$B = (2r)^2 = 4$$

$$\frac{A}{B} = \frac{\pi}{4}$$

Se scegliamo un punto a caso la probabilità che cada dentro al cerchio è  $\frac{\pi}{4}$ . Consideriamo una serie di eventi  $E_1, E_2, \dots$  (punti scelti a caso nel quadrato):

$$E_i = (p_i \text{ appartiene al cerchio})$$

Per la legge dei grandi numeri

$$\frac{\sum_{i=1}^N E_i}{N} \simeq \frac{\pi}{4}$$

Per  $N$  molto grande.

Con la catena di Markov non abbiamo una successione indipendente, ma appunto una catena di Markov.

Supponiamo di avere un insieme  $S$  finito e di avere ??? di probabilità

$$u_s = k\pi_s$$

$k$  è la costante di normalizzazione. Se  $s$  è molto grande, calcolare  $k$  può essere quasi impossibile. Vogliamo approssimarlo in modo numerico. Per farlo, possiamo generare una catena di Markov ergodica reversibile con  $u_s$  come distribuzione invariante.

Partiamo da una catena di Markov su  $S$  con probabilità di transizione  $q_{s,s'}$  irriducibile e tale che se  $q_{s,s'} > 0$  anche  $q_{s',s} > 0$ :

$$p_{s,s'} = \alpha_{s,s'} q_{s,s'} \quad \text{per } s' \neq s \quad 0 \leq \alpha_{s,s'} \leq 1$$

$$p_{s,s} = q_{s,s} + \sum_{s'} (1 - \alpha_{s,s'}) q_{s,s'}$$

$$\sum_{s'} p_{s,s'} = 1$$

Ora voglio scegliere  $s'$  in modo tale che la catena abbia come distribuzione invariante  $u_s$ :

$$\begin{aligned} u_s p_{s,s'} &= u_{s'} p_{s',s} & s' \neq s & \text{ altrimenti la condizione è banalmente verificata} \\ \pi_s p_{s,s'} &= \pi_{s'} p_{s',s} & u_s &= k \pi_s \\ \pi_s \alpha_{s,s'} q_{s,s'} &= \pi_{s'} \alpha_{s',s} q_{s',s} \\ \frac{\alpha_{s,s'}}{\alpha_{s',s}} &= \frac{\pi_{s'} q_{s',s}}{\pi_s q_{s,s'}} & \alpha_{s,s'} \leq 1, \quad \alpha_{s',s} \leq 1 \end{aligned}$$

Definiamo

$$\begin{aligned} \alpha_{s,s'} &= \phi \left( \frac{\pi_{s'} q_{s',s}}{\pi_s q_{s,s'}} \right) \\ \alpha_{s',s} &= \phi \left( \frac{\pi_s q_{s,s'}}{\pi_{s'} q_{s',s}} \right) \end{aligned}$$

Che proprietà deve avere questa funzione?

1.  $0 \leq \phi(x) \leq 1$
2.  $\frac{\phi(x)}{\phi\left(\frac{1}{x}\right)} = x$

Una possibile scelta per  $\phi$  è:

$$\phi(x) = \min(1, x)$$

Supponiamo  $0 < x \leq 1$ ; allora  $\phi(x) = x$  e  $\phi(\frac{1}{x}) = 1$ .

Se invece  $x > 1$ :

$$\begin{aligned} \phi(x) &= 1 \\ \phi\left(\frac{1}{x}\right) &= \frac{1}{x} \\ \frac{1}{\frac{1}{x}} &= 1 \end{aligned}$$

Definiamo

$$\alpha_{s,s'} = \min \left( \frac{\pi_s q_{s,s'}}{\pi_{s'} q_{s',s}}, 1 \right)$$

### 3.6. METODO MONTECARLO BASATO SULLE CATENE DI MARKOV (MCMC) 57

**Nota** Abbiamo definito  $\phi$  in un certo modo, ma non è l'unico: un'altra possibile funzione è  $\phi(x) = \frac{x}{1+x}$ .

#### 3.6.1 Algoritmo di Metropolis-Hastings

**Modello di Ising** È usato per la meccanica statica, ma può essere applicato al riconoscimento delle immagini.

$$S = \{-1, 1\}^\Lambda \quad \Lambda \subset \mathbb{Z}^2$$

Ho un sottoinsieme di  $\mathbb{Z}^2$  che posso vedere come un quadrato:

$$\begin{array}{ccc} -1 & -1 & 1 \\ \bullet & \bullet & \bullet \\ 1 & -1 & 1 \\ \bullet & \bullet & \bullet \\ 1 & 1 & -1 \\ \bullet & \bullet & \bullet \end{array}$$

Se per ognuno di questi punti assegno una quantità, ottengo una configurazione:

$$u_s = k \exp \left[ -j \sum_{\langle i,j \rangle} s_i s_j \right] \quad i \in \Lambda$$

$\langle i, j \rangle$ :  $i, j$  sono vicini, considero solo coppie di  $i, j$  vicini.

Per calcolare  $k$  (la costante di normalizzazione) devo fare la somma di tutti i possibili  $s$ :  $|s| = 2^{|\Lambda|}$

Se prendo un quadrato di lato 10, il numero di possibili configurazioni è  $2^{100}$ .

??? Scelgo a caso un punto; se era uguale a 1, lo faccio diventare -1 e viceversa.

Questa  $q$  verifica le proprietà che voglio.

Questa catena di Markov è irriducibile: se cambio uno alla volta i valori della configurazione posso passare da una configurazione all'altra.

$$\pi_s = \exp \left( j \sum_{\langle i,j \rangle} s_i s_j \right)$$

$$\alpha_{s,s'} = \min \left( \frac{\pi_s q_{s,s'}}{\pi_{s'} q_{s',s}}, 1 \right)$$

$\alpha_{s,s'}$  è diverso da 0 solo se  $q_{s,s'} > 0$ , cioè solo se  $s, s'$  sono definiti cambiando di 1 ???

$$\frac{\pi}{\pi_{s'}} = \frac{u_s}{u_{s'}} = \frac{\exp \left[ -j \sum_{\langle i,j \rangle} s_i s_j \right]}{\exp \left[ -j \sum_{\langle i,j \rangle} s'_i s'_j \right]}$$

So che  $s, s'$  sono uguali tranne in un punto. Quindi, tranne che per quel punto, tutte le coppie si cancellano.

I vicini di un punto in 2 dimensioni sono 4, quindi rimane una somma di 4 termini.

Qual è la probabilità che un certo punto sia uguale a 1? ???

### 3.6.2 Campionatore di Gibbs (Gibbs sampler)

È una versione dell'algoritmo di Metropolis-Hastings. Supponiamo che l'insieme degli stati sia un prodotto cartesiano di  $n$  spazi:

$$S = X_1 \times \dots \times X_n$$

$$\underline{s} \in S \quad \underline{s} = (x_1, \dots, x_n)$$

Scegliamo un punto a caso:  $\underline{s} = (\underbrace{s}_1, s_2)$

$s_1$  è la configurazione in tutti gli altri punti,  $s_2$  è la configurazione in quel punto.

Possiamo scegliere  $\underline{s}' = (\underline{s}_1, s'_2)$  tale che ???

Qual è la probabilità di passare da  $s$  a  $s'$ ?

$$p_{\underline{s}, \underline{s}'} = \frac{1}{n} p(s'_2 | \underline{s}_1)$$

$$\alpha_{s, s'} = \min \left( \frac{\pi(\underline{s}' q_{\underline{s}', s})}{\pi(\underline{s}) q_{s, s'}}, 1 \right)$$

Osserviamo che

$$k\pi(s') q_{\underline{s}', \underline{s}} = k\pi(\underline{s}) q_{\underline{s}, s'}$$

Questo perché

$$k\pi(s) = k\pi(s'_1) p(s'_2, \underline{s}_1)$$

e quindi

$$\alpha_{s, s'} = \min \left( \frac{\pi(s_1) p(s'_2 | \underline{s}_1)}{\pi(s_1) p(s_2 | \underline{s}_1)}, 1 \right)$$

$q_{\underline{s}', s}$  a meno di un fattore è data da  $p(s_2 | s_1)$ ; quindi  $\alpha_{s, s} = 1$ .  
???

# Chapter 4

## Catene di Markov con tempo continuo

### 4.1 Distribuzione esponenziale

Nelle catene di Markov con tempo continuo assume importanza la **distribuzione esponenziale**.

Questa distribuzione ha un parametro  $\lambda > 0, \lambda \in \mathbb{R}$ .

Consideriamo una variabile aleatoria  $X$  con distribuzione esponenziale di parametro  $\lambda$ . La funzione di ripartizione ci dà la probabilità che  $X < x$ :

$$F_\lambda(X) = P(X < x) = \begin{cases} 1 - e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

La distribuzione esponenziale è continua. Dal punto di vista probabilistico, vuol dire che la probabilità che  $X$  sia esattamente uguale a un certo valore è 0. Se la funzione fosse continua, la probabilità sarebbe il salto che fa da un punto ???

Introduciamo ora la **densità di probabilità**  $f_\lambda$ :

$$F_\lambda(x) = \int_{-\infty}^x f_\lambda(y) dy$$
$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

La densità non è univocamente definita, potremmo cambiare un punto della funzione e l'integrale non cambierebbe.

Se abbiamo una variabile aleatoria, possiamo calcolare l'attesa e la varianza.

Supponiamo di avere  $X$  con distribuzione esponenziale di parametro  $\lambda$ . L'attesa è data da

$$\begin{aligned}
 P(X) &= \int_{-\infty}^{+\infty} x f_x(x) dx \\
 &= \int_0^{+\infty} x \lambda e^{-\lambda x} dx \\
 &= \left[ -x e^{-\lambda x} \right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx \\
 &= 0 + \frac{1}{\lambda} \int_0^{+\infty} \lambda e^{-\lambda x} dx \\
 &= \frac{1}{\lambda} \left[ -e^{-\lambda x} \right]_0^{+\infty} \\
 &= \frac{1}{\lambda}
 \end{aligned}$$

La varianza è:

$$\text{var}(X) = P(X^2) - P(X)^2$$

$$\begin{aligned}
 P(X^2) &= \int_{-\infty}^{+\infty} x^2 f_\lambda(x) dx \\
 &= \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx \\
 &= \left[ -x^2 e^{-\lambda x} \right]_0^{+\infty} + 2 \int_0^{+\infty} x e^{-\lambda x} dx \\
 &= 0 + \frac{2}{\lambda^2}
 \end{aligned}$$

$$\Rightarrow \text{var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Lo **scarto quadratico medio** è  $\sigma(X) = \sqrt{\text{var}(X)} = \frac{1}{\lambda}$ . Questa quantità è la grandezza con cui  $X$  si allontana dalla sua attesa.



### 4.1.1 Assenza di memoria

La variabile aleatoria esprime un tempo aleatorio. In questo caso la distribuzione ha una proprietà detta **assenza di memoria**, collegata alla proprietà di Markov.

Supponiamo di avere  $X$  con distribuzione esponenziale di parametro  $\lambda$ . Supponiamo  $x > 0, y > 0$ .

$$P(X > x + y | X > x)$$

L'interpretazione della probabilità appena enunciata è la seguente: sapendo che  $X > x$ , ovvero: sapendo che al tempo  $x$  il fatto  $X$  non si è ancora verificato, ci chiediamo quale sia la probabilità che non si verifichi al tempo  $x + y$ :

$$P(X > x + y | X > x) = \frac{P(X > x + y)}{P(X > x)}$$

perché dobbiamo fare l'intersezione di due eventi, ma  $P(X > x + y) < P(X > x)$

$$\begin{aligned} &= \frac{1 - F_\lambda(x + y)}{1 - F_\lambda(x)} \\ &= \frac{1 - (1 - e^{-\lambda(x+y)})}{1 - (1 - e^{-\lambda x})} \\ &= \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} \\ &= e^{-\lambda y} \\ &= P(X > y) \end{aligned}$$

Interpretazione: se al tempo  $x$  non si è ancora verificato l'evento, qual è la probabilità che si verifichi dopo un ulteriore tempo  $y$ ? È la stessa cosa che chiedersi direttamente se l'evento si verifica dopo il tempo  $y$ , come se ripartisse dall'inizio.

Vediamo altre proprietà. Supponiamo di avere due tempi indipendenti  $X, Y$  con distribuzioni esponenziali di parametri rispettivamente  $\lambda_1, \lambda_2$ . Consideriamo

$$Z = \min(X, Y)$$

Qual è la distribuzione di  $Z$ ?  $Z$  avrà ancora distribuzione esponenziale di parametro  $\lambda_1 + \lambda_2$ . Consideriamo un  $x > 0$ . Calcoliamo

$$\begin{aligned}
 P(Z < x) &= 1 - P(> Z) \\
 &= 1 - P(\min(X, Y) > x) \\
 &= 1 - P((X > x)(Y > x)) \\
 &= 1 - P(e^{-\lambda_1 x} e^{-\lambda_2 x}) \\
 &= 1 - e^{-(\lambda_1 + \lambda_2)x}
 \end{aligned}$$

Chiediamoci ora  $P(X < Y)$  (nota: possiamo chiederci anche  $P(X > Y)$ , ma  $P(X = Y)$  è 0 perché essendo la distribuzione continua la probabilità che assuma un certo valore è 0):

$$\begin{aligned}
 &= \int_0^{+\infty} (1 - e^{-\lambda_1 y}) \lambda_2 e^{-\lambda_2 y} dy \\
 &= \int_0^{+\infty} \lambda_2 e^{-\lambda_1 y} dy - \lambda_2 \int_0^{+\infty} e^{-(\lambda_1 + \lambda_2)y} dy \\
 &= 1 - \frac{\lambda_2}{\lambda_1 + \lambda_2} \\
 &= \frac{\lambda_1}{\lambda_1 + \lambda_2} \qquad \Rightarrow P(Y < X) = \frac{\lambda_1}{\lambda_1 + \lambda_2}
 \end{aligned}$$

$$P(Y < X) + P(X < Y) = 1$$

Possiamo estenderlo al caso di  $n$  variabili aleatorie indipendenti  $X_1, X_2, \dots, X_n$  di parametri  $\lambda_1, \lambda_2, \dots, \lambda_n$ :

$$Z = \min(X_1 + \dots + X_n)$$

Possiamo fare il minimo tra  $X_1$  e  $X_2$ , poi il minimo con  $X_3$  etc.  $Z$  ha una distribuzione esponenziale di parametro  $\lambda_1 + \lambda_2 + \dots + \lambda_n$  e

$$\begin{aligned}
 P(Z = X_i) &= P(X_i < \min(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)) \\
 &= \frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_n}
 \end{aligned}$$

### 4.1.2 Tasso di rischio

Quando abbiamo una variabile aleatoria con una distribuzione esponenziale si considera una funzione detta **tasso di rischio**.

Sia  $X$  una variabile aleatoria con densità  $f(x)$ ,  $x > 0$  e  $f(x) = 0$  per  $x < 0$ .

Consideriamo  $P(t \leq X \leq t + 1 | X > t)$ ; il tasso di rischio  $r(t)$  è:

$$r(t) = \lim_{\Delta \rightarrow 0} \frac{P(t \leq X \leq t + 1 | X > t)}{\Delta}$$

Che relazione c'è tra la densità di probabilità e il tasso di rischio?

$$r(t) = \lim_{\Delta \rightarrow 0} \frac{\int_t^{t+\Delta} f(x) dx}{(1 - F(t))\Delta}$$

Siccome  $\Delta$  tende a 0, l'integrale è calcolato nel punto  $t$ :

$$= \frac{f(t)}{1 - F(t)}$$

Osservazione:

$$\begin{aligned} r(t) &= -\frac{\partial}{\partial t} \log(1 - F(t)) \\ &= -\frac{-f(t)}{1 - F(t)} \\ &= \frac{f(t)}{1 - F(t)} \end{aligned}$$

Siccome  $F(0) = 0$  e  $\log(1 - F(0)) = 0$  possiamo scrivere che

$$\begin{aligned} \log(1 - F(t)) &= -\int_0^t r(s) ds \\ 1 - F(t) &= \exp\left(-\int_0^t r(s) ds\right) \\ F(t) &= 1 - \exp\left(-\int_0^t r(s) ds\right) \end{aligned}$$

Questa è la formula che ci permette di calcolare la funzione di ripartizione per una variabile aleatoria sapendo il tasso di rischio. Se il tasso di rischio è pari a  $\lambda$ :

$$r(s) = \lambda \qquad F(t) = 1 - e^{-\lambda t}$$

### 4.1.3 Distribuzione iperesponenziale

Partendo dalla distribuzione esponenziale possiamo ottenere un'altra distribuzione che si ottiene facendo una somma di tante distribuzioni esponenziali: supponiamo di avere  $X_1, X_2, \dots, X_n$  variabili aleatorie indipendenti con distribuzione esponenziale di parametri  $\lambda_1, \dots, \lambda_n$  rispettivamente.

Supponiamo di avere una variabile aleatoria  $X$  indipendente da  $X_1, \dots, X_n$ .

$$P(X = j) = p_j \quad j = 1, \dots, n$$

$X_T$  è iperesponenziale.

$$P(X_T > t)???$$

$$\begin{aligned} P(X_T) &= \sum_{j=1}^n P(X_T > t | T = j) P(T = j) \\ &= \sum_{j=1}^n e^{-\lambda_j t} p_j \end{aligned}$$

Questa non ha la proprietà di assenza di memoria e non è una distribuzione esponenziale.

Qual è il tasso di rischio di questa?

$$r(t) = \frac{\sum_{i=1}^n p_i \lambda_i e^{-\lambda_i t}}{\sum_{i=1}^n p_i e^{-\lambda_i t}}$$

Il tasso di rischio non è costante. Vogliamo vedere cosa succede al limite. Supponiamo che  $\lambda_1$  sia il minimo dei  $\lambda_i$ ,  $\lambda_i = \lambda_1 + \dots + i + 1$

$$r(t) = \sum_{i=1}^n p_i (\lambda_i e^{(-\lambda_i - \lambda_1)t} \xrightarrow[t \rightarrow \infty]{} \lambda_1)$$

???

La distribuzione esponenziale decade più lentamente delle altre. ???

Comunque non è costante, dipende dal tempo.

Vediamo qual è la somma di due variabili indipendenti con distribuzione esponenziale. Presentiamo prima un esempio nel caso discreto. Supponiamo di avere due variabili  $X, Y$  indipendenti con distribuzione di Poisson di parametri rispettivamente  $\lambda_1, \lambda_2$ . Sia  $Z = X + Y$ . Qual è la distribuzione di  $Z$ ? Dobbiamo vedere qual è la probabilità che  $Z$  sia uguale ad un certo

valore.

$$\begin{aligned}
 P(Z = u) &= \sum_{k=0}^n P(X = k)P(Y = n - k) && n \geq 0 \\
 &= \sum_{k=0}^n \frac{\lambda_1^k e^{-\lambda_1}}{k!} \frac{\lambda_2^{n-k} e^{-\lambda_2}}{(n-k)!} \\
 &= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^n \underbrace{\binom{n}{k} \lambda_1^k \lambda_2^{n-k}}_{\text{binomio di Newton}} \\
 &= \frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1+\lambda_2)}
 \end{aligned}$$

Vediamo ora il caso continuo.  $X$  e  $Y$  sono due variabili aleatorie indipendenti con densità  $f(t)$ ,  $g(t)$  rispettivamente. Vogliamo analizzare  $Z = X + Y$ . Sia  $h$  la densità di  $Z$ .

$$\begin{aligned}
 h(t) &= \int f(u)g(t-u)du \\
 &= f * g(t) \quad \text{convoluzione di } f \text{ e } g
 \end{aligned}$$

La **convoluzione** è un'operazione che si può fare in generale e che date due funzioni ne restituisce una.

Proprietà della convoluzione:

1. proprietà commutativa:  $f * g = g * f$
2. bilinearità:  $f * (ag + bh) = af * g + bf * h$  ???

Vogliamo applicare la convoluzione per studiare la somma di due variabili indipendenti con distribuzione esponenziale di parametro  $\lambda$ :  $Z = X + Y$ . La densità di  $X$  e  $Y$  è  $g_\lambda$ . Calcoliamo la densità di  $Z$ :

$$\begin{aligned}
 h(t) &= g_\lambda * g_\lambda(t) \\
 &= \int g_\lambda(u)g_\lambda(t-u)du
 \end{aligned}$$

La densità esponenziale è positiva se ??? è positiva, quindi se  $u$  è negativa è 0 e se  $t - u$  è negativa è 0:

$$\begin{aligned}
&= \int_0^t \lambda e^{-\lambda u} \lambda e^{-\lambda(t-u)} du \\
&= \lambda^2 e^{-\lambda t} \int_0^t du \\
&= \lambda^2 t e^{-\lambda t} \qquad t \geq 0
\end{aligned}$$

#### 4.1.4 Distribuzione di Erlanger

In generale cosa succede se facciamo la somma di  $n$  variabili aleatorie  $X_1, \dots, X_n$  indipendenti, con distribuzioni esponenziali di parametro  $\lambda$ ?

Supponiamo  $Z = X_1 + \dots + X_n$ ;  $Z$  ha densità  $h(t)$  con  $t > 0$ :

$$h(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}$$

Questa è chiamata distribuzione di Erlang di parametri  $(\lambda, n)$ .

È possibile dimostrare l'uguaglianza soprastante per induzione.

Supponiamo sia vera fino a  $n-1$ . Sappiamo che  $Z = X_1 + \dots + X_n$ ; possiamo riscriverla come  $Z = (X_1 + \dots + X_{n-1}) + X_n$

Siccome  $X_1, \dots, X_n$  sono indipendenti, anche la somma evidenziata dalle parentesi tonde è indipendente da  $n$  e possiamo usare la convoluzione:

$$\begin{aligned}
h(t) &= \int_0^t \frac{\lambda^{n-1} u^{n-2}}{(n-2)!} e^{-\lambda u} \lambda e^{-\lambda(t-u)} du \\
&= \frac{\lambda^n}{(n-2)!} e^{-\lambda t} \int_0^t u^{n-2} du \\
&= \frac{\lambda^n}{(n-2)!} e^{-\lambda t} \frac{t^{n-1}}{n-1} \\
&= \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t} ???
\end{aligned}$$

La distribuzione di Erlang, a differenza della distribuzione esponenziale, non ha la proprietà di assenza di memoria. Però posso rappresentarla come una somma di variabili aleatorie indipendenti.

Vediamo ora il caso di  $n$  variabili aleatorie con parametri  $\lambda_1, \dots, \lambda_n$  tutti diversi tra loro. Iniziamo studiando la somma tra due variabili  $X$  e  $Y$ .

Sia  $Z = X + Y$  e  $h(t)$  la densità di  $Z$ :

$$\begin{aligned}
 h(t) &= \int g_{\lambda_1}(u)g_{\lambda_2}(t-u)du \\
 &= \int_0^t \lambda_1 e^{-\lambda_1 u} \lambda_2 e^{-\lambda_2(t-u)} du \\
 &= \lambda_1 \lambda_2 e^{-\lambda_2 t} \int_0^t e^{-u(\lambda_1 - \lambda_2)} du \\
 &= \lambda_1 \lambda_2 e^{-\lambda_2 t} \left[ \frac{e^{-u(\lambda_1 - \lambda_2)}}{\lambda_2 - \lambda_1} \right]_0^t \\
 &= \frac{\lambda_1 \lambda_2 e^{-\lambda_2 t}}{\lambda_2 - \lambda_1} e^{-(\lambda_1 - \lambda_2)t} + \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} \\
 &= \frac{\lambda_2}{\lambda_2 - \lambda_1} \lambda_1 e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_2 t} \\
 \Rightarrow g_{\lambda_1} * g_{\lambda_2} &= \frac{\lambda_2}{\lambda_2 - \lambda_1} g_{\lambda_1} + \frac{\lambda_1}{\lambda_1 - \lambda_2} g_{\lambda_2}
 \end{aligned}$$

Consideriamo ora effettivamente  $n$  distribuzioni esponenziali  $X_1, X_2, \dots, X_n$  indipendenti con parametri  $\lambda_1, \dots, \lambda_n$  diversi tra loro. Bisogna fare una convoluzione di  $n$  parametri, nell'ordine che si preferisce.

???

Mi basta il coefficiente.

$$\begin{array}{ll}
 \text{Convoluzione con} & g_{\lambda_1} : \frac{\lambda_1}{\lambda_1 - \lambda_i} \\
 & g_{\lambda_2} : \frac{\lambda_2}{\lambda_2 - \lambda_1} \\
 & \dots \\
 & g_{\lambda_n} : \frac{\lambda_n}{\lambda_n - \lambda_i}
 \end{array}$$

La densità di  $X_1 + \dots + X_n$  è:

$$h(t) = \sum_{i=1}^n \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} g_{\lambda_i}$$

Sembra analoga alla distribuzione iperesponenziale, ma non è così perché i coefficienti non sono necessariamente tutti positivi. Infatti il denominatore ha la differenza  $\lambda_j - \lambda_i$ , quindi può assumere valori negativi. Questa distribuzione si chiama **ipoesponenziale**.

Il fattore di rischio non è costante. Cosa succede mandando il tempo all'infinito? Supponiamo  $\lambda_1$  sia il più piccolo di tutti:  $\lambda_1 = \min(\lambda_1, \dots, \lambda_n)$

$$\prod_{i \neq 1, j \neq 1}^n \left( \frac{\lambda_j}{\lambda_j - \lambda_i} \right) g_{\lambda_1} ???$$

Tasso di rischio:

$$r(t) = \frac{\sum_{i=1}^n \prod_{j \neq i}^n \left( \frac{\lambda_j}{\lambda_j - \lambda_i} \right) g \lambda_i}{\sum_{i=1}^n \prod_{j \neq i}^n \left( \frac{\lambda_j}{\lambda_j - \lambda_i} \right) e^{-\lambda_i t}} ???$$

$$\lambda_1 = \min(\lambda_1, \dots, \lambda_n)$$

$$r(t) \xrightarrow[t \rightarrow \infty]{} \lambda_1$$

## 4.2 Processo di Poisson

Fa parte di una famiglia di processi detti processi di conteggio, un tipo di processo che descrive il numero di volte in cui si verifica un certo fatto al variare del tempo:  $N_t, t \geq 0$  indica quante volte si è verificato un fatto (es. decadimento radioattivo, arrivo di un cliente allo sportello...) dal tempo 0 al tempo  $t$ .

### Proprietà dei processi di conteggio

1.  $N_0 = 0$ ;
2.  $N_t$  ha valori naturali;
3.  $N_t$  è non decrescente (quindi può rimanere uguale).

Il processo di Poisson con parametro  $\lambda$  è un processo di conteggio  $N_t$  con in più le seguenti proprietà:

1. per  $0 \leq t_1 \leq t_2$ ,  $N_{t_2} - N_{t_1}$  (numero di fatti verificatisi tra  $t_1$  e  $t_2$ ) ha distribuzione di Poisson di parametro  $\lambda(t_2 - t_1)$ ;
2. ha incrementi indipendenti.

L'ultimo punto vuol dire che se si prendono tanti intervalli disgiunti  $(0, t_1], (t_1, t_2], \dots, (t_{n-1}, t_n]$  con  $0 \leq t_1 < t_2 < \dots < t_n$ , allora  $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$  sono indipendenti.

Osservazione di compatibilità: se prendo  $(0, t_1], (t_1, t_2], \dots, (0, t_2]$ :

$$N_{t_2} - N_0 = (N_{t_2} - N_{t_1}) + (N_{t_1} - N_0)$$

Queste distribuzioni (di parametri rispettivamente  $\lambda(t_2 - 0), \lambda(t_2 - t_1), \lambda(t_1 - 0)$ ) sono indipendenti.



Supponiamo di prendere  $u_1, \dots, u_n$  indipendenti con distribuzioni esponenziali di parametro  $\lambda$ . Immaginiamo che  $u_1, \dots, u_n$  descrivano il tempo d'arrivo di clienti ad uno sportello. Allora

$$T_k = u_1 + \dots + u_n$$

ovvero  $T_k$  è il tempo dell'arrivo del cliente k-esimo. Posso descrivere un processo di conteggio (quanti clienti sono arrivati entro il tempo  $t$ ?)

$$N_t = \max\{k | T_k \leq t\}, \quad T_k \leq t \leq T_{k+1}$$

Allora  $N_t$  è un processo di Poisson di parametro  $\lambda$ .

Quanto vale quindi  $P(N_t = k)$ ?

$$\begin{aligned} P(N_t = k) &= P(T_k \leq t < T_{k+1}) \\ &= P((T_k \leq t) \cap (T_{k+1} > t)) \\ &= P(T_k \leq t) - P(T_{k+1} \leq t) \end{aligned}$$

Possiamo dire che

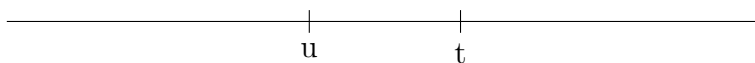
$$P(X_k \leq t) - P(X_{k+1} \leq t) = \int_0^t \frac{\lambda^k s^{k-1}}{(k-1)!} e^{-\lambda s} ds - \int_0^t \frac{\lambda^{k+1} s^k}{k!} e^{-\lambda s} ds$$

Integro per parti:

$$\begin{aligned} &= \lambda^k \left[ \frac{s^k}{k!} e^{-\lambda s} \right]_0^t + \int_0^t \frac{\lambda^{k+1} s^k}{k!} e^{-\lambda s} ds - \int_0^t \frac{\lambda^{k+1} s^k}{k!} e^{-\lambda s} ds \\ &= \frac{\lambda^k t^k}{k!} e^{-\lambda t} \end{aligned}$$

Ma questa è la probabilità del processo di Poisson ???

Adesso devo verificare l'indipendenza: ??? intervalli separati sono indipendenti.



Prendiamo l'intervallo tra 0 e  $u$ ; ci chiediamo quanto vale

$$P(N_t - N_u = k | N_s, s \leq u)$$

???

$$\frac{\lambda(t-u)^k}{k!} e^{-\lambda(t-u)}$$

Questo rapporto tra distribuzione esponenziale e processo di Poisson ci permette di trovare un metodo per simulare il processo di Poisson.

Consideriamo una successione di ???

Come si fa a generare una successione di variabili aleatorie con distribuzione esponenziale? Si usa questo fatto: sia  $Z$  una variabile con distribuzione uniforme in  $[0, 1]$  e sia  $F$  la funzione di ripartizione dell'esponenziale:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

$F$  è una funzione strettamente crescente. Posso considerare la funzione inversa  $F^{-1}$ , e quello che si vede è che  $F^{-1}(Z) = W$  è una variabile aleatoria con distribuzione esponenziale di parametro  $\lambda$ .

Esempio:

$$\begin{aligned} P(W \leq x = F^{-1})(Z \leq x) \\ = P(Z \leq F(x)) \\ = F(x) \end{aligned}$$

In questo caso è facile anche calcolare l'inversa:

$$\begin{aligned} x &= F^{-1}(y) \\ F(x) &= y \\ y &= 1 - e^{-\lambda x} \\ e^{-\lambda x} &= 1 - y \\ -\lambda x &= \log(1 - y) \\ x &= -\frac{1}{\lambda} \log(1 - y) \end{aligned}$$

Se prendo una variabile aleatoria  $U$  con distribuzione uniforme su  $[0, 1]$  e definisco  $W$ :

$$W = -\frac{1}{\lambda} \log(1 - U)$$

$W$  è una distribuzione esponenziale di parametro  $\lambda$ .

Abbiamo visto il caso di una sola variabile; se vogliamo generare una successione di variabili  $U_1, U_2, \dots$  indipendenti con distribuzioni uniformi in  $[0, 1]$  definisco  $W_1, W_2, \dots$  nel seguente modo:

$$W_i = \frac{1}{\lambda} \log(1 - U_i)$$

Osservazione: in base alla definizione di processo di Poisson è un conteggio a valori interi non decrescenti.

In linea di principio un processo di Poisson potrebbe fare salti maggiori di 1, in realtà la probabilità che accada è 0.

Consideriamo un intervallo  $[0, t]$  e dividiamolo in intervalli di lunghezza  $\frac{1}{2^N}$ .

Qual è la probabilità che in uno di questi intervalli il processo di Poisson faccia un salto più grande di 1? Siccome ??? questo è uguale alla probabilità che nel primo salto non faccia salti più grandi di 1, elevato al numero degli intervalli:

$$\begin{aligned} (P(N_{\frac{1}{2}N} \leq 1))^{2^N} &= (P(N_{\frac{1}{2}N} = 0) + P(N_{\frac{1}{2}N} = 1))^{2^N} \\ &= (e^{-\frac{\lambda}{2^N}} + \frac{\lambda}{2^N} e^{-\frac{\lambda}{2^N}})^{2^N} \\ &= e^{-\lambda} (1 + \frac{\lambda}{2^N})^{2^N} \end{aligned}$$

$$\lim_{N \rightarrow \infty} e^{-\lambda} (1 + \frac{\lambda}{2^N})^{2^N} = 1$$

??? Quindi non ci sono salti più grandi di 1. ??? Quindi o non ha salti oppure ha tutti salti di grandezza 1.

### 4.3 Decomposizione di un processo di Poisson

È un'altra proprietà dei processi di Poisson. Supponiamo di avere un processo di Poisson di parametro  $\lambda$  che descrive gli arrivi dei clienti ad un supermercato (quanti clienti sono arrivati prima di un tempo  $t$ ). Dividiamo i clienti in due categorie, ad esempio maschi e femmine. Il cliente è in una certa categoria, ad esempio "maschi", con probabilità  $p$  ed è nell'altra categoria con probabilità  $1 - p$ . È come se ogni volta che arriva un cliente tirassimo una moneta, se esce testa finisce nella prima categoria, se esce croce nella seconda.

Possiamo allora scrivere il nostro processo come due processi di conteggio:

$$N_t = N_t^{(1)} + N_t^{(2)}$$

$N_t^{(1)}$  e  $N_t^{(2)}$  sono due processi di Poisson indipendenti.

Vogliamo calcolare ora:

$$\begin{aligned} P(N_t^{(1)} = k_1, N_t^{(2)} = k_2) &= P(N_t^{(1)} = k_1, N_t^{(2)} = k_2 | N_t = k_1 + k_2) P(N_t = k_1 + k_2) \\ &= p \binom{k_1 + k_2}{k_1} p^{k_1} (1-p)^{k_2} - \frac{(\lambda t)^{k_1 + k_2}}{(k_1 + k_2)!} e^{-\lambda t} \\ &= \frac{(\lambda p t)^{k_1}}{k_1!} \frac{[\lambda(1-p)t]^{k_2}}{k_2!} e^{-\lambda p t} e^{-\lambda(1-p)t} \\ &= \frac{(\lambda p t)^{k_1}}{k_1!} e^{-\lambda p t} \frac{[\lambda(1-p)t]^{k_2}}{k_2!} e^{-\lambda(1-p)t} \end{aligned}$$

$N_t^{(1)}$  è processo di Poisson di parametro  $\lambda p$

$N_t^{(2)}$  è processo di Poisson di parametro  $\lambda(1-p)$

In questo modo abbiamo mostrato per un solo intervallo che sono indipendenti, dobbiamo ancora verificare che gli incrementi su intervalli separati siano indipendenti.

Ma siccome partiamo già da variabili aleatorie indipendenti, otterremo ancora ???

Questo si può generalizzare nella decomposizione in  $n$  processi di Poisson: consideriamo

$$\begin{aligned} P_1, P_2, \dots, P_n \quad & P_1 + P_2 + \dots + P_n = 1 \quad 0 \leq p_i \leq 1 \\ N_t &= N_t^{(1)} + \dots + N_t^{(n)} \end{aligned}$$

I processi  $N_t^{(1)}, \dots, N_t^{(n)}$  sono anche loro di Poisson indipendenti di parametri rispettivamente  $\lambda_{p_1}, \dots, \lambda_{p_n}$ .

Lo stesso argomento si può mostrare al contrario: supponiamo di avere  $N_t^{(1)}, \dots, N_t^{(n)}$  processi di Poisson indipendenti di parametri  $\lambda_1, \dots, \lambda_n$ . Definiamo:

$$N_t = N_t^{(1)} + \dots + N_t^{(n)}$$

Allora  $N_t$  è processo di Poisson di parametro  $\lambda = \lambda_1 + \dots + \lambda_n$ .

Supponiamo  $N_t$  sia un processo di Poisson di parametro  $\lambda = \lambda_{p_1} + \dots + \lambda_{p_n}$ ; definiamo

$$p_1 = \frac{\lambda_1}{\lambda_1 + \dots + \lambda_n}, \dots, p_n = \frac{\lambda_n}{\lambda_1 + \dots + \lambda_n}$$

$N_t^{(1)}, \dots, N_t^{(n)}$  sono indipendenti di parametro  $\lambda_{p_i} = \lambda_i$

## 4.4 Definizione equivalente di processo di Poisson

Processo di conteggio con incrementi indipendenti

$$P(N_{t+h} = N_t) = 1 - \lambda h + o(h)$$

$$P(N_{t+h} = N_t + 1) = \lambda h + o(h)$$

??? uniforme in  $t$

Cioè anziché ???  $o(h)$ : infinitesimo di ordine superiore ad  $h$ , ovvero quantità tale che:

$$\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$$

Vogliamo mostrare che le due definizioni sono equivalenti.

Qual è la probabilità di nessun incremento?

$$\begin{aligned} P(N_{t+h} = N_t) &= e^{-\lambda h} \\ &= 1 - \lambda h + o(h) \end{aligned}$$

Incremento di 1?

$$\begin{aligned} P(N_{t+h} = N_t + 1) &= \lambda h e^{-\lambda h} \\ &= \lambda h (1 - \lambda h + o(h)) \\ &= \lambda h - \lambda^2 h^2 + o(h^2) \\ &= \lambda h + o(h) \end{aligned}$$

Incremento  $k \geq 2$ ?

$$\begin{aligned} P(N_{t+h} = N_t + k) &\leq 1 - P(N_{t+h} = N_t) - P(N_{t+h} = N_t + 1) \\ &= 1 - (1 - \lambda h + o(h)) - \lambda h + o(h) \\ &= o(h) \end{aligned}$$

Quindi la probabilità di fare un salto più grande di 2 è un infinitesimo di ordine superiore a  $h$ .

Il processo di Poisson soddisfa queste condizioni. ??? Per concludere la dimostrazione che le due definizioni sono equivalenti, dobbiamo mostrare anche l'altro verso dell'implicazione, cioè ???

$$P(N_{t+h} = 0) = P(N_t = 0)P(N_{t+h} - N_t = 0)$$

$$P(N_{t+h} = 0) - P(N_t = 0) = P(N_{t+h} - N_t = 0 - 1)P(N_t = 0)$$

$$u_t = P(N_t = 0)$$

$$\begin{aligned} u_{t+h} - u_t &= 0 \\ &= ((1 - \lambda h + o(h)) - 1)u_t \end{aligned}$$

$$\frac{u_{t+h} - u_t}{h} = \lambda u_t + \frac{o(h)}{h}$$

$$\lim_{h \rightarrow 0} u'_t = \lambda u_t \qquad u_0 = 1$$

$$\begin{aligned} &???ke^{-\lambda t} & k = 1 \\ u_t = e^{-\lambda t} & P(N_t = 0)e^{-\lambda t} \end{aligned}$$

Questo corrisponde al processo di Poisson. Devo farlo vedere anche per valori diversi.

Consideriamo  $P(N_{t+h} = k), k \geq 1$