1 Catene di Markov a stati continui

In questo caso abbiamo ancora una successione di variabili casuali $X_0, X_1, X_2, \ldots$ ma lo spazio degli stati è un insieme più che numerabile. Nel seguito supporremo che lo spazio degli stati sia $\mathbb{R}$ o, più in generale $\mathbb{R}^d$. Perché la successione sia una catena di Markov, deve risultare

$$\mathbb{P}(X_{n+1} \in I \mid X_n = x, X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_1 = x_1, X_0 = x_0) = \mathbb{P}(X_{n+1} \in I \mid X_n = x) \quad (1)$$

e supporremo che le probabilità di transizione

$$p(x, I) = \mathbb{P}(X_{n+1} \in I \mid X_n = x)$$

non dipendano dal tempo (da n). Dovremo dare anche la distribuzione iniziale (di X_0)

$$\mu(I) = \mathbb{P}(X_0 \in I).$$

Si ha che

$$p^{(2)}(x, I) = \mathbb{P}(X_{n+2} \in I \mid X_n = x) = \int_{\mathbb{R}} p(x, \mathrm{d}y)\, p(y, I)$$

e, più in generale, induttivamente

$$\begin{aligned} p^{(k)}(x, I) = \mathbb{P}(X_{n+k} \in I \mid X_n = x) &= \int_{\mathbb{R}} p^{(k-1)}(x, \mathrm{d}y)\, p(y, I) \\ &= \int_{\mathbb{R}} p(x, \mathrm{d}y)\, p^{(k-1)}(y, I) \\ &= \int_{\mathbb{R}} p^{(l)}(x, \mathrm{d}y)\, p^{(k-l)}(y, I) \end{aligned}$$

Tra le distribuzioni (di probabilità) iniziali è importante distinguere la distribuzione invariante (o stazionaria), definita dalla proprietà

$$\mu(I) = \int_R p(x, I)\, \mu(\mathrm{d}x)$$

1.0.1 Successione di variabili casuali indipendenti ed equidistribuite

È una catena di Markov con $p(x, I) = \mathbb{P}(X_{n+1} \in I \mid X_n = x) = \mathbb{P}(X_{n+1} \in I) = \pi(I)$, e la distribuzione π è anche la distribuzione stazionaria.
Qui $\pi(I)$ è la distribuzione di ogni X_n (equidistribuite). Inoltre $\mathbb{P}(X_{n+1} \in I \mid X_n = x) = \mathbb{P}(X_{n+1} \in I)$ discende dall'indipendenza.

1.0.2 Processo AR(1) (*AutoRegressive*)

Data una successione di variabili casuali indipendenti ed equidistribuite come normali $N(0,1)$

$$Z_1, Z_2, Z_3, \dots, \tag{2}$$

ogni variabile casuale X_n (della catena di Markov AR(1)) è definita induttivamente da

$$X_n = \beta X_{n-1} + \sigma Z_n \tag{3}$$

e da una scelta (a priori come si vuole) della distribuzione iniziale $\nu(I) = \mathbb{P}(X_0 \in I)$, che sceglieremo indipendente da tutta la successione (2) delle Z.

Risulta da (3) che

$$\mathbb{P}(X_n \in I \mid X_{n-1} = x) = \mathbb{P}(\beta x + \sigma Z_n \in I \mid X_{n-1} = x) = \mathbb{P}(\beta x + \sigma Z_n \in I)$$

quindi

$$p(x, I) = \int_I \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\beta x)^2}{2\sigma^2}} \mathrm{d}y$$

Inoltre, induttivamente si ricava che

$$X_n = \beta^n X_0 + \sigma(\beta^{n-1} Z_1 + \beta^{n-2} Z_2 + \cdots + \beta Z_{n-1} + Z_n)$$

Ma $\sigma(\beta^{n-1} Z_1 + \beta^{n-2} Z_2 + \cdots + \beta Z_{n-1} + Z_n)$ è ancora una gaussiana di media zero e varianza la somma delle varianze, cioè

$$\sigma^2(\beta^{2(n-1)} + \beta^{2(n-2)} + \beta^{2(n-3)} + \cdots + \beta^2 + 1) = \sigma^2 \frac{1-\beta^{2n}}{1-\beta^2}$$

Se denotiamo con $\mathcal{Z}_n = \sigma(\beta^{n-1} Z_1 + \beta^{n-2} Z_2 + \cdots + \beta Z_{n-1} + Z_n)$, possiamo scrivere $X_n = \beta^n X_0 + \mathcal{Z}_n$.
Se scegliamo $|\beta| < 1$ abbiamo che la legge di $\mathcal{Z}_n$ tende (in legge) ad una gaussiana di media zero e varianza

$$v^2 = \sigma^2 \frac{1}{1-\beta^2}.$$

Scegliamo allora come distribuzione iniziale una gaussiana di media zero e varianza v^2; allora anche X_n ha una distribuzione gaussiana di media zero e varianza

$$\beta^{2n} v^2 + \sigma^2 \frac{1-\beta^{2n}}{1-\beta^2} = v^2.$$

Questo significa che la normale $N(0, v^2)$ è la distribuzione stazionaria: infatti ciò risulta anche da

$$\int_R \frac{1}{\sqrt{2\pi v^2}} e^{-\frac{x^2}{2v^2}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\beta x)^2}{2\sigma^2}} \mathrm{d}x = \frac{1}{\sqrt{2\pi v^2}} e^{-\frac{y^2}{2v^2}}$$

Infine notiamo che X_n tende in legge ad una normale $N(0, v^2)$.

1.0.3 Processo AR(2)

Data una successione di variabili casuali indipendenti ed equidistribuite come normali $N(0,1)$

$$Z_2, Z_3, Z_4, \ldots, \tag{4}$$

ogni variabile casuale X_n (che non formano una catena di Markov) è definita induttivamente da

$$X_n = \beta_1 X_{n-1} + \beta_2 X_{n-2} + \sigma Z_n \tag{5}$$

e da una scelta (a priori come si vuole) della distribuzione congiunta di X_0, X_1, che sceglieremo indipendente da tutta la successione (4) delle Z.

Supponiamo che esista una distribuzione stazionaria gaussiana $N(0, v^2)$, con v^2 da determinare. Deve risultare

$$v^2 = \mathbb{E}(X_n^2) = \mathbb{E}((\beta_1 X_{n-1} + \beta_2 X_{n-2})^2) + \sigma^2$$

Sviluppando si ha

$$v^2 = \beta_1^2 v^2 + \beta_2^2 v^2 + 2\beta_1\beta_2\, \mathbb{E}(X_{n-1}X_{n-2}) + \sigma^2 \tag{6}$$

È bene osservare che anche la distribuzione congiunta di X_n, X_{n-1} è invariante con n; allora, moltiplicando la (5) per X_{n-1} e prendendo il valor medio si ottiene

$$\mathbb{E}(X_n X_{n-1}) = \beta_1 v^2 + \beta_2 \mathbb{E}(X_{n-1}X_{n-2})$$

che per l'invarianza dei valori di aspettazione si ha

$$\mathbb{E}(X_n X_{n-1}) = \frac{\beta_1 v^2}{1-\beta_2}$$

con cui dalla (6) possiamo ricavare

$$v^2 = \sigma^2 \frac{1}{1 - \beta_1^2 - \beta_2^2 - 2\beta_1\beta_2 \frac{\beta_1}{1-\beta_2}}$$

Deve risultare

$$1 - \beta_1^2 - \beta_2^2 - 2\beta_1\beta_2 \frac{\beta_1}{1-\beta_2} = \frac{(1+\beta_2)(\beta_1+\beta_2-1)(\beta_2-\beta_1-1)}{1-\beta_2} > 0$$

2 Catene di Markov a tempi continui

2.0.4 Leggi gamma

Introduciamo la legge Gamma di parametri α e λ, come la densità di probabilità

$$\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$$

per $x > 0$, altrimenti 0; qui $\Gamma(\alpha)$ è la classica funzione gamma definita per $\alpha > 0$ da

$$\Gamma(\alpha) = \int_0^\infty e^{-\xi}\xi^{\alpha-1}\mathrm{d}\xi$$

La funzione caratteristica della distribuzione Gamma di parametri α e λ è data da

$$\varphi(t) = \left(\frac{\lambda}{\lambda - \mathtt{i}t}\right)^\alpha$$

dove la radice complessa è presa in modo tale che per valori reali è reale. Subito questo fatto ci dice che la somma di due Gamma di parametri rispettivamente α_1, λ e α_2, λ ed indipendenti è ancora una distribuzione Gamma di parametri $\alpha_1 + \alpha_2, \lambda$. Per $\alpha = 1$ la distribuzione Gamma si riduce alla distribuzione esponenziale di parametro λ.

2.1 Processo di Poisson

Sia $T_1, T_2, T_3, \ldots$ una successione di variabili casuali indipendenti ed equidistribuite come esponenziali di parametro λ; allora risulta che

$$\mathbb{P}(T_1 + T_2 + \cdots + T_k \le t) = \int_0^t \lambda^k \frac{s^{k-1}}{(k-1)!} e^{-\lambda s}\mathrm{d}s$$

Per ogni t esiste un solo intero X_t aleatorio tale che

$$T_1 + T_2 + \cdots + T_{X_t} \le t < T_1 + T_2 + \cdots + T_{X_t} + T_{X_t+1}$$

Questo processo X_t si chiama processo di Poisson. Risulta

$$\mathbb{P}(X_t = k) = \mathbb{P}(T_1 + T_2 + \cdots + T_k \le t < T_1 + T_2 + \cdots + T_k + T_{k+1}) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$$

Inoltre X_s e $X_t - X_s$ sono indipendenti e

$$\mathbb{P}(X_t - X_s = k) = e^{-\lambda(t-s)}\frac{(\lambda(t-s))^k}{k!}$$

2.2 Ancora sulle leggi esponenziali

Siano $\{T_1, T_2, T_3, \ldots, T_n\}$ una famiglia di variabili casuali indipendenti e ogni T_i con legge esponenziale di parametro λ_i. Allora la variabile casuale

$$\tau_n = \min\{T_1, T_2, T_3, \ldots, T_n\}$$

è ancora una variabile casuale di tipo esponenziale di parametro $\lambda_1 + \lambda_2 + \lambda_3 + \cdots + \lambda_n$.

Date due variabili casuali indipendenti T_1, T_2 distribuite in modo esponenziale di parametri rispettivamente λ_1, λ_2, calcolare

$$\mathbb{P}(T_1 < T_2 \mid \min(T_1, T_2) = t)$$

Possiamo calcolare la probabilità precedente come limite di

$$\frac{\mathbb{P}(T_1 < T_2, t \leq \min(T_1, T_2) \leq t + \Delta t)}{\mathbb{P}(t \leq \min(T_1, T_2) \leq t + \Delta t)} = \frac{\displaystyle\int_t^{t+\Delta t} \mathbb{P}(T_2 > s)\mathbb{P}(T_1 \in \mathrm{d}s)}{\displaystyle\int_t^{t+\Delta t} \mathbb{P}(\min(T_1, T_2) \in \mathrm{d}s)}$$

Ora

$$\frac{\displaystyle\int_t^{t+\Delta t} \mathbb{P}(T_2 > s)\mathbb{P}(T_1 \in \mathrm{d}s)}{\displaystyle\int_t^{t+\Delta t} \mathbb{P}(\min(T_1, T_2) \in \mathrm{d}s)} = \frac{\displaystyle\int_t^{t+\Delta t} e^{-\lambda_2 s}\, \lambda_1 e^{-\lambda_1 s} \mathrm{d}s}{\displaystyle\int_t^{t+\Delta t} (\lambda_1 + \lambda_2) e^{-(\lambda_1+\lambda_2)s} \mathrm{d}s} = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

quindi

$$p = \mathbb{P}(T_1 < T_2 \,|\, \min(T_1, T_2) = t) = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$

$$q = 1 - p = \mathbb{P}(T_1 > T_2 \,|\, \min(T_1, T_2) = t) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

Un secondo modo di formulare il risultato è di considerare una variabile casuale T distribuita in modo esponenziale di parametro λ $(= \lambda_1 + \lambda_2)$, e una seconda variabile casuale indipendente dalla prima che assume solo due valori. per esempio $+1$ e -1, con probabilità rispettivamente p e $q = 1 - p$. L'equivalenza tra le due formulazioni passa attraverso la posizione $\lambda_1 = p\lambda$ e $\lambda_2 = q\lambda$.

2.3 Processi di nascita e morte

Sono processi di Markov a tempi continui e spazio degli stati $S = \{0, 1, 2, 3, \ldots\}$. Tra i vari parametri che determinano questi processi ci sono la successione dei *tassi di nascita* (positivi)

$$\{\lambda_0, \lambda_1, \lambda_2, \ldots\}$$

e la successione dei *tassi di morte*

$$\{\mu_1, \mu_2, \mu_3, \ldots\}$$

Il processo X_t rappresenta la popolazione (gli elementi della popolazione possono essere della natura più varia) al tempo t. Si può partire da $X_0 = 0$ ma non necessariamente. Il passaggio da 0 ad 1 può avvenire per esempio per *immigrazione.*

Supponiamo di partire da una distribuzione iniziale per X_0 uguale per esempio a $\{\pi_0, \pi_1, \pi_2, \ldots\}$. Il tempo T_1 del primo cambiamento (a $k-1$ o a $k+1$ se eravamo in k) è distribuito come un esponenziale, ma non indipendente da X_0; infatti, *sapendo che* $X_0 = k$, T_1 ha parametro uguale a $\lambda_k + \mu_k$ (se $k = 0$ allora il parametro è solo λ_0) e, all'istante T_1 il processo salta in $k - 1$ con probabilità

$$\mathbb{P}(X_{T_1} = k - 1 \,|\, X_0 = k) = \frac{\mu_k}{\lambda_k + \mu_k}$$

mentre salta in $k+1$ con probabilità

$$\mathbb{P}(X_{T_1} = k+1 \,|\, X_0 = k) = \frac{\lambda_k}{\lambda_k + \mu_k}$$

Nel caso che $X_0 = 0$ allora

$$\mathbb{P}(X_{T_1} = 1 \,|\, X_0 = 0) = 1$$

Dopo il primo salto, supporremo che l'evoluzione proceda nello stesso modo; dopo un tempo T_2, relativo al tempo T_1, avviene un secondo cambiamento (a partire da 0 il tempo trascorso è $S_2 = T_1 + T_2$): supporremo che T_2 sia indipendente da T_1 ma dipendente da X_{T_1}. Infatti supporremo T_2 abbia distribuzione esponenziale con parametro, *sapendo che* $X_{T_1} = k$, uguale a $\lambda_k + \mu_k$. All'istante $S_2 = T_1 + T_2$ il processo salta con la stessa modalità precedente.

Vogliamo determinare la distribuzione di probabilità di X_t al tempo (deterministico) t, che potremo denotare con

$$p(t) = \{p_0(t), p_1(t), p_2(t), \ldots\}$$

A questo scopo consideriamo quello che può succedere al tempo $t + \Delta t$. Ebbene può non esserci stato nessun cambiamento, oppure solo uno, oppure più di uno. Nel primo caso

$$\mathbb{P}(X_{t+\Delta t} = k \,|\, X_t = k) = \mathbb{P}(T > \Delta t \,|\, X_t = k) = e^{-(\lambda_k + \mu_k)\Delta t} = 1 - (\lambda_k + \mu_k)\Delta t + o(\Delta t);$$

nel secondo

$$\mathbb{P}(X_{t+\Delta t} = k-1 \,|\, X_t = k) = \mathbb{P}(T \le \Delta t < T + T' \,|\, X_t = k)\frac{\mu_k}{\lambda_k + \mu_k} = \mu_k \Delta t + o(\Delta t)$$

e

$$\mathbb{P}(X_{t+\Delta t} = k+1 \,|\, X_t = k) = \mathbb{P}(T \le \Delta t < T + T' \,|\, X_t = k)\frac{\lambda_k}{\lambda_k + \mu_k} = \lambda_k \Delta t + o(\Delta t);$$

ed infine nel terzo

$$\mathbb{P}(X_{t+\Delta t} = j, |j-k| > 1 \,|\, X_t = k) = \mathbb{P}(T + T' \le \Delta t \,|\, X_t = k) = o(\Delta t).$$

Quindi

$$\begin{aligned} p_k(t+\Delta t) = \mathbb{P}(X_{t+\Delta t} = k) &= \mathbb{P}(X_{t+\Delta t} = k \,|\, X_t = k)p_k(t) + \\ &\quad \mathbb{P}(X_{t+\Delta t} = k \,|\, X_t = k-1)p_{k-1}(t) + \\ &\quad \mathbb{P}(X_{t+\Delta t} = k \,|\, X_t = k+1)p_{k+1}(t) + \\ &\quad \mathbb{P}(X_{t+\Delta t} = k \,|\, X_t = j, |k-j| > 1) \end{aligned}$$

da cui, per $k \ge 1$

$$p_k'(t) = \lambda_{k-1}p_{k-1}(t) + \mu_{k+1}p_{k+1}(t) - (\lambda_k + \mu_k)p_k(t)$$

e

$$p_0'(t) = \mu_1 p_1(t) - \lambda_0 p_0(t)$$

Se le $p_k(t)$ sono indipendenti dal tempo (misura invariante) uguale a π_k, allora abbiamo che

$$0 = \lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1} - (\lambda_k + \mu_k)\pi_k$$

e

$$0 = \mu_1\pi_1 - \lambda_0\pi_0$$

da cui

$$\pi_1 = \frac{\lambda_0}{\mu_1}\pi_0, \quad \pi_2 = \frac{\lambda_0\lambda_1}{\mu_1\mu_2}\pi_0, \quad \pi_2 = \frac{\lambda_0\lambda_1\lambda_2}{\mu_1\mu_2\mu_3}\pi_0, \ldots$$

Deve quindi risultare che la serie

$$1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0\lambda_1}{\mu_1\mu_2} + \frac{\lambda_0\lambda_1\lambda_2}{\mu_1\mu_2\mu_3} + \cdots < \infty$$

Denotando con Π la sua somma otteniamo che la misura invariante è

$$\pi_k = \frac{\frac{\lambda_0\lambda_1\lambda_2\cdots\lambda_{k-1}}{\mu_1\mu_2\mu_3\cdots\mu_k}}{\Pi}$$