

1. Catene di Markov

Lo studio delle catene di Markov è piuttosto avvincente e conduce a risultati di utilità pratica, anche perché mette in relazione diverse strutture matematiche:

- i) i grafi,
- ii) le matrici,
- iii) i processi stocastici.

Come preliminare, inquadrriamo il discorso parlando più in generale dei processi stocastici a tempo discreto e stati discreti.

1.1. Processi a tempo discreto e stati discreti

Chiamiamo processo stocastico a tempo discreto e stati discreti una successione di v.a.

$$X_0, X_1, \dots, X_n, \dots$$

dove ciascuna X_n è una v.a. discreta a valori in un insieme S , detto *spazio degli stati*.

L'insieme S è finito o al più numerabilmente infinito. Senza perdere in generalità, supponiamo che S sia un sottoinsieme dei numeri interi relativi \mathbb{Z} .

Penseremo all'indice n di X_n come al *tempo*; chiameremo poi *stati* i possibili valori di X_n .

Example 1. *consideriamo la coda che si forma ad uno sportello che offre un servizio. Osserviamo la coda solo ad istanti di tempo discreti, ad esempio ogni minuto. Al minuto n -esimo ci sarà in coda un certo numero di persone, aleatorio, che indichiamo con N_n . Qui la v.a. N_n può assumere tutti i valori interi non negativi. Al variare di n abbiamo quindi un esempio di processo stocastico del tipo descritto sopra.*

Al trascorrere del tempo (discreto), il processo può saltare da uno stato all'altro. Se ad un certo istante n si trova in uno stato i ed all'istante successivo $n+1$ si trova in uno stato $j \neq i$, diremo che c'è stata una *transizione*. Useremo, a volte, questo linguaggio anche quando non è necessariamente $j \neq i$, solo per unità espositiva.

Dato un processo stocastico (X_n) , ci interessa calcolare certe probabilità ad esso associate. Le più semplici sono quelle a tempo fissato, della forma

$$P(X_n = i)$$

che rappresentano la probabilità di osservare al tempo n il sistema (descritto dal processo) nello stato i . Indicheremo questi numeri col simbolo $p_i^{(n)}$ ed useremo il simbolo $p^{(n)}$ per indicare il *vettore* di componenti $p_i^{(n)}$. $p^{(n)}$ è la *distribuzione* della v.a. X_n . Per ogni istante di tempo n vale $\sum_{i \in \mathbb{Z}} p_i^{(n)} = 1$.

Oltre a queste semplici probabilità ci può interessare il calcolo di probabilità più complesse, ad esempio che coinvolgono più istanti temporali contemporaneamente. Ad esempio può essere interessante calcolare la probabilità di trovarsi nello stato j al tempo $n + 1$ sapendo di essere nello stato i al tempo n :

$$P(X_{n+1} = j | X_n = i)$$

detta *probabilità di transizione* da i a j al tempo n . Usando la definizione di probabilità condizionale, questa si riscrive nella forma

$$\frac{P(X_n = i, X_{n+1} = j)}{P(X_n = i)}.$$

Pertanto, per questo calcolo è sufficiente conoscere $p_i^{(n)}$ e le *probabilità congiunte* della forma $P(X_n = i, X_{n+1} = j)$. Si capisce che per calcolare espressioni più complesse serve la conoscenza delle generiche probabilità congiunte della forma

$$P(X_0 = i_0, \dots, X_n = i_n)$$

al variare di tutti gli n e tutti gli $i_0, \dots, i_n \in \mathbb{Z}$. In un certo senso, queste probabilità esauriscono le informazioni possibili: il processo stocastico è completamente determinato statisticamente quando sono note tutte le densità (discrete) congiunte, ovvero le densità di tutte le variabili discrete multiple (X_1, \dots, X_n) al variare di tutti gli n e tutti gli $i_0, \dots, i_n \in \mathbb{Z}$. Non dimostriamo questo fatto non banale.

La conoscenza di queste probabilità congiunte è in generale un problema molto arduo. Bisogna introdurre una qualche semplificazione e rinunciare alla totale generalità, per ottenere risultati pratici. Il concetto che viene maggiormente in aiuto è quello di catena di Markov. L'idea, dal punto di vista matematico, può partire ad esempio dal seguente calcolo. Usando il concetto di probabilità condizionale, vale

$$\begin{aligned} & P(X_0 = i_0, \dots, X_{n+1} = i_{n+1}) \\ = & P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) \cdot P(X_0 = i_0, \dots, X_n = i_n). \end{aligned}$$

Pertanto, se fossero note le probabilità condizionali $P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n)$, potremmo calcolare le probabilità congiunte $P(X_0 = i_0, \dots, X_{n+1} = i_{n+1})$ in modo *ricorsivo*, cioè tramite $P(X_0 = i_0, \dots, X_n = i_n)$, e così via. Ma le probabilità condizionali scritte sopra sono oggetti altrettanto complicati, salvo che per le catene di Markov, specialmente quelle omogenee.

1.2. Catene di Markov omogenee

Un processo stocastico discreto $(X_n)_{n \geq 0}$ si dice di *Markov* se

$$P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = P(X_{n+1} = i_{n+1} | X_n = i_n). \quad (1.1)$$

Questa è detta *proprietà di Markov*. Si intende che questa relazione deve valere al variare di tutti i parametri, e qualora siano ben definite le probabilità condizionate in oggetto.

Si dice poi *omogeneo* se le probabilità di transizione $P(X_{n+1} = j | X_n = i)$ non dipendono da n , ma solo da i e j . Quando questo accade, posto

$$p_{ij} = P(X_{n+1} = j | X_n = i)$$

possiamo calcolare tutte le probabilità congiunte conoscendo i soli numeri p_{ij} più la distribuzione iniziale $p_i^{(0)} = P(X_0 = i)$ (distribuzione del processo al tempo zero). Le probabilità p_{ij} si dicono *probabilità di transizione*, e precisamente p_{ij} è la probabilità di transizione da i a j in un passo temporale.

Proposition 1. *Se (X_n) è una catena di Markov omogenea con probabilità di transizione p_{ij} e distribuzione iniziale $p_i^{(0)} = P(X_0 = i)$, allora*

$$P(X_0 = i_0, \dots, X_n = i_n) = p_{i_0}^{(0)} \cdot p_{i_0 i_1} \cdots p_{i_{n-1} i_n}.$$

Dimostrazione. Usando la nota formula $P(A \cap B) = P(A|B)P(B)$ e le ipotesi, vale

$$\begin{aligned} & P(X_0 = i_0, \dots, X_n = i_n) \\ &= P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \cdot P(X_n = i_n | X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \\ &= P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \cdot P(X_n = i_n | X_{n-1} = i_{n-1}) \\ &= P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \cdot p_{i_{n-1} i_n}. \end{aligned}$$

Riapplicando in maniera ricorsiva questo calcolo al termine $P(X_0 = i_0, \dots, X_{n-1} = i_{n-1})$ e così via, otteniamo

$$\begin{aligned} &= P(X_0 = i_0, \dots, X_{n-2} = i_{n-2}) \cdot p_{i_{n-2}i_{n-1}} \cdot p_{i_{n-1}i_n} = \dots \\ &= P(X_0 = i_0) \cdot p_{i_0i_1} \cdots p_{i_{n-1}i_n}. \end{aligned} \tag{1.2}$$

La dimostrazione è completa.

Remark 1. *L'interpretazione della proprietà di Markov è la seguente: la previsione statistica dello stato $X_{n+1} = i_{n+1}$ che si può effettuare conoscendo la storia passata e presente $X_0 = i_0, \dots, X_n = i_n$, è identica alla previsione statistica effettuabile conoscendo il solo stato presente $X_n = i_n$. E' una forma di perdita di memoria del passato, una volta noto lo stato presente.*

Remark 2. *L'omogeneità temporale ha un'interpretazione immediata: la dinamica stocastica che determina i salti del processo di Markov non dipende dal tempo, ovvero agisce sempre con le stesse modalità (aleatorie ma tempo-invarianti dal punto di vista statistico).*

A livello di notazioni, esistono vari usi in letteratura per le probabilità di transizione. Ciò che sopra abbiamo indicato con p_{ij} viene a volte indicato con $p(i, j)$, con $p(j|i)$, con $p_i(j)$, con $p_{i \rightarrow j}$ e possibilmente altre.

Conoscendo le probabilità di transizione p_{ij} e la distribuzione iniziale $p^{(0)}$ è possibile anche il calcolo delle $p_i^{(n)}$.

Proposition 2. *Se (X_n) è una catena di Markov omogenea con probabilità di transizione p_{ij} e distribuzione iniziale $p^{(0)}$, allora*

$$p_i^{(n)} = \sum_{i_0, i_1, \dots, i_{n-1}} p_i^{(0)} \cdot p_{i_0i_1} \cdots p_{i_{n-1}i}.$$

Dimostrazione. Per la formula delle probabilità totali vale

$$\begin{aligned} p_i^{(n)} &= P(X_n = i) \\ &= \sum_{i_0, i_1, \dots, i_{n-1}} P(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \end{aligned}$$

e quindi, usando la proposizione precedente,

$$p_i^{(n)} = \sum_{i_0, i_1, \dots, i_{n-1}} p_i^{(0)} \cdot p_{i_0i_1} \cdots p_{i_{n-1}i}.$$

La dimostrazione è completa.

1.3. Grafi e matrici

Lo studio delle catene di Markov omogenee diventa particolarmente semplice ed efficace utilizzando la rappresentazione tramite grafi e la descrizione con le matrici. In particolare, la formula espressa dalla proposizione precedente diventa molto più leggibile. Rappresentiamo gli stati i, j, \dots della catena con dei piccoli cerchi nel piano, detti *nodì* o *vertici* del grafo, e tracciamo una freccia dal nodo i al nodo j se $p_{ij} > 0$. Da ogni punto esce almeno una freccia. Alcuni punti possono non essere connessi da frecce, alcune coppie di punti possono essere connessi da frecce in ambo i sensi, alcuni punti possono essere connessi a se stessi, ed eventualmente solo a se stessi.

Ad ogni connessione $i \rightarrow j$ associamo il numero p_{ij} , la probabilità di transizione dallo stato i allo stato j . I numeri p_{ij} soddisfano le proprietà:

$$p_{ij} \in [0, 1], \quad \sum_j p_{ij} = 1 \text{ per ogni } i.$$

Si può ribaltare il discorso: partendo da un grafo munito di numeri p_{ij} con queste proprietà è possibile costruire una catena di Markov con quelle probabilità di transizione. Introduciamo anche la matrice quadrata

$$P = (p_{ij})$$

detta *matrice di transizione*. Essa ha elementi non negativi, con somma uno su ciascuna riga.

Usando la matrice P possiamo riscrivere una formula importante vista sopra. Osserviamo che vale

$$(P \cdot P)_{ij} = \sum_{i_1} p_{i, i_1} p_{i_1 j}$$

e più in generale

$$(P^n)_{ij} = \sum_{i_1, \dots, i_{n-1}} p_{i i_1} \cdots p_{i_{n-1} j}.$$

Pertanto vale:

Proposition 3. *Se (X_n) è una catena di Markov omogenea con probabilità di transizione p_{ij} e distribuzione iniziale $p^{(0)}$, allora*

$$p_j^{(n)} = \sum_i p_i^{(0)} \cdot (P^n)_{ij}.$$

In termini vettoriali

$$p^{(n)} = p^{(0)} \cdot P^n$$

dove intendiamo i vettori $p^{(n)}$ e $p^{(0)}$ come vettori riga, ed il prodotto $p^{(0)} \cdot P^n$ come prodotto di vettore riga per matrice.

In modo analogo si può dimostrare la seguente relazione tra i vettori delle probabilità a due istanti di tempo generici $k < n$:

$$p^{(n)} = p^{(k)} P^{n-k}.$$

Questa relazione dice che se si parte dalla distribuzione $p^{(k)}$ (a cui si era arrivati in k passi, ma che ora viene vista come distribuzione iniziale) e si eseguono $n - k$ passi, si arriva alla distribuzione $p^{(n)}$. Il passato prima del tempo k non serve per determinare $p^{(n)}$, quando si conosca $p^{(k)}$.

Remark 3. A volte è utile considerare la probabilità di transizione a n passi

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i).$$

Essa è la probabilità che il sistema si trovi al tempo n nello stato j , sapendo che è partito all'istante $t = 0$ dallo stato i . Si noti che essa non è la probabilità che il sistema resti in i fino al tempo $n - 1$ e poi all'istante n avvenga la transizione $i \rightarrow j$: il passaggio da i a j in n passi può avvenire attraverso vari passi intermedi. $p_{ij}^{(n)}$ sarà la somma delle probabilità che, sapendo che il sistema è partito all'istante $t = 0$ dallo stato i , successivamente abbia seguito questo o quel cammino $i \rightarrow i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_{n-1} \rightarrow i_n = j$ arrivando in j al tempo n . Allora

$$\begin{aligned} p_{ij}^{(n)} &= P(X_n = j | X_0 = i) \\ &= \sum_{i_1, i_2, \dots, i_{n-1}} P(X_n = j, X_{n-1} = i_{n-1}, \dots, X_1 = i_1 | X_0 = i) \\ &= \sum_{i_1, i_2, \dots, i_{n-1}} P(X_n = j, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i) \frac{1}{P(X_0 = i)}. \end{aligned}$$

Usando la relazione trovata sopra per le probabilità congiunte, troviamo

$$p_{ij}^{(n)} = \sum_{i_1, i_2, \dots, i_{n-1}} p_{i, i_1} p_{i_1, i_2} \dots p_{i_{n-2}, i_{n-1}} p_{i_{n-1}, i_n}.$$

Usando il simbolismo matriciale, abbiamo dimostrato che

$$p_{ij}^{(n)} = (P^n)_{ij}$$

dando così anche un'interpretazione agli elementi della matrice P^n .

Remark 4. Un commento di natura “fisica” sulla proprietà di Markov. Essa afferma che, noto il presente, il futuro non dipende dal passato. Questa proprietà è simile a quanto avviene in fisica per il concetto di stato: esso indica (di solito) l’insieme delle informazioni ad un certo istante che permettono di determinare le analoghe informazioni ad istanti successivi (tramite le leggi del moto), senza che si debba conoscere il passato. Ad esempio, per il moto di un punto materiale, la conoscenza di posizione e velocità permette di calcolare posizione e velocità agli istanti successivi (lo stato di un punto materiale che si muove nello spazio è il vettore a 6 componenti dato dalla sua posizione e velocità). Invece, se ad un certo istante conosciamo solo la posizione, diversa sarebbe la nostra previsione sul futuro se in aggiunta ci fosse data la posizione nel passato: da quest’ultima (derivando in t) potremmo ricavare la velocità all’istante presente e poi predire lo stato futuro. Quindi la sola posizione non è lo stato di un punto materiale. Da queste analogie si capisce perché chiamiamo stati i punti i, j, \dots delle catene di Markov, e soprattutto perché è naturale in certi problemi la proprietà di Markov.

Remark 5. Senza dirlo esplicitamente, sopra abbiamo considerato matrici infinite, cioè con infinite righe e colonne. Le operazioni matriciali sono inalterate. Nei casi più complicati, facendo un prodotto di due matrici infinite si devono eseguire delle serie (si pensi al significato di $(AB)_{ij} = \sum_k a_{ik}b_{kj}$). Tuttavia, nella maggior parte degli esempi tali serie si riducono a somme finite, perché quasi tutti i termini sono nulli. Quindi anche algebricamente la situazione non è molto diversa.

Exercise 1. Data la catena a tre stati 1,2,3, con probabilità di transizione $p_{11} = \frac{1}{2}$, $p_{12} = \frac{1}{4}$, $p_{13} = \frac{1}{4}$, $p_{21} = 1$, $p_{32} = 1$, sapendo che $P(X_0 = 1) = 1$, calcolare la distribuzione di X_3 .

Soluzione n. 1. Ovviamente è corretto usare la formula $p^{(3)} = p^{(0)}P^3$. Conviene però dare una risoluzione grafica, che presentando meno calcoli è anche più sicura. A tal scopo, prendiamo ad esempio $p_1^{(3)} = P(X_3 = 1)$. Dobbiamo vedere in quali modi può il sistema, partendo da 1 (questa è la condizione data per ipotesi), trovarsi in 1 dopo tre passi. Un cammino è quello in cui passa da 1 ad 1 per tre volte di seguito, che ha probabilità $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$. Un secondo cammino va da 1 a 2, poi torna in 1, e poi va di nuovo in 1; la sua probabilità è $\frac{1}{4} \cdot 1 \cdot \frac{1}{2} = \frac{1}{8}$. Un terzo è $1 \rightarrow 1 \rightarrow 2 \rightarrow 1$, con probabilità $\frac{1}{2} \cdot \frac{1}{4} \cdot 1 = \frac{1}{8}$. Infine, un quarto è $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$ che ha probabilità $\frac{1}{4} \cdot 1 \cdot 1 = \frac{1}{4}$. Complessivamente, $P(X_3 = 1) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{4} = \frac{5}{8}$. Si calcolino analogamente $P(X_3 = 2)$ e $P(X_3 = 3)$.

Soluzione n. 2. A titolo di esempio, verifichiamo il risultato precedente tramite la formula $p^{(3)} = p^{(0)}P^3$. Vale

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

e quindi

$$\begin{aligned} P \cdot P &= \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 1 & \frac{1}{2} \cdot \frac{1}{4} & \frac{1}{2} \cdot \frac{1}{4} \\ & 1 \cdot \frac{1}{4} & 1 \cdot \frac{1}{4} \\ & 1 \cdot \frac{1}{4} & 1 \cdot \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \end{aligned}$$

$$P^3 = P^2 \cdot P = \begin{pmatrix} \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{5}{8} & \frac{3}{16} & \frac{3}{16} \\ \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

Infine, essendo $p^{(0)} = (1, 0, 0)$, vale

$$p^{(3)} = p^{(0)}P^3 = (1, 0, 0) \cdot \begin{pmatrix} \frac{5}{8} & \frac{3}{16} & \frac{3}{16} \\ \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} = \left(\frac{5}{8} \quad \frac{3}{16} \quad \frac{3}{16} \right).$$

In particolare $P(X_3 = 1) = p_1^{(3)} = \frac{5}{8}$, come con l'altro metodo.

Exercise 2. Si consideri la catena di Markov a due stati 1, 2 e probabilità di transizione

$$p_{11} = p, p_{12} = 1 - p, p_{22} = 1$$

dove $p \in (0, 1)$ è un numero dato. Calcolare le probabilità di transizione ad n passi.

Soluzione. La matrice di transizione è semplicemente

$$P = \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix}.$$

Le sue potenze sono

$$P^2 = \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p^2 & pq + q \\ 0 & 1 \end{pmatrix},$$

$$P^3 = \begin{pmatrix} p & q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p^2 & pq + q \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p^3 & p^2q + pq + q \\ 0 & 1 \end{pmatrix}$$

ecc., da cui si capisce che

$$P^n = \begin{pmatrix} p^n & p^{n-1}q + p^{n-2}q + \dots + q \\ 0 & 1 \end{pmatrix}.$$

Da essa si leggono le probabilità di transizione ad n passi. Osserviamo però che la somma sulle righe è uno, quindi accorgiamoci che

$$p^{n-1}q + p^{n-2}q + \dots + q = 1 - p^n.$$

Si verifichi algebricamente questa identità. Inoltre, si osservi che potevamo stabilire l'uguaglianza

$$p_{11}^{(n)} = p^n$$

senza fare la potenza n -esima di P , ma tramite semplici considerazioni grafiche. Quindi, in definitiva, si poteva pervenire alle probabilità di transizione ad n passi in modo grafico e con semplici considerazioni.

Exercise 3. Calcolare inoltre la probabilità che il sistema prima o poi si trovi in 2, sapendo che all'istante $t = 0$ si trova in 1. (Attenzione: la probabilità di trovarsi prima o poi in 2 (condizionata...) non è la somma su n delle probabilità di trovarsi in 2 al tempo n , sapendo che all'istante $t = 0$ si trova in 1, ovvero $\sum_{n=1}^{\infty} p_{12}^{(n)}$. Infatti questi non sono eventi disgiunti. Provate a verificare che $\sum_{n=1}^{\infty} p_{12}^{(n)} = +\infty$.)

2. Probabilità invarianti

Data una catena di Markov, il calcolo delle probabilità congiunte o di $p_i^{(n)}$ è in linea di principio possibile ed agevole usando la matrice di transizione P e la distribuzione iniziale. Con l'uso di un calcolatore non è troppo dispendioso calcolare queste probabilità anche per n alto. Se però cerchiamo risultati senza calcolatore, abbiamo visto già negli esercizi precedenti che le possibilità di calcolo delle $p_i^{(n)}$ è limitato o a pochi esempi facili oppure ad n molto basso.

Ciò che invece è spesso facile da calcolare sono le *probabilità invarianti*.

Definizione 1. Un vettore di probabilità π (cioè tale che $\sum \pi_i = 1$, $\pi_i \in [0, 1]$ per ogni i), si dice invariante per la catena di Markov descritta dalla matrice di transizione P se $\pi = \pi P$.

In termini dinamici questo significa che partendo da una situazione iniziale aleatoria con distribuzione di probabilità π , dopo un passo (e quindi dopo un numero arbitrario di passi, essendo

$$\pi P^n = \pi P P^{n-1} = \pi P^{n-1} = \dots = \pi)$$

la distribuzione di probabilità è sempre π . Si può dimostrare inoltre che in questo caso il processo stocastico $\{X_n\}$ è *stazionario* in senso stretto. Per questo le componenti di π si dicono anche *probabilità stazionarie*.

Si usano vari altri nomi equivalenti, come distribuzione invariante, massa invariante.

In larga misura l'interesse per le probabilità invarianti deriva dallo studio di $p_i^{(n)}$ nel limite per $n \rightarrow \infty$. Supponiamo che questo limite esista ed indichiamolo con π_i :

$$\pi_i = \lim_{n \rightarrow \infty} p_i^{(n)}. \quad (2.1)$$

Siccome vale la relazione

$$p^{(n+1)} = p^{(n)} P$$

(come caso particolare della $p^{(n)} = p^{(k)} P^{n-k}$), passando al limite per $n \rightarrow \infty$ otteniamo

$$\pi = \pi P.$$

Quindi π è una probabilità invariante. Abbiamo dimostrato:

Lemma 1. Se le probabilità $p^{(n)}$ al passo n -esimo convergono ad un vettore di probabilità π , allora π è invariante.

Il viceversa non è vero (lo discuteremo in un'altro paragrafo): esistono esempi con una certa caratteristica di periodicità in cui ci sono soluzioni di $\pi = \pi P$ ma $p^{(n)}$ non converge. Inoltre, un po' articolato è il viceversa nel caso in cui l'equazione $\pi = \pi P$ ha più di una soluzione (il limite delle $p^{(n)}$ dipende dalle condizioni iniziali).

Nel caso di matrici infinite questo teorema richiede delle precisazioni, ma per ora restiamo ad un livello lievemente non rigoroso.

L'interpretazione che diamo dei numeri π_i è la seguente: π_i è approssimativamente la probabilità di trovare il sistema nello stato i , dopo che è trascorso un tempo sufficientemente lungo. Allora la relazione $\pi = \pi P$ rappresenta un enorme vantaggio di calcolo, in quanto il calcolo di $p^{(n)}$ era molto complicato mentre spesso non lo è la risoluzione di questo sistema lineare.

Notiamo che l'equazione $\pi = \pi P$ è equivalente all'equazione $(I - P)^T \pi^T = 0$.

Nota. Una equazione del tipo $(I - P)^T \pi^T = 0$ ha sempre la soluzione $\pi^T = 0$, che però non ci interessa (cerchiamo una distribuzione di probabilità π). In generale, tale equazione potrebbe non avere altre soluzioni (se la matrice $I - P$ ha rango massimo, sempre che stiamo studiando un sistema con un numero finito di stati, altrimenti le cose sono ancora più complicate). Se invece esistono soluzioni $\pi \neq 0$, ne esistono infinite: il nucleo della matrice $I - P$ è uno spazio vettoriale. Bisogna allora ricordarsi di imporre le altre condizioni generali che deve soddisfare una distribuzione di probabilità: $\pi_i \in [0, 1]$, $\sum_i \pi_i = 1$. In casi "buoni" (quelli che incontreremo lo sono), queste condizioni individuano una unica soluzione.

Exercise 4. *Calcolare tutte le probabilità invarianti per una generica catena a due stati.*

In pratica, per calcolare π a partire da P si ragiona nel seguente modo. Prima di tutto si effettua la cosiddetta classificazione degli stati, che descriveremo tra un attimo. Dopo essersi ricondotti, tramite essa, ad una classe irriducibile (in breve un insieme di stati in cui tutti gli stati comunicano tra di loro), si cerca una π in tale classe. Per calcolare π si può ovviamente affrontare il sistema $\pi = \pi P$, ma, almeno in casi semplici, conviene adottare un metodo grafico detto del *bilancio di flusso*. Anche questo verrà descritto tra breve.

2.1. Classificazione degli stati

2.1.1. Stati comunicanti

Si dice che lo stato i comunica con lo stato j se esiste un $n > 0$ tale che $p_{ij}^{(n)} > 0$. Questo si indica con $i \rightarrow j$. Se i non comunica con j scriveremo $i \nrightarrow j$.

2.1.2. Classe chiusa

Indicando con S l'insieme degli stati di una catena, un sottoinsieme $A \subset S$ è una classe chiusa se gli stati di A non comunicano con gli stati del complementare A^c .

2.1.3. Classe chiusa irriducibile

Una classe chiusa A si dice irriducibile se tutti i suoi stati comunicano fra loro.

Se uno stato costituisce da solo una classe irriducibile, esso si dice *assorbente*.

Una catena di Markov si dice irriducibile se tutti i suoi stati comunicano, ovvero se la classe di tutti i suoi stati è una classe irriducibile.

2.1.4. Stati transitori, stati ricorrenti

Uno stato i si dice transitorio o transiente se con probabilità strettamente positiva la catena che parte da i non ci torna più.

Al contrario, uno stato i si dice ricorrente se vale 1 la probabilità che la catena, partendo da esso, ci ritorni. Per definizione, uno stato è o ricorrente, o transitorio.

Uno stato assorbente è ricorrente.

2.1.5. Alcune proprietà generali

Enunciamo, senza dimostrazione, alcune proprietà che aiutano nella classificazione degli stati. Chiamiamo catena finita una catena il cui insieme S degli stati è finito.

Proposition 2. *Se per lo stato i vale la condizione*

$$\exists j \text{ tale che } i \rightarrow j \text{ ma } j \not\rightarrow i$$

allora i è transiente. Se la catena è finita, allora la condizione precedente è necessaria e sufficiente per la transienza di i .

Proposition 3. *Una catena finita ha almeno uno stato ricorrente.*

E' su questa proprietà che si basa il seguente fatto, che spiegheremo meglio più avanti: una catena finita ha sempre almeno una misura invariante.

La proprietà di ricorrenza è contagiosa, come esprime la seguente proposizione.

Proposition 4. *Se i è ricorrente e $i \rightarrow j$, allora anche j è ricorrente.*

Da questo discende:

Proposition 5. *Se una catena è irriducibile, allora gli stati sono o tutti ricorrenti, o tutti transitori.*

Da due di queste proprietà discende:

Proposition 6. *Se una catena è finita ed irriducibile, allora tutti i suoi stati sono ricorrenti.*

2.2. Probabilità invarianti (continuazione)

Indichiamo con S l'insieme degli stati di una catena di Markov omogenea $(X_n)_{n \geq 0}$. Per tutto questo capitolo supporremo che S sia un **insieme finito**. Molti risultati si generalizzano o direttamente o con alcune modifiche al caso in cui S ha un'infinità numerabile di elementi; altri risultati non si generalizzano; rimandiamo ai testi specializzati per una discussione dettagliata.

Ricordiamo che le probabilità invarianti sono importanti per almeno due ragioni. Da un lato esse sono un'approssimazione delle probabilità al tempo n , le $p_i^{(n)} = P(X_n = i)$. Infatti, come vedremo, sotto opportune ipotesi il vettore $p^{(n)}$ converge a π . Quando questo accade, il complicato calcolo delle $p^{(n)}$ si può rimpiazzare col calcolo molto più agevole della π , che richiede "solo" la risoluzione del problema algebrico $(I - P)^T \pi^T = 0$. Dall'altro, la conoscenza di π fornisce informazioni sullo stato del sistema fisico che stiamo descrivendo con la catena di Markov, se quando cominciamo ad osservare il sistema (cioè al tempo che per noi è $t = 0$) il sistema è già da tempo soggetto ad evoluzione casuale. In altre parole, quello che per noi è il tempo $t = 0$, per il sistema può essere un tempo già molto elevato, e quindi il sistema può già trovarsi nella situazione invariante (stazionaria) descritta dalla probabilità π .

Precisiamo un fatto: nel seguito per probabilità invariante intendiamo una qualsiasi soluzione dell'equazione algebrica $(I - P)^T \pi^T = 0$ (indipendentemente dal fatto che si ottenga come limite delle $p^{(n)}$).

Ci chiediamo ora:

- 1) data una catena di Markov, esiste sempre almeno una probabilità invariante π ?
- 2) è unica?
- 3) vale la *convergenza all'equilibrio*, ovvero la proprietà

$$p^{(n)} \rightarrow \pi$$

che abbiamo già discusso in modo non rigoroso? Per quali distribuzioni iniziali $p^{(0)}$ vale?

3. Esistenza ed unicità

Per rispondere ad alcune di queste domande ricordiamo i concetti sulla classificazione degli stati, precedentemente introdotti: stati comunicanti, classi chiuse, classi irriducibili, stati transitori (o transienti), stati ricorrenti. Si può dimostrare,

partendo dalle varie proprietà descritte in precedenza per questi concetti, che l'insieme S degli stati si decompone in alcuni sottoinsiemi, come segue. Indichiamo con T l'insieme degli stati transienti, e con R l'insieme degli stati ricorrenti. Vale per definizione

$$S = T \cup R.$$

A sua volta, l'insieme R si decompone nell'unione di classi irriducibili. Esistono cioè delle classi irriducibili C_1, \dots, C_N , disgiunte, con unione pari ad R :

$$R = C_1 \cup \dots \cup C_N.$$

Quindi in definitiva vale la decomposizione

$$S = T \cup C_1 \cup \dots \cup C_N.$$

Non esistono altri sottoinsiemi di stati. Invitiamo il lettore a raffigurare questa decomposizione.

Invitiamo inoltre il lettore a farsi un'idea intuitiva e dinamica di come evolve nel tempo la massa in questi insiemi. Al tempo $n = 0$ abbiamo la distribuzione di probabilità $p^{(0)}$, che dobbiamo pensare come una distribuzione di massa (discreta, cioè concentrata negli stati del sistema). Al tempo $n = 1$ questa si è trasformata nella distribuzione $p^{(1)}$, e così via. Nel limite $n \rightarrow \infty$ questa tende alla distribuzione di massa π (nota: per ora questa convergenza va presa intuitivamente, perchè non abbiamo chiarito sotto quali ipotesi valga). Ecco allora un'idea intuitiva: la massa che inizialmente si trova nell'insieme transiente T , prima o poi esce da esso, entrando nelle varie classi irriducibili. Di conseguenza, ci aspettiamo che la distribuzione limite π dia massa zero ai punti dell'insieme T . Questo si può dimostrare rigorosamente:

$$\begin{aligned} &\text{ogni probabilità invariante } \pi \text{ ha la proprietà} \\ \pi_i &= 0 \text{ per ogni stato transiente } i. \end{aligned}$$

Esprimeremo questo dicendo che T *non supporta* massa invariante.

Riprendiamo l'immagine intuitiva. La massa che si trova in una classe irriducibile C_k da questa classe non può uscire. Questo fa pensare che ci sia una massa invariante in questa classe. Questo si può dimostrare rigorosamente. Inoltre, il fatto che la classe sia irriducibile, ovvero che ogni stato comunichi con

ogni altro, fa sì che la massa invariante sia unica. Si può dimostrare rigorosamente il seguente risultato non banale:

in ogni classe irriducibile C_k esiste
una ed una sola probabilità invariante $\pi^{(k)}$.

Precisiamo il significato di questa frase. Esiste una probabilità invariante $\pi^{(k)}$ supportata da C_k , ovvero tale che

$$\pi_i^{(k)} = 0 \text{ per ogni stato } i \text{ che non appartiene a } C_k.$$

Inoltre, ne esiste una sola con questa proprietà. Si può ulteriormente dimostrare, sempre grazie alla comunicatività tra tutti gli stati della classe irriducibile, che $\pi^{(k)}$ attribuisce massa strettamente positiva ad ogni stato di C_k .

Siamo quindi arrivati a questa immagine: l'insieme transiente T non contiene massa invariante; ogni classe irriducibile C_k contiene una ed una sola massa invariante. Queste però non sono tutte le possibili probabilità invarianti della catena nel suo complesso. Detto altrimenti, se ci restringiamo ad una classe irriducibile, lì c'è una sola probabilità invariante, ma se guardiamo il sistema nel suo complesso, ce ne sono altre. Queste altre sono tutte e sole le *combinazioni convesse* delle $\pi^{(k)}$, cioè i vettori di probabilità π che si possono esprimere nella forma

$$\pi = \alpha_1 \pi^{(1)} + \dots + \alpha_N \pi^{(N)}$$

dove i coefficienti soddisfano

$$\alpha_i \in [0, 1], \quad \sum_{i=1}^N \alpha_i = 1.$$

Infatti è immediato verificare che, essendo le $\pi^{(k)}$ invarianti, lo sono anche queste combinazioni convesse (l'invarianza è una condizione lineare, quindi si conserva per le combinazioni lineari; la restrizione di considerare combinazioni convesse deriva dal fatto che π deve essere un vettore di probabilità). Si dimostra poi che non ci sono altre probabilità invarianti oltre a queste.

Resta comunque chiaro che le probabilità invarianti $\pi^{(k)}$ sono le più importanti, e che ci si deve concentrare sulla loro ricerca ed indagine.

3.1. Convergenza all'equilibrio

Quanto detto fino ad ora esaurisce le domande 1) e 2) poste sopra, su esistenza ed unicità delle distribuzioni invarianti. La domanda 3) si collega a quanto appena detto, con un'ulteriore complicazione.

Innanzitutto osserviamo cosa intuitivamente accade alla massa iniziale $p^{(0)}$ (riprendiamo l'idea intuitiva descritta sopra). La parte di questa massa che occupa T a poco a poco fluisce fuori da T ed entra, definitivamente, in qualche classe irriducibile, aggiungendosi così alla massa già presente in quella classe irriducibile (massa che non può uscire dalla classe). Il ragionamento si riconduce allora a capire cosa succede al crescere del tempo alla massa che occupa una certa classe irriducibile. Prima di far questo osserviamo che è possibile (noi non lo faremo) studiare varie proprietà di questo *assorbimento* di massa da parte delle classi irriducibili: probabilità di finire in una o un'altra classe, tempo medio di entrata in quella classe, ecc.

Restringiamoci quindi ad una classe irriducibile. In altre parole, studiamo una *catena di Markov irriducibile*, avendo capito che a questo studio ci si riconduce dal caso generale.

La difficoltà aggiuntiva a cui abbiamo accennato è che, pur restringendosi ad una catena irriducibile, la proprietà di convergenza all'equilibrio può non valere. Ecco un esempio.

Esempio. Consideriamo una catena a due stati, detti 1 e 2. Supponiamo che sia

$$p_{11} = 0, \quad p_{12} = 1, \quad p_{21} = 1, \quad p_{22} = 0.$$

In altre parole, la catena dallo stato 1 passa allo stato 2, e viceversa. Partiamo con tutta la massa iniziale concentrata nello stato 1. Dopo un passo, tutta la massa è concentrata nello stato 2, e così via alternativamente. Al variare del tempo, la massa oscilla periodicamente tra i due stati. Quindi è chiaro che questa massa non converge ad alcuna massa limite π . Quindi non si ha convergenza all'equilibrio delle $p^{(n)}$. Nonostante questo, una massa invariante esiste (ed è unica): è la massa che vale $\frac{1}{2}$ in ciascun stato.

Riflettendo su varianti più complicate di questo esempio si arriva a capire che la difficoltà è causata da qualche forma di periodicità della catena, che produce delle oscillazioni periodiche di massa, non convergenti ad un equilibrio. Una condizione

perché questo non avvenga è la seguente:

esiste n tale che $p_{ij}^{(n)} > 0$ per ogni i, j .

Questa condizione afferma che esiste un n fissato relativamente al quale tutti gli stati comunicano tra loro. Si dimostra (non è banale) che questo rompe le possibili periodicità ed implica la convergenza all'equilibrio. Il teorema afferma:

sotto questa ipotesi vale $p^{(n)} \rightarrow \pi$ per ogni distribuzione iniziale $p^{(0)}$.

Si noti che π esiste ed è unica, perché abbiamo supposto dall'iniziazione che la catena fosse irriducibile.

Osservazione. Ci si può chiedere se nei casi “periodici” non si salvi qualche forma debole di convergenza all'equilibrio. In effetti, pensando all'esempio precedente, vediamo che la massa oscilla tra i due stati senza convergere, ma in un certo senso medio (una media temporale) essa tende alla massa uguale a $\frac{1}{2}$. In assoluta generalità si può dimostrare il seguente *teorema ergodico*:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p^{(k)} = \pi.$$

In altre parole, se anche le $p^{(k)}$ non tendono a π , almeno ci tendono le medie temporali.

3.2. Calcolo delle probabilità invarianti

Mettiamoci nel caso irriducibile. Il calcolo delle probabilità invarianti, come già detto, equivale alla ricerca dell'autovettore sinistro di P con autovalore uno

$$\pi = \pi P$$

preoccupandosi di trovare, tra tutti i multipli possibili, quello che è un vettore di probabilità. Equivalentemente si cerca il vettore di probabilità che risolve l'equazione algebrica

$$(I - P)^T \pi^T = 0$$

avendo l'accortezza di imporre sempre le condizioni che sia un vettore di probabilità. Le metodologie note dell'algebra lineare, e quelle numeriche corrispondenti,

sono ovviamente di aiuto. Vogliamo però indicare alcune scorciatoie o di tipo grafico o valide in casi particolari.

1) La prima osservazione è che possiamo riscrivere la generica componente j -esima dell'equazione $\pi = \pi P$,

$$\pi_j = \sum_i \pi_i p_{ij}$$

nella forma

$$(1 - p_{jj}) \pi_j = \sum_{i \neq j} \pi_i p_{ij} \quad (3.1)$$

(basta isolare nella sommatoria il termine con $i = j$ e portarlo a sinistra). Questa equazione si presta ad una utilissima interpretazione grafica. Osservando che p_{jj} è la probabilità che la massa π_j non fluisca fuori dallo stato j . Allora $(1 - p_{jj}) \pi_j$ è la probabilità che la massa π_j fluisca fuori dallo stato j , e pertanto $(1 - p_{jj}) \pi_j$ è la proporzione di massa che esce dallo stato j . D'altra parte, per ogni stato $i \neq j$, $\pi_i p_{ij}$ è la massa che arriva a j dallo stato i , e la somma $\sum_{i \neq j} \pi_i p_{ij}$ rappresenta la massa che giunge allo stato j dagli altri stati. Per queste ragioni l'equazione (3.1) viene detta *equazione di bilancio del flusso*. Essa si può leggere facilmente sul grafo della catena e può essere di grande aiuto nello studio della probabilità invariante.

2) A volte, basandosi sull'intuizione, di solito dovuta a qualche particolare simmetria del grafo della catena, si riesce ad intuire a priori quanto dovrebbe valere il vettore π . In tal caso, è sufficiente *verificare* l'uguaglianza $\pi = \pi P$. Questa verifica si può svolgere algebricamente, sostituendo nell'equazione $\pi = \pi P$, oppure appoggiandosi al grafo, e verificando l'equazione di bilancio del flusso (3.1) descritta sopra.

3) In aiuto a questa semplice strategia viene anche il concetto di *equilibrio dettagliato*. Diciamo che, per un vettore π (di cui non si richiede a priori che sia invariante) vale l'equilibrio dettagliato se per ogni i, j vale

$$\pi_i p_{ij} = \pi_j p_{ji}.$$

Il significato intuitivo di questa condizione (riprendendo l'idea intuitiva usata più volte sopra) è che la percentuale p_{ij} della massa π_i che si sposta in un passo dallo

stato i allo stato j è esattamente uguale alla percentuale p_{ji} della massa π_j che si sposta in un passo dallo stato j allo stato i . Di conseguenza, in un passo, gli stati i e j si trasmettono l'un l'altro la stessa quantità di massa. E' quindi una versione molto particolare e semplice dell'equazione di bilancio del flusso.

La condizione di equilibrio dettagliato implica che π è invariante. Quest'affermazione è chiara a livello intuitivo per quanto appena detto, e si verifica algebricamente in modo semplice (si veda l'esercizio seguente).

Dal punto di vista pratico, spesso è immediato verificare la condizione di equilibrio dettagliato per un vettore π che pensiamo possa essere invariante. Oppure, a volte si riesce a scoprire una massa invariante cercandone una che verifica l'equilibrio dettagliato. In questo modo si trovano esplicitamente le probabilità invarianti di molti esempi, evitando calcoli laboriosi. Si veda il secondo esercizio dato sotto.

Esercizio. Dimostrare (algebricamente) che la condizione di equilibrio dettagliato implica la condizione di invarianza $\pi = \pi P$.

Esercizio. Si consideri una catena con 4 stati, tutti comunicanti tra loro, con tutte le probabilità di transizione uguali tra loro. Trovare la probabilità invariante.

4. Processo di nascita e morte a tempo discreto

Sviluppiamo ora un esempio che esula un poco dalla trattazione teorica precedente, in quanto il numero degli stati è infinito. Tuttavia si riesce a calcolare in modo esplicito una distribuzione invariante e si capisce che è unica. Questo esempio è fondamentale in teoria delle code.

Supponiamo che la catena (X_n) abbia come stati tutti i numeri interi non negativi, ovvero $0, 1, \dots$

Supponiamo che da ogni stato i siano possibili solo le transizioni ad $i, i + 1$ ed $i - 1$, con l'eccezione dello stato 0 da cui si può transire solo in 1 o restare in 0.

Indichiamo con λ_0 la probabilità di transire da 0 a 1, per cui $1 - \lambda_0$ è la probabilità di restare in 0.

Poniamo poi, per ogni stato $i > 0$,

$$\begin{aligned}\lambda_i &= p_{i,i+1} \\ \mu_i &= p_{i,i-1}\end{aligned}$$

per cui vale

$$p_{i,i} = 1 - \lambda_i - \mu_i.$$

La matrice di transizione, infinita, è

$$P = \begin{pmatrix} 1 - \lambda_0 & \lambda_0 & 0 & 0 & \dots \\ \mu_1 & 1 - \lambda_1 - \mu_1 & \lambda_1 & 0 & \dots \\ 0 & \mu_2 & 1 - \lambda_2 - \mu_2 & \lambda_2 & \dots \\ 0 & 0 & \mu_3 & 1 - \lambda_3 - \mu_3 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Supponiamo infine che tutti i numeri λ_i (incluso $i = 0$) e μ_i siano strettamente positivi.

Il significato di questa catena in teoria delle code il seguente. La catena conta il numero di persone presenti in una coda. Ad ogni istante n di tempo, essendo i il numero di utenti in coda, può non accadere nulla, ovvero restare i tale numero, con probabilità $1 - \lambda_i - \mu_i$. Oppure può venir servito un utente, che esce dalla coda, per cui il numero di utenti diventa $i - 1$, e questo accade con probabilità μ_i . Oppure infine può arrivare un nuovo utente, quindi il numero aumenta di uno, e ciò avviene con probabilità λ_i .

Ci chiediamo: nel regime stazionario, ovvero quando è passato un tempo abbastanza grande e la coda si è statisticamente assestata, qual'è la probabilità di avere i persone in coda? Dobbiamo calcolare la distribuzione invariante.

4.1. Preliminare: stati $0, \dots, N$

Come preliminarmente, consideriamo lo stesso problema con un numero finito di stati $0, \dots, N$, in modo da applicare fatti più semplici. Per le probabilità di transizione usiamo le stesse notazioni del caso infinito, salvo il fatto che dallo stato N non si può transire in $N + 1$ (quindi $\lambda_N = 0$).

Ogni stato comunica con ogni altro, avendo presupposto strettamente positivi tutti i λ_i e μ_i . Quindi $\{0, \dots, N\}$ è un'unica classe irriducibile, che supporta un'unica misura invariante π .

Il bilancio di flusso nel generico stato $i = 1, \dots, N - 1$ è

$$\pi_i (\lambda_i + \mu_i) = \pi_{i-1} \lambda_{i-1} + \pi_{i+1} \mu_{i+1}.$$

Il bilancio in $i = 0$ è invece

$$\pi_0 \lambda_0 = \pi_1 \mu_1.$$

Quindi

$$\pi_1 = \pi_0 \frac{\lambda_0}{\mu_1}.$$

Il bilancio in $i = 1$ è

$$\pi_2 \mu_2 = \pi_1 (\lambda_1 + \mu_1) - \pi_0 \lambda_0$$

e quindi, sostituendo la formula precedente,

$$\pi_2 \mu_2 = \pi_0 \frac{\lambda_0}{\mu_1} (\lambda_1 + \mu_1) - \pi_0 \lambda_0$$

ovvero

$$\pi_2 = \pi_0 \frac{\lambda_0 (\lambda_1 + \mu_1) - \mu_1 \lambda_0}{\mu_1 \mu_2} = \pi_0 \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2}.$$

Così di seguito si riconosce la validità della seguente formula (la dimostrazione rigorosa andrebbe svolta per induzione).

Theorem 1. *Posto*

$$a_i = \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i},$$

la distribuzione invariante per la catena ora descritta è

$$\pi_i = \frac{a_i}{\sum_{i=0}^N a_i}$$

Infatti dai calcoli precedenti si trova

$$\pi_i = \pi_0 a_i$$

e poi imponendo $\sum_{i=0}^N \pi_i = 1$ si trova

$$\pi_0 = \frac{1}{\sum_{i=0}^N a_i}.$$

4.2. Il caso infinito

Tornando alla catena di nascita e morte con infiniti stati, i calcoli sono gli stessi. L'unica differenza è che la serie $\sum_{i=0}^{\infty} a_i$ potrebbe divergere a più infinito (è a termini positivi, quindi o converge o diverge). Quando questo accade, la condizione $\sum_{i=0}^{\infty} \pi_i = 1$ non può essere soddisfatta con alcuna scelta di π_0 , quindi non si

trova, con questi calcoli, alcuna misura invariante. Può restare il dubbio teorico che ce ne sia una rintracciabile in altro modo, ma non è così, come si intuisce ovviamente sulla base del caso con stati finiti e con semplice buon senso (omettiamo però la dimostrazione). Abbiamo quindi il seguente risultato.

Theorem 2. *Posto*

$$a_i = \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i},$$

se risulta

$$\sum_{i=0}^{\infty} a_i < \infty$$

allora la distribuzione invariante per la catena di nascita e morte è

$$\pi_i = \frac{a_i}{\sum_{i=0}^{\infty} a_i}.$$

Se invece $\sum_{i=0}^{\infty} a_i = +\infty$, allora non ci sono distribuzioni invarianti.

Esaminiamo un caso particolare in cui possono accadere entrambe le situazioni.

Example 2. *Supponiamo che sia $\lambda_i = \lambda$ e $\mu_i = \mu$ per ogni i . Allora*

$$a_i = \left(\frac{\lambda}{\mu}\right)^i$$

e quindi

$$\sum_{i=0}^{\infty} a_i = \sum_{i=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^i$$

converge se e solo se $\frac{\lambda}{\mu} < 1$ (ricordiamo che $\frac{\lambda}{\mu} > 0$) ovvero se e solo se

$$\lambda < \mu.$$

Nelle code di servizio questo si interpreta dicendo che se il tasso di servizio è superiore al tasso di arrivo di nuovi utenti, allora c'è una distribuzione invariante, ovvero il sistema raggiunge una situazione di equilibrio. Se invece il tasso di servizio è inferiore al tasso di arrivo, allora la coda non riesce a smaltire gli arrivi, aumenta indefinitamente e non si assesta in un regime di equilibrio.