

# Processi Stocastici.

G. Sanfilippo

4 giugno 2010

*“Insegnare non è riempire un vaso, ma accendere un fuoco.”*  
(Teofrato)

**P.S.** Queste dispense (in stesura provvisoria) non comprendono tutti gli argomenti trattati durante il corso e possono contenere delle sviste. Pertanto non possono sostituire i libri di testo consigliati.

# Indice

<b>1</b>	<b>Introduzione ai processi</b>	<b>4</b>
1.1	Introduzione ai Processi . . . . .	4
<b>2</b>	<b>Random Walk</b>	<b>6</b>
2.1	Processo di Bernoulli-Passeggiata aleatoria semplice . . . . .	6
2.1.1	Problema della rovina di un giocatore - Passeggiata aleatoria semplice con barriere assorbenti . . . . .	9
2.1.2	Capitale infinito . . . . .	16
<b>3</b>	<b>Catene di Markov discrete</b>	<b>18</b>
3.1	Catene di Markov a tempo discreto DTMC. . . . .	18
3.1.1	Introduzione ed esempi DTCM. . . . .	18
3.1.2	Equazione di Chapman-Kolmogoroff. . . . .	27
3.2	Catena di Markov a 2 stati . . . . .	30
3.3	Classificazione degli stati . . . . .	36
3.3.1	Stati ricorrenti e transitori . . . . .	38
3.4	Distribuzioni invarianti o stazionarie . . . . .	43
3.4.1	Catene di Markov Finite . . . . .	47
<b>4</b>	<b>Processi di Rinnovo e Processi di Poisson</b>	<b>50</b>
4.1	Processo dei rinnovi e processo dei conteggi . . . . .	50
4.1.1	Processo dei rinnovi . . . . .	50
4.1.2	Processo dei conteggi . . . . .	50
4.1.3	Dal processo dei rinnovi a quello di conteggio . . . . .	51
4.1.4	Dal processo di conteggio a quello dei rinnovi . . . . .	51
4.1.5	Processo di conteggio a incrementi indipendenti e stazionari . . . . .	52
4.2	Processo di Poisson . . . . .	52
4.3	Processo Uniforme ([1],[12],[9]) . . . . .	58
4.3.1	Processo di Poisson come limite di un Processo Uniforme . . . . .	62
4.3.2	Processo Uniforme come Processo di Poisson condizionato . . . . .	63
4.4	Alcune proprietà del Processo di Poisson . . . . .	63
4.4.1	Superposition . . . . .	64
4.4.2	Splitting . . . . .	64
4.4.3	Richiami sul valore atteso condizionato . . . . .	66

4.4.4	Splittin non omogeneo - Campionamento da un Processo di Poisson . . . . .	69
4.4.5	Altra Definizione di un Processo di Poisson . . . . .	71
4.4.6	Auto correlazione e Auto covarianza . . . . .	72
4.5	Processo di Poisson non omogeneo . . . . .	72
4.5.1	Tempi di attesa . . . . .	73
4.6	Processo di Poisson Composto . . . . .	74
4.7	Processo di Poisson Randomizzato . . . . .	76
4.8	Processi di Punto: cenni . . . . .	76
<b>5</b>	<b>Esercizi</b>	<b>77</b>
5.1	Esercizi (svolti e non) sui processi di conteggio . . . . .	77
5.1.1	Esercizi . . . . .	77
<b>6</b>	<b>Catene di Markov continue: cenni</b>	<b>80</b>
6.1	Catene di Markov continue (CTMC): Cenni . . . . .	80
6.2	Processi di Nascita e di Morte . . . . .	82
<b>7</b>	<b>Classificazione di un processo e cenni al moto Browniano</b>	<b>84</b>
7.1	Caratterizzazione dei processi stocastici . . . . .	84
7.1.1	Funzione media, correlazione e covarianza . . . . .	84
7.1.2	Processi stazionari . . . . .	85
7.1.3	Processi a incrementi indipendenti e stazionari . . . . .	86
7.2	Funzione caratteristica congiunta . . . . .	88
7.3	Processi Normali: cenni . . . . .	88
7.3.1	Processo di Wiener-Moto Browniano . . . . .	88
7.3.2	Moto Browniano e passeggiata aleatoria . . . . .	89
7.3.3	Moto Browniano standard . . . . .	90

# Capitolo 1

## Introduzione ai processi

### 1.1 Introduzione ai Processi

(vedi [6]) Un processo aleatorio o stocastico è un famiglia di v.a. che caratterizzeremo mediante un indice:

$$\{X_t : t \in T\}$$

**Classificazione in base al parametro** . Si hanno i seguenti casi

- Se  $T$  ha la potenza del continuo, (un intervallo di una retta), allora il processo si dirà a *parametro continuo*. Esempio  $X_t : t \in [0, 1]$
- Se  $T$  è al più numerabile (es.  $X_n : n = 1, 2, \dots$ ), allora il processo si dirà a *parametro discreto*.
- Più in generale  $T$  può essere un insieme qualsiasi.

**Classificazione in base ai valori** . Un'ulteriore classificazione si fa pure in base ai valori assunti da  $X_t$ .

- Se  $X_t$  assume un insieme di valori che hanno la potenza del continuo, (un intervallo di una retta), allora il processo si dirà a *valori continui*.
- Se  $X_t$  assume un insieme di valori numerabile, allora il processo si dirà a *valori discreti*.

Come fatto per i vettori aleatori, le v.a.  $X_t$  di un processo sono funzioni  $P$ -misurabili definite su uno stesso spazio di probabilità  $(\Omega, \mathcal{A}, P)$ .

Supponiamo, ad esempio, di lanciare infinite volte una moneta. Indichiamo con  $E_i$  l'evento "esce testa all' $i$ -esimo lancio", per  $i \in \mathbb{N}$ . Sia  $X_i = |E_i|$  per  $i \in \mathbb{N}$ . In tal caso il processo

$$X_i \in \{0, 1\}$$

è un processo discreto a valori discreti e ogni punto di  $\Omega$  rappresenta una generica sequenza  $CCT \cdots T \cdots$ .

Pertanto ogni  $\omega \in \Omega$  si può vedere come il possibile risultato di un esperimento in cui si rilevano i valori di **tutte** le variabili aleatorie.

Un processo aleatorio si presenta quindi come una funzione a due variabili  $(\omega, t)$  a valori in  $R$ , cioè

$$X(\omega, t) : \Omega \times T \longrightarrow R.$$

In generale si potrebbe avere

$$X(\omega, t) : \Omega \times T \longrightarrow R^k.$$

Se fissiamo la variabile  $t$  si ottiene la singola variabile aleatoria  $X_t$ . Se invece fissiamo la variabile  $\omega = \bar{\omega}$  si ottiene la funzione  $X(\bar{\omega}, t) = f_{\bar{\omega}(t)}$  che rappresenta i valori risultanti in una singola prova al variare di  $t \in T$  (come detto prima si deve immaginare ad una prova in cui si rilevano i valori di tutte le variabili aleatorie).

La funzione  $X(\bar{\omega}, t)$  (di  $t \in T$ ) viene chiamata *realizzazione* del processo o *traiettoria* o *storia*. Se  $t$  rappresenta la variabile tempo, l'interpretazione che nasce è quella di un punto che si muove con il passare del tempo.

Riprendendo l'esempio di prima, fissato  $\bar{\omega} = CCT \cdots T \cdots$  una traiettoria sarà la seguente funzione

$$X_1(\bar{\omega}) = 0, X_2(\bar{\omega}) = 0, X_3(\bar{\omega}) = 1, \dots$$

Distinguiamo due casi

1. *Indagine longitudinale*: se si studia l'evoluzione di una caratteristica di una popolazione, fissare  $\bar{\omega}$  significa scegliere una unità e seguirla nel tempo ( $X(\bar{\omega}, t)$ ).
2. *Indagine trasversale*: se invece si è interessati allo studio dell'intera popolazione in un dato momento, allora la variabile in oggetto è data da  $X_t(\omega)$  con  $t$  fissato.

**Esempio 1.** Alcuni processi.

1. Il prezzo giornaliero di una certa merce osservabile su una certa piazza in un dato periodo
2. il numero di particelle emesse da una sostanza radioattiva registrato in un intervallo di tempo
3. il numero di individui presenti in una data popolazione che evolve secondo certe leggi

## Capitolo 2

# Random Walk

### 2.1 Processo di Bernoulli-Passeggiata aleatoria semplice

(vedi [12, 6])

Sia  $E_1, E_2, \dots, E_n, \dots$  una successione di eventi stocasticamente indipendenti ed equiprobabili. Se chiamiamo successo alla  $n$ -esima prova il verificarsi dell'evento  $E_n$  allora  $p = P(E_n)$  rappresenta la probabilità di successo in ogni prova. Sia  $q = 1 - p$  la probabilità di insuccesso. Per ogni  $n \in \mathbb{N}$  poniamo  $X_n = |E_n|$ , pertanto si ha

$$X_n = \begin{cases} 1, & \text{con } P(X_n = 1) = p \\ 0, & \text{con } P(X_n = 0) = 1 - p = q. \end{cases}$$

In tal caso una realizzazione del processo è rappresentata da una successione di numeri binari. La successione  $\{X_n\}$  definisce un processo stocastico discreto a valori discreti, detto *processo di Bernoulli*.

Indichiamo con

$$S_n = X_1 + X_2 + \dots + X_n$$

il numero di successi su  $n$  prove. Per ogni  $n \in \mathbb{N}$  si ha

$$S_n \sim \text{Bin}(n, p)$$

A volte anche il processo  $\{S_n\}$  viene detto processo di Bernoulli. In tal caso è possibile risalire al processo  $\{X_n\}$  semplicemente osservando che

$$X_1 = S_1; \quad X_n = S_n - S_{n-1}, n > 1.$$

Introduciamo adesso i seguenti numeri aleatori

$$Y_n = |E_n| - |E_n^c| = 2X_n - 1,$$

si ha

$$Y_n = \begin{cases} +1, & \text{con } P(Y_n = 1) = p \\ -1, & \text{con } P(Y_n = -1) = 1 - p = q. \end{cases}$$

La variabile aleatoria  $Y_n$  può rappresentare il guadagno aleatorio in ogni prova. Si suppone cioè di guadagnare 1 o  $-1$ , rispettivamente, se si verifica o meno l'evento  $E_n$ .

Nasce, pertanto, il guadagno relativo alle prime  $n$  prove definito come

$$G_n = Y_1 + Y_2 + \dots + Y_n.$$

La successione  $\{G_n\}$  definisce un processo stocastico discreto a valori discreti, detto *passeggiata aleatoria semplice*.

Osservando che  $Y_i = 2X_i - 1$  si ha  $G_n = 2 \sum_{i=1}^n X_i - n = 2S_n - n$ .

Vogliamo trovare la distribuzione di  $G_n$  che si può ricavare da quella di  $S_n$ . Innanzitutto vediamo quali sono i possibile valori del guadagno. Fissiamo  $n = 4$  otteniamo

$$\begin{array}{rcccccc} S_4 = & & 0 & 1 & 2 & 3 & 4 \\ G_4 = 2S_4 - 4 = & -4 & -2 & 0 & 2 & 4. \end{array}$$

Per  $n = 5$  invece si ha

$$\begin{array}{rcccccc} S_5 = & & 0 & 1 & 2 & 3 & 4 & 5 \\ G_5 = 2S_5 - 5 = & -5 & -3 & -1 & 1 & 3 & 5. \end{array}$$

In generale, distinguiamo il caso  $n$  pari e dispari.

- Per  $n = 2m$ ,  $m \in \mathbb{N}$  i valori di  $G_n$  saranno tutti pari

$$G_n \in \{-2m, -2m + 2, -2, \dots, 0, 2, \dots, 2m - 2, 2m\}$$

- Per  $n = 2m - 1$ ,  $m \in \mathbb{N}$  i valori di  $G_n$  saranno tutti dispari

$$G_n \in \{-2m - 1, -2m + 1, -1, \dots, 1, \dots, 2m - 3, 2m - 1\}$$

- In generale

$$G_n \in \{-n, -n - 2, \dots, \dots, n - 2, n\}.$$

Pertanto per  $k \in \{-n, -n - 2, \dots, \dots, n + 2, n\}$  si ha

$$P(G_n = k) = P(2S_n - n = k) = P\left(S_n = \frac{n+k}{2}\right) = \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}$$

Per quanto visto prima, se  $n$  è pari (dispari) allora  $k$  è pari (dispari), quindi il numero  $\frac{n+k}{2}$  è sempre un intero.

In particolare su  $n$  prove la quantità  $\frac{n+k}{2}$  rappresenta il numero di passi positivi (successi) per arrivare al risultato mentre  $\frac{n-k}{2}$  rappresenta il numero di passi negativi insuccessi.

Inoltre si ha

$$\begin{aligned}\mathbb{E}(G_n) &= \mathbb{E}(2S_n - n) = 2np - n = n(2p - 1) = n(2p - p - q) = n(p - q), \\ \text{var}(G_n) &= \text{var}(2S_n - n) = 4npq.\end{aligned}$$

Osserviamo che la previsione  $\mathbb{E}(G_n)$  assume, per ogni  $n$ , lo stesso segno di  $p - q$ . Se  $p = q$  si ha  $\mathbb{E}(G_n) = 0$ .

Ricordando che per la legge dei grandi numeri si ha

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n - np}{n}\right| < \epsilon\right) = 1,$$

cioè  $\frac{S_n}{n} \xrightarrow{P} p$ , pertanto la frequenza relativa dei guadagni converge in probabilità a  $2p - 1 = p - q$ , cioè

$$\frac{G_n}{n} \xrightarrow{P} p - q.$$

Inoltre, poichè (per l'altro teorema di Bernoulli)

$$\forall k > 0 \lim_{n \rightarrow \infty} P(|S_n - np| > k) = 0,$$

si ha

$$\forall k > 0 \lim_{n \rightarrow \infty} P(|G_n - n(p - q)| > k) = 0.$$

ovvero, la v.a.  $|G_n - \mathbb{E}(G_n)|$  tende in probabilità all'infinito, cioè con alta probabilità i valori assoluti del guadagno saranno lontani dal suo valor medio. In particolare possiamo osservare che la  $\text{var}(G_n) = 4npq$  diverge al crescere di  $n$  (e pertanto lo s.q.m diverge al crescere di  $\sqrt{n}$ ). Ciò non accade invece per il guadagno relativo, infatti si ha che  $\text{var}\left(\frac{G_n}{n}\right) = \frac{4pq}{n}$  tende a zero.

Introduciamo il punto iniziale come  $G_0 = a$ .

**Lemma 1.** La passeggiata aleatoria semplice è omogenea nello spazio, cioè

$$P(G_n = j | G_0 = a) = P(G_n = j + b | G_0 = a + b)$$

Infatti entrambi sono uguali a  $P(\sum_{i=1}^n Y_i = j - a)$

**Lemma 2.** La passeggiata aleatoria semplice è omogenea nel tempo, cioè

$$P(G_n = j | G_0 = a) = P(G_{n+m} = j + b | G_m = a + b)$$

Infatti si ha

$$P\left(\sum_{i=1}^n Y_i = j - a\right) = P\left(\sum_{i=m+1}^{m+n} Y_i = j - a\right)$$

**Lemma 3.** La passeggiata aleatoria semplice gode della proprietà di Markov, cioè

$$P(G_{n+m} = j | G_0, G_1, \dots, G_n) = P(G_{n+m} = j | G_n)$$

### 2.1.1 Problema della rovina di un giocatore - Passeggiata aleatoria semplice con barriere assorbenti

(vedi [12, 3]). Immaginiamo un giocatore che gioca al casinò con un certo capitale iniziale  $a$  (ad esempio in mila euro) contro il banco vincendo o perdendo ad ogni giocata 1 (1000 euro). Indicando con  $E_i$  l'evento vincita alla  $i$ -esima prova, e supponiamo gli eventi indipendenti ed equiprobabili di probabilità  $p$ .

Supponiamo inoltre che il capitale totale del giocatore e del Banco sia pari ad  $T$ . Pertanto il capitale iniziale del banco sarà  $b = T - a$ . Il gioco termina quando si verifica una delle seguenti due situazioni

1. il giocatore con capitale iniziale  $a$  si rovina, ovvero il suo capitale si riduce a zero;
2. il giocatore ha sbancato il banco, ovvero il capitale del giocatore diviene pari ad  $T$ .

Indicando sull'asse  $x$  il numero di giocate e sull'asse  $y$  il capitale del giocatore, graficamente il gioco termina quando il capitale ( $y$ ) del giocatore interseca l'asse delle  $x$  oppure la retta di equazione  $y = T$ . A differenza della passeggiata aleatoria semplice, in questo caso, il processo può avere fine, pertanto si chiama pure *passeggiata aleatoria semplice con barriere assorbenti*.

Indichiamo con  $F_a = P(R_a)$  la probabilità che il giocatore con capitale  $a$  (prima o poi) si rovini. Si ha

$$F_a = P(R_a) = P(R_a \wedge \Omega) = P(R_a \wedge (E_1 \vee E_1^c)) = P(R_a|E_1)p + P(R_a|E_1^c)q.$$

Per l'indipendenza degli eventi  $E_i$  l'evento  $R_a|E_1$  è ugualmente probabile all'evento  $R_{a+1}$  che il giocatore con capitale iniziale  $a + 1$  si rovini, mentre  $R_a|E_1^c$  è equiprobabile all'evento  $R_{a-1}$ . Pertanto si ha

$$F_a = pF_{a+1} + qF_{a-1}, \quad a \in \{2, 3, \dots, T - 2\}.$$

In particolare se poniamo per convenzione  $F_0 = P(R_1|E_1^c) = 1$ , cioè la probabilità che un giocatore con capitale iniziale pari a 1 perdendo la prima prova si rovini è 1, e  $F_T = P(R_{T-1}|E_1) = 0$ , cioè la probabilità che un giocatore con capitale iniziale pari a  $T-1$  vincendo la prima prova si rovini è 0, si ha

$$F_a = pF_{a+1} + qF_{a-1}, \quad a \in \{1, 2, 3, \dots, T - 1\}.$$

Inoltre, (supponendo  $0 < p < 1$ ) essendo  $p + q = 1$ , per  $a \in \{1, 2, 3, \dots, T - 1\}$  si ha

$$\begin{aligned} (p + q)F_a &= pF_{a+1} + qF_{a-1} && \Rightarrow \\ pF_a + qF_a &= pF_{a+1} + qF_{a-1} && \Rightarrow \\ p(F_{a+1} - F_a) &= q(F_a - F_{a-1}) && \Rightarrow \\ (F_{a+1} - F_a) &= \frac{q}{p}(F_a - F_{a-1}) = \left(\frac{q}{p}\right)^2 (F_{a-1} - F_{a-2}) = \dots \end{aligned}$$

$$(F_{a+1} - F_a) = \left(\frac{q}{p}\right)^a (F_1 - F_0), \quad a \in \{1, 2, 3, \dots, T - 1\} \quad (2.1)$$

Gli incrementi  $F_{a+1} - F_a$  sono tutti negativi, in particolare abbiamo

$$F_T - F_0 = 0 - 1 = -1.$$

Calcoliamo  $(F_1 - F_0)$ . Iterando l'equazione 2.1 si ottiene

$$\begin{aligned} (F_T - F_0) &= (F_T - F_{T-1}) + (F_{T-1} - F_0) + \dots + (F_1 - F_0) = \\ &= \left(\frac{q}{p}\right)^{T-1} (F_1 - F_0) + \dots + \left(\frac{q}{p}\right)^0 (F_1 - F_0) = \\ &= (F_1 - F_0) \sum_{n=0}^{T-1} \left(\frac{q}{p}\right)^n \end{aligned}$$

ovvero

$$F_1 - F_0 = \frac{-1}{\sum_{n=0}^{T-1} \left(\frac{q}{p}\right)^n}. \quad (2.2)$$

Procedendo in modo analogo si ottiene

$$(F_a - F_0) = (F_1 - F_0) \sum_{n=0}^{a-1} \left(\frac{q}{p}\right)^n = -\frac{\sum_{n=0}^{a-1} \left(\frac{q}{p}\right)^n}{\sum_{n=0}^{T-1} \left(\frac{q}{p}\right)^n}$$

In definitiva si ha

$$F_a = 1 - \frac{\sum_{n=0}^{a-1} \left(\frac{q}{p}\right)^n}{\sum_{n=0}^{T-1} \left(\frac{q}{p}\right)^n}, \quad a = 1, 2, \dots, T \quad (2.3)$$

Distinguiamo 2 casi,  $p = q = \frac{1}{2}$  e  $p \neq q$ .

1. Caso del gioco equo  $p = q = \frac{1}{2}$ . Dalla (2.3) si ottiene

$$F_a = 1 - \frac{a}{T} = \frac{T-a}{T} = \frac{b}{T}.$$

Possiamo pertanto dire che la probabilità di rovina è proporzionale al capitale iniziale dell'avversario.

2. Caso del gioco non equo  $p \neq q$ . Sempre dalla (2.3), ricordando che

$$\sum_{n=0}^{a-1} \left(\frac{q}{p}\right)^n = \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \frac{q}{p}} \text{ si ottiene}$$

$$F_a = 1 - \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^T} = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^T}{1 - \left(\frac{q}{p}\right)^T}$$

Consideriamo adesso la probabilità  $H_b$  che il banco con capitale iniziale  $b$  si rovini. In maniera analoga a quanto visto per  $F_a$ , scambiando  $p$  con  $q$  e mettendo  $b$  al posto di  $a$  si ottiene

$$H_b = \frac{\left(\frac{p}{q}\right)^b - \left(\frac{p}{q}\right)^T}{1 - \left(\frac{p}{q}\right)^T}$$

Si può verificare che  $F_a + H_b = 1$ , cioè la probabilità che prima o poi il giocatore o il banco si rovini è 1. Quindi la probabilità che il gioco duri all'infinito è 0. Inoltre si ha che la probabilità di vincita  $V_a = 1 - F_a = H_b$  si può esprimere come

$$V_a = \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^T}$$

**Durata del gioco** Indichiamo con  $D_a$  la previsione della durata del gioco nel caso in cui il capitale iniziale del giocatore sia pari ad  $a$  (e quello totale sia pari a  $T$ ). Poichè con probabilità 1 il gioco prima o poi termina si può assumere che  $D_a$  sia un numero finito.

Ricordiamo la definizione di previsione condizionata. Dato il vettore aleatorio  $(X, Y)$ , consideriamo la previsione condizionata di  $Y$  dato  $X = x$ , cioè

$$\mathbb{E}(Y|X = x). \quad (2.4)$$

Pertanto,  $\mathbb{E}(Y|X)$  è una variabile aleatoria funzione di  $X$ , che indicheremo con  $\psi(X)$ .

**Definizione 1.** Dato il vettore aleatorio  $(X, Y)$ , si definisce previsione condizionata (o valore atteso condizionato) di  $Y$  dato  $X$  la seguente variabile aleatoria

$$\mathbb{E}(Y|X) = \psi(X). \quad (2.5)$$

Essendo  $\psi(X)$  una v.a. possiamo calcolarne la previsione. Si dimostra il seguente

**Teorema 1.** La previsione condizionata  $\psi(X) = \mathbb{E}(Y|X)$  soddisfa la seguente uguaglianza

$$\mathbb{E}(\psi(X)) = \mathbb{E}(Y).$$

Nel caso discreto si ottiene la seguente formula utile a volte per calcolare la previsione di  $Y$ .

$$\mathbb{E}(Y) = \sum_x \mathbb{E}(Y|X = x)P(X = x). \quad (2.6)$$

Nel nostro caso se indichiamo con  $Z_a$  la variabile aleatoria durata del gioco, quindi  $D_a = \mathbb{E}(Z_a)$ , applicando la formula (2.6) si ottiene per  $1 < a < T - 1$

$$\begin{aligned} D_a &= \mathbb{E}(Z_a) = \mathbb{E}(Z_a|E_1)P(E_1) + \mathbb{E}(Z_a|E_1^c)P(E_1^c) = \\ &= \mathbb{E}(Z_{a+1} + 1)p + \mathbb{E}(Z_{a-1} + 1)q. \end{aligned}$$

Ponendo  $D_0 = D_T = 0$ , si ha che la seguente equazione

$$D_a = pD_{a+1} + qD_{a-1} + 1 \quad (2.7)$$

deve essere soddisfatta per  $a = 1, \dots, T - 1$ . Tale equazione non è omogenea in quanto compare il termine 1 a secondo membro.

Consideriamo il caso  $p \neq q$ . Una soluzione è data da  $D_a = a/(q - p)$ .

Sia  $\Delta_a$  la differenza tra due soluzioni dell'equazione (2.7). Si osserva che  $\Delta_a$  deve soddisfare l'equazione

$$\Delta_a = p\Delta_{a+1} + q\Delta_{a-1} \quad (2.8)$$

per  $0 < a < 1$ . Osserviamo che  $\Delta'_a = 1$  e  $\Delta''_a = (q/p)^a$  sono due soluzioni della (2.8). Pertanto, tutte le soluzioni della (2.8) sono combinazioni lineari di 1 e di  $(q/p)^a$ , ovvero del tipo

$$\alpha + \beta(q/p)^a,$$

con  $\alpha$  e  $\beta$  reali positivi.

Poichè,  $D_a = a/(q - p)$  è una soluzione della (2.7) si ha che tutte le soluzioni della (2.7) sono date da

$$D_a = \frac{a}{(q - p)} + \alpha + \beta(q/p)^a.$$

Imponendo le condizioni  $D_0 = D_T = 0$ , si ha

$$D_0 = \alpha + \beta = 0, \quad D_T = \frac{T}{(q - p)} + \alpha + \beta(q/p)^T = 0.$$

Risolvendo per  $\alpha$  e  $\beta$  si ottiene (per  $p \neq q$ )

$$D_a = \frac{a}{(q - p)} - \frac{T}{(q - p)} \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^T}. \quad (2.9)$$

Consideriamo il caso  $p = q = 1/2$ . In tal caso, si ha che  $-a^2$  è una particolare soluzione della (2.7). In tal caso le due precedenti soluzioni  $\Delta'_a = 1$  e  $\Delta''_a = (q/p)^a$  coincidono. Osserviamo che  $\Delta'''_a = a$  è una soluzione della (2.8), pertanto tutte le soluzioni sono del tipo

$$-a^2 + \alpha + a\beta.$$

Imponendo le condizioni  $D_0 = 0$  e  $D_T = 0$  si ottiene

$$D_0 = \alpha = 0, \quad D_T = -T^2 + \alpha + \beta T = 0.$$

Si ricava che  $\alpha = 0$  e  $\beta = T$ . Pertanto, per  $p = q = 1/2$ , la soluzione è

$$D_a = -a^2 + aT = a(T - a) = ab.$$

Ovvero, nel caso di gioco equo, la durata media del gioco è data dal prodotto dei capitali. Va osservato che la durata media è particolarmente lunga rispetto a quanto ci si possa aspettare. Infatti se 2 giocatori con 500 euro ciascuno giocano lanciando una moneta sino a quando uno dei 2 si rovini, la durata media del gioco è pari a 250 mila lanci. Se il giocatore ha 1 euro e il banco 1000 la durata media del gioco è di 1000 prove.

Se  $p = 0.45$  e  $q = 0.55$  si ottiene

$p$	$q$	$a$	$T$	b	$F_a$	$H_b$	$\mathbb{E}(G)$	$D_a$
0.5	0.5	9	10	1	0.1	0.9	0	9
0.45	0.55	9	10	1	0.210	0.790	-1.1	11
0.5	0.5	90	100	10	0.1	0.9	0	900
0.45	0.55	90	100	10	0.866	0.134	-76.6	765.6
0.45	0.55	99	100	1	0.182	0.818	-17.2	171.8

Tabella 2.1: Problema della rovina di un giocatore. Capitale iniziale= $a$ .  $F_a$  =probabilità di rovina.  $H_b$  =probabilità di sbancare il banco. Il gioco termina quando il capitale diviene 0 oppure  $T$ .

**Riformulazione** (vedi [7]) Riformuliamo il gioco in un'altra veste. Supponiamo che un giocatore con capitale iniziale  $a$  gioca contro il banco con capitale infinito che è sempre disposto a giocare, nonostante il giocatore abbia il privilegio di fermarsi a suo piacimento. Il giocatore adotta la strategia di giocare sino a quando non si verifica una delle seguenti condizioni

1. perde tutto il suo capitale.
2. il suo capitale raggiunge il valore  $T$  (con guadagno  $T - a$ ).

Pertanto  $F_a$  è la probabilità di rovina e  $1 - F_a$  è la probabilità di vincita. Sotto queste condizioni il guadagno finale  $G$  del giocatore, per quanto abbiamo visto, sarà

$$G = \begin{cases} -a, & \text{se si rovina} \\ T - a, & \text{se vince} \end{cases}$$

Il valor medio del guadagno è

$$\mathbb{E}(G) = -aF_a + (T - a)(1 - F_a) = T(1 - F_a) - a.$$

Quindi avremo  $\mathbb{E}(G) = 0$  se e solo se  $p = q = \frac{1}{2}$ .

**Esempio 2.** Un giocatore con capitale iniziale  $a = 999$  euro vuole raggiungere  $T = 1000$  euro, cioè vuole guadagnare un euro. Se  $p = q$  allora la probabilità che guadagni un euro (cioè raggiunge quota  $T = 1000$ ) sarà

$$V_a = 1 - F_a = \frac{999}{1000} = 0.999.$$

Se invece  $p = 0.4$  e  $q = 0.6$  tale probabilità si abbassa, ma è ancora più alta del 50%, infatti

$$1 - F_a = \frac{\left(\frac{2}{3}\right)^{1000-999} - \left(\frac{2}{3}\right)^{1000}}{1 - \left(\frac{2}{3}\right)^{1000}} \simeq \frac{2}{3}.$$

Il guadagno medio sarà  $\mathbb{E}(G) = 1000\left(\frac{2}{3}\right) - 999 = -332$ . In generale un giocatore con un grande capitale iniziale  $a$  ha buone chance di guadagnare un piccola cifra  $T - a$  prima di rovinarsi.

**Esercizio 1.** (vedi [10]) Alice e Bob decidono di giocare con le monete da un euro. Chi lancia la moneta più vicino al muro vince un euro. Supponiamo che Alice sia più brava di Bob e che la sua probabilità di vincere sia pari a 0.6. Se Alice inizia a giocare con 5 monete da un 1 euro e Bob con 10 qual è la probabilità che Alice metterà fuori gara Bob. Calcolare tale probabilità nel caso in cui Alice inizia con 10 e Bob con 20.

**Esempio 3** (Passeggiata aleatoria, probabilità di raggiungimento). Siano  $a$  e  $b$  due interi positivi e sia  $G_n$  una passeggiata aleatoria semplice (di posizione iniziale  $G_0 = 0$ ). Vogliamo calcolare la probabilità  $p(a)$  che  $G_n$  raggiunga il valore  $-a$  prima del valore  $b$ .

**Esempio 4** (Applicazione del problema della R.G. ai Test sui farmaci). (vedi [10]) Consideriamo una applicazione del problema della rovina di un giocatore in campo sanitario. Supponiamo che due nuovi farmaci  $F_1, F_2$  sono stati sviluppati per curare un certa malattia. Siano  $r_1, r_2$  i tassi di guarigione (incogniti), rispettivamente, di  $F_1, F_2$ , nel senso che ciascun paziente curato con  $F_i$  guarirà con probabilità  $r_i$ . Si è interessati a trovare un metodo che permette di stabilire quale farmaco ha il tasso di guarigione più alto. Cioè  $H_1 = (r_1 > r_2)$  o  $H_2 = (r_2 > r_1)$ . Per scegliere una tra le 2 alternative si considera il seguente test. Coppie di pazienti sono curati sequenzialmente uno con il farmaco  $F_1$  e l'altro con il farmaco  $F_2$ . I risultati di ciascuna coppia sono noti e il test finisce quando il numero di pazienti guariti con uno dei 2 farmaci supera l'altro di una certa quantità  $M$ . Indichiamo con

$A_i$  = "il paziente della coppia  $i$ -esima curato con il farmaco  $F_1$  guarisce"  
 $B_i$  = "il paziente della coppia  $i$ -esima curato con il farmaco  $F_2$  guarisce"

e con

$$\begin{aligned} S_n &= |A_1| + |A_2| + \dots + |A_n|, \\ T_n &= |B_1| + |B_2| + \dots + |B_n| \\ Z_n &= S_n - T_n. \end{aligned}$$

Il test finisce quando  $|Z_n| = M$ . Se si verifica  $Z_n = M$  si asserisce  $H_1$  invece nel caso in cui  $Z_n = -M$  si asserisce  $H_2$ . Per vedere se tale scelta è fatta bene, siamo interessati a calcolare la probabilità che essa ci conduca ad una decisione incorretta. Supponendo noti i valori di  $r_1$  e di  $r_2$ , con  $r_1 > r_2$ , cioè l'ipotesi  $H_1$  vera, vogliamo calcolare la probabilità che il test asserirà incorrettamente  $H_2$ . Cominciamo col porre

$$\Delta_i = |A_i| - |B_i|, \quad i = 1, 2, \dots, n.$$

Pertanto si avrà

$$Z_n = S_n - T_n = \sum_{i=1}^n \Delta_i.$$

Osserviamo che  $\Delta_i$  si può scrivere nel seguente modo

$$\Delta_i = 0|A_i B_i \vee A_i^c B_i^c| + 1|A_i B_i^c| - 1|A_i^c B_i|$$

Se supponiamo gli eventi  $A_1, A_2, \dots, B_1, B_2, \dots$  stocasticamente indipendenti, gli eventi  $A_i$  equiprobabili con probabilità  $r_1$  e gli eventi  $B_i$  equiprobabili con probabilità  $r_2$  si ha (per ogni intero  $i$ )

$$\begin{aligned} P(\Delta_i = 0) &= r_1 r_2 + (1 - r_1)(1 - r_2), \\ P(\Delta_i = 1) &= r_1(1 - r_2), \\ P(\Delta_i = -1) &= r_2(1 - r_1). \end{aligned} \tag{2.10}$$

Quindi, possiamo affermare che  $Z_n$  rappresenta una passeggiata aleatoria in cui gli incrementi  $\Delta_i$  assumono uno dei tre valori  $\{-1, 0, 1\}$  con probabilità fornite dalla (2.10). Noi vogliamo calcolare la probabilità che  $Z_n$  raggiunga  $-M$  prima di  $M$ . Questa probabilità si può vedere nell'ambito del problema della rovina di un giocatore. Ovvero, essa coincide con la probabilità che un giocatore con capitale iniziale  $M$  si rovini prima di raddoppiare il suo capitale ( $2M$ ). E' ovvio che l'evento "il giocatore si rovina prima di vincere  $2M$ " è equivalente all'evento "il giocatore si rovina" (da notare che siamo nel caso in cui sono presenti le barriere assorbenti). L'unica differenza tra  $Z_n$  e il guadagno nel problema della rovina di un giocatore è che in  $Z_n$  si ha  $P(\Delta_i = 0) > 0$ . Fortunatamente, nel caso in cui  $\Delta_i = 0$  il valore di  $Z_n$  rimane inalterato, pertanto non cambia nulla se ci restringiamo al caso in cui si hanno delle variazioni, cioè  $\Delta_i \neq 0$ . Cioè definiamo dei nuovi incrementi condizionati  $\Delta'_i = (\Delta_i | \Delta_i \neq 0)$ . Si ha

$$P(\Delta'_i = 1) = P(\Delta_i = 1 | \Delta_i \neq 0) = \frac{r_1(1 - r_2)}{r_1(1 - r_2) + r_2(1 - r_1)}$$

e ovviamente

$$P(\Delta'_i = -1) = 1 - p = P(\Delta_i = -1 | \Delta_i \neq 0).$$

Pertanto  $Z'_n = \sum_{i=1}^n \Delta'_i$  rappresenta il guadagno aleatorio in un problema della rovina di un giocatore con  $p = P(\Delta'_i = 1)$  e  $q = 1 - p = P(\Delta'_i = -1)$ . Quindi la probabilità che  $Z_n$  raggiunga  $-M$  prima di  $M$  coincide con la probabilità che  $Z'_n$  raggiunga  $-M$  prima di  $M$ . Concludendo, la probabilità che il test asserirà  $H_2$  è uguale alla probabilità che un giocatore, con probabilità di vincita  $p$  e con capitale iniziale  $M$  si rovinerà prima di raddoppiare il suo capitale ( $2M$ ). Pertanto si ha

$$P(\text{Il test asserisce } H_2 | H_1) = 1 - \frac{1 - \left(\frac{q}{p}\right)^M}{1 - \left(\frac{q}{p}\right)^{2M}}$$

ovvero

$$P(\text{Il test asserisce } H_2 | H_1) = \frac{1}{1 + \left(\frac{p}{q}\right)^M}$$

Se ad esempio  $r_1 = 0.6, r_2 = 0.4$  (non è detto che  $r_1 + r_2 = 1$ ) allora la probabilità di una decisione incorretta sarà 0.017 per  $M = 5$  e si riduce a 0.003 per  $M = 20$ .

**Osservazione 1.** (vedi [7]) Supponiamo che in una situazione sfavorevole,  $q > p$ , il giocatore giochi un importo di 0.5 al posto di 1. In tal caso il gioco è

equivalente a considerare un importo iniziale pari a  $2a$  e un importo totale pari a  $2T$ . Pertanto se indichiamo con  $F_a^*$  la probabilità di rovina del giocatore che utilizza tale strategia si ha

$$F_a^* = F_{2a} = \frac{\left(\frac{q}{p}\right)^{2a} - \left(\frac{q}{p}\right)^{2T}}{1 - \left(\frac{q}{p}\right)^{2T}} = F_a \frac{\overbrace{\left(\frac{q}{p}\right)^a}^{>1} + \left(\frac{q}{p}\right)^T}{1 + \left(\frac{q}{p}\right)^T} > F_a.$$

Cioè la probabilità di rovina del giocatore aumenta. Viceversa se il giocatore gioca un importo doppio, la sua probabilità di rovina diminuisce. Ad esempio se  $p = 0.45$  e  $a = 90$  euro e  $T = 100$  euro (cioè  $b = 10$  euro), allora giocando importi di un euro si ha

$$F_{90} = 0.866$$

mentre se si giocano importi di 10 euro si ha (indicando con  $F'_{90}$  tale probabilità)

$$F'_{90} = 0.210.$$

In generale, se vengono giocati  $k$  euro ad ogni puntata, indicando con  $F_{a,k}$  la probabilità di rovina, si ha

$$F_{a,k} = \frac{\left(\frac{q}{p}\right)^{a/k} - \left(\frac{q}{p}\right)^{T/k}}{1 - \left(\frac{q}{p}\right)^{T/k}}$$

e la probabilità di rovina decresce al crescere di  $k$ .

### 2.1.2 Capitale infinito

(vedi [7]) Analizziamo adesso il caso in cui il capitale del banco sia molto più grande di quello del giocatore. Ovvero supponiamo  $a$  finito e  $b$  infinito. In tal caso si ha pure  $T = a + b = \infty$ . Analizziamo tre casi differenti.

1.  $p = q = \frac{1}{2}$ . In tal caso si ha

$$\lim_{T \rightarrow \infty} F_a = \lim_{T \rightarrow \infty} 1 - \frac{a}{T} = 1. \quad (2.11)$$

Cioè se il gioco è equo, asintoticamente la probabilità che prima o poi il giocatore si rovini è 1. Inoltre si ha

$$\lim_{T \rightarrow \infty} D_a = \lim_{T \rightarrow \infty} a(T - a) = \infty.$$

Quindi mediamente il gioco dura all'infinito.

2. Sia  $p < q$ . Si ha  $\frac{q}{p} > 1$ , pertanto

$$\lim_{T \rightarrow \infty} F_a = \lim_{T \rightarrow \infty} \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^T}{1 - \left(\frac{q}{p}\right)^T} = 1. \quad (2.12)$$

Ovvero, se in ogni prova la probabilità di vincita del giocatore è più bassa di quella del banco, allora, anche in questo caso, asintoticamente la probabilità che prima o poi il giocatore si rovini è 1. Per quanto riguarda la durata di ottiene

$$\lim_{T \rightarrow \infty} D_a = \lim_{T \rightarrow \infty} \frac{a}{(q-p)} - \frac{T}{(q-p)} \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^T} = \frac{a}{(q-p)}.$$

Ovvero, la durata media del gioco è finita.

3. Sia  $p > q$ . Si ha  $\frac{q}{p} < 1$

$$\lim_{T \rightarrow \infty} F_a = \lim_{T \rightarrow \infty} \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^T}{1 - \left(\frac{q}{p}\right)^T} = \left(\frac{q}{p}\right)^a \quad (2.13)$$

Ovvero, se in ogni prova la probabilità di vincita del giocatore è più alta di quella del banco, allora asintoticamente la probabilità che prima o poi il giocatore con capitale iniziale  $a$  si rovini (nonostante il banco abbia capitale infinito) è data da  $\left(\frac{q}{p}\right)^a < 1$ . Per quanto riguarda la durata di ottiene

$$\lim_{T \rightarrow \infty} D_a = \lim_{T \rightarrow \infty} \frac{a}{(q-p)} - \frac{T}{(q-p)} \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^T} = +\infty.$$

Ovvero, la durata media del gioco è infinita.

**Esempio 5** (Compagnia di Assicurazione). Consideriamo una compagnia di assicurazioni che incassa costantemente \$1 al giorno e che potrebbe pagare per risarcimento \$2 al giorno con probabilità  $q = 1 - p$ . Quindi dopo  $n$  giorni il guadagno netto corrisponde al guadagno aleatorio nel problema della rovina di un giocatore in cui la probabilità giornaliera di guadagnare \$1 è  $p$  e quella di guadagnare \$ - 1 è  $q$ . Se supponiamo che la compagnia abbia un capitale iniziale \$ $a$  e che il capitale che può guadagnare sia infinito allora le probabilità di rovina della compagnia di assicurazione sono date dalle formule (2.11, 2.12, 2.13). Pertanto è evidente che nessuna compagnia è disposta ad accettare la condizione  $p \leq q$  poichè in tal caso sarebbe pari a 1 la probabilità di rovina ed inoltre il guadagno medio giornaliero sarebbe pari a  $p - q \leq 0$ . Se invece si ha  $p > q$  il guadagno medio giornaliero è pari a  $p - q > 0$  e la probabilità di rovina è  $\left(\frac{q}{p}\right)^a$ . Quest'ultima diminuisce all'aumentare del capitale iniziale.

## Capitolo 3

# Catene di Markov discrete

### 3.1 Catene di Markov a tempo discreto DTMC.

#### 3.1.1 Introduzione ed esempi DTCM.

(vedi [9],[10],[2]) Sia  $\{X_n, n = 0, 1, 2, \dots\}$  un processo stocastico a parametro discreto e a valori discreti. Supponiamo inoltre che, per ogni  $n \in \mathbb{N}_0$ , sia  $X_n$  una variabile aleatoria con valori in  $S$  (con  $S$  un insieme finito o al più numerabile). L'insieme  $S$  viene chiamato *spazio degli stati*, il quale, se non diversamente specificato, rappresenta l'insieme dei numeri naturali compreso lo zero, cioè

$$S = \{0, 1, 2, \dots\}.$$

Nell'ipotesi in cui il processo rappresenta l'evoluzione aleatoria di un sistema nel tempo (discreto) la variabile  $X_n$  denota lo *stato del sistema al tempo  $n$* , ovvero l'evento  $(X_n = h)$  significa che il sistema al tempo  $n$  si trova nello stato  $h$ . Se l'istante  $n$  rappresenta il *presente*, allora  $X_n$  rappresenta lo *stato attuale* del sistema, mentre la sequenza  $\{X_{n+1}, X_{n+2}, \dots\}$  rappresenta gli stati *futuri* (in breve il futuro) del sistema e la sequenza  $\{X_0, X_1, \dots, X_{n-1}\}$  rappresenta gli stati *passati* (in breve il passato) del sistema. In questo paragrafo restringiamo la nostra attenzione allo studio dei processi che godono della seguente proprietà, detta *Proprietà di Markov*,

(M) "Noto lo stato presente del sistema, il futuro del sistema è (stocasticamente) indipendente dal passato".

In altre parole, in un sistema avente la suddetta proprietà, il passato influisce sulle valutazioni probabilistiche relative al futuro solo attraverso il presente, ovvero lo stato presente del sistema contiene tutte le informazioni necessarie per le valutazioni probabilistiche sul futuro.

Un processo stocastico  $\{X_n, n = 0, 1, 2, \dots\}$  che gode della proprietà (M) si dice *Catena di Markov a tempo discreto* (discrete-time Markov chain DTMC). Più precisamente, in una DTCM la distribuzione di probabilità condizionata di  $X_{n+1}$  dato il passato  $\{X_0, X_1, \dots, X_{n-1}\}$  e il presente  $X_n$  dipende solo da  $X_n$ . Formalmente,

**Definizione 2.** Il processo  $\{X_n, n = 0, 1, 2, \dots\}$  dicesi DTCM con spazio degli stati  $S$  se gode delle seguenti proprietà:

- (i) per ogni  $n \geq 0, X_n \in S$ ;
- (ii) per ogni  $n \geq 0, i, j, i_{n-1}, \dots, i_0 \in S$ , si ha

$$P(X_{n+1} = j | X_n = i, (X_{n-1} = i_{n-1}, \dots, X_0 = i_0)) = P(X_{n+1} = j | X_n = i). \quad (3.1)$$

Indichiamo con

$$p_{ij}(n) = P(X_{n+1} = j | X_n = i), \quad (3.2)$$

la probabilità di transizione (ad un passo) dallo stato  $i$  allo stato  $j$  al tempo  $n$ .

**Definizione 3.** Una Catena di Markov a tempo discreto  $\{X_n, n \in \mathbb{N}_0\}$  dicesi *omogenea* se  $p_{ij}(n)$  non dipende dal tempo  $n$ , ovvero se una DTCM gode della seguente proprietà

$$p_{ij}(n) = p_{ij}, \quad n = 0, 1, \dots \quad (3.3)$$

**Da questo momento ci occuperemo solo delle catene di Markov a tempo discreto omogenee.** In tali catene le probabilità di transizione, come detto, non dipendono dal tempo  $n$  pertanto possono tutte ricondursi alle probabilità di transizione dall'istante iniziale a quello successivo, cioè

$$p_{ij}(n) = P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i), \quad \forall n \in \mathbb{N}_0.$$

Indichiamo con  $p_{ij}$  la generica probabilità di transizione dallo stato  $i$  allo stato  $j$ . Le  $p_{ij}$  possono essere rappresentate in forma matriciale. Indichiamo con  $P = [p_{ij}]$  la matrice in cui  $P_{ij} = p_{ij}$ . Tale matrice viene detta *matrice di transizione ad un passo*. Osserviamo che  $P$  può anche essere una matrice con infinite righe e infinite colonne. Se  $S$  è finito, ad esempio, se  $S = \{0, 1, 2, \dots, m\}$ , la matrice  $P$  è la seguente

$$P = \begin{array}{c} 0 \\ 1 \\ \vdots \\ m \end{array} \left\| \begin{array}{cccc} 0 & 1 & \dots & m \\ p_{00} & p_{01} & \dots & p_{0m} \\ p_{10} & p_{11} & \dots & p_{1m} \\ \vdots & \vdots & \dots & \vdots \\ p_{m0} & p_{m1} & \dots & p_{mm} \end{array} \right\|$$

Si può facilmente verificare che la matrice  $P$  soddisfa le seguenti proprietà

1.  $p_{ij} \geq 0$  per ogni  $i, j \in S$ ;
2.  $\sum_{j \in S} p_{ij} = 1$  per ogni  $i \in S$ .

La Proprietà 1. ci dice che la matrice  $P$  è costituita da termini non negativi, mentre la Proprietà 2. (che è valida solo sotto l'ipotesi di  $\sigma$ -additività) ci dice che ciascuna riga ha somma (che in generale è la somma di una serie) pari ad 1.

**Osservazione 2.** La Proprietà 2. delle matrici di transizione si può esprimere dicendo che la matrice  $P$  ammette sempre l'autovalore 1. Infatti indicando con  $\mathbf{1}_u$  un vettore unitario si ha

$$P \cdot \mathbf{1}_u = \mathbf{1}_u.$$

Per poter descrivere interamente una DTCM non è sufficiente conoscere la matrice  $P$ , in quanto in essa sono presenti solo le probabilità condizionate. L'ulteriore conoscenza della distribuzione di probabilità iniziale  $\mathbf{a}^{(0)} = \mathbf{a} = (a_i, i \in S)$ , con

$$a_i = P(X_0 = i), \quad \forall i \in S,$$

consente, come mostrato dal seguente teorema, di poter descrivere completamente una DTCM.

**Teorema 2.** Una DTCM omogenea  $\{X_n, n = 0, 1, 2, \dots\}$  è completamente caratterizzata dalla sua distribuzione iniziale  $\mathbf{a}$  e dalla matrice di transizione  $P$

*Dimostrazione.* Siano date la distribuzione di probabilità iniziale  $a_{i_0} = P(X_0 = i_0)$  e la matrice di transizione  $P = [p_{ij}]$  di  $X_n, n \in \mathbb{N}_0$ . Allora, la distribuzione marginale di  $X_0$  è data da  $a_{i_0} = P(X_0 = i_0)$ . Consideriamo la distribuzione congiunta di  $(X_0, X_1)$ . Si ha

$$P(X_1 = i_1, X_0 = i_0) = P(X_1 = i_1 | X_0 = i_0)P(X_0 = i_0) = p_{i_0, i_1} a_{i_0}. \quad (3.4)$$

Pertanto, anche la distribuzione congiunta di  $(X_0, X_1)$  è calcolabile a partire dalle ipotesi tramite la (3.4). Supponiamo per  $k = 1, 2, \dots, n-1$  che valga la seguente equazione

$$P(X_k = i_k, \dots, X_1 = i_1, X_0 = i_0) = a_{i_0} p_{i_0, i_1} p_{i_1, i_2} \cdots p_{i_{k-1}, i_k} \quad (3.5)$$

e mostriamo che essa vale per  $k = n$ .

$$\begin{aligned} P(X_n = i_n, \dots, X_1 = i_1, X_0 = i_0) &= \\ & \underbrace{P(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_0 = i_0)}_{\text{Prop. Markov} \downarrow} P(X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \\ & P(X_n = i_n | X_{n-1} = i_{n-1}) \underbrace{P(X_{n-1} = i_{n-1}, \dots, X_0 = i_0)}_{\text{ipotesi ind.} \downarrow} = \\ & = p_{i_{n-1}, i_n} \cdot a_{i_0} p_{i_0, i_1} p_{i_1, i_2} \cdots p_{i_{n-2}, i_{n-1}} = \\ & = a_{i_0} p_{i_0, i_1} p_{i_1, i_2} \cdots p_{i_{n-1}, i_n} \quad (3.6) \end{aligned}$$

Abbiamo mostrato, per induzione, che, per ogni  $n$ , la distribuzione congiunta  $(X_1, X_2, \dots, X_n)$  è determinata a partire dalle probabilità iniziali  $a_i$  e dalla matrice di transizione  $P$ .  $\square$

**Diagramma di transizione degli stati.** Una DTCM (omogenea) può essere rappresentata tramite il *diagramma di transizione degli stati*. Esso è un grafo orientato con un nodo per ogni stato in  $S$  e un arco orientato dal nodo  $i$  al

nodo  $j$  se  $p_{ij} > 0$ . Un loop (un arco dal nodo  $i$  al nodo  $i$ ) è presente se  $p_{ii} > 0$ . Il comportamento dinamico della DTCM può essere ben visualizzato tramite il diagramma di transizione degli stati immaginando una particella che con una certa probabilità iniziale si trova in un nodo e che si muove da un nodo ad un altro scegliendo gli archi tramite le corrispondenti probabilità di transizione.

**Esempio 6.** Supponiamo che la probabilità che domani piova o no dipenda dalle condizioni meteorologiche passate solo relativamente allo stato presente, cioè al fatto che oggi piova o no. Supponiamo inoltre che se oggi piove, allora la probabilità che domani non piova sia  $\alpha$ , e, se oggi non piove, la probabilità che domani piova sia pari a  $\beta$ . Diremo che il sistema si trova nello stato 0 quando *non piove* e nello stato 1 quando *piove*,  $S = \{0, 1\}$ . Il precedente sistema forma una catena di Markov a 2 stati con la seguente matrice di transizione

$$\begin{array}{c} 0 \\ 1 \end{array} \left\| \begin{array}{cc} 0 & 1 \\ 1 - \beta & \beta \\ \alpha & 1 - \alpha \end{array} \right\|$$

Supposto che piova oggi, si vuole calcolare la probabilità della durata di una sequenza di pioggia. Consideriamo il numero aleatorio condizionato dei giorni di attesa, a partire da un giorno di pioggia, sino all'arrivo di un giorno senza pioggia. Indicando con  $T$  tale numero aleatorio la probabilità dell'evento ( $T = k$ ), per ogni  $k \in \mathbb{N}$ , è la probabilità che, partendo da un giorno in cui piove, il primo giorno in cui non pioverà si verificherà dopo  $k$  giorni,

$$P(T = k) = P(X_k = 0, X_{k-1} = 1, \dots, X_1 = 1 | X_0 = 1) = (1 - \alpha)^{k-1} \alpha.$$

Infatti, indicando con  $E_n$  l'evento "Piove il giorno  $n$ -esimo", per  $n \in \mathbb{N}_0$ , si ha  $(T = k) = (E_k^c E_{k-1} \cdots E_1 | E_0)$ , per  $k = 1, 2, \dots$ , pertanto

$$\begin{aligned} P(T = k) &= P(E_k^c | E_{k-1} \cdots E_1 E_0) P(E_{k-1} | E_{k-2} \cdots E_1 E_0) \cdots P(E_1 | E_0) = \\ &= P(E_k^c | E_{k-1}) P(E_{k-1} | E_{k-2}) \cdots P(E_2 | E_1) P(E_1 | E_0) = \\ &= P(E_1^c | E_0) [P(E_1 | E_0)]^{k-1} = (1 - \alpha)^{k-1} \alpha. \end{aligned}$$

Si verifica facilmente che  $T$  ha distribuzione geometrica di parametro  $\alpha$  ( $T \simeq G(\alpha)$ ), ovvero

$$\begin{cases} T \in \{\mathbb{N} \vee \infty\}, \\ P(T = k) = (1 - \alpha)^{k-1} \alpha, \quad k \in \mathbb{N}, \\ P(T = \infty) = 0. \end{cases}$$

**Esercizio 2.** Relativamente all'Esempio 6 calcolare

- $P(T = 10 | T = 5)$ ;
- $\mathbb{E}(T)$ ;
- $P(T > 10 | T > 5)$ .

**Esempio 7** (Trasformazione di un processo in una catena di Markov). (Vedi [10]). Supponiamo il fatto che domani piova o no dipenda dalle condizioni meteorologiche sia di oggi che di ieri. Indichiamo con  $E_n$  l'evento "Piove il giorno n-esimo", per  $n \in \mathbb{N}_0$  valutiamo le seguenti probabilità:

- se ha piovuto sia oggi sia ieri allora, la probabilità che domani piovierà è 0.7,  $P(E_2|E_1E_0) = 0.7$ ;
- se ha piovuto oggi ma non ieri allora, la probabilità che domani piovierà è 0.5,  $P(E_2|E_1E_0^c) = 0.5$ ;
- se ha piovuto ieri ma non oggi allora, la probabilità che domani piovierà è 0.4,  $P(E_2|E_1^cE_0) = 0.4$ ;
- se non ha piovuto nè ieri nè oggi allora, la probabilità che domani piovierà è 0.2,  $P(E_2|E_1^cE_0^c) = 0.2$ .

Per poter rappresentare il sistema con una catena di Markov, indicando con  $n$  il presente, consideriamo i seguenti stati

1. ha piovuto sia oggi sia ieri, ovvero l'evento  $E_nE_{n-1}$ ,
2. ha piovuto oggi ma non ieri, ovvero l'evento  $E_nE_{n-1}^c$ ,
3. ha piovuto ieri ma non oggi, ovvero l'evento  $E_n^cE_{n-1}$ ,
4. non ha piovuto nè ieri nè oggi, ovvero l'evento  $E_n^cE_{n-1}^c$ .

In tale rappresentazione degli stati il sistema forma una Catena di Markov con la seguente matrice di transizione

$$\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0.7 & 0 & 0.3 & 0 \\ 2 & 0.5 & 0 & 0.5 & 0 \\ 3 & 0 & 0.4 & 0 & 0.6 \\ 4 & 0 & 0.2 & 0 & 0.8 \end{array} \Bigg\|$$

Ad esempio,  $p_{21} = P(E_{n+1}E_n|E_nE_{n-1}^c)$  indica la probabilità che piova sia domani sia oggi (stato 1) condizionata al fatto che abbia piovuto oggi ma non ieri (stato 2), mentre  $p_{11} = P(E_{n+1}E_n|E_nE_{n-1})$  indica la probabilità che piova sia domani sia oggi (stato 1) condizionata al fatto che abbia piovuto sia oggi sia ieri (stato 1).

**Esercizio.** Valutando  $a_0(1) = P(E_1E_0) = 0.1$  calcolare la probabilità di  $E_3^cE_2E_1E_0$ . Si ha

$$\begin{aligned} P(E_3^cE_2E_1E_0) &= P(E_3^cE_2|E_2E_1E_0)P(E_2E_1|E_1E_0)P(E_1E_0) = \\ &= P(X_2 = 3|X_1 = 1)P(X_1 = 1|X_0 = 1)P(E_1E_0) = \\ &= p_{13}p_{11}a_0(1) = 0.5 \cdot 0.7 \cdot 0.1 = 0.035. \end{aligned}$$

**Esempio 8** (Passeggiata aleatoria). (vedi [9]) Consideriamo una particella che si muove su una retta con salti discreti ed in cui i punti sono  $0, \pm 1, \pm 2, \dots$ . Supponiamo che il moto della particella sia il seguente: se la particella è nella posizione  $i$  al tempo  $n$ , allora al tempo  $n + 1$  essa si sposterà nella posizione  $i + 1$  con probabilità  $p_i$ , nella posizione  $i - 1$  con probabilità  $q_i$  o rimarrà nella posizione  $i$  con probabilità  $r_i = 1 - p_i - q_i$ . Se indichiamo con  $X_n$  la posizione della particella al tempo  $n$ , allora  $\{X_n, n = 0, 1, 2, \dots\}$  con spazio degli stati  $S = \mathbb{Z}$  forma una catena di Markov con la seguente matrice di transizione

$$\begin{array}{cccccccc} & \dots & -2 & -1 & 0 & 1 & 2 & \dots \\ \vdots & \left\| \begin{array}{cccccccc} \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ \dots & r_{-2} & p_{-2} & 0 & 0 & 0 & 0 & \dots \\ \dots & q_{-1} & r_{-1} & p_{-1} & 0 & 0 & 0 & \dots \\ \dots & 0 & q_0 & r_0 & p_0 & 0 & 0 & \dots \\ \dots & 0 & 0 & q_1 & r_1 & p_1 & 0 & \dots \\ \dots & 0 & 0 & 0 & q_2 & r_2 & p_2 & \dots \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \end{array} \right\| & \dots \\ -2 & & & & & & & \\ -1 & & & & & & & \\ 0 & & & & & & & \\ 1 & & & & & & & \\ 2 & & & & & & & \\ \vdots & & & & & & & \end{array}$$

Analizziamo alcuni casi particolari

- In particolare se  $r_i = 0$ ,  $p_i = p$  e  $q_i = 1 - p$  per ogni  $i \in S$ , allora  $\{X_n, n = 0, 1, 2, \dots\}$  viene detta passeggiata aleatoria semplice.
- Se  $q_0 = 0$  e  $P(X_0 \geq 0) = 1$ , allora la passeggiata aleatoria ha spazio degli stati  $S = \{0, 1, 2, \dots\}$ .
- Se  $q_0 = 0$  e  $r_0 = 0$  (e quindi  $p_0 = 1$ ) allora,  $\{X_n, n = 0, 1, 2, \dots\}$  si dirà avere una *barriera riflettente* in 0.
- Se  $q_0 = 0$  e  $r_0 = 1$  (e quindi  $p_0 = 0$ ) allora,  $\{X_n, n = 0, 1, 2, \dots\}$  si dirà avere una *barriera assorbente* in 0.
- Se  $q_0 = 0$ ,  $p_M = 0$  e  $P(0 \leq X_n \leq M) = 1$ , allora la passeggiata aleatoria  $\{X_n, n \geq 0\}$  avrà spazio degli stati  $S = \{0, 1, 2, \dots, M\}$ . La barriera in  $M$  si dirà riflettente o assorbente a seconda che  $q_M = 1$  o  $q_M = 0$ .
- Le catene di nascita e morte ([2]) sono un ulteriore caso particolare.

**Esempio 9** (Passeggiata aleatoria semplice). Una catena di Markov in cui lo spazio degli stati è dato da  $S = \{0, \pm 1, \pm 2, \dots\}$  viene detta Passeggiata aleatoria semplice, se per qualche fissato  $0 < p < 1$ , si ha

$$p_{i,i+1} = p = 1 - p_{i,i-1}, \quad i = 0, \pm 1, \pm 2, \dots$$

Pertanto, la matrice (infinita) di transizione è la seguente

$$\begin{array}{c} \vdots \\ -2 \\ -1 \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} \left\| \begin{array}{cccccc} \dots & -2 & -1 & 0 & 1 & 2 & \dots \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ \dots & 0 & p & 0 & 0 & 0 & \dots \\ \dots & 1-p & 0 & p & 0 & 0 & \dots \\ \dots & 0 & 1-p & 0 & p & 0 & \dots \\ \dots & 0 & 0 & 1-p & 0 & p & \dots \\ \dots & 0 & 0 & 0 & 1-p & 0 & \dots \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \end{array} \right\|$$

**Esempio 10** (Rovina di un giocatore). Un giocatore ad ogni giocata guadagna 1 euro con probabilità  $p$  e perde un euro con probabilità  $1-p$ . Supponiamo che le partite siano stocasticamente indipendenti e che il gioco finisce o quando il capitale del giocatore arriva a 0 o a  $M$ . In tal caso se indichiamo con  $X_n$  il capitale del giocatore al tempo  $n$  allora  $\{X_n, n \geq 0\}$  forma una catena di Markov con la seguente matrice di transizione.

$$\begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ M \end{array} \left\| \begin{array}{cccccc} 0 & 1 & 2 & \dots & M-1 & M \\ 1 & p & 0 & \dots & 0 & 0 \\ 1-p & 0 & p & \dots & 0 & 0 \\ 0 & 1-p & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \end{array} \right\|$$

Ovvero,

$$p_{00} = p_{MM} = 1, \quad p_{i,i+1} = p = 1 - p_{i,i-1}, \quad i = 1, 2, \dots, M-1.$$

Quindi  $\{X_n, n \geq 0\}$  è una passeggiata aleatoria su  $S = \{0, 1, 2, \dots, M\}$  con barriere assorbenti in 0 e in  $M$ .

Una situazione più divertente che porta alla precedente catena di Markov è quella di un ubriaco che con probabilità  $q$  fa un passo verso casa (posizione 0) e con probabilità  $p = 1 - q$  fa un passo verso il bar (posizione  $M$ ). Non appena egli raggiunge la casa o il bar vi rimane per sempre. Questo tipo di random walk viene detta *drunkard's walk*.

**Esempio 11** (Bonus Malus System). Consideriamo un automobilista che ogni anno paga un'assicurazione di tipo bonus-malus. Supponiamo che ci siano 3 classi di sconto. Classe 2 con 50% di sconto, classe 1 con 25% di sconto e classe 0 con 0% di sconto. Ogni anno l'automobilista sale di una classe se non ha fatto incidenti in quell'anno (eccetto il caso in cui si trova già nella classe 2) oppure scende nella classe 0 nel caso in cui abbia fatto almeno un incidente (eccetto il caso in cui si trova già nella classe 0). Supponiamo che la probabilità di fare almeno un incidente in un anno sia  $\pi$  indipendentemente dagli incidenti compiuti negli anni precedenti e indipendentemente dalla classe

in cui l'automobilista si trova. Se indichiamo con  $X_n$  la classe di appartenenza dell'automobilista nell'anno  $n$ , allora  $\{X_n, n \geq 0\}$  forma una catena di Markov con spazio degli stati  $S = \{0, 1, 2\}$  e matrice di transizione  $P$  è la seguente

$$\begin{array}{c} \phantom{0} \phantom{\left\| \right.} \phantom{\phantom{0}} \phantom{\phantom{1}} \phantom{\phantom{2}} \\ 0 \left\| \begin{array}{ccc} \pi & 1 - \pi & 0 \\ \pi & 0 & 1 - \pi \\ \pi & 0 & 1 - \pi \end{array} \right\| \\ 1 \\ 2 \end{array}$$

**Esercizio 3.** Relativamente all'Esempio 11, data l'assegnazione  $P(X_0 = 0) = 1$  calcolare la probabilità  $P(X_n = j)$ , per  $j = 0, 1, 2$ .

Osserviamo che si ha

$$P^2 = \begin{bmatrix} \pi & (1-\pi)\pi & (1-\pi)^2 \\ \pi & (1-\pi)\pi & (1-\pi)^2 \\ \pi & (1-\pi)\pi & (1-\pi)^2 \end{bmatrix}.$$

Per ogni  $n \geq 0$  poniamo  $\mathbf{a}^{(n)} = (P(X_n = j), j = 0, 1, 2)$ . Si ha  $\mathbf{a}^{(0)} = (1, 0, 0)$ ,  $\mathbf{a}^{(1)} = (\pi, 1 - \pi, 0)$  e

$$\mathbf{a}^{(n)} = (\pi, (1 - \pi)\pi, (1 - \pi)^2), \quad \forall n \geq 2.$$

Se, per esempio,  $\pi = 0.1$ , si ha  $\mathbf{a}^{(1)} = (0.1, 0.9, 0)$  e

$$\mathbf{a}^{(n)} = (0.1, 0.09, 0.81), \quad \forall n \geq 2.$$

**Esempio 12** (Assicurazione sulla vita). Consideriamo il seguente spazio degli stati  $S = \{H, S, D\}$  ( $H$ =healthy (in salute),  $S$ =sick (malato),  $D$ =dead (morto)). Indicando con  $X_n$  lo stato di salute di un individuo nell'anno  $n$ , se supponiamo (irrealisticamente) che  $\{X_n, n \geq 0\}$  formi una catena di Markov omogenea (quindi non dipendente dall'età), allora la matrice di transizione è data da

$$\begin{array}{c} H \\ S \\ D \end{array} \left\| \begin{array}{ccc} H & S & D \\ p_{HH} & p_{HS} & p_{HD} \\ p_{SH} & p_{SS} & p_{SD} \\ 0 & 0 & 1 \end{array} \right\|$$

Osserviamo che poichè  $p_{DD} = 1$ ,  $D$  è uno stato assorbente.

**Esempio 13** (Brand Switching). Consideriamo un cliente che acquista ad intervalli regolari (ad esempio una volta a settimana) un prodotto e ad ogni acquisto egli può scegliere tra tre marchi (brands) differenti:  $M_1, M_2$  ed  $M_3$ . Supponiamo che la scelta del prodotto alla  $n + 1$ -esima settimana dipenda solo dall'acquisto effettuato alla  $n$ -esima settimana, senza tener conto dei precedenti acquisti. Se indichiamo con  $X_n$  il marchio acquistato alla  $n$ -esima settimana allora,  $\{X_n, n \geq 0\}$  costituisce una catena di Markov con spazio  $S = \{M_1, M_2, M_3\}$  e una matrice di transizione potrebbe essere la seguente

$$\begin{array}{c} M_1 \\ M_2 \\ M_3 \end{array} \left\| \begin{array}{ccc} M_1 & M_2 & M_3 \\ 0.1 & 0.2 & 0.7 \\ 0.2 & 0.4 & 0.4 \\ 0.1 & 0.3 & 0.6 \end{array} \right\|$$

Questo tipo di catena viene utilizzato nel campo commerciale per predire gli acquisti.

**Esempio 14.** Per altri esempi interessanti consultare [9]. In particolare l'esempio di genetica a pag.30 sez. 2.1 e di sociologia (branching process) a pag.34 sez. 2.2.

### 3.1.2 Equazione di Chapman-Kolmogoroff.

Nella sezione precedente abbiamo definito la probabilità di transizione  $p_{ij}$  ad un passo. Data una catena di Markov omogenea  $\{X_n, n \geq 0\}$  con spazio degli stati  $S$  definiamo *probabilità di transizione  $p_{ij}^{(n)}$  ad  $n$  passi* la probabilità che il processo passi dallo stato  $i$  allo stato  $j$  dopo  $n$  passi, ovvero la probabilità che il processo si trovi nello stato  $j$ , dato che si trovava allo stato  $i$ , nell' $n$ -esimo istante precedente. In formule abbiamo

$$p_{ij}^{(n)} = P(X_{n+m} = j | X_m = i), \quad \forall m$$

Poichè la catena di Markov in considerazione è omogenea,  $p_{ij}^{(n)}$  non dipende da  $m$ , pertanto possiamo scrivere

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i). \quad (3.7)$$

In particolare la (3.7) per  $n = 1$  ci dà le probabilità di transizione ad un passo, ovvero

$$p_{ij}^{(1)} = p_{ij}.$$

**Teorema 3** (Chapman-Kolmogorov). Le probabilità di transizione ad  $n$  passi  $p_{ij}^{(n)}$  soddisfano le seguenti equazioni

$$p_{ij}^{(n+m)} = \sum_{r \in S} p_{ir}^{(n)} p_{rj}^{(m)}, \quad (3.8)$$

$$\forall n, m \in \mathbb{N}, \quad \forall i, j \in S,$$

o in maniera equivalente, le seguenti equazioni

$$p_{ij}^{(n)} = \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)} \quad (3.9)$$

$$\forall n \in \mathbb{N}, \forall k \in \{1, \dots, n\}, \quad \forall i, j \in S.$$

*Dimostrazione.* Fissato  $k \in \{1, \dots, n\}$  si ha

$$\begin{aligned} p_{ij}^{(n)} &= P(X_n = j | X_0 = i) = P[X_n = j, \bigvee_{r \in S} (X_k = r) | X_0 = i] = \\ (\sigma \text{ additività}) &= \sum_{r \in S} P(X_n = j, X_k = r | X_0 = i) = \\ &= \sum_{r \in S} P(X_n = j | X_k = r, X_0 = i) P(X_k = r | X_0 = i) = \\ (\text{Markov}) &= \sum_{r \in S} P(X_n = j | X_k = r) P(X_k = r | X_0 = i) = \\ (\text{omog.}) &= \sum_{r \in S} P(X_{n-k} = j | X_0 = r) P(X_k = r | X_0 = i) = \\ &= \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)} \end{aligned}$$

□

Le equazioni (3.9) o (3.8) sono chiamate equazioni di *Chapman-Kolmogorov*. Le suddette equazioni ci dicono che la probabilità che la catena vada dallo stato

$i$  allo stato  $j$  in  $n$ -passi è data dalla somma, al variare di  $r$ , dei prodotti ottenuti moltiplicando la probabilità di andare dallo stato  $i$  allo stato intermedio  $r$  nei primi  $k$  passi per la probabilità di andare dallo stato  $r$  allo stato  $n$  nei successivi  $n - k$  passi.

Le probabilità  $p_{ij}^{(n)}$  formano una matrice detta matrice di transizione ad  $n$  passi ed indicata come segue

$$P^{(n)} = [p_{ij}^{(n)}].$$

Per  $n = 1$  si ha  $P^{(1)} = P$ . Indichiamo con  $\pi^{(n)} = (\pi_j^{(n)}, j \in S)$  la distribuzione di probabilità di  $X_n$ , ovvero

$$\pi_j^{(n)} = P(X_n = j), \quad \forall j \in S.$$

Attenzione, per  $n = 0$  il vettore  $\pi^{(0)}$  ci dà le probabilità iniziali  $\mathbf{a} = (a_j, j \in S)$ . L'utilità di tale notazione matriciale si evince dal teorema seguente che ci fornisce un metodo per calcolare la matrice di transizione ad  $n$  passi a partire dalla matrice di transizione ad un passo.

**Teorema 4.** Si ha

$$P^{(n)} = P^n, \quad (3.10)$$

dove  $P^n$  rappresenta la potenza  $n$ -esima di  $P$ .

*Dimostrazione.* Per induzione. Dalle equazioni (3.9) si ha per  $n = 2$

$$P^{(2)} = P^{(1)} \cdot P^{(1)} = P \cdot P = P^2.$$

Supponiamo che  $P^{(k)} = P^k$  per  $k = 1, \dots, n - 1$  e dimostriamo che  $P^{(n)} = P^n$ . Dalle equazioni 3.9 si ha

$$\begin{aligned} P^{(n)} &= P^{(n-1)} P^{(1)} = \\ \text{dall'ipotesi induttiva} &= P^{n-1} P = P^n. \end{aligned}$$

□

**Corollario 1.** Si ha

$$\pi^{(n)} = \pi^{(0)} \cdot P^{(n)} = \mathbf{a} \cdot P^n. \quad (3.11)$$

Ovvero, in altri termini

$$\pi_j^{(n)} = P(X_n = j) = \sum_{i \in S} P(X_n = j | X_0 = i) P(X_0 = i) = \sum_{i \in S} p_{ij}^{(n)} a_i.$$

**Esempio 15.** Riprendiamo l'Esempio 6 (pioggia o non pioggia) in cui  $\alpha = 0.4$  e  $\beta = 0.3$ . Calcoliamo la matrice delle probabilità di transizione dopo 2 passi. Si ha

$$P = \begin{array}{c|cc} & 0 & 1 \\ 0 & 0.7 & 0.3 \\ 1 & 0.4 & 0.6 \end{array} \parallel$$

Quindi  $P^{(2)} = P^2$ , con

$$P^2 = \begin{matrix} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \left\| \begin{matrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{matrix} \right\| \end{matrix}$$

Pertanto se si vuole calcolare la probabilità che piova tra due giorni dato che piove oggi, essa è pari a  $p_{11}^{(2)} = 0.48$ . Se supponiamo che le probabilità iniziali siano  $P(X_0 = 0) = 0.6$  e  $P(X_0 = 1) = 0.4$  allora la probabilità che piova tra due giorni sarà data da

$$\begin{aligned} P(X_2 = 1) &= P(X_2 = 1|X_0 = 0)P(X_0 = 0) + P(X_2 = 1|X_0 = 1)P(X_0 = 1) \\ &= P(X_0 = 0)p_{01}^{(2)} + P(X_0 = 1)p_{11}^{(2)} = \\ &= 0.6 \cdot 0.39 + 0.4 \cdot 0.48 = 0.426. \quad (3.12) \end{aligned}$$

Calcolare la probabilità (non condizionata) che piova il 4<sup>o</sup> giorno, la probabilità (non condizionata) che nei primi 4 giorni piova almeno 1 volta.

### 3.2 Catena di Markov a 2 stati

In questa sezione analizzeremo in dettaglio una catena di Markov a 2 stati e riprenderemo in considerazione l'esempio 6 (Vedi [4]). Possiamo vedere uno stato come *successo* (denotato con 1) e uno come *insuccesso* denotato con 0. Pertanto abbiamo un esempio di prove Bernoulliane dipendenti nelle quali la probabilità di successo in ogni prova dipende dal risultato della prova precedente (e solo da quella). Indicheremo con  $\alpha$  la probabilità di successo alla  $n+1$ -esima prova supposto che alla  $n$ -esima si sia avuto insuccesso e con  $\beta$  la probabilità di insuccesso alla  $n+1$ -esima prova supposto che alla  $n$ -esima si sia avuto successo. Pertanto se indichiamo con  $X_n$  il risultato aleatorio della  $n$ -esima prova allora  $\{X_n, n \geq 0\}$  con  $S = \{0, 1\}$  forma una catena di Markov con la seguente matrice  $P$  di transizione

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-\alpha & \alpha \\ \hline 1 & \beta & 1-\beta \end{array}$$

Le probabilità  $\alpha, \beta$  rappresentano le probabilità di cambiamento di stato mentre  $1-\alpha$  e  $1-\beta$  quelle di permanenza. Escludiamo in questa trattazione i casi in cui  $\alpha + \beta = 0$  e  $\alpha + \beta = 2$ . Infatti, nel caso  $\alpha + \beta = 0$ , cioè  $\alpha = 0, \beta = 0$  il sistema rimane con probabilità 1 nel suo stato iniziale. Nel caso  $\alpha + \beta = 2$ , cioè  $\alpha = 1, \beta = 1$  il sistema cambia alternativamente stato con probabilità 1.

**Esempio 16** (Rainfall in Tel Aviv, riformulazione dell'esempio 6). In un classico articolo Gabriel e Neumann (1962) proposero di utilizzare una catena di Markov per descrivere l'alternanza di giorni di pioggia e asciutti nei mesi di Dicembre Gennaio e Febbraio a Tel Aviv. La catena ha quindi due soli stati: piove o asciutto, e i tempi si riferiscono ai giorni. Attribuiamo il valore zero allo stato giorno asciutto e il valore 1 allo stato giorno di pioggia. Gabriel e Neumann usarono le frequenze relative in 27 anni come valutazioni delle probabilità di transizione. La tabella delle frequenze in 2473 giorni era

	Asciutto	Pioggia	<i>Totale</i>
Asciutto	1049	350	1399
Pioggia	351	687	1038

Pertanto, la matrice di transizione è la seguente

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0.750 & 0.250 \\ \hline 1 & 0.338 & 0.662 \end{array}$$

Consideriamo la matrice  $P^{(5)}$  ottenuta dal prodotto  $P^2 \cdot P^3$ ,

$$P^5 = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0.580 & 0.420 \\ \hline 1 & 0.568 & 0.432 \end{array}.$$

Essa tra le altre cose ci permette di calcolare la probabilità che piova il 6 Febbraio supposto che il 1 Febbraio non abbia piovuto (0.420).

Sia  $\pi^{(n)} = (\pi_0^{(n)}, \pi_1^{(n)})$  il vettore delle probabilità di trovare il sistema nello stato 0 o 1 al tempo  $n$  e  $\pi^{(0)} = (\pi_0^{(0)}, \pi_1^{(0)})$  il vettore delle probabilità iniziali del sistema. Consideriamo gli eventi  $X_n = 0$  e  $X_n = 1$ . Dal Corollario 1 si ha

$$\pi^{(n)} = \pi^{(0)} \cdot P^{(n)} \quad (3.13)$$

Con un ragionamento analogo si ha pure

$$\pi^{(n)} = \pi^{(n-1)} \cdot P \quad (3.14)$$

Ci si chiede se in un periodo sufficientemente lungo di tempo il sistema possa evolvere in una situazione statistica di equilibrio in cui le probabilità  $\pi^{(n)}$  siano indipendenti dalle probabilità iniziali  $\pi^{(0)}$ . Se ciò accade (e non è sempre detto) allora ci sarà una distribuzione di probabilità di equilibrio  $\pi = (\pi_0, \pi_1)$  la quale, facendo tendere  $n \rightarrow \infty$  nella 3.14, dovrà soddisfare la seguente equazione

$$\pi = \pi \cdot P \quad (3.15)$$

ovvero

$$\pi \cdot (I - P) = 0 \quad (3.16)$$

In tal caso dalla (3.13) si ha che la matrice  $P^n$  per  $n \rightarrow \infty$  diventa altrettanto stazionaria.

L'equazione matriciale (3.16) equivale al seguente sistema omogeneo

$$(\mathbf{S}) \begin{cases} \pi_0 \alpha - \pi_1 \beta & = 0 \\ -\pi_0 \alpha + \pi_1 \beta & = 0 \end{cases} \quad (3.17)$$

Tale sistema (essendo omogeneo) avrà soluzioni in  $\pi$  diverse da quella banale se e solo se il determinante  $|I - P|$  è nullo. In tal caso imponendo  $\pi_0 + \pi_1 = 1$  si ottiene un' unica soluzione. In particolare si ottiene

$$\pi_0 = \frac{\beta}{\alpha + \beta}, \quad \pi_1 = \frac{\alpha}{\alpha + \beta} \quad (3.18)$$

In particolare, osserviamo che se la distribuzione iniziale  $\pi^{(0)}$  è  $\pi$  allora, si ha

$$\begin{aligned} \pi^{(1)} &= \pi^{(0)} \cdot P = \pi \cdot P = \pi, \\ \pi^{(2)} &= \pi^{(1)} \cdot P = \pi \cdot P = \pi, \\ &\vdots \\ \pi^{(n)} &= \pi \cdot P = \pi, \\ &\vdots \end{aligned}$$

Cioè, se  $\pi^{(0)} = \pi$  la distribuzione  $\pi^{(n)}$  è stazionaria nel tempo.

**Esempio 17** (segue Rainfall in Tel Aviv). Dalla matrice  $P$  di tale esempio, si ricava che

$$\pi_0 = 0.575, \quad \pi_1 = 0.425.$$

Infatti, calcolando  $P^{10}$ , essa differisce da  $P^5$  solo nella seconda cifra decimale,

$$P^{10} = \begin{matrix} & & 0 & 1 \\ 0 & \left\| \begin{matrix} 0.575 & 0.425 \\ 0.575 & 0.425 \end{matrix} \right\| \\ 1 & & & \end{matrix}.$$

Poichè le due righe sono (quasi) uguali moltiplicando  $P^{10}$  per un qualsiasi vettore  $\pi^{(0)} = (\pi_0^{(0)}, \pi_1^{(0)})$  con  $\pi_0^{(0)} + \pi_1^{(0)} = 1$  si ha

$$\begin{aligned} (\pi_0^{(0)}, \pi_1^{(0)}) \cdot \left\| \begin{matrix} 0.575 & 0.425 \\ 0.575 & 0.425 \end{matrix} \right\| &= (0.575(\pi_0^{(0)} + \pi_1^{(0)}), 0.425(\pi_0^{(0)} + \pi_1^{(0)})) \\ &= (0.575, 0.425). \end{aligned} \quad (3.19)$$

Pertanto, se chiamiamo con 0 il 31 Dicembre e con 10 il 10 Gennaio, qualsiasi sia la distribuzione iniziale  $\pi^{(0)}$  al tempo 0, dalla equazione (3.13) si ottiene

$$\pi^{(10)} = (0.575, 0.425).$$

Illustriamo adesso un metodo che permette di calcolare le potenze della matrice  $P$  in maniera semplice. Calcoliamo gli autovalori  $\lambda_1, \lambda_2$  della matrice  $P$ . Ricordiamo che gli autovalori di una matrice  $P$  sono le soluzioni della seguente equazione (nel determinante)

$$|P - \lambda I| = 0.$$

Si ottiene  $\lambda_1 = 1$  e  $\lambda_2 = 1 - \alpha - \beta$ . Avendo supposto  $\alpha + \beta \neq 0$  si ha  $\lambda_1 \neq \lambda_2$ . Inoltre dalla teoria algebrica si ha che la matrice  $P$  è diagonalizzabile, ovvero esiste una matrice  $Q$  invertibile tale che

$$P = Q \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q^{-1}. \quad (3.20)$$

Le colonne  $q_i$  di  $Q$  sono le soluzioni del sistema

$$Pq_i = \lambda_i q_i$$

e rappresentano gli autovettori associati all'autovalore  $\lambda_i$ . Dalla (3.20) si ottiene la seguente formula

$$P^n = Q \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} Q^{-1}, \quad (3.21)$$

che permette di calcolare facilmente la matrice  $P^n$ . La matrice  $Q$  si ottiene come soluzione della seguente equazione

$$P \cdot Q = Q \cdot D \quad (3.22)$$

dove con  $D$  abbiamo indicato la matrice diagonale

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

. L'equazione 3.22 si può risolvere manualmente (siamo nel caso di matrici  $2 \times 2$ ) oppure tramite opportuni software (quali  $R^1$ , *Matlab*, *Scilab*<sup>2</sup>). Una delle soluzioni che si ottiene è

$$Q = \begin{bmatrix} 1 & \alpha \\ 1 & -\beta \end{bmatrix}, \quad Q^{-1} = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \alpha \\ 1 & -1 \end{bmatrix}.$$

Infine dalla (3.21), dopo opportune semplificazioni si ha

$$P^n = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{(1 - \alpha - \beta)^n}{\alpha + \beta} \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix}. \quad (3.23)$$

Inoltre, osserviamo che la (3.23) oltre ad essere utile per il calcolo della matrice  $P$  ci mostra che per  $n \rightarrow \infty$  (poichè abbiamo supposto  $\alpha + \beta \neq 0, 2$ ) si ha

$$P^n \rightarrow \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} = \begin{bmatrix} \pi_0 & \pi_1 \\ \pi_0 & \pi_1 \end{bmatrix}.$$

Infine possiamo dire che in tal caso esiste la distribuzione limite  $\pi$  ed essa è indipendente dalle condizioni iniziali rappresentate da  $\pi^{(0)}$ , infatti si ha

$$\pi^{(n)} \rightarrow \pi^{(0)} \begin{bmatrix} \pi_0 & \pi_1 \\ \pi_0 & \pi_1 \end{bmatrix} = (\pi_0, \pi_1).$$

**Esempio 18** (segue Rainfall in Tel Aviv). In questo esempio la rappresentazione (3.23) di  $P^n$  diviene

$$P^n = \begin{bmatrix} 0.575 & 0.425 \\ 0.575 & 0.425 \end{bmatrix} + (0.412)^n \begin{bmatrix} 0.425 & -0.425 \\ -0.575 & 0.575 \end{bmatrix}. \quad (3.24)$$

Il fattore  $(0.412)^n$  tende a 0 al crescere di  $n$ .

. Consideriamo adesso la quantità  $X_1 + X_2 + \dots + X_n$  che rappresenta il numero aleatorio di successi in  $n$  passi e ne calcoliamo la previsione condizionata a  $X_0 = i$ , si ha (ricordiamo che la previsione di una variabile bernoulliana è la probabilità di successo, cioè se  $X \in \{0, 1\}$ , si ha  $\mathbb{E}(X) = P(X = 1)$ )

$$\begin{aligned} \mathbb{E}(X_1 + X_2 + \dots + X_n | X_0 = i) &= \sum_{k=1}^n \mathbb{E}(X_k | X_0 = i) = \\ &= \sum_{k=1}^n P(X_k = 1 | X_0 = i) = \sum_{k=1}^n p_{i1}^{(k)}. \end{aligned}$$

Pertanto, considerando la previsione della frequenza relativa, si ha

$$\mathbb{E}\left(\frac{X_1 + X_2 + \dots + X_n}{n} | X_0 = i\right) = \frac{1}{n} \sum_{k=1}^n p_{i1}^{(k)}.$$

Dall'analisi è noto che se una successione  $\{a_n\}$  è convergente ad  $a$  allora anche la successione delle medie aritmetiche

$$\left\{ \frac{a_1 + a_2 + \dots + a_n}{n} \right\}$$

<sup>1</sup><http://cran.r-project.org/>

<sup>2</sup><http://www.scilab.org/>

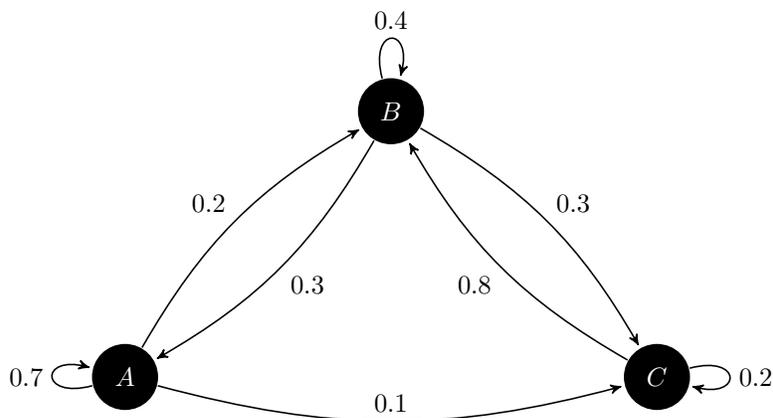


Figura 3.1: Diagramma degli stati di una catena di Markov.

converge ad  $a$ . Quindi, sapendo che  $p_{i1}^{(n)}$  converge a  $\pi_1$  segue che la previsione della frequenza relativa converge a  $\pi_1$ , ovvero

$$\mathbb{E}\left(\frac{X_1 + X_2 + \dots + X_n}{n} \mid X_0 = i\right) \rightarrow \pi_1.$$

In maniera analoga se introduciamo  $Y_n = 1 - X_n$  si ha

$$\mathbb{E}\left(\frac{Y_1 + Y_2 + \dots + Y_n}{n} \mid Y_0 = i\right) \rightarrow \pi_0.$$

Va notato inoltre che gli ultimi due limiti sono indipendenti dallo stato iniziale.

Indichiamo con  $f_{00}^{(n)}$  la probabilità di ottenere per la prima volta insuccesso al tempo  $n$  supposto che ci sia stato insuccesso al tempo 0, cioè

$$f_{00}^{(n)} = P(X_n = 0, X_{n-1} = 1, \dots, X_1 = 1 \mid X_0 = 0).$$

Si ha, per  $n = 1$

$$f_{00}^{(1)} = P(X_1 = 0 \mid X_0 = 0) = 1 - \alpha$$

e per  $n \geq 2$

$$f_{00}^{(n)} = \alpha\beta(1 - \beta)^{n-2}.$$

Indicando con  $T$  la variabile aleatoria primo insuccesso dopo  $n$  passi calcolare la previsione condizionata di  $T \mid X_0 = 0$ .

**Esercizio 4.** Costruire la matrice di transizione corrispondente al seguente diagramma degli stati.

**Esercizio 5.** Disegnare il diagramma degli stati relativo alla seguente matrice di transizione

$$P = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0.25 & 0.15 & 0.2 & 0.4 \\ 2 & 0 & 0.5 & 0.5 & 0 \\ 3 & 0 & 0 & 1 & 0 \\ 4 & 0.3 & 0.4 & 0.1 & 0.2 \end{array}$$

**Esercizio 6** (Bonus Malus System2). Consideriamo un automobilista che ogni anno paga un'assicurazione di tipo bonus-malus. Supponiamo che ci siano 4 classi di rischio  $C_1, C_2, C_3, C_4$  con premi rispettivamente €250,300,350,450.

- Nel primo periodo l'automobilista è inserito nella classe  $C_4$ .
- Ogni anno l'automobilista scende di una classe se non ha fatto incidenti in quell'anno; sale di una classe se ha fatto un incidente; sale di due classi se ha fatto due incidenti; viene inserito nella classe  $C_4$  se ha fatto tre o più incidenti.
- Supponiamo che la variabile aleatoria  $Z_n$  di incidenti nell' $n$ -esimo periodo abbia distribuzione di Poisson di parametro  $\lambda = \frac{2}{3}$  e che le  $Z_n$  siano stocasticamente indipendenti.
- Indicando con  $X_n$  la classe di rischio dell'automobilista nel periodo  $n + 1$ -esimo ( $X_0$  è il primo anno) determinare la corrispondente matrice di transizione. Inoltre, calcolare il valore atteso del premio nel quarto periodo.
- Scrivere l'equazione ricorsiva per  $X_n$  nella forma  $X_n = \phi(X_{n-1}, Z_n), n \geq 1$ .

### 3.3 Classificazione degli stati

In questa sezione introdurremo il concetto di classe comunicante e di catena irriducibile.

**Definizione 4.** Uno stato  $j$  si dice *raggiungibile* dallo stato  $i$  se per qualche  $n \geq 0$  si ha  $p_{ij}^{(n)} > 0$ .

Se lo stato  $j$  è raggiungibile da  $i$  scriveremo  $i \rightarrow j$ . Osserviamo che uno stato  $j$  è raggiungibile da uno stato  $i$  se e solo se nel diagramma di transizione degli stati esiste un percorso che va da  $i$  a  $j$ . Infatti la condizione  $p_{ij}^{(n)} > 0$  è equivalente ad avere una sequenza  $i_0, i_1, \dots, i_n$  tale che  $i_0 = i$ ,  $i_n = j$  e  $p_{i_k, i_{k+1}} > 0$  per  $k = 0, 1, \dots, n-1$ . Ciò è equivalente ad avere il percorso

$$i = i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_n = j$$

Per convenzione poniamo  $p_{ii}^{(0)} = P(X_0 = i | X_0 = i) = 1$ , pertanto si ha  $i \rightarrow i$  per ogni stato  $i \in S$ . Osserviamo che la definizione di raggiungibilità non dipende da quando sia grande  $p_{ij}^{(n)} > 0$ , ma solamente dal fatto che esso sia o no zero.

**Definizione 5.** Diciamo che gli stati  $i, j$  *comunicano* se  $i \rightarrow j$  e  $j \rightarrow i$ . Cioè se essi sono raggiungibili vicendevolmente ( $p_{ij}^{(n)} > 0$  per qualche  $n$  e  $p_{ji}^{(m)} > 0$  per qualche  $m$ .)

La relazione di comunicazione è una relazione di equivalenza, poichè essa è riflessiva, simmetrica e transitiva. Infatti si ha

- (i)  $i \leftrightarrow i$  (riflessività),
- (ii)  $i \leftrightarrow j \Leftrightarrow j \leftrightarrow i$  (simmetria),
- (iii)  $i \leftrightarrow j, j \leftrightarrow k \Rightarrow i \leftrightarrow k$  (transitività) .

Le proprietà (i) e (ii) sono immediate dalla definizione. Proviamo la (iii). Le relazioni  $i \rightarrow j, j \rightarrow k$  implicano che  $p_{ij}^{(n)} > 0$  per qualche  $n \geq 0$  e  $p_{jk}^{(m)} > 0$  per qualche  $m \geq 0$ . Pertanto dall'equazione di Chapman-Kolmogorov si ha

$$p_{ik}^{(n+m)} = \sum_{r \in S} p_{ir}^{(n)} p_{rk}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0.$$

Pertanto si ha  $i \rightarrow k$ . In maniera analoga si prova che  $k \rightarrow i$ . Quindi si ha  $i \leftrightarrow k$ .

**Esempio 19.** Nella passeggiata aleatoria con spazio degli stati  $S = \{0, 1, \dots, M\}$  avente barriere assorbenti in 0 e in  $M$ , lo stato 0 e lo stato  $M$  comunicano solo con loro stessi. Lo stato 0 non è raggiungibile dallo stato  $M$  e lo stato  $M$  non è raggiungibile dallo stato 0. Gli stati  $1, \dots, M-1$  comunicano tra di loro, possono raggiungere gli stati 0 e  $M$ , ma non possono essere raggiunti da 0 e  $M$ .

Due stati che comunicano tra di loro si dicono appartenere alla stessa classe (di equivalenza).

**Definizione 6.** Una classe  $C \subseteq S$  si dice *comunicante* se

1.  $i \in C, j \in C \Rightarrow i \leftrightarrow j$
2.  $i \in C, i \leftrightarrow j \Rightarrow j \in C$

La 1 assicura che due qualsiasi stati nella classe comunicano tra di loro. La 2 impone alla classe di essere massimale, cioè  $C$  non può essere contenuta in una ulteriore classe comunicante. Osserviamo che può esistere uno stato  $j$  non in  $C$  raggiungibile da qualche stato  $i$ . Cioè possono esistere  $j \notin C$  e  $i \in C$  tali che  $i \rightarrow j$ , e ovviamente non può essere  $j \rightarrow i$  altrimenti si avrebbe  $j \in C$ . In maniera analoga possono esistere  $j \notin C$  e  $i \in C$  tali che  $j \rightarrow i$ , ma ovviamente non può essere  $i \rightarrow j$ . Per tale motivo si considera la seguente

**Definizione 7.** Una classe  $C \subset S$  si dice *chiusa* se ogni stato fuori da  $C$  non è raggiungibile da alcun stato in  $C$ . Cioè  $p_{ij}^{(n)} = 0, n \geq 0$ , per ogni  $i \in C$  e  $j \notin C$ .

**Definizione 8.** Una classe  $C \subset S$  comunicante e chiusa si dice *irriducibile*.

Osserviamo che se una catena di Markov visita una classe comunicante e chiusa (irriducibile)  $C$ , allora essa non uscirà più da  $C$ . Cioè

$$X_n \in C \Rightarrow X_m \in C, \forall m > n.$$

Due classi comunicanti devono essere disgiunte. Quindi nasce una partizione di  $S$  (potenzialmente infinita anche se la rappresentiamo finita)

$$S = C_1 \cup C_2 \dots C_k \cup T$$

dove  $C_i$  sono classi irriducibili e  $T$  è l'unione delle altre classi comunicanti.

**Definizione 9.** Una DTCM si dice *irriducibile* se l'intero spazio degli stati è una classe irriducibile, cioè  $S = C$ . Altrimenti si dice *riducibile*.

Ovviamente in una DTCM irriducibile tutti gli stati comunicano tra di loro.

**Esempio 20.** Sia  $S = \{0, 1\}$  e

$$P^{10} = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0.2 & 0.8 \\ 1 & 0.3 & 0.7 \end{array}.$$

La classe  $C = S$  è irriducibile.

**Esempio 21.** Sia  $S = \{0, 1\}$  e

$$P^{10} = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 0.7 & 0.3 \end{array}.$$

Si ha  $0 \leftrightarrow 0, 1 \rightarrow 0$  e  $1 \leftrightarrow 1$ , ma  $0 \not\leftrightarrow 1$ . Pertanto  $C_1 = \{0\}$  è irriducibile, mentre  $T = \{1\}$  non lo è.

vedi Ex. 3.5 pag 74 [9]

**Definizione 10.** Uno stato da solo che forma una classe irriducibile è detto stato *assorbente*.

**Definizione 11.** Una DTCM che ammette uno stato assorbente raggiungibile dagli altri stati non necessariamente in un solo passo si dice *assorbente*.

### 3.3.1 Stati ricorrenti e transitori

In una DTCM per ogni  $i \in S$  poniamo

$$T_i = \min\{n > 0 : X_n = i\}.$$

La variabile aleatoria  $T_i$  dicesi *tempo di primo passaggio* della catena nello stato  $i$ . Va osservato che  $T_i$  può assumere valori interi positivi oppure il valore  $+\infty$ . Si ha

$$P(T_i = n) = P(X_n = i, X_k \neq i, k = 1, \dots, n-1)$$

Pertanto, per calcolare la distribuzione di probabilità del tempo di primo passaggio occorre conoscere oltre alla matrice di transizione  $P$  anche la distribuzione iniziale. Consideriamo adesso la variabile aleatoria condizionata  $T_i|X_0 = i$ . Poniamo

$$f_i = P(T_i < \infty | X_0 = i) = \sum_{n=1}^{\infty} P(T_i = n | X_0 = i).$$

Quando  $f_i = 1$ , la variabile  $T_i|X_0 = i$  dicesi propria o finita con probabilità 1. Quando  $f_i < 1$ , dicesi impropria e si pone

$$P(T_i = \infty | X_0 = i) = P(X_n \neq i, \forall n > 0) = 1 - f_i > 0.$$

Poniamo

$$m_i = \mathbb{E}(T_i | X_0 = i).$$

Quindi  $m_i$ , quando è definito, è il tempo medio di primo passaggio, o *tempo medio di ritorno*. Osserviamo che se  $f_i < 1$  si ha necessariamente  $m_i = \infty$ , cioè il tempo di primo passaggio ha valor medio infinito. Comunque, quando  $f_i = 1$ ,  $m_i$  potrebbe anche essere infinito.

**Esempio 22.** Se

$$P(T_i = n | X_0 = i) = \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}, \quad n = 1, 2, \dots$$

allora si ha

$$f_i = \sum_{n=1}^{\infty} P(T_i = n | X_0 = i) = 1$$

, ma

$$m_i = \mathbb{E}(T_i | X_0 = i) = \sum_{n=1}^{\infty} \frac{n}{n(n+1)} = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty$$

**Definizione 12.** Uno stato  $i$  si dice

- *ricorrente* se  $f_i = 1$ ,
- *transitorio* se  $f_i < 1$ .

Inoltre uno stato ricorrente può essere classificato a seconda del valore finito o infinito del valore atteso del tempo di primo passaggio  $m_i$ .

$$\text{ricorrente } (f_i = 1) \begin{cases} \text{ricorrente positivo,} & \text{se } m_i \text{ è finito} \\ \text{ricorrente nullo,} & \text{se } m_i = \infty. \end{cases}$$

Poichè  $f_i$  denota la probabilità che una catena, partendo dallo stato  $i$ , prima o poi vi ritorni, allora una catena di Markov che parte da uno stato ricorrente vi ritornerà con probabilità 1. Pertanto, per la definizione di catena di Markov, essa ripartirà di nuovo da tale stato per ritornarvi di nuovo. Continuando in questo modo si evince che se lo stato  $i$  è ricorrente allora, partendo da  $i$ , la catena rivisiterà  $i$  infinite volte. D'altro canto, se  $i$  è transitorio, partendo da  $i$ , la catena ha una probabilità non nulla  $1 - f_i$  di non ritornarvi. Quindi, partendo da  $i$  transitorio, la probabilità che la catena visita lo stato  $i$  esattamente  $n$  volte è  $f_i^{n-1}(1 - f_i)$ . Indichiamo con  $(N_i|X_0 = i)$  il numero aleatorio di visite, partendo da  $i$ , che fa la catena nello stato  $i$  (compreso quello iniziale). Si ha che se  $i$  è transitorio  $(N_i|X_0 = i) \in \{1, 2, \dots\}$  ha una distribuzione geometrica modificata di parametro  $1 - f_i$  e valore atteso  $\frac{1}{1-f_i}$ . In particolare si ha

- $P(N_i = 1|X_0 = i) = (1 - f_i)$  prob. che la catena visiti  $i$  solo nello stato iniziale,
- $P(N_i = 2|X_0 = i) = f_i(1 - f_i)$  prob. che la catena visiti  $i$  esattamente 2 volte
- $\vdots$
- $P(N_i = n|X_0 = i) = f_i^{n-1}(1 - f_i)$  prob. che la catena visiti  $i$  esattamente  $n$  volte

Pertanto se  $i$  è transitorio abbiamo

$$\mathbb{E}(N_i|X_0 = i) = \frac{1}{1 - f_i}$$

Conseguentemente, la catena di Markov, partendo da uno stato transitorio lo visiterà mediamente  $\frac{1}{1-f_i}$ . Se invece  $i$  è ricorrente si ha

$$\mathbb{E}(N_i|X_0 = i) = \infty$$

**Osservazione 3.** La variabile aleatoria  $(N_i|X_0 = i)$  conta il numero di visite, compreso il primo, e non il numero di volte che la catena rivisita tale stato. Quest'ultimo potrebbe definirsi come  $(N_i - 1|X_0 = i)$ . Inoltre, va osservato che la variabile aleatoria  $(N_i|X_0 = i)$  è condizionata all'ipotesi che la catena parta dallo stato  $i$ . Se la catena non parte dallo stato  $i$ , allora potrebbe accadere che uno stato ricorrente non sia mai visitato

**Esempio 23.** Sia

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}.$$

Entrambi gli stati sono ricorrenti ed essi formano due classi irriducibili. Pertanto, se la catena parte da 0 essa vi rimane per sempre.

Inoltre si ha

$$\begin{aligned} \mathbb{E}(N_i|X_0 = i) &= \mathbb{E}\left(\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n = i\}} \mid X_0 = i\right) = \\ &= \sum_{n=0}^{\infty} \mathbb{E}\left(\mathbb{1}_{\{X_n = i\}} \mid X_0 = i\right) \\ &= \sum_{n=0}^{\infty} P(X_n = i \mid X_0 = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)}. \end{aligned} \quad (3.25)$$

Pertanto, segue che

**Teorema 5.** Lo stato  $i$  è ricorrente se e solo se  $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$ .

Lo stato  $i$  è transitorio se e solo se  $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1-f_i} < \infty$ .

**Osservazione 4.** Va notato che in una catena di Markov finita non tutti gli stati possono essere transitori, cioè deve esistere almeno uno stato ricorrente. Infatti, supponiamo che  $S = \{1, \dots, M\}$  e supponiamo che tutti gli stati sono transitori. Allora, dopo un tempo finito  $T_1$  la catena non visiterà lo stato 1, dopo un tempo  $T_2$  la catena non visiterà lo stato 2 ecc. Quindi, dopo un tempo finito  $T = \max\{T_1, T_2, \dots, T_M\}$  la catena non visiterà nessuno stato, che è un assurdo.

Un importante corollario del Teorema 5 è il seguente

**Corollario 2.**

- (i) Se lo stato  $i$  è ricorrente e comunica con lo stato  $j$ , allora anche  $j$  è ricorrente. Ovvero

$$i \leftrightarrow j, i \text{ ricorrente} \Rightarrow j \text{ ricorrente.}$$

- (ii) se lo stato  $i$  è transitorio e comunica con lo stato  $j$ , allora anche  $j$  è transitorio. Ovvero,

$$i \leftrightarrow j, i \text{ transitorio} \Rightarrow j \text{ transitorio.}$$

*Dimostrazione (i).* Dal fatto che  $i \leftrightarrow j$  segue che esistono  $m, n$  tali che  $p_{ij}^{(m)} > 0$  e  $p_{ji}^{(n)} > 0$ , pertanto dall'equazione di Chapman-Kolmogoroff si ha che per ogni intero  $k$

$$p_{jj}^{(n+k+m)} \geq p_{ji}^{(n)} p_{ii}^{(k)} p_{ij}^{(m)}.$$

Ciò accade perchè,  $p_{jj}^{(n+k+m)}$  è la probabilità di andare da  $j$  a  $j$  in  $n+k+m$  passi, mentre  $p_{ji}^{(n)} p_{ii}^{(k)} p_{ij}^{(m)}$  è la probabilità di andare da  $j$  a  $j$  in  $n+k+m$  passi **attraverso** un percorso (tra tutti i possibili) che va da  $j$  a  $i$  in  $n$  passi, quindi da  $i$  a  $i$  in ulteriori  $k$  passi e da  $i$  a  $j$  in ulteriori  $m$  passi. Dalla precedente formula segue

$$\sum_{k=0}^{\infty} p_{jj}^{(n+k+m)} \geq \sum_{k=0}^{\infty} p_{ji}^{(n)} p_{ii}^{(k)} p_{ij}^{(m)} = \underbrace{p_{ji}^{(n)} p_{ij}^{(m)}}_{>0} \underbrace{\sum_{k=0}^{\infty} p_{ii}^{(k)}}_{=\infty} = \infty. \quad (3.26)$$

Quindi, dal Teorema 5 si ha che  $j$  è ricorrente.  $\square$

*Dimostrazione (ii).* Sia  $i$  transitorio e  $i \leftrightarrow j$ . Se fosse  $j$  ricorrente allora per la parte (i) lo stato  $i$  sarebbe ricorrente, che è un assurdo. Quindi  $j$  è transitorio.  $\square$

Inoltre si ha (non lo dimostriamo)

### Corollario 3.

- (i) Se lo stato  $i$  è ricorrente positivo e comunica con lo stato  $j$ , allora anche  $j$  è ricorrente positivo
- (ii) se lo stato  $i$  è ricorrente nullo e comunica con lo stato  $j$ , allora anche  $j$  è ricorrente nullo.

Pertanto, ne segue un importante proprietà delle classi di comunicazione.

**Proprietà 1.** Tutti gli stati di una classe di comunicazione sono dello stesso tipo (o tutti ricorrenti (tutti positivi o tutti nulli) o tutti transitori).

La precedente proprietà ci conduce alla seguenti definizioni.

**Definizione 13.** Una classe di comunicazione  $C$  dicesi *ricorrente* (nulla o positiva) se tutti gli stati in essa sono ricorrenti (nulli o positivi).

**Definizione 14.** Una classe di comunicazione  $C$  dicesi *transitoria* se tutti gli stati in essa sono transitori.

In particolare si ha che in una catena di Markov irriducibile tutti gli stati sono o ricorrenti o transitori.

Inoltre, per le catene di Markov finite e irriducibili vale il seguente

**Corollario 4.** In una catena di Markov finita e irriducibile tutti gli stati sono ricorrenti (positivi).

*Dimostrazione.* Poichè la catena è finita dall'Osservazione 4 segue che deve esistere almeno uno stato ricorrente, inoltre essendo la catena irriducibile lo spazio degli stati forma un'unica classe di comunicazione. Quindi, dalla Proprietà 1 segue che tutti gli stati devono essere ricorrenti. Che tali stati siano positivi non lo dimostriamo.  $\square$

Un analogo teorema vale per le classi irriducibili finite.

**Teorema 6.** Sia  $C$  una classe di comunicazione chiusa finita. Allora, tutti gli stati in  $C$  sono ricorrenti (positivi).

**Teorema 7.** Sia  $C$  una classe di comunicazione finita e non chiusa. Allora, tutti gli stati in  $C$  sono transitori.

Pertanto ne segue che per le catene di Markov finite la seguente partizione in classi di

$$S = C_1 \cup C_2 \dots C_k \cup T$$

comporta che le classi  $C_k$  (irriducibili) sono le classi di comunicazione ricorrenti e  $T$  rappresenta l'unione delle classi di comunicazione transitorie. Quando il sistema parte da uno stato transitorio in  $T$  allora il sistema (essendo lo spazio degli stati finito) prima o poi deve arrivare in una classe irriducibile e vi rimane per sempre. Se invece la catena fosse infinita allora partendo da uno stato transitorio in  $T$  la catena potrebbe anche rimanere per sempre in  $T$ , potendo  $T$  essere infinito.

Per le catene con un numero finito di stati abbiamo un criterio (che non dimostriamo) che ci permette di decidere se uno stato è transitorio senza dover studiare il comportamento della serie delle probabilità di transizione.

**Teorema 8.** Per una catena di Markov con spazio degli stati finiti uno stato  $i$  è transitorio se e solo se esiste uno stato  $j$  tale che  $j$  è raggiungibile da  $i$  ma  $i$  non è raggiungibile da  $j$ , in formule

$$i \text{ è transitorio} \Leftrightarrow \exists j : i \rightarrow j, j \nrightarrow i.$$

**Esempio 24.** Nel problema della rovina del giocatore con spazio  $S = \{0, 1, \dots, M\}$ , gli stati 0 e 1 sono ricorrenti e assorbenti, infatti  $p_{00}^{(n)} = 1$  e  $p_{MM}^{(n)} = 1$ . Tutti gli altri stati per il teorema precedente sono transitori. Pertanto, con probabilità 1 la catena partendo da qualsiasi stati prima o poi arriverà a 0 o a  $M$ . Inoltre quella che abbiamo chiamato  $F_a$ , ovvero le probabilità che un giocatore con capitale iniziale  $a$  prima o poi si rovini, rappresenta la probabilità di assorbimento nella classe  $[0]$  partendo dallo stato  $a$ , mentre  $H_a$ , ovvero la probabilità che un giocatore con capitale  $a$  possa sbancare il banco, rappresenta la probabilità che la probabilità di assorbimento della catena nella classe  $[M]$  partendo dallo stato 0.

### 3.4 Distribuzioni invarianti o stazionarie

Lo studio del comportamento a lungo termine del sistema descritto da una catena di Markov è importante per due motivi. In primo luogo perchè si cerca di determinare una distribuzione “di equilibrio” che descriva lo stato del sistema dopo molto tempo. In questo caso l’interesse sta nel capire se il sistema ammette una tale distribuzione di equilibrio. In secondo luogo perchè tale distribuzione è una buona approssimazione del comportamento del sistema anche per istanti grandi ma finiti ed è di più immediata comprensione che non la legge del sistema all’istante fissato. In questa sezione troveremo la distribuzione invariante (i cui sinonimi sono distribuzione stazionaria o distribuzione di equilibrio) per catene di Markov con struttura particolare che andremo tra poco a definire. Diamo invece ora la definizione di distribuzione invariante. Supponiamo data una catena di Markov con spazio degli stati  $S$ , finito o infinito e matrice di transizione  $P$ .

**Definizione 15.** Una distribuzione  $\pi$  su  $S$  è detta distribuzione invariante (o stazionaria o di equilibrio) se soddisfa la relazione

$$\pi = \pi P.$$

L’esistenza di una distribuzione invariante implica che il comportamento del sistema descritto dalla catena sia molto stabile. Inoltre, se la distribuzione iniziale della catena di Markov è una distribuzione invariante, cioè se  $\pi^{(0)} = \pi$ , allora la distribuzione all’istante  $n$  è sempre  $\pi$  per ogni  $n$ , ovvero

$$\pi^{(0)} = \pi \Rightarrow \pi^{(n)} = \pi \quad \forall n > 0.$$

Tale risultato è stato ampiamente illustrato nella Sezione 3.2 relativa alle catene di Markov a 2 stati, ovvero nel caso in cui  $S = \{1, 2\}$  e la matrice  $P$  ha tutti gli elementi positivi. A priori non è detto che la distribuzione invariante esista, infatti la soluzione del sistema

$$\pi = \pi P$$

è soggetta ai vincoli

$$\pi_i \geq 0, \quad \sum_{i \in S} \pi_i = 1.$$

Inoltre se esiste una tale distribuzione, non è detto che questa sia unica. Ad esempio, nel problema della rovina del giocatore, le distribuzioni  $\pi^* = (1, 0, \dots, 0)$  e  $\pi^{**} = (0, 0, \dots, 1)$  sono entrambe invarianti. Inoltre se considero una loro combinazione lineare convessa, ovvero

$$\lambda \pi^* + (1 - \lambda) \pi^{**}$$

si ha

$$\pi^* = \pi^* P$$

$$\pi^{**} = \pi^{**} P$$

$$\lambda \pi^* + (1 - \lambda) \pi^{**} = \lambda \pi^* P + (1 - \lambda) \pi^{**} P = [\lambda \pi^* + (1 - \lambda) \pi^{**}] P$$

che è invariante. Pertanto, se esistono due distribuzioni invarianti, allora ne esistono infinite. In tal caso non ha senso parlare di comportamento limite. Restringiamo il nostro studio al caso in cui esiste un'unica distribuzione invariante.

A tal proposito è necessario introdurre il concetto di periodicità di uno stato ([10], pag. 204, 9th ed.).

**Definizione 16** (Periodicità). Uno stato  $i$  per il quale esiste un  $n$  tale che  $p_{ii}^{(n)} > 0$  ha periodo  $d$  se  $p_{ii}^{(k)} = 0$  quando  $k$  non è divisibile per  $d$ ,  $p_{ii}^{(k)} \geq 0$  (maggiore o uguale a zero) quando  $k$  è divisibile per  $d$  e  $d$  è il più grande intero con questa proprietà.

Osserviamo che se uno stato  $i$  ha periodo  $d_i$  allora  $p_{ii}^{(k)}$  può essere positivo solo se  $k$  è multiplo di  $d_i$ , ma  $p_{ii}^{(k)}$  non necessariamente deve essere positivo se  $k$  è multiplo di  $d_i$ . Una definizione analoga è la seguente ([2], pag.234; [5], pag.261)

**Definizione 17** (Periodicità). Consideriamo l'insieme

$$R_i = \{n \geq 1, p_{ii}^{(n)} > 0\}$$

formato dagli istanti  $n$  per i quali la probabilità di tornare in  $i$  in  $n$  passi sia strettamente positiva. Indichiamo con  $d_i$  il massimo comune divisore dei numeri che stanno in  $R_i$ , ovvero

$$d_i = MCD(R_i).$$

In altre parole il periodo  $d_i$  di uno stato  $i$  è il massimo comune divisore degli stati  $n$  per i quali  $p_{ii}^{(n)}$  è positiva. Si può provare ([5]) che  $p_{ii}^{(n)}$  è **definitivamente** positiva se e solo se  $n$  è multiplo di  $d_i$ .

**Definizione 18** (Aperiodicità). Uno stato  $i$  si dice aperiodico se  $d = 1$ .

Per esempio se partendo da  $i$  è possibile ritornare in  $i$  solo dopo  $2, 4, 6, \dots$  passi allora  $i$  ha periodo 2 (basti pensare alla passeggiata aleatoria con barriere riflettenti). Vedi inoltre Esempio 26.

Si dimostra che la periodicità è una proprietà di classe. Cioè se uno stato di una classe di comunicazione ha periodo  $d$  allora tutti gli stati della classe hanno periodo  $d$ . In particolare in una catena di Markov irriducibile tutti gli stati hanno lo stesso periodo e quindi si può parlare di periodo della catena.

**Esempio 25.** Data una catena di Markov con il diagramma degli stati di Figura 3.2 osserviamo che si può ritornare nello stato 1 solo dopo  $n = 8, 10, 16, 18, 20$  passi e per ogni numero pari  $n \geq 24$ . Allora lo stato 1 ha periodo  $d_1 = 2$ , infatti  $MCD\{8, 10, 16, 18, 20, 24 + 2k, k \in \mathbb{N}_0\} = 2$ . Osserviamo che in questo caso  $p_{11}^n$  non è positiva per ogni numero pari  $n$ , infatti si ha  $p_{11}^{(k)} = 0$  per  $k = 2, 4, 6, 12, 14, 22$ .

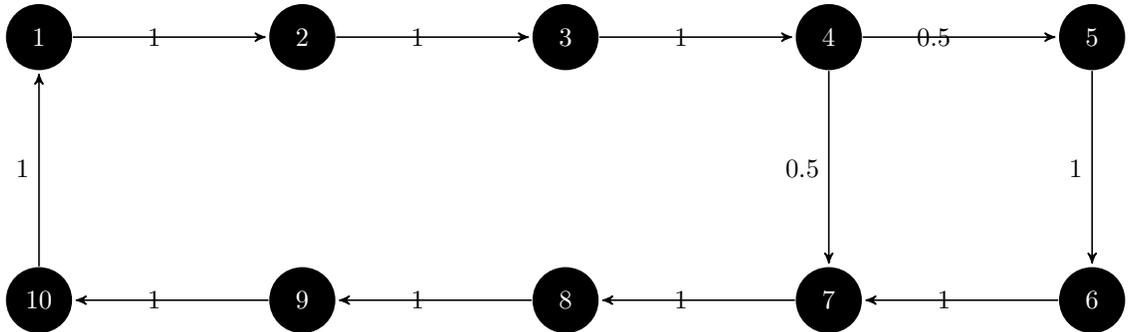


Figura 3.2: Diagramma degli stati.

**Esercizio 7.** Data la seguente catena di Markov

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \left\| \begin{matrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{matrix} \right\| \end{matrix}$$

stabilire se è irriducibile. Calcolare la periodicità di ogni stato.

Il diagramma degli stati di tale catena riportato in Figura 3.3 mostra che

$$S = C \cup T$$

dove  $C = \{1, 2, 3\}$  è una classe irriducibile di stati ricorrenti positivi e  $T = \{4, 5\}$  una classe di stati transitori. Inoltre per  $i \in C$  si ha  $p_{ii}^{(3k)} = 1 > 0$  per ogni intero  $k$  e  $p_{ii}^{(n)} = 0$  per ogni  $n$  non multiplo di 3. Pertanto  $d_i = 3$  se  $i \in C$ . Per gli stati transitori il periodo non è definito.

Ricordiamo che uno stato  $i$  è ricorrente positivo se il tempo medio di ritorno in  $i$  è finito mentre esso è ricorrente nullo se tale tempo è infinito. Si dimostra che nel caso di catene di Markov finite, se un stato è ricorrente allora esso è ricorrente positivo.

**Definizione 19 (Catena Ergodica).** Una catena di Markov con stati ricorrenti positivi (cio con tempo medio di ritorno  $m$  finito) e aperiodici ( $d = 1$ ) dicesi *ergodica*.

Si ha, il seguente importante teorema del quale non diamo la dimostrazione

**Teorema 9.** Per una catena di Markov irriducibile ed ergodica (cioè ricorrente positiva e aperiodica) esiste un'unica distribuzione stazionaria  $\pi$  tale che

$$\begin{aligned} \pi &= \pi P, \\ \pi_i &\geq 0, \quad \sum_{i \in S} \pi_i = 1. \end{aligned}$$

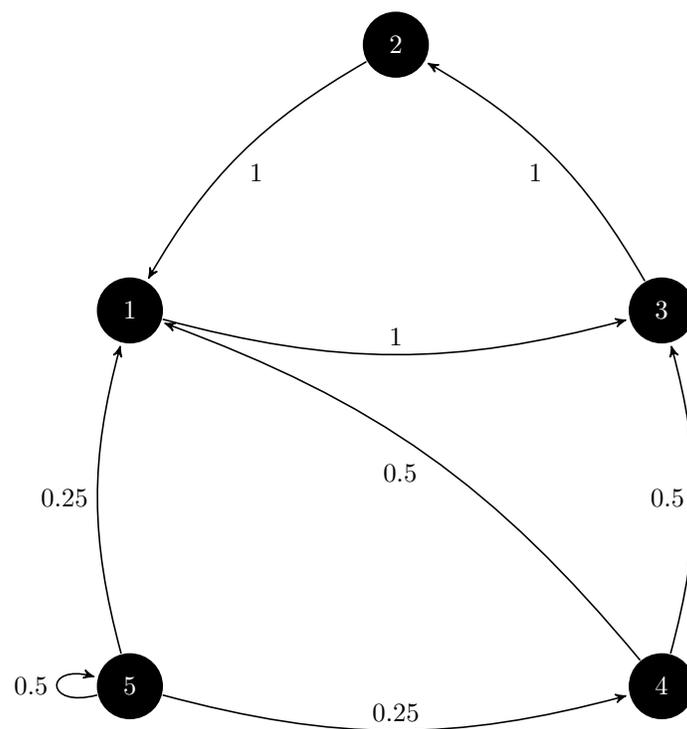


Figura 3.3: Diagramma degli stati.

e inoltre si ha

$$\exists \lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j = \frac{1}{m_j} > 0, \quad \forall i, j \in S,$$

cioè il  $\lim_{n \rightarrow \infty} p_{ij}^{(n)} > 0$  esiste, è indipendente da  $i$ , coincide con  $\pi_j$  ed è l'inverso del tempo medio di ritorno  $m_j$ .

### 3.4.1 Catene di Markov Finite

(vedi [4] par. 3.10-11, vedi Perron-Frobenius(appunti) ) Sia  $A$  una matrice (finita) con elementi non negativi. Se tramite delle opportune permutazioni delle righe e delle colonne di  $A$  (nel caso di matrice di transizione ciò significa rinominare gli stati) si ottiene una matrice della forma

$$A^* = \left\| \begin{array}{cc} B & \mathbf{0} \\ C & D \end{array} \right\|. \quad (3.27)$$

dove  $B$  è una matrice quadrata e  $\mathbf{0}$  indica una matrice di elementi nulli, allora  $A$  dicesi *riducibile*. Altrimenti  $A$  dicesi *irriducibile*. In particolare una matrice  $A > 0$  (i cui elementi sono tutti positivi) è irriducibile .

Notiamo che ciò corrisponde al concetto di irriducibilità per una catena di Markov. Per una catena finita la matrice di transizione, rinominando opportunamente gli stati, ha la forma (3.27) quando gli stati corrispondenti a  $B$  formano una classe chiusa. Va osservato inoltre che in una matrice della forma (3.27) la sottomatrice  $\mathbf{0}$  si preserva in tutte le potenze della matrice. Si dice che una matrice quadrata  $A$  di elementi non negativi è *primitiva* se ammette l'autovalore 1 con molteplicità 1 (semplice), ed esso è il più grande autovalore in modulo. Osserviamo che una matrice stocastica di transizione  $P$  ammette sempre l'autovalore 1. Cioè, se indichiamo con  $\mathbf{1}$  il vettore colonna con tutti uno, si ha

$$P \cdot \mathbf{1} = \mathbf{1}.$$

Però, non è sempre detto che l'autovalore 1 sia semplice. Si dimostra il seguente teorema

**Teorema 10.** Una catena di Markov finita è ergodica (cioè ricorrente positiva e aperiodica) se e solo se la sua matrice di transizione è primitiva e irriducibile, cioè  $P$  ha l'autovalore semplice 1 che è massimo in modulo e inoltre  $P$  non si può scrivere nella forma (3.27). Inoltre in tal caso, se indichiamo con  $S = \{1, 2, \dots, s\}$ , si ha ([4] pag.124 par. 3.11)

$$\exists \lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j = \frac{1}{m_j} > 0$$

$$\lim_{n \rightarrow \infty} P^n = \left\| \begin{array}{cccc} \pi_1 & \pi_2 & \dots & \pi_s \\ \vdots & \vdots & \dots & \vdots \\ \pi_1 & \pi_2 & \dots & \pi_s \end{array} \right\| = \mathbf{1}\pi.$$

**Osservazione 5.** Ricordiamo che  $m_j$  rappresenta il tempo medio di ritorno. Pertanto, si ha che la probabilità limite e stazionaria di  $p_{ij}$  coincide con l'inverso del tempo medio di primo passaggio  $m_j$ . Come, a voler dire la frequenza (intesa come probabilità) è l'inverso del periodo. In altri termini, qualunque sia la distribuzione di probabilità iniziale, il valor medio della percentuale di tempo spesa dal processo nello stato  $j$  è asintoticamente uguale a  $\pi_j$ .

Quindi si ha che  $\pi$  è una distribuzione invariante, ovvero

$$\pi = \pi P.$$

Ricordiamo che data una matrice  $A$  quadrata essa è diagonalizzabile se si può scrivere nella forma

$$A = QDQ^{-1}$$

con  $D$  matrice diagonale degli autovalori. Inoltre dato un autovalore  $\lambda_i$  allora la colonna  $q_i$  di  $Q$  rappresenta l'autovettore associato a  $\lambda_i$ , quindi la soluzione dell'equazione

$$Aq_i = \lambda_i q_i.$$

Nel nostro caso abbiamo che  $P$  ammette l'autovalore 1. Inoltre si ha

$$\pi = \pi P$$

scrivendo tale equazione nella forma trasposta, ricordando che  $(AB)' = (B'A')$  ove  $A'$  indica la matrice trasposta di  $A$ , si ha

$$P'\pi' = \pi'.$$

Cioè  $\pi^t$  è un autovettore associato all'autovalore 1 della matrice trasposta  $P^t$ .

Detto questo, un algoritmo utile per calcolare la distribuzione stazionaria consiste nel calcolare un autovettore  $\pi^*$  associato all'autovalore 1 della matrice trasposta  $P^t$ , quindi normalizzarlo (somma 1), per ottenere la distribuzione stazionaria.

Il seguente codice definisce una funzione in  $R$  che traduce l'algoritmo precedente.

**Algoritmo 1** (Calcolo della distribuzione invariante per catene di Markov finite ed ergodiche).

```

invar <- function(P){
  A<-eigen(t(P))
  #scelgo l'indice relativo all'a.valore 1 (il max)
  indice <- which(A$values==max(A$values))
  # normalizzo l'a.vettore corrispondente
  A$vector[,indice]/sum(A$vector[,indice])
}

```

```

P <- matrix(c(0.45,0.48,0.07,0.05,0.7,

```

```

0.25,0.01,0.5,0.49),3,3, byrow=TRUE)
invar(P)
[1] 0.06238859 0.62344029 0.31417112

```

**Definizione 20** (Catena regolare). Una DTCM è detta *regolare* se esiste un indice  $\bar{n}$  tale che  $P^{(n)} > 0$  per ogni  $n > \bar{n}$ . Ovvero, la matrice di transizione è una matrice regolare.

Per una catena regolare tutti gli stati comunicano, pertanto essa è irriducibile, il viceversa non vale sempre, infatti

**Esempio 26.** Consideriamo la catena associata alla matrice

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

osserviamo che essa è irriducibile ma non regolare. Infatti si ha

$$P^{2n-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, n = 1, 2, \dots$$

e

$$P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, n = 1, 2, \dots$$

In realtà, ciò accade perchè  $P$  ha periodo pari a 2.

In altri termini una catena regolare, nel caso finito, sarà una catena

- irriducibile (per quanto visto prima)
- ricorrente positiva, poichè essendo finita in cui tutti gli stati comunicano allora tutti gli stati devono essere ricorrenti positivi.
- aperiodica, altrimenti non si avrebbe  $P^n > 0$  per qualche  $n$ .

**Pertanto, per una catena finita regolare esiste sempre la distribuzione limite e valgono i teoremi per le catene ergodiche irriducibili.**

Una condizione sufficiente per verificare la regolarità di una catena è la seguente

**Osservazione 6** (Criterio di regolarità ([2] pag. 204)). Se una catena finita è irriducibile ed esiste un  $h \in S$  tale che  $p_{hh} > 0$  allora la catena è regolare.

Pertanto se tutti gli stati comunicano ed esiste un elemento positivo nella diagonale della matrice di transizione allora la catena è regolare.

## Capitolo 4

# Processi di Rinnovo e Processi di Poisson

### 4.1 Processo dei rinnovi e processo dei conteggi

Prerequisiti: Funzione caratteristica, distribuzione esponenziale, gamma, beta, poisson, binomiale negativa, binomiale, geometrica.

#### 4.1.1 Processo dei rinnovi

Sia  $\{X_n\}$  una famiglia di v.a. non negative indipendenti ed identicamente distribuite. Poniamo

$$\begin{aligned}T_1 &= X_1 \\T_2 &= X_1 + X_2 \\&\dots \\T_n &= X_1 + X_2 + \dots + X_n \\&\dots\end{aligned}$$

Se, ad esempio, le  $X_n$  rappresentano i tempi di vita di particolari dispositivi (ovvero i tempi di inter-arrivo), allora le  $T_n$  rappresentano i tempi di attesa sino all'  $n$ -esimo rinnovo (ovvero sino all' $n$ -esimo arrivo). Ad esempio  $T_1 = X_1$  indica il tempo di attesa sino al primo rinnovo,  $T_2$  il tempo di attesa sino al secondo ecc. In altre parole l' $n$ -esimo rinnovo si avrà al tempo  $T_n$ . Il processo a parametro discreto  $\{T_n, n \in \mathbb{N}\}$  dicesi *processo dei rinnovi*.

#### 4.1.2 Processo dei conteggi

Un processo discreto a parametro continuo  $\{N_t, t \geq 0\}$  dicesi *processo di conteggio* se

- (i)  $N_0 = 0$  (per convenzione);
- (ii)  $N_t \in \{0, 1, 2, \dots\}$ ;

(iii) se  $t_1 < t_2$  allora  $N_{t_1} \leq N_{t_2}$ .

### 4.1.3 Dal processo dei rinnovi a quello di conteggio

Strettamente legato al processo dei rinnovi presentiamo il processo dei conteggi, che a differenza del primo, è un processo discreto a parametro continuo. Considerato un intervallo di tempo  $(0, t]$ ,  $t > 0$  e indichiamo con  $N_t$  il numero dei rinnovi che cadono in questo intervallo. In simboli

$$N_t = \sup\{n : T_n \leq t\}.$$

Il processo  $\{N_t, t \geq 0\}$  è un processo di conteggio.

L'evento  $(N_t = 0)$  significa che sino all'istante  $t$  non si è effettuato alcun rinnovo e pertanto coincide con l'evento  $(T_1 > t)$ . In generale l'evento  $(N_t = n)$ , cioè sino all'istante  $t$  si sono effettuati  $n$  rinnovi, coincide con l'evento  $(T_n \leq t) \wedge (T_{n+1} > t)$ , cioè l' $n$ -esimo rinnovo si è effettuato in  $(0, t]$  e l' $n + 1$ -esimo in un istante maggiore di  $t$ . Quindi si ha

$$(N_t = n) = (T_n \leq t) \wedge (T_{n+1} > t) = (T_n \leq t < T_{n+1})$$

Inoltre si ha

$$\begin{aligned} (T_{n+1} > t) &= [(T_{n+1} > t) \wedge (T_n \leq t)] \vee [(T_{n+1} > t) \wedge (T_n > t)] = \\ &= (N_t = n) \vee (T_n > t). \end{aligned}$$

Infatti,  $(T_n > t)$  implica  $(T_{n+1} > t)$ , cioè se al tempo  $t$  non si sono avuti  $n$  rinnovi, allora al tempo  $t$  non si avranno  $n + 1$  rinnovi. Pertanto, si ha

$$\begin{aligned} P(N_t = n) &= P(T_{n+1} > t) - P(T_n > t) = S_{T_{n+1}}(t) - S_{T_n}(t) = \\ &= F_{T_n}(t) - F_{T_{n+1}}(t) \end{aligned}$$

dove  $S_{T_n}$  e  $F_{T_n}$  sono, rispettivamente, la funzione di sopravvivenza ( $S_{T_n} = 1 - F_{T_n}$ ) e la funzione di ripartizione di  $T_n$ .

### 4.1.4 Dal processo di conteggio a quello dei rinnovi

Sia  $\{N_t, t \geq 0\}$  un processo di conteggio, ad esempio per ogni  $t$  positivo sia  $N_t$  il numero di arrivi (dei rinnovi, dei guasti ecc.) nell'intervallo  $(0, t]$ . Per ogni intero positivo  $n$  sia  $T_n$  il tempo di attesa sino al  $n$ -esimo arrivo e sia  $X_n = T_n - T_{n-1}$  il tempo di interarrivo. Se le  $X_n, n \in \mathbb{N}$ , sono variabili aleatorie indipendenti e identicamente distribuite allora si dice che  $\{T_n, n \in \mathbb{N}\}$  è un **processo dei rinnovi**.

**Esempio 27.** Data una successione di eventi  $E_1, E_2, \dots, E_n, \dots$ , indipendenti ed equiprobabili, con  $P(E_i) = p, q = 1 - p$ . Diremo “successo alla  $i$ -esima prova” se si verifica l'evento  $E_i$ . Introduciamo la seguente successione di variabili aleatorie

- $X_1 =$  “ numero di prove sino al primo successo” ;

- $X_2 =$  “numero di prove dopo il primo successo sino al secondo successo”;
- ...
- $X_n =$  “numero di prove dopo l’ $(n - 1)$ -esimo successo sino all’ $n$ -esimo successo”.
- ...

Si ha che  $\{X_n\}_{n \in \mathbb{N}}$  è una successione di variabili aleatorie indipendenti con distribuzione geometrica di parametro  $p$ . Cioè,

$$X_n \in \mathbb{N}, \\ P(X_n = k) = pq^{k-1}, k = 1, 2, \dots$$

La variabile aleatoria  $T_n = \sum_{i=1}^n X_i$  rappresenta il numero di prove sino all’ $n$ -esimo successo. Si ha che  $T_n \in \{n, n + 1, \dots\}$  ed ha distribuzione di Pascal, ovvero

$$P(T_n = k) = \binom{n-1}{k-1} p^k q^{n-k}, k = n, n + 1, \dots$$

Consideriamo la variabile aleatoria  $N_k$  del numero di successi nelle  $k$ -esime prove, in tal caso si ha ovviamente che  $N_k$  ha distribuzione  $Bin(k, p)$ . Concludendo si ha che  $\{T_n\}_{n \in \mathbb{N}}$  è il processo dei rinnovi ed  $\{N_k\}_{k \in \mathbb{N}}$  è il processo dei conteggi. In tal caso entrambi i processi sono discreti a parametri discreti.

**Esercizio 8.** Sia  $X_1, X_2, \dots$  una successione di variabili aleatorie Bernoulliane di parametro  $p$  e indipendenti. Calcolare le distribuzioni di probabilità delle variabili  $T_n = \sum_{k=1}^n X_k$  e  $N_k = \sup\{n, T_n \leq k\}$ .

#### 4.1.5 Processo di conteggio a incrementi indipendenti e stazionari

- Un processo di conteggio  $\{N_t, t \geq 0\}$  ha **incrementi indipendenti** se il numero degli arrivi che si verificano in intervalli di tempo disgiunti sono indipendenti. In particolare, il numero degli arrivi  $N_t - N_s$ , in  $(s, t]$ , e il numero degli arrivi  $N_s$ , in  $(0, s]$ , sono stocasticamente indipendenti.
- Un processo di conteggio  $\{N_t, t \geq 0\}$  ha **incrementi stazionari** se la distribuzione del numero degli arrivi in qualsiasi intervallo dipende solo dalla lunghezza dell’intervallo. In altre parole  $N_{t_2+s} - N_{t_1+s}$  ha la stessa distribuzione di  $N_{t_2} - N_{t_1}$  per ogni  $t_1 < t_2$  e  $s > 0$ .

## 4.2 Processo di Poisson

Il processo dei conteggi  $\{N_t, t \geq 0\}$  dicesi *processo di Poisson*, e lo indichiamo con  $PP(\lambda)$ , se le variabili  $X_n$  sono stocasticamente indipendenti ed identicamente distribuite con distribuzione esponenziale di parametro  $\lambda > 0$ . Ricordiamo

che una v.a.  $X$  ha distribuzione esponenziale di parametro  $\lambda > 0$ , in breve  $X \sim Exp(\lambda)$ , se la densità di probabilità di  $X$  è

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0; \\ 0, & \text{altrimenti.} \end{cases}$$

Sia  $\{X_1, X_2, \dots\}$  una successione di variabili aleatorie indipendenti con  $X_n \sim Exp(\lambda), n \in \mathbb{N}$ . La distribuzione della v.a.  $T_n = X_1 + X_2 + \dots + X_n$  si ottiene effettuando  $n - 1$  convoluzioni di distribuzioni esponenziali (si pone  $T_0 = 0$  e  $N_0 = 0$ ). Ricordiamo che

$$\overbrace{Exp(\lambda) * Exp(\lambda) * \dots * Exp(\lambda)}^n = G_{n,\lambda}$$

dove la densità gamma  $G_{n,\lambda}$  è data da

$$G_{n,\lambda}(t) = \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t}, \quad t > 0,$$

con la funzione  $\Gamma(c)$  definita come segue

$$\Gamma(c) = \int_0^{+\infty} x^{c-1} e^{-x} dx, \quad \forall c \in R^+. \quad (4.1)$$

Infatti, la funzione caratteristica di una  $G_{\alpha,\lambda}$  è data da

$$\phi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx = \left( \frac{\lambda}{\lambda - it} \right)^\alpha.$$

Ricordando che la distribuzione esponenziale di parametro  $\lambda$  è una  $G_{1,\lambda}$ , si ha che la funzione caratteristica di  $T_n$  (somma di  $n$  esponenziali indipendenti) è data da

$$\phi_{T_n} = \phi_{X_1} \phi_{X_2} \dots \phi_{X_n} = \left( \frac{\lambda}{\lambda - it} \right)^n,$$

cioè i tempi di attesa  $T_n$  hanno distribuzione  $G_{n,\lambda}$ .

Calcoliamo adesso la distribuzione dei conteggi. Ricordiamo il legame tra  $N_t$  e  $T_n$

$$P(N_t = n) = P(T_{n+1} > t) - P(T_n > t).$$

In tal caso, abbiamo

$$P(T_{n+1} > t) = \int_t^{+\infty} G_{n+1,\lambda}(x) dx; \quad P(T_n > t) = \int_t^{+\infty} G_{n,\lambda}(x) dx.$$

Se indichiamo con  $I_n(t) = \int_t^{+\infty} G_{n,\lambda}(x) dx$  si ha

$$P(N_t = n) = I_{n+1}(t) - I_n(t).$$

Integrando per parti, otteniamo

$$\begin{aligned}
I_{n+1}(t) &= \int_t^{+\infty} \frac{\lambda^{n+1}}{\Gamma(n+1)} x^n e^{-\lambda x} dx = \\
&= \frac{\lambda^{n+1}}{\Gamma(n+1)} \left\{ \left[ \frac{-x^n e^{-\lambda x}}{\lambda} \right]_t^{+\infty} + \int_t^{+\infty} \frac{nx^{n-1}}{\lambda} e^{-\lambda x} dx \right\} = \\
&= \frac{\lambda^{n+1}}{n!} \frac{t^n e^{-\lambda t}}{\lambda} + \int_t^{+\infty} \frac{\lambda^{n+1}}{n!} \frac{nx^{n-1}}{\lambda} e^{-\lambda x} dx = \\
&= \frac{\lambda^n}{n!} t^n e^{-\lambda t} + \int_t^{+\infty} \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} dx = \\
&= \frac{\lambda^n}{n!} t^n e^{-\lambda t} + I_n(t).
\end{aligned}$$

Pertanto

$$P(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t},$$

cioè il numero aleatorio di rinnovi (arrivi) in  $(0, t]$  ha distribuzione di Poisson di parametro  $\lambda t$ , in simboli

$$N_t \sim \mathcal{P}(\lambda \cdot t).$$

Osservando che  $E(N_t) = \lambda t$ , possiamo dire che il valore atteso di ogni numero aleatorio  $N_t$  è una misura lineare dell'ampiezza dell'intervallo.

Consideriamo il numero dei rinnovi nell'intervallo  $(t, t+h]$ , dato da  $N_{t+h} - N_t$ . Il primo rinnovo nell'intervallo  $(t, \infty]$  si ha quando cessa di funzionare il dispositivo che era in funzione all'istante  $t$ , pertanto esso rappresenta la vita residua del dispositivo. Ma, poichè i tempi di vita sono esponenziali di parametro  $\lambda$  essi godono della proprietà di assenza di memoria, cioè la vita residua del dispositivo è ancora con distribuzione esponenziale di ugual parametro. Così facendo, all'istante  $h$ , nasce un nuovo processo di Poisson indipendente dal passato, cioè la v.a.  $N_t^h = N_{t+h} - N_t$ , ha distribuzione di Poisson di parametro  $\lambda h$ . Quindi la v.a.  $N_t^h = N_{t+h} - N_t$ , per ogni  $h \in R^+$ , è indipendente da  $N_t$  e ha la stessa distribuzione di  $N_h$  (vedi [6]).

Analiticamente (vedi Fig. 5.4 in [9]), osserviamo che  $N_t^h = N_{t+h} - N_t$  conta il numero di rinnovi in  $(t, t+h]$ . Pertanto  $T_{N_{t+1}}$  è il tempo di attesa del primo rinnovo dopo l'istante  $t$  e  $T_{N_t}$  è il tempo di attesa dell'ultimo rinnovo prima di  $t$ . Sfruttando la proprietà di assenza di memoria delle variabili aleatorie  $X_n$  si ha

$$\begin{aligned}
&P(T_{N_{t+1}} - t > y | N_t = k, T_{N_t} = x, N_u : 0 \leq u \leq t) = \\
&= P(T_k + X_{k+1} - t > y | N_t = k, T_{N_t} = x, N_u : 0 \leq u \leq t) = \\
&= P(X_{k+1} > t + y - x | X_{k+1} > t - x) = \\
&= e^{-\lambda y}
\end{aligned}$$

Quindi il primo evento contato da  $N_t^h$  si verifica dopo un tempo dato da una variabile aleatoria di tipo esponenziale (vita residua) di parametro  $\lambda$ . I successivi tempi di interarrivo sono i.i.d. con distribuzione esponenziale di parametro  $\lambda$ . Pertanto il processo  $\{N_t^h, h \geq 0\}$  è un processo di poisson di intensità  $\lambda$  ed è indipendente da  $\{N_u, u \leq t\}$ .

In conclusione il Processo di Poisson è caratterizzato dalle seguenti proprietà

1. Gli incrementi  $N_{t+h} - N_t$  sono indipendenti (cioè incrementi relativi ad intervalli disgiunti sono indipendenti) e omogenei nel tempo, la loro distribuzione non dipende da  $t$  ma solo dalla lunghezza  $h$  dell'intervallo.
2. Per ogni fissato  $t$  la v.a.  $N_t$  ha distribuzione di Poisson di parametro  $\lambda t$

Si può dimostrare anche il viceversa del risultato precedente, pertanto si ha

**Teorema 11.** Un processo stocastico  $\{N_t, t \geq 0\}$  è un processo di Poisson di intensità  $\lambda$  se e solo se

- (i)  $\{N_t, t \geq 0\}$  è a incrementi indipendenti e stazionari (omogenei);
- (ii)  $P(N_t = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}$ , per ogni  $t \geq 0$  e per  $k \in \mathbb{N}_0$ .

Nell'Osservazione successiva si dimostra la parte necessaria del precedente Teorema.

**Osservazione 7.** Sia  $\{N_t, t \geq 0\}$  un processo che soddisfi le proprietà (i), (ii). Innanzitutto, dalla (ii) si ottiene che  $N_0 = 0$  con probabilità 1. Sia  $T_1$  il tempo di attesa sino al primo arrivo, ovvero

$$T_1 = \{\inf t \geq 0, N_t = 1\}.$$

Poichè  $(T_1 > t) \equiv (N_t < 1) \equiv (N_t = 0)$ , la funzione di sopravvivenza di  $T_1$  è

$$S_1(t) = P(T_1 > t) = P(N_t = 0) = e^{-\lambda t}.$$

Pertanto  $T_1 \sim Exp(\lambda)$ , cioè il tempo di attesa sino al primo arrivo ha una distribuzione esponenziale di parametro  $\lambda$ .

Sia  $T_n$  il tempo di attesa sino all' $n$ -esimo arrivo. Si ha

$$(T_n > t) \equiv (N_t < n). \quad (4.2)$$

Pertanto, la funzione di sopravvivenza  $S_n(t)$  di  $T_n$  è

$$S_n(t) = P(T_n > t) = P(N_t < n) = \sum_{k=0}^{n-1} \frac{e^{-\lambda t} (\lambda t)^k}{k!}.$$

E' facile verificare che

$$\begin{aligned} \int_t^{+\infty} G_{n,\lambda}(x) dx &= \int_t^{+\infty} \frac{(\lambda x)^{n-1} \lambda e^{-\lambda x}}{\Gamma(n)} dx = \\ &= \frac{e^{-\lambda t} (\lambda t)^{n-1}}{(n-1)!} + \int_t^{+\infty} \frac{(\lambda x)^{n-2} \lambda e^{-\lambda x}}{\Gamma(n-1)} dx = \\ &= \sum_{k=0}^{n-1} \frac{e^{-\lambda t} (\lambda t)^k}{k!}. \end{aligned}$$

Pertanto  $T_n \sim G_{n,\lambda}$ .

Sia  $X_n = T_n - T_{n-1}$  l' $n$ -esimo tempo di inter-arrivo. Si ha

$$P(X_n > t) = P(T_n - T_{n-1} > t) = P(T_n > T_{n-1} + t)$$

Condizionando si ottiene

$$\begin{aligned} P(X_n > t) &= P(T_n - T_{n-1} > t) = P(T_n > T_{n-1} + t) = \\ &= \int_0^{+\infty} P(T_n > t_{n-1} + t | T_{n-1} = t_{n-1}) G_{n-1, \lambda}(t) dt. \end{aligned}$$

Poichè gli incrementi sono indipendenti e stazionari, si ha

$$\begin{aligned} P(T_n > t_{n-1} + t | T_{n-1} = t_{n-1}) &= \\ &= P(N_{t_{n-1}+t} - N_{t_{n-1}} = 0 | \inf\{t \geq 0, N_t = n-1\} = t_{n-1}) = \\ &= P(N_{t_{n-1}+t} - N_{t_{n-1}} = 0) = e^{-\lambda t}. \end{aligned}$$

Quindi la distribuzione di  $X_n$  non dipende da  $n$ , pertanto

$$P(X_n > t) = \int_0^{+\infty} e^{-\lambda t} G_{n-1, \lambda}(t) dt = e^{-\lambda t} \cdot 1 = e^{-\lambda t},$$

ovvero  $X_n \sim Exp(\lambda)$ , per ogni  $n$ . Si può verificare che le variabili  $X_1, X_2, \dots$ , sono stocasticamente indipendenti. Pertanto  $\{N_t, t \geq 0\}$  è un processo di Poisson. In particolare verificiamo che  $X_1, X_2$  sono stocasticamente indipendenti. Siano  $t_1, t_2$  numeri positivi. Si ha

$$\begin{aligned} P(X_2 > t_2, X_1 > t_1) &= P(T_2 > T_1 + t_2, T_1 > t_1) = \\ &= \int_0^{+\infty} P(T_2 > t_2 + t, T_1 > t_1 | T_1 = t) \lambda t e^{-\lambda t} dt = \\ &= \int_{t_1}^{+\infty} P(T_2 > t_2 + t | T_1 = t) \lambda t e^{-\lambda t} dt = \\ &= \int_{t_1}^{+\infty} P(X_2 > t_2) \lambda t e^{-\lambda t} dt = \int_{t_1}^{+\infty} e^{-\lambda t_2} \lambda t e^{-\lambda t} dt = \\ &= e^{-\lambda t_1} e^{-\lambda t_2}. \end{aligned}$$

**Osservazione 8.** In un processo di Poisson

$$N_t \sim \mathcal{P}(\lambda t),$$

con  $\mathbb{E}(N_t) = \lambda t = \frac{t}{\mathbb{E}(X_n)}$ . Cioè il numero medio dei rinnovi in  $(0, t]$  è proporzionale alla lunghezza  $t$  dell'intervallo e inversamente proporzionale alla vita media dei singoli dispositivi.

**Esempio 28.** Supponiamo che i tempi tra due nascite successive in un ospedale siano con distribuzione esponenziale con valore atteso pari a 1 giorno. Qual è la probabilità  $p$  che la 10<sup>a</sup> nascita (in un calendario annuale) avvenga dopo il 15 Gennaio ?

Notiamo che non ha importanza quando è avvenuta l'ultima nascita nell'anno precedente in quanto la proprietà di assenza di memoria implica che i tempi fino alla nascita nel corrente anno sono ancora con distribuzione esponenziali. Quindi il tempo di attesa sino alla decima nascita sarà  $T_{10} = \sum_{i=1}^{10} X_i$ , con  $X_n \sim Exp(1)$ . Pertanto, poichè  $T_{10} \sim G_{10,1}$ , la probabilità richiesta sarà

$$p = P(T_{10} > 15).$$

Osserviamo che nei processi di conteggio vale la seguente relazione

$$(T_n > t) = (N_t < n)$$

pertanto

$$p = P(N_{15} < 10) = \sum_{k=0}^9 \frac{(15)^k}{k!} e^{-15}.$$

**Esempio 29.** [9] Supponiamo che i clienti arrivano in un ufficio postale secondo un  $PP(\lambda)$  con  $\lambda=10$  clienti l'ora.

1. Qual è il numero medio di clienti che va all'ufficio postale durante un periodo di 8 ore?
2. Qual è la probabilità che nessuno arrivi all'ufficio postale dalle 13.00 alle 13.06?

Risposte

1. Supponiamo che il giorno inizia al tempo 0. Allora si ha  $\mathbb{E}(N_8) = 8\lambda = 80$  clienti;
2. Dalla omogeneità dei tempi, la probabilità richiesta coincide con la probabilità che nessuno arrivi da  $t = 0$  sino a  $t = 0.1$  ore. Quindi abbiamo

$$P(N_{0.1} = 0) = e^{-\lambda t} = e^{-10 \cdot 0.1} = 0.3678.$$

### 4.3 Processo Uniforme ([1],[12],[9])

L'esperimento di estrarre o generare  $n$  punti a caso in un intervallo  $[0, a]$  dicesi processo uniforme. Per ogni fissato  $n$  scegliere  $n$  punti equivale a scegliere una  $n$ -upla di numeri reali  $(x_1, x_2, \dots, x_n)$  tutti in  $[0, a]$ . In tal caso lo spazio campione uniforme  $n$ -dimensionale è l'ipercubo di  $\mathbb{R}^n$  di lato  $a$ . Si può verificare che il processo uniforme di campionare  $n$  punti in  $[0, a]$  è una successione di  $n$  variabili aleatorie indipendenti  $X_1, X_2, \dots, X_n$  con distribuzione uniforme in  $[0, a]$ . Indichiamo con  $I_x$  l'intervallo  $[0, x]$ , con  $0 < x \leq a$ . Sia  $N_x$  il numero aleatorio di punti che cadono in  $I_x$ , ovvero

$$N_x = \sum_{i=1}^n |X_i \in I_x| = \sum_{i=1}^n |X_i \leq x|.$$

Si ha  $N_x \sim \text{Bin}(n, \frac{x}{a})$ .

Sia  $E_x^k =$  "almeno  $k$  degli  $n$  punti cadono in  $I_x$ ", calcolare  $P(E_x^k)$ . Poniamo  $X_{(1)} = \min(X_1, X_2, \dots, X_n)$ . Verificare che la densità di probabilità di  $X_{(1)}$  è

$$f_{(1)}(x) = \begin{cases} \frac{n}{a} \left(1 - \frac{x}{a}\right)^{n-1}, & x \in [0, a]; \\ 0, & \text{altrimenti.} \end{cases}$$

La funzione di ripartizione  $F_{(1)}(x)$  è

$$F_{(1)}(x) = \begin{cases} 0, & x < 0; \\ 1 - \left(1 - \frac{x}{a}\right)^n, & x \in [0, a]; \\ 1, & x > a. \end{cases}$$

Verificare inoltre che  $\mathbb{E}(X_{(1)}) = \frac{a}{n+1}$ . Nel caso particolare in cui  $a = 1$  si ottiene che  $X_{(1)}$  ha una distribuzione Beta di parametri  $r = 1, s = n$ . Ricordiamo che dati due parametri  $r, s$  entrambi positivi, un n.a. continuo  $X$  con densità di probabilità data da

$$B_{r,s}(x) = \begin{cases} \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} x^{r-1} (1-x)^{s-1}, & \text{se } x \in (0, 1), \\ 0, & \text{altrimenti} \end{cases} \quad (4.3)$$

si dice che ha distribuzione beta, di parametri  $r$  ed  $s$ , e si indica nel seguente modo:  $X \sim B_{r,s}(x)$ .

Indicando con  $X_{(n)} = \max(X_1, X_2, \dots, X_n)$  calcolarne la densità e verificare che nel caso  $a = 1$  tale densità è una beta di parametri  $r = n, s = 1$ .

Piu' in generale introduciamo la seguente notazione

- $X_{(1)} = \min(X_1, X_2, \dots, X_n)$ ,
- $X_{(2)} =$  punto di ascissa immediatamente dopo  $X_1$ ,
- $X_{(3)} =$  punto di ascissa immediatamente dopo  $X_2$ ,
- ...

- $X_{(n)} = \max(X_1, X_2, \dots, X_n)$ .

Diremo che  $X_{(k)}$ , con  $1 \leq k \leq n$ , è la  $k$ -esima statistica d'ordine di  $X_1, X_2, \dots, X_n$ . Ovviamente si ha che  $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$  è una permutazione di  $X_1, X_2, \dots, X_n$  tale che

$$X_{(1)} \leq X_{(2)} \leq \dots \leq \dots, X_{(n)}.$$

Sia  $F_{(k)}(x)$  la funzione di ripartizione di  $X_{(k)}$ , ovvero  $F_{(k)}(x) = P(X_{(k)} \leq x)$ , dove l'evento  $(X_{(k)} \leq x)$ , per  $x \geq 0$ , significa "almeno  $k$  degli  $n$  punti cadano in  $[0, x]$ ". Pertanto per  $x \in [0, a]$  si ha

$$F_{(k)}(x) = \sum_{j=k}^n \binom{n}{j} \left(\frac{x}{a}\right)^j \left(1 - \frac{x}{a}\right)^{n-j},$$

mentre  $F_{(k)}(x) = 0$  se  $x < 0$  e  $F_{(k)}(x) = 1$  se  $x > a$ . Derivando  $F_{(k)}(x)$  si ottiene la seguente densità, di Dirichlet di parametri  $a, n, k$ , della  $k$ -esima statistica d'ordine

$$f_{(k)}(x) = \begin{cases} \frac{n!}{(k-1)!(n-k)!} \frac{1}{a^n} x^{k-1} (a-x)^{n-k}, & x \in [0, a] \\ 0, & \text{altrimenti.} \end{cases}$$

Per  $a = 1$  si ha  $X_{(k)} \sim \text{Beta}(r = k, s = n - k + 1)$ . Un altro modo per calcolare  $f_{(k)}(x)$  è il seguente. Consideriamo la probabilità che il  $k$ -esimo punto cada in  $(x, x + h]$ . Supponiamo che l'intervallo  $[0, a]$  sia suddiviso in 3 intervalli di lunghezze rispettivamente  $x, h, a - (x + h)$ . La probabilità che  $k - 1$  punti cadano in  $[0, x]$ , uno e uno solo cada in  $(x, x + h]$  ed i rimanenti  $n - h$  cadano in  $(x + h, a]$  è

$$\frac{n!}{(k-1)!(n-k)!} \left(\frac{x}{a}\right)^{k-1} \left(\frac{h}{a}\right)^1 \left(1 - \frac{x+h}{a}\right)^{n-k}$$

Osservando che  $P(x < X_{(k)} \leq x + h)$  è pari alla probabilità precedente (più un eventuale termine comprendente  $(\frac{h}{a})^r$  per qualche cosa nel caso in cui  $r \geq 2$  statistiche cadano in  $(x, x + h]$ ) possiamo calcolare la densità  $f_{(k)}(x)$  mediante il limite (destra) del rapporto incrementale. Cioè, per  $x \in [0, a]$ , si ha

$$\begin{aligned} f_{(k)}(x) &= \lim_{h \rightarrow 0} \frac{P(x < X_{(k)} \leq x + h)}{h} + o\left(\frac{h}{a}\right) = \\ &= \lim_{h \rightarrow 0} \frac{\frac{n!}{(k-1)!(n-k)!} \left(\frac{x}{a}\right)^{k-1} \left(\frac{h}{a}\right)^1 \left(1 - \frac{x+h}{a}\right)^{n-k}}{\frac{n!}{(k-1)!(n-k)!} \left(\frac{x}{a}\right)^{k-1} \left(\frac{1}{a}\right)^1 \left(1 - \frac{x}{a}\right)^{n-k}}. \end{aligned}$$

Con un ragionamento simile a quanto fatto per  $f_{(k)}(x)$  si può verificare che la densità congiunta della coppia  $(X_{(j)}, X_{(k)})$  è

$$f_{(j,k)}(x, y) = \begin{cases} \frac{n!}{(j-1)!(k-j-1)!(n-k)!} \frac{x^{j-1} (y-x)^{k-j-1} (a-y)^{n-k}}{a^n}, & \text{se } 0 \leq x < y \leq a, \\ 0, & \text{altrimenti.} \end{cases} \quad (4.4)$$

**Esercizio 9.** Calcolare  $f_{(1,2)}(x, y)$ ,  $f_{(n-1,n)}(a - y, a - x)$ ,  $f_{(n+1-k, n+1-j)}(a - y, a - x)$ .

Se consideriamo il sottovettore  $X_{(1)}, X_{(2)}, \dots, X_{(k)}$  delle prime  $k$  statistiche d'ordine per  $0 \leq x_1 < \dots < x_k \leq a$  si ha

$$f_{(1,2,\dots,k)}(x_1, x_2, \dots, x_k) = \frac{n!}{(n-k)!} \frac{(a-x_k)^{n-k}}{a^n}.$$

Infine la densità congiunta di tutte le statistiche d'ordine è data da

$$f_{(1,2,\dots,n)}(x_1, x_2, \dots, x_n) = \begin{cases} \frac{n!}{a^n} & \text{se } 0 \leq x_1 < x_2 < \dots < x_n \leq a, \\ 0 & \text{altrimenti.} \end{cases} \quad (4.5)$$

Introduciamo le  $n+1$  lacune di un processo uniforme così definite

$$\begin{aligned} L_1 &= X_{(1)} - 0 \\ L_2 &= X_{(2)} - X_{(1)} \\ &\dots \\ L_{n+1} &= a - X_{(n)} \end{aligned}$$

e osserviamo che a sua volta le statistiche d'ordine possono esprimersi in funzione delle lacune

$$X_{(k)} = L_1 + L_2 + \dots + L_k.$$

Si può verificare che le Lacune non sono stocasticamente indipendenti, infatti se una è più grande allora le altre devono essere più piccole. Inoltre si può verificare che esse sono equidistribuite ([12],[1]) e poiché  $L_1 = X_{(1)}$  allora tutte hanno distribuzione di Dirichelet, ovvero indicando con  $l_k$  la densità della  $k$ -esima lacuna si ha

$$l_{(k)}(x) = \begin{cases} \frac{n}{a} \left(1 - \frac{x}{a}\right)^{n-1}, & x \in [0, a] \\ 0, & \text{altrimenti.} \end{cases}$$

Osservando che  $\mathbb{E}(L_k) = \mathbb{E}(L_1)$  e che  $L_1 + L_2 + \dots + L_n = a$  si ha  $a = \sum_{k=1}^{n+1} \mathbb{E}(L_k) = \mathbb{E}(L_1)(n+1)$ , pertanto

$$\mathbb{E}(L_k) = \frac{a}{n+1}.$$

Inoltre poiché  $X_{(k)} = L_1 + L_2 + \dots + L_k$  si ha

$$\mathbb{E}(X_{(k)}) = \frac{ak}{n+1}.$$

**Esercizio 10.** Dati 5 numeri aleatori  $X_1, \dots, X_5$ , indipendenti e con distribuzione uniforme nell'intervallo  $[0, 5]$ , siano  $X_{(1)}, \dots, X_{(5)}$  le corrispondenti statistiche d'ordine. Calcolare la probabilità  $p$  dell'evento  $(X_{(3)} \leq 3) \wedge (X_{(4)} > 4)$ ,

La probabilità  $p$  può essere calcolata in diversi modi. Illustriamone alcuni

1) Calcoliamo  $p$  a partire dalla densità di  $X_{(1)}, \dots, X_{(5)}$  data dalla densità (4.5). Si ha

$$\begin{aligned}
P(X_{(3)} \leq 3, X_{(4)} > 4) &= \int_0^3 \int_{x_1}^3 \int_{x_2}^3 \int_4^5 \int_{x_4}^5 \frac{5!}{5^5} dx_1 \cdots dx_5 = \\
&= \frac{5!}{5^5} \left( \int_0^3 \int_{x_1}^3 \int_{x_2}^3 dx_1 dx_2 dx_3 \right) \left( \int_4^5 \int_{x_4}^5 dx_4 dx_5 \right) = \\
&= \frac{5!}{5^5} \left( \int_0^3 \int_{x_1}^3 (3 - x_2) dx_1 dx_2 \right) \left( \int_4^5 (5 - x_4) dx_4 \right) = \\
&= \frac{5!}{5^5} \left( \int_0^3 \left[ 3x_2 - \frac{x_2^2}{2} \right]_{x_1}^3 dx_1 \right) \left( \left[ 5x_4 - \frac{x_4^2}{2} \right]_4^5 \right) = \\
&= \frac{5!}{5^5} \left( \int_0^3 \left( 9 - \frac{9}{2} - 3x_1 + \frac{x_1^2}{2} \right) dx_1 \right) (25 - \frac{25}{2} - 20 + 8) = \\
&= \frac{5!}{5^5} \left( \left[ \frac{9}{2}x_1 - 3\frac{x_1^2}{2} + \frac{x_1^3}{3!} \right]_0^3 \right) \left( \frac{1}{2!} \right) = \frac{5!}{5^5} \left( \frac{27}{2} - \frac{27}{2} + \frac{3^3}{3!} \right) \frac{1}{2} = \\
&= \frac{5!}{5^5} \frac{3^3}{3!} \frac{1}{2!} = \frac{5!}{3!2!} \left( \frac{3}{5} \right)^3 \left( \frac{1}{5} \right)^2.
\end{aligned}$$

2) In maniera analoga si può calcolare  $p$  a partire dalla densità data in (4.4). Infatti, si ha

$$\begin{aligned}
P(X_{(3)} \leq 3, X_{(4)} > 4) &= \int_0^3 \int_4^5 \frac{5!}{2!0!1!} \frac{x^2(y-x)^0(5-y)^1}{5^5} dx dy = \\
&= \frac{5!}{2!5^5} \int_0^3 \int_4^5 x^2(5-y) dx dy = \\
&= \frac{5!}{2!5^5} \left( \int_0^3 x^2 dx \right) \left( \int_4^5 (5-y) dy \right) = \\
&= \frac{5!}{2!5^5} \frac{x^3}{3} \Big|_0^3 \left( 5y - \frac{y^2}{2} \right) \Big|_4^5 = \\
&= \frac{5!}{2!5^5} \frac{3^3}{3} \frac{1}{2} = \frac{5!}{5^5} \frac{3^3}{3!} \frac{1}{2!} = \frac{5!}{3!2!} \left( \frac{3}{5} \right)^3 \left( \frac{1}{5} \right)^2.
\end{aligned}$$

3) Osserviamo che l'evento  $(X_{(3)} \leq 3, X_{(4)} > 4)$  coincide con  $(N_3 \geq 3 \wedge N_4 < 4)$  che a sua volta coincide con  $(N_{[0,3]} = 3 \wedge N_{(3,4]} = 0 \wedge N_{(4,5]} = 2)$ . Dove con  $N_{[a,b]}$  indichiamo il numero di punti che cadono in  $[a, b]$ . Consideriamo, per  $i = 1, \dots, 5$ , gli eventi  $A_i = |X_i \leq 3|$ ,  $B_i = |3 < X_i \leq 4|$  e  $C_i = |4 < X_i \leq 5|$ . Per ogni  $i$  si ha che gli eventi  $A_i, B_i, C_i$  formano una partizione di  $\Omega$ . Osserviamo che gli eventi relativi a partizioni distinte sono indipendenti (ciò è dovuto all'indipendenza tra le  $X_i$ ). Infine, per  $i = 1 \dots 5$ , si ha  $P(A_i) = \frac{3}{5}$ ,  $P(B_i) = \frac{1}{5}$  e  $P(C_i) = \frac{1}{5}$  (cioè gli eventi  $A_i$  sono equiprobabili, idem per  $B_i$  e  $C_i$ ). Pertanto la distribuzione del vettore aleatorio  $(N_{[0,3]}, N_{(3,4]}, N_{(4,5]})$  è una multinomiale di parametri  $n = 5$  e  $p_a = \frac{3}{5}, p_b = \frac{1}{5}, (p_c = 1 - p_a - p_b = \frac{1}{5})$ . Quindi

$$\begin{aligned}
P(X_{(3)} \leq 3, X_{(4)} > 4) &= P(N_{[0,3]} = 3 \wedge N_{(3,4]} = 0 \wedge N_{(4,5]} = 2) = \\
&= \frac{5!}{3!2!} \left( \frac{3}{5} \right)^3 \left( \frac{1}{5} \right)^2.
\end{aligned}$$

**Esercizio 11.** Quattro amici si incontrano in una città  $\mathcal{C}$  dove arrivano a caso nell'intervallo (misurato in ore)  $[0, 5]$ . Calcolare la probabilità  $p$  che 3 amici arrivino con un ritardo, rispetto all'istante 0, inferiore a 2 ore e uno arrivi con un ritardo superiore a 3 ore. Inoltre, indicando con  $(X_{(1)}, X_{(2)})$  il vettore aleatorio relativo agli istanti del primo e secondo arrivo, calcolare la densità congiunta  $\phi_{12}(x, y)$  e la probabilità  $\alpha$  dell'evento condizionato  $(X_{(2)} \leq 3 | X_{(1)} \leq 2)$ .

$$p = \quad ; \quad \phi_{12}(x, y) = \quad ; \quad \alpha =$$

Si ha

$$p = P(X_3 < 2 \wedge X_4 > 3) = \int_0^2 \int_3^5 \frac{5!}{2!0!1!} \frac{x^2(y-x)^0(5-y)^1}{5^4} dx dy = \dots$$

oppure, sfruttando la distribuzione multinomiale,

$$p = \frac{5!}{3!0!2!} \left(\frac{2}{5}\right)^3 \left(\frac{1}{5}\right)^0 \left(\frac{2}{5}\right)^2.$$

### 4.3.1 Processo di Poisson come limite di un Processo Uniforme

Ricordiamo che le lacune del processo uniforme giocano il ruolo dei tempi di interarrivo nel processo di Poisson, con la differenza che le lacune non sono stocasticamente indipendenti. Se si porta all'infinito l'ampiezza dell'intervallo, affinché il processo Uniforme continui ad avere senso occorre che si generino un numero infinito di punti (non ha senso parlare di un campionamento finito da un intervallo di ampiezza infinita). Per far ciò si può costruire un processo limite in cui l'ampiezza  $a$  (o  $a_n$ ) dell'intervallo sia infinito dello stesso ordine di  $n$  o in maniera più semplice si può pensare che il rapporto  $\lambda = \frac{n}{a}$  si mantenga costante al crescere di  $n$ . Ciò significa che non varia il numero di punti generati per unità di intervallo. In tal caso si ha che la densità della  $k$ -esima lacuna converge alla densità di una distribuzione esponenziale di parametro  $\lambda$ , infatti per  $t > 0$  si ha

$$\lim_{n \rightarrow +\infty} l_k(t) = \lim_{n \rightarrow +\infty} \frac{n}{a} \left(1 - \frac{t}{a}\right)^{n-1} = \lim_{n \rightarrow +\infty} \frac{\lambda \left(1 - \frac{\lambda t}{n}\right)^n}{\left(1 - \frac{\lambda t}{n}\right)} = \lambda e^{-\lambda t}.$$

cioè le lacune nel processo limite hanno distribuzione esponenziale. Si può verificare ([1, 12]) che al limite tale lacune sono anche stocasticamente indipendenti. Pertanto la statistica d'ordine  $X_{(k)}$  al limite avrà una distribuzione gamma di parametri  $k, \lambda$  e se consideriamo il limite della variabile conteggio  $N_t$  nel processo uniforme è facile verificare che ha una distribuzione di Poisson di parametro  $\lambda t$ , infatti

$$\begin{aligned} \lim_{n \rightarrow \infty} P(N_t = h) &= \lim_{n \rightarrow \infty} \binom{n}{h} \left(\frac{t}{a}\right)^h \left(1 - \frac{t}{a}\right)^{n-h} = \\ &= \lim_{n \rightarrow \infty} \frac{n(n-1)\dots(n-h+1)}{h!} \left(\frac{\lambda t}{n}\right)^h \frac{\left(1 - \frac{\lambda t}{n}\right)^n}{\left(1 - \frac{\lambda t}{n}\right)^h} = \frac{(\lambda t)^h}{h!} e^{-\lambda t}. \end{aligned}$$

Riassumendo un processo di Poisson di intensità  $\lambda$  si può vedere come processo limite di un processo Uniforme, in cui il rapporto  $\frac{n}{a} = \lambda$  si mantiene costante.

P. Uniforme	Distribuzione	P. Poisson	Distribuzione
Lacune $L_k$	Dirichelet $a, n, 1$	Tempi di interarrivo $X_k$	$\text{Exp}(\lambda), \lambda = \frac{n}{a}$
Statistiche $X_{(k)}$	Dirichelet $a, n, k$	Tempi di attesa $T_k$	$\text{Gamma}(k, \lambda)$
Conteggio $N_t$	$\text{Bin}(n, \frac{t}{a})$	Conteggio $N_t$	Poisson $\lambda t = n \frac{t}{a}$

Tabella 4.1: Analogie tra il processo Uniforme e il Processo di Poisson

P. Bernoulli	Distribuzione
Lacune $L_k$	$G(p)$
Tempi di attesa $T_k$	Pascal $(k, p)$
Conteggio $S_n$	$\text{Bin}(n, p)$ .

Tabella 4.2: Processo di Bernoulli

### 4.3.2 Processo Uniforme come Processo di Poisson condizionato

Fissiamo un intervallo  $(0, a]$  e indichiamo con  $N_a$  il numero aleatorio di arrivi in  $[0, a]$ . Supponiamo di sapere che  $N_a = n$  e consideriamo il n.a. condizionato

$$N_t | (N_a = n), t \in (0, a].$$

Innanzitutto osserviamo che  $N_t | (N_a = n) \in \{0, 1, 2, \dots, n\}$ . Al più gli arrivi che si sono verificati in  $(0, t]$  sono  $n$ . Dimostriamo che la variabile aleatoria condizionata ( $N_t = h | N_a = n$ ) ha distribuzione binomiale di parametri  $n, \frac{t}{a}$ . Infatti,

$$\begin{aligned} P(N_t = h | N_a = n) &= \frac{P(N_t = h \wedge N_a = n)}{P(N_a = n)} = \frac{P(N_t = h \wedge N_{(t,a]} = n - h)}{P(N_a = n)} = \\ &= \frac{P(N_t = h) P(N_{(t,a]} = n - h)}{P(N_a = n)} = \frac{\frac{(\lambda t)^h}{h!} e^{-\lambda t} \frac{(\lambda(a-t))^{n-h}}{(n-h)!} e^{-\lambda(a-t)}}{\frac{(\lambda a)^n}{n!} e^{-\lambda a}} = \\ &= \binom{n}{h} \left(\frac{t}{a}\right)^h \left(1 - \frac{t}{a}\right)^{n-h}. \end{aligned}$$

Cioè la variabile conteggio di un processo di Poisson condizionato ha una distribuzione binomiale. Pertanto, il processo Uniforme può essere visto come un processo di Poisson condizionato. Osserviamo in particolare che tale risultato **non dipende** dall'intensità  $\lambda$  del processo. Pertanto quando si condiziona il processo di Poisson a ( $N_a = n$ ) il risultato è sempre il processo Uniforme di estrarre  $n$  punti da  $[0, a]$ , qualunque sia l'intensità del processo di Poisson di partenza. Si suole dire che il processo di Poisson è il processo di spruzzare punti a caso sulla semiretta positiva ([1]).

## 4.4 Alcune proprietà del Processo di Poisson

Indichiamo per brevità con  $N_t$  un processo di Poisson  $\{N_t, t \geq 0\}$ .

#### 4.4.1 Superposition

Siano dati due processi di Poisson  $N'_t, N''_t$  di intensità rispettive  $\lambda', \lambda''$ , il processo  $N_t = N'_t + N''_t$  è un processo di Poisson di intensità  $\lambda' + \lambda''$ . Ricordiamo che la variabile aleatoria somma di variabili aleatorie di Poisson indipendenti è ancora una variabile aleatoria di Poisson. Infatti, siano  $X, Y$  due v. a. indipendenti con distribuzione di Poisson, di parametro, rispettivamente,  $\alpha$  e  $\beta$ , ovvero

$$X \sim \mathcal{P}(\alpha), \quad Y \sim \mathcal{P}(\beta).$$

Calcoliamo la distribuzione di probabilità di  $Z = X + Y$  ottenuta dalla convoluzione delle distribuzioni di  $X$  e di  $Y$ . Fissato  $n \in \mathbb{N}$ , si ha

$$\begin{aligned} P(Z = n) &= P[\bigvee_{i=0}^n (X = i, Y = n - i)] = \\ &= \sum_{i=0}^n P(X = i, Y = n - i) = \\ &= \sum_{i=0}^n e^{-\alpha} \frac{\alpha^i}{i!} \cdot e^{-\beta} \frac{\beta^{n-i}}{(n-i)!} = \dots = \\ &= e^{-(\alpha+\beta)} \frac{(\alpha+\beta)^n}{n!}. \end{aligned}$$

Quindi  $Z \sim \mathcal{P}(\alpha + \beta)$ .

**Esercizio 12.** Verificare che  $(X|Z = n) \sim \text{Bin}(n, \frac{\lambda_1}{\lambda_1 + \lambda_2})$

Pertanto, poichè i processi  $N'_t, N''_t$  sono indipendenti, per ogni fissato  $t > 0$ , si ha che la distribuzione di  $N_t$  è la convoluzione delle distribuzioni di  $N'_t, N''_t$  e quindi è una distribuzione di Poisson di parametro  $\lambda t = (\lambda' + \lambda'')t$ . Ovvero  $N_t = N'_t + N''_t$  è un  $PP(\lambda' + \lambda'')$ .

Iterando il risultato precedente si ha ([9] pag.214 )

**Teorema 12** (Superposition). Sia  $N_t^{(i)}$ , per  $i = 1 \dots, r$ , un  $PP(\lambda_i)$ . Supponiamo  $N_t^{(1)}, N_t^{(2)}, \dots, N_t^{(r)}$  indipendenti. Allora, il processo

$$N_t = N_t^{(1)} + N_t^{(2)} + \dots + N_t^{(r)}$$

è un  $PP(\lambda)$ , con  $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_r$ .

#### 4.4.2 Splitting

Sia  $N_t$  un  $PP(\lambda)$  e supponiamo che ogni volta che si verifica un arrivo esso venga classificato come arrivo di tipo  $I$  o di tipo  $II$ . Si supponga inoltre che il meccanismo di classificazione degli arrivi sia di tipo Bernoulliano. In tale meccanismo ogni arrivo viene classificato con probabilità  $p_1 = p$  di tipo  $I$ , e con probabilità  $p_2 = 1 - p$  di tipo  $II$ , con  $p_1, p_2$  tali che  $p_1, p_2 > 0$  e  $p_1 + p_2 = 1$ . Inoltre, indicando con  $E_i$  l'evento "l' $i$ -esimo" arrivo viene classificato di tipo  $I$ , per  $i \in \mathbb{N}$ , supponiamo che  $E_1, E_2, \dots$  siano stocasticamente indipendenti.

Se indichiamo con  $N'_t$  e  $N''_t$  il numero degli arrivi, rispettivamente, di tipo  $I$  e di tipo  $II$  in  $(0, t]$  si ha

**Teorema 13.** I processi  $N'_t$  e  $N''_t$  sono processi di Poisson di intensità, rispettivamente,  $\lambda' = \lambda p_1$  e  $\lambda'' = \lambda p_2$ . Inoltre  $N'_t$  e  $N''_t$  sono indipendenti.

*Dimostrazione.* Sia  $p = p_1$ . In primo luogo calcoliamo la probabilità  $P(N'_t = k, N''_t = m)$ . Si ha

$$P(N'_t = k, N''_t = m) = \sum_{n=0}^{\infty} P(N'_t = k, N''_t = m | N_t = n) P(N_t = n).$$

Osserviamo che

$$P(N'_t = k, N''_t = m | N_t = n) = 0, \text{ se } k + m \neq n.$$

Pertanto

$$\begin{aligned} P(N'_t = k, N''_t = m) &= P(N'_t = k, N''_t = m | N_t = k + m) P(N_t = k + m) = \\ &= P(N'_t = k, N''_t = m | N_t = k + m) \frac{e^{-\lambda t} (\lambda t)^{k+m}}{(k+m)!}. \end{aligned}$$

Poichè gli eventi vengono classificati con lo schema bernoulliano si ha che la probabilità che  $k$  di  $k + m$  eventi siano classificati di tipo  $I$  è

$$P(N'_t = k, N''_t = m | N_t = k + m) = \binom{k+m}{k} p^k (1-p)^m.$$

Pertanto,

$$\begin{aligned} P(N'_t = k, N''_t = m) &= \binom{k+m}{k} p^k (1-p)^m \frac{e^{-\lambda t} (\lambda t)^{k+m}}{(k+m)!} = \\ &= \frac{(k+m)!}{k!m!} p^k (1-p)^m \frac{e^{-\lambda t} (\lambda t)^{k+m}}{(k+m)!} = \\ &= \frac{e^{-\lambda p t} (\lambda p t)^k}{(k)!} \frac{e^{-\lambda(1-p)t} (\lambda(1-p)t)^m}{(m)!}. \end{aligned}$$

Quindi

$$P(N'_t = k) = \sum_{m=0}^{\infty} \frac{e^{-\lambda p t} (\lambda p t)^k}{(k)!} \frac{e^{-\lambda(1-p)t} (\lambda(1-p)t)^m}{(m)!} = \frac{e^{-\lambda p t} (\lambda p t)^k}{(k)!}$$

e

$$P(N''_t = m) = \sum_{k=0}^{\infty} \frac{e^{-\lambda p t} (\lambda p t)^k}{(k)!} \frac{e^{-\lambda(1-p)t} (\lambda(1-p)t)^m}{(m)!} = \frac{e^{-\lambda(1-p)t} (\lambda(1-p)t)^m}{(m)!}.$$

Cioè,  $N'_t, N''_t$  sono due processi di Poisson indipendenti di intensità, rispettivamente,  $\lambda p$  e  $\lambda(1-p)$ .  $\square$

Il precedente risultato si può generalizzare al caso in cui ogni arrivo può essere classificato in uno tra  $r$  modi diversi, ognuno con probabilità  $p_r$ , con  $p_r > 0$ ,  $p_1 + p_2 + \dots + p_r = 1$ , e le classificazioni tra i vari arrivi sono tra loro stocasticamente indipendenti. In tal caso, indicando con  $N_t^i$  il processo di conteggio degli arrivi classificati di tipo  $i$ , si ha che  $N_t^i$  è un  $PP(\lambda p_i)$ ,  $i = 1, \dots, r$  e i processi  $N_t^i$  sono tra loro indipendenti.

**Esercizio 13.** Consideriamo una compagnia di assicurazioni che offre due tipi di polizze: Polizza  $A$  e Polizza  $B$ . Supponiamo che il numero totale di richieste di risarcimento segua un processo di Poisson con intensità pari a 9 al giorno. Inoltre supponiamo che la probabilità che una richiesta di risarcimento si riferisca ad una polizza di tipo  $A$  sia pari a  $\frac{1}{3}$ . Inoltre, supponiamo che gli eventi  $E_i$  l'iesima richiesta di risarcimento è di tipo  $A$  siano tra loro stocasticamente indipendenti.

1. Calcolare la probabilità che il numero di richieste di risarcimento relativi alla polizza  $A$ , in un dato giorno, sia inferiore a 2.
2. Calcolare la probabilità che il numero di richieste di risarcimento relativi alla polizza  $B$ , in un dato giorno, sia inferiore a 2.
3. Calcolare la probabilità che il numero di richieste di risarcimento totali, in un dato giorno, sia inferiore a 2.

Indichiamo con  $N_t$  il numero di richieste di risarcimento in  $[0, t]$  con  $t$  espresso in giorni. Si ha  $\{N_t, t \geq 0\}$  è un  $PP(9)$ . Indichiamo con  $A_t$  il numero di richieste di risarcimento in  $[0, t]$  di polizze di tipo  $A$  e con  $B_t$  il numero di richieste di risarcimento in  $[0, t]$  di polizze di tipo  $B$ . Si ha che  $\{A_t, t \geq 0\}$  è un  $PP(3)$  e che  $\{B_t, t \geq 0\}$  è un  $PP(6)$ . Pertanto, si ha

1.  $P(A_1 < 2) = P(A_1 = 0) + P(A_1 = 1) = e^{-3} + 3e^{-3} = 4e^{-3} \simeq 0.199$ ;
2.  $P(B_1 < 2) = P(B_1 = 0) + P(B_1 = 1) = e^{-6} + 6e^{-6} = 7e^{-6} \simeq 0.017$ ;
3.  $P(N_1 < 2) = P(N_1 = 0) + P(N_1 = 1) = e^{-9} + 9e^{-9} = 9e^{-9} \simeq 0.0011$ ;

**Esercizio 14.** Con riferimento all'Esercizio 13 supponiamo che con meccanismo Bernoulliano ogni richiesta di risarcimento di tipo  $A$  con probabilità pari a  $\frac{2}{3}$  superi i \$10000 dollari. Inoltre che con meccanismo Bernoulliano ogni richiesta di risarcimento di tipo  $B$  con probabilità pari a  $\frac{2}{9}$  superi i \$10000 dollari.

1. Calcolare il valor medio del numero di richieste di risarcimento in un dato giorno che superino i \$10,000 dollari. day.
2. Calcolare la probabilità che su un dato giorno si abbiano meno di 2 richieste di risarcimento che superino i \$10,000 dollari.

Indichiamo con  $A'_t$  il numero di richieste di risarcimento in  $[0, t]$  di polizze di tipo  $A$  che superino i \$10000 dollari e con  $B'_t$  il numero di richieste di risarcimento in  $[0, t]$  di polizze di tipo  $B$  che superino i \$10000 dollari. Si ha che  $\{A'_t, t \geq 0\}$  è un  $PP(2)$  e che  $\{B'_t, t \geq 0\}$  è un  $PP(\frac{4}{3})$ . Inoltre i processi  $\{A'_t, t \geq 0\}$  e  $\{B'_t, t \geq 0\}$  sono stocasticamente indipendenti (poichè lo sono  $\{A_t, t \geq 0\}$  e  $\{B_t, t \geq 0\}$ ). Pertanto il processo  $\{A'_t + B'_t, t \geq 0\}$  è un  $PP(\frac{10}{3})$ . Quindi si ha

1.  $\mathbb{E}(A'_1 + B'_1) = \frac{10}{3}$ ;
2.  $P(A'_1 + B'_1 < 2) = e^{-\frac{10}{3}} + \frac{10}{3}e^{-\frac{10}{3}} = \frac{13}{3}e^{-\frac{10}{3}} \simeq 0.155$ .

#### 4.4.3 Richiami sul valore atteso condizionato

Sia  $(X, Y)$  un vettore aleatorio bidimensionale, con densità  $f(x, y)$ . Siano  $f_X = \int_{\mathbb{R}} f(x, y)dy$  e  $f_Y = \int_{\mathbb{R}} f(x, y)dx$  le densità marginali, rispettivamente, di  $X$  e di

$Y$ . Sia  $f_{X|y} = \frac{f(x,y)}{f_Y(y)}$ , per ogni fissato  $y \neq 0$ , la densità marginale condizionata di  $(X|Y = y)$ . Il valore atteso condizionato di  $(X|Y = y)$  è definito come

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{+\infty} x f_{X|y}(x, y) dx . \quad (4.6)$$

Per ogni fissato valore  $y$  di  $Y$  si ottiene un valore di  $\mathbb{E}(X|Y = y)$ . Pertanto si può definire  $Z = \mathbb{E}(X|Y)$  come una funzione del numero aleatorio  $Y$  che per ogni valore  $y$  di  $Y$  assume valore  $\mathbb{E}(X|Y = y)$ . Nel caso discreto si ha

$$\mathbb{E}(X|Y = y_k) = \sum_{x_h} x_h P(X = x_h | Y = y_k) .$$

In tal caso possiamo dire che la variabile  $Z$  assume il valore  $\mathbb{E}(X|Y = y_k)$  con probabilità  $P(Y = y_k)$ .

Più in generale si ha, data una funzione  $g(X)$  del numero aleatorio  $X$ , si ha

$$\mathbb{E}(g(X)|Y = y) = \begin{cases} \sum_{x_h} g(x_h) P(X = x_h | Y = y), & \text{caso discreto;} \\ \int_{(R)} g(x) f_{X|y}(x, y) dx, & \text{caso continuo} \end{cases}$$

Inoltre, si ottiene la seguente formula di decomposizione

$$\begin{aligned} \mathbb{E}(Z) &= \mathbb{E}(\mathbb{E}(X|Y)) = \sum_{y_k} \sum_{x_h} x_h P(X = x_h | Y = y_k) P(Y = y_k) = \\ &= \sum_{y_k} \sum_{x_h} x_h P(X = x_h, Y = y_k) = \\ &= \sum_{x_h} x_h \sum_{y_k} P(X = x_h, Y = y_k) = \\ &= \sum_{x_h} x_h P(X = x_h) = \mathbb{E}(X). \end{aligned}$$

Si dimostra che anche nel caso continuo si ha

$$\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X).$$

Pertanto, la formula di decomposizione (disintegrazione) nel continuo diviene

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} \mathbb{E}(X|Y = y) f_Y(y) dy.$$

Osserviamo che ogni probabilità del tipo  $P(X \in A)$  può essere calcolata come il valore atteso dell'evento  $|X \in A|$ , cioè

$$\mathbb{E}(|X \in A|) = P(X \in A).$$

Quindi, dato sottoinsieme misurabile  $A$  di  $(R)$  si ha

$$P(X \in A) = \mathbb{E}(|X \in A|) = \int_A 1 \cdot f_X(x) dx.$$

Pertanto, dato sottoinsieme misurabile  $A$  di  $(R)$  ed un vettore aleatorio  $(X, Y)$ , mediante la formula di decomposizione si ottiene il seguente risultato molto utile nelle applicazioni.

$$\begin{aligned} P(X \in A) &= \mathbb{E}(|X \in A|) = \mathbb{E}\{\mathbb{E}(|X \in A| | Y)\} = \\ &= \int_{-\infty}^{+\infty} \mathbb{E}(|X \in A| | Y = y) f_Y(y) dy = \\ &= \int_{-\infty}^{+\infty} P(X \in A | Y = y) f_Y(y) dy. \end{aligned} \quad (4.7)$$

Se  $Y$  è discreta si riottiene la già nota formula di decomposizione

$$P(X \in A) = \sum_{y_k} P(X \in A|Y = y_k)P(Y = y_k).$$

La formula 4.7 è molto utile quando la distribuzione di una variabile aleatoria dipende da un “parametro” a sua volta aleatorio.

**Esempio 30.** Sia  $T = X_1 + X_2 + \dots + X_N$  dove le variabili aleatorie  $X_r$  sono indipendenti esponenziali di ugual parametro  $\lambda$  ed  $N$  è una variabile aleatoria, indipendente dalle  $X_r$ , che ha una distribuzione geometrica di parametro  $p$ . Si ha

$$\begin{aligned} \mathbb{E}(T) &= \sum_{n=1}^{+\infty} \mathbb{E}(X_1 + X_2 + \dots + X_N|N = n)P(N = n) = \\ &= \sum_{n=1}^{+\infty} \mathbb{E}(X_1 + X_2 + \dots + X_n)P(N = n) = \\ &= \sum_{n=1}^{+\infty} \frac{n}{\lambda} P(N = n) = \frac{1}{\lambda} \mathbb{E}(N) = \frac{1}{\lambda} \frac{1}{p}. \end{aligned}$$

Inoltre si può verificare che  $T \sim Exp(\lambda p)$ . Notare che il risultato

$$\mathbb{E}(X_1 + \dots + X_N) = \mathbb{E}(X_i)\mathbb{E}(N)$$

vale più in generale sotto le ipotesi che le  $X_i$  siano con stesso valore atteso  $\mathbb{E}(X)$  e la v.a.  $N$  sia (discreta) e indipendente dalle  $X_i$ . Infatti

$$\begin{aligned} \mathbb{E}(T) &= \sum_n \mathbb{E}(X_1 + X_2 + \dots + X_N|N = n)P(N = n) = \\ &= \sum_n \mathbb{E}(X_1 + X_2 + \dots + X_n)P(N = n) = \\ &= \mathbb{E}(X) \sum_n nP(N = n) = \mathbb{E}(X)\mathbb{E}(N). \end{aligned}$$

Un altro modo per giungere allo stesso risultato è il seguente. Osserviamo che  $\mathbb{E}(X_1 + \dots + X_N|N = n) = n\mathbb{E}(X)$ , pertanto la variabile aleatoria valore atteso condizionato è definita nel seguente modo

$$\underline{Z = \mathbb{E}(X_1 + \dots + X_N|N) = N\mathbb{E}(X)}.$$

Come si vede  $Z$  è funzione di  $N$  e si ha

$$\mathbb{E}(X_1 + \dots + X_N) = \mathbb{E}(Z) = \mathbb{E}(N\mathbb{E}(X)) = \mathbb{E}(X)\mathbb{E}(N)$$

**Esempio 31.** Supponiamo che il numero di persone che arrivano alla banchina di una fermata di una metropolitana al tempo  $t$  (espresso in ore) segua un processo di Poisson di intensità  $\lambda$ . Supponiamo inoltre, che il Tempo aleatorio  $T$  di arrivo del primo treno alla fermata sia con distribuzione uniforme in  $(0, b)$ . Calcolare il valore atteso del numero di passeggeri che salgono sul treno.

Indichiamo con  $N(t)$  il numero di passeggeri che arrivano in  $(0, t)$ , si ha che  $N(t)$  è un  $PP(\lambda)$ . La variabile aleatoria di interesse è

$$N(T)$$

dove  $T$  è il tempo aleatorio di arrivo del primo treno e vogliamo calcolare  $\mathbb{E}(N(T))$ . Osserviamo che

$$\mathbb{E}(N(T)|T = t) = \mathbb{E}(N(t)) = \lambda t$$

Si ha che la variabile aleatoria valore atteso condizionato  $\mathbb{E}(N(T)|T)$  è pari a  $\lambda T$ . Pertanto

$$\mathbb{E}(N(T)) = \mathbb{E}(\lambda T) = \lambda \mathbb{E}(T) = \lambda \frac{b}{2}.$$

**Esempio 32.** Siano  $X, Y$  due variabili aleatorie indipendenti con densità, rispettivamente,  $f_X, f_Y$ . Calcolare  $P(X < Y)$ .

Un modo per calcolare rapidamente  $P(X < Y)$  è il seguente

$$\begin{aligned} P(X < Y) &= \int_{(R)} P(X < Y | Y = y) f_Y(y) dy = \\ &= \int_{-\infty}^{+\infty} P(X < y) f_Y(y) dy = \\ &= \int_{-\infty}^{+\infty} F_X(y) f_Y(y) dy, \end{aligned}$$

dove  $F_X(y)$  è la funzione di ripartizione di  $X$  calcolata in  $y$ . Se in particolare  $X, Y$  sono indipendenti esponenziali di parametro, rispettivamente,  $\lambda_1$  e  $\lambda_2$  si ha

$$\begin{aligned} P(X < Y) &= \int_{-\infty}^{+\infty} F_X(y) f_Y(y) dy = \\ &= \int_0^{+\infty} (1 - e^{-\lambda_1 y}) \lambda_2 e^{-\lambda_2 y} dy = \\ &= \int_0^{+\infty} \lambda_2 e^{-\lambda_2 y} dy + \int_0^{+\infty} -e^{-\lambda_1 y} \lambda_2 e^{-\lambda_2 y} dy = \\ &= 1 - \frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{\lambda_2}{\lambda_1 + \lambda_2} = \frac{\mathbb{E}(\min(X, Y))}{\mathbb{E}(X)}. \end{aligned}$$

**Esempio 33.** Siano  $X, Y$  due variabili aleatorie indipendenti con densità, rispettivamente,  $f_X, f_Y$ . Determinare la distribuzione di  $Z = X + Y$ .

$$\begin{aligned} F_Z(z) &= P(X + Y \leq z) = \int_{(R)} P(X + Y < z | Y = y) f_Y(y) dy = \\ &= \int_{-\infty}^{+\infty} P(X + y < z) f_Y(y) dy = \\ &= \int_{-\infty}^{+\infty} F_X(z - y) f_Y(y) dy. \end{aligned}$$

Se in particolare  $X, Y$  sono indipendenti esponenziali di ugual parametro  $\lambda$ , per  $z > 0$  si ha

$$\begin{aligned} F_Z(z) &= P(X + Y \leq z) = \\ &= \int_{-\infty}^{+\infty} F_X(z - y) f_Y(y) dy \\ &= \int_0^z (1 - e^{-\lambda(z-y)}) e^{-\lambda y} \lambda dy \\ &= \int_0^z e^{-\lambda y} \lambda dy + \int_0^z (-e^{-\lambda z}) \lambda dy \\ &= 1 - e^{-\lambda z} - e^{-\lambda z} \lambda z = 1 - S_Z(z), \end{aligned}$$

dove  $S_Z(z) = e^{-\lambda z} + e^{-\lambda z} \lambda z$  è la funzione di sopravvivenza di una gamma di parametri 2 e  $\lambda$ . Pertanto  $Z \sim G_{2, \lambda}$ .

#### 4.4.4 Splittin non omogeneo - Campionamento da un Processo di Poisson

Dato un processo di Poisson  $\{X_t, t \geq 0\}$  sia  $p : (0, +\infty) \rightarrow [0, 1]$  una funzione che ad ogni istante  $t$  del processo assegna la probabilità  $p(t)$  che un arrivo all'istante

$t$  venga classificato di tipo  $I$ . Il meccanismo di classificazione è simile a quello Bernoulliano con la differenza che la probabilità di classificare (campionare) un arrivo dipende, in questo caso, dal tempo (*Bernoulli non-homogeneous splitting*). Se indichiamo con  $N'_t$  il numero di arrivi campionati in  $(0, t]$  si ha ([9],[10])

**Teorema 14.** Per ogni fissato  $t \in ]0, +\infty)$ , la variabile  $N'_t$  ha una distribuzione di Poisson di parametro

$$\lambda \cdot \int_0^t p(s) ds.$$

*Dimostrazione.* Calcoliamo dapprima

$$P(N'_t = k | N_t = m), \quad m \geq k.$$

Supposto che in  $(0, t]$  si siano verificati  $m$  arrivi allora i tempi di attesa degli  $m$  arrivi  $T_1, \dots, T_m$  condizionatamente a  $(N_t = m)$  hanno la stessa distribuzione di probabilità delle statistiche d'ordine di  $m$  punti generati a caso in  $(0, 1]$  (vedi il paragrafo Processo Uniforme come Processo di Poisson condizionati). Pertanto, il tempo aleatorio  $T$  in cui si verifica un generico arrivo (non ordinato) in  $(0, t]$  ha una distribuzione uniforme in  $(0, t]$ ,  $T \sim U((0, t])$ . Sia  $E(T)$ , l'evento "uno (generico) degli  $m$  arrivi in  $[0, t]$  viene campionato" (ovvero classificato di tipo  $I$ ). Ovviamente se  $T = t$  allora  $P(E|T = t) = p(t)$ . Osservando che la densità di  $T$  è  $f(s) = \frac{1}{t}$ , per in  $s \in (0, t]$ , si ha

$$P(E|N_t = m) = \int_0^t P(E|T = s, N_t = m) \frac{1}{t} ds = \frac{1}{t} \int_0^t p(s) ds.$$

Poniamo  $\alpha(t) = \frac{1}{t} \int_0^t p(s) ds$  Pertanto, per  $m \geq k$ , la probabilità che esattamente  $k$  degli  $m$  arrivi si verificano in  $[0, t]$  è data da

$$P(N'_t = k | N_t = m) = \binom{m}{k} \alpha(t)^k (1 - \alpha(t))^{m-k}.$$

Mediante la decomposizione si ottiene

$$\begin{aligned} P(N'_t = k) &= \sum_{m=k}^{+\infty} \binom{m}{k} \alpha(t)^k [1 - \alpha(t)]^{m-k} \frac{e^{-\lambda t} (\lambda t)^m}{m!} = \\ &= \frac{\alpha(t)^k}{k!} \sum_{m=k}^{+\infty} [1 - \alpha(t)]^{m-k} \frac{e^{-\lambda t} (\lambda t)^m}{(m-k)!} = \\ &= e^{-\lambda t} \frac{[\lambda \alpha(t) t]^k}{k!} \sum_{m=k}^{+\infty} \frac{(\lambda(1-\alpha(t))t)^{m-k}}{(m-k)!} = \\ &= e^{-\lambda t} \frac{[\lambda \alpha(t) t]^k}{k!} \sum_{n=0}^{+\infty} \frac{(\lambda(1-\alpha(t))t)^n}{(n)!} = \\ &= e^{-\lambda t} \frac{[\lambda \alpha(t) t]^k}{k!} e^{\lambda(1-\alpha(t))t} = \\ &= \frac{e^{-\lambda \alpha(t)t} (\lambda \alpha(t) t)^k}{k!} \end{aligned}$$

Quindi

$$P(N'_t = k) = \frac{e^{-\lambda \int_0^t p(s) ds} (\lambda \int_0^t p(s) ds)^k}{k!}.$$

□

**Osservazione 9.** Notare che sebbene  $N'_t$ , per ogni  $t > 0$ , abbia una distribuzione di Poisson di parametro  $\lambda \int_0^t p(s) ds$  e che  $\{N'_t, t \geq 0\}$  ha incrementi indipendenti, il processo  $\{N'_t, t \geq 0\}$  non è un processo di Poisson, infatti gli incrementi non sono stazionari (omogenei). Tale processo viene chiamato processo di Poisson *Non Omogeneo*. In tal caso si ha, per  $t_2 > t_1 > 0$ , che  $N'_{t_2} - N'_{t_1}$  è una variabile con distribuzione di Poisson di parametro  $\lambda \int_{t_1}^{t_2} p(s) ds$ . Se  $p(t) = 1, t > 0$  si riottiene il processo di Poisson.

#### 4.4.5 Altra Definizione di un Processo di Poisson

Una ulteriore caratterizzazione del processo di Poisson che non dimostriamo e che in alcuni testi è utilizzata come definizione è la seguente.

**Definizione 21.** Un Processo di conteggio  $\{N_t, t \geq 0\}$  è un processo di Poisson di intensità  $\lambda > 0$  se

- (i)  $N_0 = 0$ ;
- (ii)  $\{N_t, t \geq 0\}$  ha incrementi indipendenti e stazionari;
- (iii)  $P(N_t = 1) = \lambda t + o(t)$  ( $o(t)$  leggasi o “ piccolo di  $t$ ”);
- (iv)  $P(N_t \geq 2) = o(t)$ .

Le proprietà di stazionarietà può essere eliminata dalla definizione precedente mediante la seguente

**Definizione 22.** Un Processo di conteggio  $\{N_t, t \geq 0\}$  è un processo di Poisson di intensità  $\lambda > 0$  se

- (i)  $N_0 = 0$ ;
- (ii)  $\{N_t, t \geq 0\}$  ha incrementi indipendenti.
- (iii)  $P(N_{t+h} - N_t = 1) = \lambda h + o(h)$ ;
- (iv)  $P(N_{t+h} - N_t \geq 2) = o(h)$ .

Commenti . . .

#### 4.4.6 Auto correlazione e Auto covarianza

La funzione di autocorrelazione di un processo  $\{X_t, t \geq 0\}$  si indica con  $R(t, s)$  ed è definita come

$$R(t, s) = E(X_t X_s).$$

La funzione di autocovarianza di un processo  $\{X_t, t \geq 0\}$  si indica con  $K(t, s)$  ed è definita come

$$K(t, s) = cov(X_t, X_s).$$

Sia  $\{N_t, t \geq 0\}$  un processo di Poisson di intensità  $\lambda > 0$  è facile verificare che

$$K(t_1, t_2) = cov(N_{t_1}, N_{t_2}) = \lambda \min\{t_1, t_2\}.$$

Infatti, supponiamo che  $t_2 > t_1$ , indicando con  $t_1 = t, t_2 = t + s$ , si ha

$$\begin{aligned} R(t, t + s) &= E(N_t \cdot N_{t+s}) = E(N_t \cdot (N_{t+s} - N_t) + N_t^2) = \\ &= \lambda t \lambda s + \lambda^2 t^2 + \lambda t \end{aligned}$$

pertanto

$$cov(N_t, N_{t+s}) = \lambda t \lambda s + \lambda^2 t^2 + \lambda t - \lambda t \lambda (t + s) = \lambda t$$

### 4.5 Processo di Poisson non omogeneo

**Definizione 23.** Un Processo di conteggio  $\{N_t, t \geq 0\}$  è un processo di Poisson non omogeneo con funzione intensità  $\lambda(t) > 0$  (in breve  $(\tilde{N}PP(\lambda(t)))$ ) se

- (i)  $N_0 = 0$ ;
- (ii)  $\{N_t, t \geq 0\}$  ha incrementi indipendenti.
- (iii)  $P(N_{t+h} - N_t = 1) = \lambda(t)h + o(h)$ ;
- (iv)  $P(N_{t+h} - N_t \geq 2) = o(h)$ .

Commenti . . .

Si dimostra il seguente risultato

**Teorema 15.** Siano  $\{N'_t, t \geq 0\}$  e  $\{N''_t, t \geq 0\}$  due indipendenti  $NPP$  di funzione intensità rispettivamente  $\lambda'(t), \lambda''(t)$ . Sia  $\{N_t, t \geq 0\}$ , con  $N_t = N'_t + N''_t$ . Si ha

- (a)  $\{N_t, t \geq 0\} \tilde{N}PP(\lambda'(t) + \lambda''(t))$
- (b) Supposto che un evento del processo  $\{N_t\}$  si verifichi al tempo  $t$ , (indipendentemente da ciò che è accaduto prima di  $t$ ), allora tale evento con probabilità  $\frac{\lambda'(t)}{\lambda'(t) + \lambda''(t)}$  proviene dal processo  $\{N'_t\}$ .

**Osservazione 10.** Dal Teorema 15 si ricava che un processo di Poisson non omogeneo  $\{N'_t, t \geq 0\}$  di funzione intensità  $\lambda'(t) \leq \lambda$  limitata può essere visto come un processo di conteggio ottenuto mediante campionamento da un processo di poisson  $\{N_t, t \geq 0\}$  di intensità  $\lambda$  in cui la probabilità di campionare è pari a

$$p(t) = \frac{\lambda'(t)}{\lambda}.$$

Infatti basta osservare che  $\{N_t = N'_t + N''_t\}$  con  $\{N''_t\}$  un processo di Poisson non omogeneo di funzione intensità  $\lambda''(t) = \lambda - \lambda'(t)$ .

Pertanto, dato un processo di Poisson non omogeneo  $\{N'_t, t \geq 0\}$  con funzione intensità  $\lambda'(t)$  si ha che, per ogni  $t > 0$ , la variabile  $N'_t$  ha una distribuzione di Poisson di parametro

$$\lambda \int_0^t \frac{\lambda'(s)}{\lambda} ds = \int_0^t \lambda'(s) ds,$$

quindi

$$E(N'_t) = \int_0^t \lambda'(s) ds.$$

La funzione

$$m(t) = \int_0^t \lambda'(s) ds$$

dicesi funzione valore medio di un processo di Poisson non omogeneo. Talvolta  $m(t)$  viene indicata con  $\Lambda(t)$  (in tal caso con  $\Lambda'(t)$ ).

Inoltre, considerando il processo di Poisson non omogeneo che parte al tempo  $t_0$  si ha che la variabile  $N'_{t+s} - N'_s$  ha una distribuzione di Poisson di parametro

$$\lambda \int_s^t \frac{\lambda'(y)}{\lambda} dy = \int_s^t \lambda'(y) dy,$$

quindi

$$E(N'_{t+s} - N'_s) = \int_s^t \lambda'(y) dy = m(t+s) - m(s).$$

#### 4.5.1 Tempi di attesa

Dato un processo di Poisson non omogeneo  $\{N_t, t \geq 0\}$  con funzione intensità  $\lambda(t)$  e indicando con  $f_n$  la densità del tempo di attesa  $T_n$  si ha, per  $t > 0$ ,

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

con  $\Lambda(t) = \int_0^t \lambda'(s) ds$ . In particolare si ha

$$P(T_1 > t) = P(X_1 > t) = e^{-\Lambda(t)}$$

e si può verificare che i tempi  $X_n$  di interarrivo non sono nè indipendenti nè identicamente distribuiti.

**Esempio 34** (Esempio 5.24 in [10]). Sia  $\{N_t, t \geq 0\}$  un processo di Poisson non omogeneo con funzione intensità  $\lambda(t)$  (numeri di clienti in un ora) data da

$$\lambda(t) = \begin{cases} 0, & 0 < t \leq 8; \\ 5 + 5(t - 8), & 8 < t \leq 11; \\ 20 & 11 < t \leq 13; \\ 20 - 2(t - 13) & 13 < t \leq 17; \\ 0, & 17 < t \leq 24; \end{cases}$$

Calcolare

La probabilità che nessun cliente arrivi tra le 8 : 30 e le 9.30.

Il numero medio di arrivi in tale fascia oraria.

**Esercizio 15** (Esempio 5.18 in [10] (Infinite Server Queue)). Supponiamo che il numero di clienti che arrivano ad una stazione di servizio seguano un processo di Poisson di intensità  $\lambda$ . Non appena un cliente arriva esso è immediatamente servito da uno tra infiniti server. Supponiamo che i tempi di servizio dei server siano variabili aleatorie i.i.d. con funzione di ripartizione  $G(t)$ . Calcolare la distribuzione di  $X(t)$ , il numero di clienti che sono stati serviti al tempo  $t$ . Calcolare la distribuzione di  $Y(t)$ , il numero di clienti che si stanno servendo al tempo  $t$ .

## 4.6 Processo di Poisson Composto

Un processo stocastico  $\{Z_t, t \geq 0\}$  dicesi processo di Poisson composto (CPP) se

$$Z_t = \sum_{i=1}^{N_t} Z_i$$

con  $Z_i$  i.i.d.,  $\{N_t\}$  è un  $PP(\lambda)$ ,  $\{N_t\}$  e  $\{Z_i\}$  indipendenti.

Se  $Z_i \equiv 1$  allora  $Z_t = N_t$  si ha che  $\{Z_t, t \geq 0\}$  è un processo di Poisson. Il processo di Poisson composto è utile nei casi in cui si verificano arrivi simultanei nello stesso istante di tempo. Ad esempio, supponiamo che degli autobus arrivino allo stadio seguano un processo di Poisson e che i numeri aleatori di tifosi che scendono da ogni autobus siano i.i.d. e siano indipendenti dal numero di autobus in arrivo. Se indichiamo con  $Z_t$  il numero di tifosi che arrivano al tempo  $t$ , con  $Z_i$  il numero di tifosi che dentro l' $i$ -esimo autobus e con  $N_t$  il numero di bus arrivati al tempo  $t$  allora  $\{Z_t, t \geq 0\}$  è un CPP.

Verificare che

$$E(Z_t) = \lambda t E(Z_i).$$

Ricordiamo che la funzione caratteristica  $\phi_Z(\theta)$  è definita come

$$\phi_Z(\theta) = E(e^{i\theta Z}).$$

Indichiamo con  $\phi_i(\theta)$  la f.c. di  $Z_i$  e verifichiamo che la funzione caratteristica di  $Z_t$  è

$$E(e^{i\theta Z_t}) = e^{-\lambda t(1-\phi_i(\theta))}$$

Si ha

$$\begin{aligned} E(e^{i\theta Z_t} | N_t = n) &= E(e^{i\theta(Z_1+Z_2+\dots+Z_N)} | N_t = n) = \\ &= E(e^{i\theta(Z_1+Z_2+\dots+Z_n)}) = \phi_i(\theta)^n \end{aligned}$$

Pertanto

$$E(e^{i\theta Z_t}) = E(E(e^{i\theta Z_t} | N_t)) = E(\phi_i(\theta)^{N_t}).$$

Ricordiamo che dato una variabile aleatoria discreta a valori interi non negativi  $N$ ,  $N \in \{0, 1, 2, \dots\}$ , la funzione generatrice della distribuzione di probabilità è definita come il valore atteso di  $Y = z^N$  per ogni  $z$ , ovvero

$$p(z) = E(z^N) = \sum_{n=0}^{+\infty} z^n P(N = n) = \sum_{n=0}^{+\infty} z^n p_n$$

dove  $p_n = P(N = n)$ . Osservando che lo sviluppo in serie di Mac-Laurin  $\sum_{n=0}^{+\infty} z^n \frac{p^{(n)}(0)}{n!}$  della  $p(z)$  è unico, nel senso che i coefficienti sono univocamente determinati, per ogni  $n$ , si ha

$$p_n = \frac{p^{(n)}(0)}{n!}.$$

Pertanto data la funzione generatrice di probabilità  $p(z)$  le probabilità  $p_n$  si possono ottenere dai coefficienti dello sviluppo in serie di  $p(z)$ .

Sia  $N \sim P(\lambda)$ , calcoliamo  $p(z)$ .

$$\begin{aligned} p(z) = E(z^N) &= \sum_{n=0}^{+\infty} z^n \frac{\lambda^n e^{-\lambda}}{n!} = \\ &= \dots e^{-\lambda} e^{\lambda z} = e^{-\lambda(1-z)} \end{aligned}$$

Ritornando al calcolo di

$$E(e^{i\theta Z_t}) = E(E(e^{i\theta Z_t} | N_t)) = E(\phi_i(\theta)^{N_t}).$$

si ha

$$E(e^{i\theta Z_t}) = E(\phi_i(\theta)^{N_t}) = e^{-\lambda t(1-\phi_i(\theta))}$$

**Esercizio 16.** Sia  $\{X_t, t \geq 0\}$  un processo definito nel seguente modo

$$X_t = \sum_{i=1}^{N_t} X_i,$$

con  $X_i$  indipendenti e con distribuzione geometrica di parametro  $0 < p < 1$ ,  $\{N_t\}$  un  $PP(\lambda)$ ,  $\{N_t\}$  e  $\{X_i\}$  indipendenti. Calcolare la funzione caratteristica di  $X_t$ .

Si ha che  $\{X_t, t \geq 0\}$  è un CPP, pertanto

$$\phi_{X_t}(\theta) = e^{-\lambda t \left(1 - p \frac{e^{i\theta}}{1 - (1-p)e^{i\theta}}\right)}.$$

**Esercizio 17.** Sia  $\{X_t, t \geq 0\}$  un processo definito nel seguente modo

$$X_t = \sum_{i=1}^{N_t} X_i,$$

con  $X_i$  indipendenti e con distribuzione geometrica di parametro  $0 < p < 1$ ,  $\{N_t\}$  un *NPP* con funzione valor medio  $m(t)$ ,  $\{N_t\}$  e  $\{X_i\}$  indipendenti. Verificare che la funzione caratteristica di  $X_t$  è

$$\phi_{X_t}(\theta) = e^{-m(t) \left(1 - \frac{pe^{i\theta}}{1 - (1-p)e^{i\theta}}\right)}.$$

## 4.7 Processo di Poisson Randomizzato

(Vedi [1, 12])

L'intensità  $\lambda$  è una variabile aleatoria con densità  $g(\lambda)$ .

$$P(N_{t+s} - N_s = n) = \int_0^\infty \frac{e^{-\lambda t} (\lambda t)^n}{n!} g(\lambda) d\lambda$$

(Trasformata di Laplace)

$$P(X_1 > t) = \int_0^\infty e^{-\lambda t} g(\lambda) d\lambda$$

## 4.8 Processi di Punto: cenni

(DA RIVEDERE) Un'altra generalizzazione del PP è quella al caso bidimensionale (o n-dimensionale). In tal caso si parla di punti di Poisson nel piano. Però, occorre introdurre una opportuna indicizzazione degli insiemi  $A$  di un piano. In alcuni casi si considera come relazione d'ordine l'area  $|A|$  di un insieme  $A$ . Brevemente, diremo che  $\{N_A, A \subset \mathbb{R}^2\}$  è un processo di punto di tipo di Poisson se

- $N_A \sim P(\lambda|A)$ ;
- $N_{A_i}$  indipendenti se  $A_i$  son disgiunti.

Ancora più in generale (caso non omogeneo) si può richiedere che

$$N_A \sim P\left(\lambda \int_A p(x) dx\right)$$

# Capitolo 5

## Esercizi

### 5.1 Esercizi (svolti e non) sui processi di conteggio

#### 5.1.1 Esercizi

Nei libri [10, 9] sono presenti svariate applicazioni del Processo di Poisson la cui lettura è largamente consigliata. Una di tale applicazioni potrebbe essere argomento a piacere da portare all'esame, mentre uno degli esercizi seguenti potrebbe essere richiesto all'esame.

**Esercizio 18.** Siano  $X_i \sim \text{Exp}(\lambda_i)$ ,  $i = 1, 2$ , due variabili aleatorie indipendenti. Indichiamo con  $F_i$  la funzione di ripartizione di  $X_i$ ,  $i = 1, 2$  e con  $S_i$  la funzione di sopravvivenza di  $X_i$ ,  $i = 1, 2$ . Verificare che

1. 
$$P(X_1 < X_2) = \int_0^{+\infty} \int_y^{+\infty} \lambda_1 \lambda_2 e^{-\lambda_1 u} e^{-\lambda_2 v} du dv = \frac{\lambda_1}{\lambda_1 + \lambda_2};$$
2.  $Z = \min(X_1, X_2) \sim \text{Exp}(\lambda_1 + \lambda_2)$ ;
3.  $S_z(z) = P(Z > z) = S_1(z)S_2(z)$ ;
4.  $\mathbb{E}(Z) \leq \frac{1}{\lambda_i}$ ,  $i = 1, 2$ ;
5.  $T = \max(X_1, X_2)$ ,  $F_t(t) = P(T \leq t) = F_1(t)F_2(t)$ .
6. se  $\lambda_1 = \lambda_2 = \lambda$ ,  $Y = X_1 + X_2 \sim G_{2,\lambda}$ .

**Esercizio 19.** Siano  $X \sim \text{Exp}(\lambda_1)$  e  $Y \sim \text{Exp}(\lambda_2)$  variabili aleatorie indipendenti. Posto  $Z = X + Y$  calcoliamo la densità di  $Z$ . Ricordando l'operatore di

convoluzione si ha

$$\begin{aligned} g(z) &= \int_{-\infty}^{+\infty} f_1(t)f_2(z-t)dt = \\ &= \int_0^z \lambda_1 e^{-\lambda_1 t} \lambda_2 e^{-\lambda_2(z-t)} dt = \dots \\ &= \lambda_1 \lambda_2 e^{-\lambda_2 z} \int_0^z e^{-(\lambda_1 - \lambda_2)t} dt \\ &= \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 z} + \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_2 z}. \end{aligned}$$

**Esercizio 20.** Le funzioni caratteristiche di due numeri aleatori  $X, Y$  indipendenti sono rispettivamente  $\phi_X(t) = e^{(e^{it}-1)}$  e  $\phi_Y(t) = e^{4(e^{it}-1)}$ . Posto  $Z = X + Y$ , calcolare la funzione caratteristica  $\phi_Z(t)$ , la previsione  $m$  di  $Z$  e la probabilità  $p$  dell'evento condizionato  $(Z \geq 2 | X \leq 1)$ .

$$\phi_Z(t) = \qquad m = \qquad p =$$

**Esercizio 21.** Dato un processo di Poisson di intensità  $\alpha = 1$ , siano  $N_s$  ed  $N_t$  rispettivamente il numero di arrivi in  $(0, s)$  e  $(0, t)$ , con  $0 < s < t$ . Calcolare la covarianza  $\sigma_{st}$  di  $N_s, N_t$ , la probabilità  $p = P(N_t = N_s)$  e la probabilità  $\gamma = P(N_s = n | N_t = n)$ .

$$\sigma_{st} = \qquad p = \qquad \gamma =$$

**Esercizio 22.** Dato un processo di Poisson di intensità  $\alpha = 1$ , sia  $t \in (0, a)$ . Calcolare la probabilità  $p_1$  dell'evento  $(N_t < 2)$ . Inoltre, supposto vero l'evento  $(N_a = 5)$ , calcolare la probabilità  $p_0$  dell'evento  $(N_t = 0)$  (condizionata a  $(N_a = 5)$ ) e la funzione generatrice  $\psi(t)$  di  $N_t | (N_a = 5)$

$$p_1 = \qquad p_0 = \qquad \psi(t) =$$

**Esercizio 23.** Siano  $X_{(1)}, X_{(2)}, X_{(3)}$  le prime tre statistiche d'ordine di cinque numeri aleatori  $X_1, \dots, X_5$ , indipendenti e con distribuzione uniforme in  $[0, 3]$ . Calcolare la probabilità  $p$  dell'evento  $(X_{(3)} \leq 1)$ , la previsione  $m$  e la densità  $g(y)$  del numero aleatorio  $Y = \frac{1}{3}X_{(3)}$ .

$$p = \qquad m = \qquad g(y) =$$

**Esercizio 24.** Dati 5 numeri aleatori  $X_1, \dots, X_5$ , indipendenti e con distribuzione uniforme nell'intervallo  $[0, 3]$ , siano  $X_{(1)}, \dots, X_{(5)}$  le corrispondenti statistiche d'ordine. Calcolare la probabilità  $p$  dell'evento  $(X_{(2)} \leq 1) \wedge (X_{(4)} \geq 2)$ , la previsione  $m$  di  $3X_{(2)} - 2X_{(3)}$  e la densità di probabilità  $g(y)$  di  $Y = X_{(4)} - X_{(2)}$ , per ogni  $y \in (0, 3)$ .

$$p = \qquad ; \quad m = \qquad ; \quad g(y) =$$

**Esercizio 25.** Con riferimento all'esercizio precedente, sia  $Z$  il numero aleatorio di punti che cadono nell'intervallo  $[1, 3]$ . Calcolare la funzione caratteristica  $\phi(t)$  di  $Z$  e la funzione caratteristica  $\psi(t)$  di  $U = 5 - Z$ .

$$\phi(t) = \qquad ; \quad \psi(t) =$$

**Esercizio 26.** Data una successione di numeri aleatori indipendenti e ugualmente distribuiti,  $T_1, \dots, T_n, \dots$ , sia  $f(t) = e^{-t}, t \geq 0$ , con  $f(t) = 0$  altrove, la densità di ciascun  $T_i$ . Calcolare la densità  $g(w)$  di  $W_n = T_1 + \dots + T_n$ , la probabilità  $p$  dell'evento  $A = (W_n \leq t, W_{n+1} > t)$  e la densità condizionata  $l(x)$  di  $T_1|A$ .

$$g(w) = \qquad \qquad \qquad p = \qquad \qquad \qquad l(t) =$$

**Esercizio 27.** Quattro amici si incontrano in una città  $\mathcal{C}$  dove arrivano a caso nell'intervallo (misurato in ore)  $[0, 5]$ . Calcolare la probabilità  $p$  che 3 amici arrivino con un ritardo, rispetto all'istante 0, inferiore a 2 ore e uno arrivi con un ritardo superiore a 3 ore. Inoltre, indicando con  $(X_{(1)}, X_{(2)})$  il vettore aleatorio relativo agli istanti del primo e secondo arrivo, calcolare la densità congiunta  $\phi_{12}(x, y)$  e la probabilità  $\alpha$  dell'evento condizionato  $(X_{(2)} \leq 3|X_{(1)} \leq 2)$ .

$$p = \qquad \qquad \qquad ; \quad \phi_{12}(x, y) = \qquad \qquad \qquad ; \quad \alpha =$$

**Esercizio 28.** Due numeri aleatori a valori interi non negativi  $X$  e  $Y$ , indipendenti e ugualmente distribuiti, hanno la funzione caratteristica  $\phi(t) = e^{e^{it} - 1}$ . Posto  $Z = X + Y$ , calcolare la probabilità  $p_n$  dell'evento  $Z = n$ , con  $n = 0, 1, \dots$ , la previsione  $\mu$  di  $Z$  e la probabilità condizionata  $\alpha = P(X = h|Z = n)$ .

$$p_n = \qquad \qquad \qquad ; \quad \mu = \qquad \qquad \qquad ; \quad \alpha =$$

**Esercizio 29.** Dati i numeri aleatori  $Z_1, \dots, Z_5$ , indipendenti e con distribuzione uniforme in  $[0, 3]$ , e posto  $X = \text{Min}Z_i, Y = \text{Max}Z_i$ , calcolare la densità congiunta  $f(x, y)$  di  $(X, Y)$ , la probabilità  $p$  dell'evento  $(X + Y \leq 2)$  e la previsione  $\mu$  della media aritmetica di  $X, Y$ .

$$f(x, y) = \qquad \qquad \qquad p = \qquad \qquad \qquad \mu =$$

**Esercizio 30.** Data una successione di numeri aleatori indipendenti e ugualmente distribuiti,  $T_1, \dots, T_n, \dots$ , sia  $f(t) = e^{-t}, t \geq 0$ , con  $f(t) = 0$  altrove, la densità di ciascun  $T_i$ . Calcolare la densità  $g(w)$  di  $W_n = T_1 + \dots + T_n$ , la probabilità  $p$  dell'evento  $A = (W_n \leq t, W_{n+1} > t)$  e la densità condizionata  $l(x)$  di  $T_1|A$ .

$$g(w) = \qquad \qquad \qquad p = \qquad \qquad \qquad l(t) =$$

## Capitolo 6

# Catene di Markov continue: cenni

### 6.1 Catene di Markov continue (CTMC): Cenni

Sia  $\{X(t), t \geq 0\}$  un processo stocastico a valori interi non negativi (in generale a valori in uno spazio degli stati discreto  $S$ ), diremo che  $\{X(t), t \geq 0\}$  è una catena di Markov continua (CTMC) (vedi [10]) se, per ogni  $s, t \geq 0$ , e per interi non negativi  $i, j, x(u)$ , con  $0 \leq u < s$ , si ha

$$\begin{aligned} P(X(t+s) = j | X(s) = i, X(u) = x(u), 0 \leq u < s) = \\ = P(X(t+s) = j | X(s) = i). \end{aligned}$$

In altre parole, una catena di Markov continua è un processo stocastico con la proprietà di Markov che la distribuzione condizionata del futuro  $X(t+s) = j$  dato il presente  $X(s)$  e il passato  $X(u) = x(u), 0 \leq u < s$  dipende solo dal presente ed è indipendente dal passato. In teoria per definire una CTMC occorre assegnare

$$P(X(t+s) = j | X(s) = i)$$

per ogni  $s, t \geq 0$  e per ogni coppia  $i, j$ . Se  $\{X(t), t \geq 0\}$  gode pure della seguente proprietà

$$P(X(t+s) = j | X(s) = i) = P(X(t) = j | X(0) = i) \text{ è indipendente da } s$$

allora diremo che  $\{X(t), t \geq 0\}$  è una catena di Markov continua stazionaria (omogenea). D'ora in poi considereremo solo CTMC stazionarie. In tal caso per definire una CTMC occorre ancora assegnare

$$p_{ij}(t) = P(X(t) = j | X(0) = i)$$

per ogni  $t \geq 0$  e per ogni coppia  $i, j$ .

Indichiamo con  $X(s:t)$  i valori assunti dalla catena  $X(u)$  per  $u$  appartenenti in  $(s, t)$ , in maniera analoga si definiscono  $X[s:t]$ ,  $X(s:t]$  e  $X[s:t)$ . Per ogni

intero  $i$  sia  $T_i =$  “il tempo aleatorio di permanenza della catena nello stato  $i$  prima di passare ad un qualsiasi altro stato  $j \neq i$ . Osserviamo che

$$P(T_i > t) = P(X(0 : t] = i | X(0) = i).$$

Osserviamo che per ogni  $t, s \geq 0$  si ha

$$\begin{aligned} P(T_i > s + t | T_i > s) &= P(X(0 : t + s] = i | X[0 : s] = i) \\ &= P(X(s : t + s] = i | X[0 : s] = i) \\ &\stackrel{\text{Mark. Prop}}{=} P(X(s : t + s] = i | X(s) = i) \\ &\stackrel{\text{Stazionarietà}}{=} P(X(0 : t] = i | X(0) = i) \\ &= P(T_i > t). \end{aligned}$$

Quindi, per ogni intero  $i$ , la variabile  $T_i$  gode della **proprietà di assenza di memoria**,

$$P(T_i > s + t | T_i > s) = P(T_i > t).$$

Pertanto tutte le variabili  $T_i$  hanno distribuzione esponenziale!. In altre parole una CTMC è un processo stocastico con le proprietà

1. ogni volta che entra nello stato  $i$  il tempo di permanenza nello stato  $i$  prima di fare una transizione in un altro stato ha distribuzione esponenziale con valore atteso  $\frac{1}{v_i}$ ;
2. quando il processo lascia lo stato  $i$ , con probabilità  $P_{ij}$  esso entra nello stato  $j$

Le  $P_{ij}$  sono ovviamente non negative e tali che  $\sum_j P_{ij} = 1$ . Quindi una CTMC è un processo stocastico che si sposta da uno stato ad un altro secondo una catena di Markov discreta,  $P_{ij}$ , ma in cui il tempo di permanenza in uno stato ha distribuzione esponenziale. Inoltre, il tempo di permanenza  $T_i$  e  $P_{ij}$  devono essere indipendenti, altrimenti se essi non fossero indipendenti allora l'informazione del tempo trascorso dalla catena nello stato  $i$  sarebbe rilevante per la probabilità che la catena passi ad un nuovo stato  $j$  e questo va in contrasto con la proprietà di Markov. Attenzione  $p_{ij}(t) \neq P_{ij}$ , infatti

$$p_{ij}(t) = P(X(t) = j | X(0) = i), P_{ij} = P(X(S_{n+1}^+) = j | X(S_n^+) = i),$$

dove  $S_n$  è “il tempo di attesa sino alla  $n$ -esima transizione da uno stato ad un altro”, quindi  $X(S_{n+1}^+)$  rappresenta lo stato della catena subito dopo l' $n$ -esima transizione.

**Osservazione 11.** In una CTMC:

- la matrice  $P = (P_{ij})$  è una matrice di transizione di una catena di markov discreta DTMC;
- $P_{ii} = 0$  per ogni  $i$ ,  $P$  ha la diagonale nulla;

- le proprietà di raggiungibilità, comunicazione, transitorietà, ricorrenza sono legate a quelle della corrispondente DTCM;
- i tempi  $T_i \sim \text{Exp}(v_i)$ , con  $\mathbb{E}(T_i) = \frac{1}{v_i}$
- $v_i$  si può interpretare come l'intensità, il tasso, di transizione dallo stato  $i$  ad un altro stato;
- poiché  $\sum_j P_{ij} = 1$  allora  $q_{ij} = v_i P_{ij}$  si può interpretare come il tasso di transizione da  $i$  a  $j$ ;
- dalle  $v_i, P_{ij}$  di possono ricavare le  $q_{ij}$  e viceversa, infatti

$$v_i = v_i \sum_j P_{ij} = \sum_{j, j \neq i} q_{ij}$$

e inoltre

$$P_{ij} = \frac{q_{ij}}{v_i}.$$

## 6.2 Processi di Nascita e di Morte

- Gli stati  $X(t) = 0, 1, \dots$  rappresentano il numero di individui nella pop.
- Le nascite e le morti dipendono dallo stato  $n$  in cui si trova il processo.
- Se  $X(t) = n$  allora si ha un nuovo arrivo (nascita) con un tasso esponenziale  $\lambda_n$  per unità di tempo.
- Se  $X(t) = n$  allora si ha una partenza (morte) con un tasso esponenziale  $\mu_n$  per unità di tempo.
- I tempi di attesa per le nascite e per le morti sono indipendenti.
- Il processo di nascita e morte è una CTMC in cui dallo stato  $n$  si può andare solamente allo stato  $n + 1$  o allo stato  $\max\{n - 1, 0\}$ .

Inoltre, sia  $i > 0$ ,

- se  $T_{B_i}$  e  $T_{D_i}$  sono i tempi di attesa, risp., per la prossima nascita e per la prossima morte, allora  $T_i = \min(T_{B_i}, T_{D_i})$ ;
- poichè  $T_{B_i}$  e  $T_{D_i}$  sono esponenziali indipendenti si ha che  $T_i$  ha un tasso  $v_i = \lambda_i + \mu_i$ ;
- quando la catena lascia lo stato  $i$ ,  $i > 0$  può andare a  $i + 1$ , se si verifica una nascita prima di una morte, o allo stato  $i - 1$  se si verifica una morte prima di una nascita;
- pertanto si ha

$$v_i = \lambda_i + \mu_i, \quad P_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad P_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}, \quad i > 0.$$

Se  $i = 0$  allora si può verificare solo una nascita, quindi

$$v_0 = \lambda_0, \quad P_{01} = 1.$$

**Esercizio 31.** Calcolare i tassi di transizione  $q_{ij}$ .

**Esempio 35** (Processo di Poisson). Consideriamo un processo di nascita e morte in cui

$$\mu_n = 0, \lambda_n = \lambda, \quad \forall n \geq 0.$$

In tal caso  $X(t)$  è un  $PP(\lambda)$ .

Un processo di nascita e morte in cui  $\mu_n = 0$  dicesi processo di nascita pura.

**Esempio 36.** Supponiamo che in una setta religiosa non ci siano abbandoni ma solo arrivi. Inoltre, supponiamo che ogni membro della setta agisce in maniera indipendentemente dagli altri e in un tempo distribuito esponenzialmente con media  $\frac{1}{\lambda}$  invita un amico esterno (tutti diversi) a far parte della setta. Se la popolazione è composta da  $n$  individui allora  $T_n$  sarà il minimo tra tutti i tempi di attesa relativi alla “nascita” da parte di ciascun individuo, pertanto  $T_n \sim Exp(n\lambda)$ . Se  $X(t)$  è il numero dei componenti della setta al tempo  $t$  allora  $\{X(t), t \geq 0\}$  sarà un processo di nascita pura con  $\lambda_n = n\lambda$ .

**Esercizio 32.** Un interessante modello di crescita lineare con immigrazione e morte è riportato in [10] (Tale esempio può essere portato all’esame come argomento a piacere). In tal caso si ha

$$\mu_n = n\mu, n \geq 1, \quad \lambda_n = n\lambda + \theta, n \geq 0.$$

**Esempio 37** (Coda  $M/M/1$ ). Una coda  $M/M/1$  è un processo di Nascita e di Morte con

$$\mu_n = \mu, n \geq 1, \quad \lambda_n = \lambda, n \geq 0.$$

In tal caso i clienti arrivano con un tasso  $\lambda$  per unità di tempo. Si ha un solo sportello, il quale serve il cliente con un tasso  $\mu$  per unità di tempo. Ovvero i tempi di inter-arrivo sono esponenziali indipendenti di parametro  $\lambda$  e i tempi di (inter-) partenza sono esponenziali indipendenti di parametro  $\mu$ . Si chiama  $M/M/1$  perchè

- $M$  gli inter-arrivi sono markoviani (processo di Poisson)
- $M$  i tempi di servizio sono markoviane (sono esponenziali)
- 1, si ha un solo sportello.

Continua ...

## Capitolo 7

# Classificazione di un processo e cenni al moto Browniano

### 7.1 Caratterizzazione dei processi stocastici

Dato un processo stocastico  $\{X(t), t \in T\}$ . Fissato il tempo  $t_1$ , poniamo  $X_1 = X_{t_1}$ . La distribuzione di  $X_1$  è detta distribuzione del primo ordine del processo  $\{X(t), t \in T\}$ . La distribuzione congiunta della coppia  $(X(t_1), X(t_2))$ , con  $t_1, t_2$  fissati, dicesi distribuzione del secondo ordine del processo  $\{X(t), t \in T\}$ . Più in generale, la distribuzione congiunta di  $(X(t_1), \dots, X(t_n))$  per fissati  $t_1, t_2, \dots, t_n$  dicesi distribuzione di ordine  $n$ -esimo di  $\{X(t), t \in T\}$ . La caratterizzazione completa di  $\{X(t), t \in T\}$  richiede la conoscenza di tutte le distribuzioni per  $n$  che tende a infinito. In realtà si dimostra che un processo è ben definito se sono date le distribuzioni di probabilità di ogni ordine  $n$  finito (Kolmogorov).

#### 7.1.1 Funzione media, correlazione e covarianza

La funzione media di un processo  $\{X(t), t \in T\}$  è definita da

$$\mu_X(t) = \mathbb{E}(X(t)).$$

Poichè  $t$  solitamente rappresenta un istante di tempo, la funzione media  $\mu_X(t)$  viene a volte detta media spaziale del processo  $\{X(t), t \in T\}$ . Notare che la funzione  $\mu_X(t)$  potrebbe non essere definita.

La funzione di autocorrelazione in un certo senso misura la dipendenza lineare tra due variabile aleatorie del processo ed è definita come segue.

$$R_X(t, s) = \mathbb{E}(X(t)X(s)).$$

Notare che

$$R_X(t, s) = R_X(s, t), R_X(t, t) = \mathbb{E}(X^2(t)).$$

La funzione di autocovarianza di un processo è definita da

$$K_X(t, s) = Cov(X(t), X(s)) = R_X(t, s) - \mu_X(t)\mu_X(s).$$

In particolare

$$\sigma_X^2(t) = K_X(t, t).$$

Si ha che  $K_X(t, s) = R_X(t, s)$  se  $\mu_X(t) = 0$  per ogni  $t$ . Notare che anche le funzioni di autocorrelazione e di autocovarianza potrebbero non essere definite.

### 7.1.2 Processi stazionari

Un processo stocastico  $\{X(t), t \geq 0\}$  dicesi omogeneo o stazionario se la distribuzione di probabilità non cambia quando si attua una traslazione sul parametro  $t$ .

**Definizione 24.** Un processo stocastico  $\{X(t)\}$  dicesi omogeneo o stazionario (*in senso stretto*) se per ogni intero  $n$  e per ogni  $t_1, t_2, \dots, t_n$  e per ogni valore di  $\tau$  si ha

$$\begin{aligned} P(X(t_1) < x_1, X(t_2) < x_2, \dots, X(t_n) < x_n) = \\ P(X(t_1 + \tau) < x_1, X(t_2 + \tau) < x_2, \dots, X(t_n + \tau) < x_n) \end{aligned} \quad (7.1)$$

$$\forall (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

In altre parole  $\{X(t), t \geq 0\}$  è omogeneo se per ogni intero  $n$  e per ogni  $t_1, t_2, \dots, t_n$  e per ogni valore di  $\tau$  i vettori  $(X(t_1), \dots, X(t_n)), (X(t_1 + \tau), \dots, X(t_n + \tau))$  hanno la stessa distribuzione di probabilità. Ed ancora, se scegliendo  $\tau$  come origine, il nuovo processo  $\{X(t + \tau), t \geq 0\}$  coincide in distribuzione con  $\{X(t), t \geq 0\}$ .

Se  $\{X(t)\}$  è stazionario per la distribuzione del primo ordine si ha che  $X(t)$  e  $X(t + \tau)$  avranno la stessa distribuzione di probabilità

$$F_{X(t)}(x) = F_{X(t+\tau)}(x), \forall x, \tau, t$$

pertanto (se esistono)  $\mu_X(t) = costante = \mu$  e  $\sigma_X(t)^2 = costante = \sigma^2$ . Inoltre se  $t_2 > t_1$  si ha

$$F_{X(t_1), X(t_2)}(x_1, x_2) = F_{X(t_1), X(t_2 - t_1)}(x_1, x_2), \forall x, \tau, t$$

**Esempio 38.** Sia  $\{X(t), t \geq 0\}$ , con  $X(t) = N(t + h) - N(t)$  per ogni  $t \geq 0$ , e con  $h > 0$  fissato. Se  $\{N(t), t \geq 0\}$  è un processo di Poisson di intensità  $\lambda$  allora  $\{X(t), t \geq 0\}$  è stazionario (individua gli incrementi di ampiezza  $h$ ). In tal caso si ha, per ogni  $t$ ,  $X(t) \sim P(\lambda h)$ .

Se la condizione di stazionarietà (7.1) non è valida per tutti i valori di  $n$ , ma solo per  $n \leq k$  si dirà che il processo è stazionario (stretto) di ordine  $k$ . Se  $\{X(t), t \geq 0\}$  è stazionario (stretto) di ordine 2 si dirà che  $\{X(t), t \geq 0\}$ . In tal caso si dimostra che

$$\begin{aligned} \mu_X(t) &= costante = \mu \\ R_X(t, s) &= \mathbb{E}(X(t)X(s)) = R_X(|t - s|) \end{aligned} \quad (7.2)$$

cioè la funzione media è costante e la funzione di autocorrelazione dipende solo dall'ampiezza  $|t - s|$ . Se valgono le condizioni (7.2) non è detto che il processo sia stazionario (in senso stretto) di ordine 2. In alcuni testi (vedi ad es. [10]) si definisce un processo stazionario di ordine 2 se valgono (7.2). Per evitare confusione diremo che un processo è stazionario *debole* di ordine 2 se valgono le (7.2). Pertanto un processo stazionario debole potrebbe non essere un processo stazionario (stretto). Se  $\{X(t), t \geq 0\}$  è stazionario debole di ordine 2 allora si definiscono le seguenti funzioni di autocovarianza e di auto correlazione.

$$R_X(s) = \mathbb{E}(X(t)X(t+s)), \quad K_X(s) = Cov(X(t), X(t+s))$$

**Esempio 39** (Processo autoregressivo). Siano  $Z_0, Z_1, Z_2, \dots$  una successione di variabili aleatorie incorrelate con  $\mathbb{E}(Z_n) = 0, n \geq 0$  e

$$Var(Z_n) = \begin{cases} \sigma^2 \frac{1}{1-\lambda^2}, & n = 0, \\ \sigma^2, & n \geq 1 \end{cases}$$

con  $\lambda^2 < 1$ . Definiamo

$$\begin{aligned} X_0 &= Z_0 \\ X_n &= \lambda X_{n-1} + Z_n. \end{aligned}$$

Il processo  $\{X_n, n \geq 0\}$  dicesi processo autoregressivo del primo ordine. Si può verificare che  $X_n = \sum_{i=0}^n \lambda^{n-i} Z_i$  e che

$$cov(X_n, X_{n+m}) = \frac{\sigma^2 \lambda^m}{1 - \lambda^2}.$$

Infatti

$$\begin{aligned} cov(X_n, X_{n+m}) &= \sum_{i=0}^n \lambda^{n-i} \lambda^{n+m-i} var(Z_i, Z_i) \\ &= \sigma^2 \lambda^{2n+m} \left( \frac{1}{1-\lambda^2} + \sum_{i=0}^n \lambda^{-2i} - 1 \right) = \\ &= \lambda^{2n+m} \left( \frac{1}{1-\lambda^2} + \frac{1-\lambda^{-2(n+1)}}{1-\lambda^{-2}} - 1 \right) = \\ &= \lambda^{2n+m} \left( \frac{1}{1-\lambda^2} + \frac{(1-\lambda^{-2(n+1)})\lambda^2}{\lambda^2-1} - 1 \right) = \\ &= \lambda^{2n+m} \left( \frac{1}{1-\lambda^2} + \frac{(1-\lambda^{-2(n+1)})\lambda^2}{\lambda^2-1} - 1 \right) \dots \end{aligned}$$

Poichè  $E(X_n) = 0$  e  $cov(X_n, X_{n+m}) = f(m)$  si ha che  $\{X_n\}$  è un processo stazionario debole del secondo ordine.

### 7.1.3 Processi a incrementi indipendenti e stazionari

Si dice che un processo aleatorio  $\{X(t), t \geq 0\}$  ha incrementi *indipendenti* se, per ogni  $0 < t_1 < t_2 \dots < t_n$  le variabili aleatorie

$$X(0), X(t_1) - X(0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$$

sono indipendenti. Se  $\{X(t), t \geq 0\}$  ha incrementi indipendenti e, per ogni  $h \geq 0, t > s > 0$ ,  $X(t) - X(s)$  e  $X(t+h) - X(s+h)$  hanno la stessa distribuzione

si dice che  $\{X(t), t \geq 0\}$  ha incrementi *indipendenti e stazionari*. Se  $\{X(t), t \geq 0\}$  ha incrementi indipendenti e stazionari e  $X(0) = 0$  si ha

$$\begin{aligned}\mathbb{E}(X(t)) &= \mu_1 t, \quad \mu_1 = \mathbb{E}(X(1)) \\ \text{Var}(X(t)) &= \sigma_1^2 t, \quad \sigma_1^2 = \text{Var}(X(1)).\end{aligned}$$

Da tali equazioni si evince che un processo a incrementi ind. e staz non è un processo stazionario (Es Poisson, Wiener). Supponiamo che  $X(0) = 0$ . Sia  $f(t) = \mathbb{E}(X(t)) = \mathbb{E}(X(t) - X(0))$ . Si ha

$$\begin{aligned}f(t+s) &= \mathbb{E}(X(t+s) - X(s) + X(s) - X(0)) = \\ &= \mathbb{E}(X(t+s) - X(s)) + \mathbb{E}(X(s) - X(0)) \\ &= \mathbb{E}(X(t) - X(0)) + \mathbb{E}(X(s) - X(0)) \\ &= f(t) + f(s).\end{aligned}$$

L'unica soluzione dell'equazione funzionale  $f(t+s) = f(t) + f(s)$  è  $f(t) = ct$ . Per  $t = 1$  si ha  $c = f(1)$ . Pertanto

$$\mathbb{E}(X(t)) = \mu_1 t, \quad \mu_1 = \mathbb{E}(X(1)).$$

In maniera analoga, sempre per il fatto che gli incrementi sono indipendenti e stazionari, sia  $g(t) = \text{Var}(X(t)) = \text{Var}(X(t) - X(0))$ . Si ha

$$\begin{aligned}g(t+s) &= \text{Var}(X(t+s) - X(s) + X(s) - X(0)) = \\ &= \text{Var}(X(t+s) - X(s)) + \text{Var}(X(s) - X(0)) \\ &= \text{Var}(X(t) - X(0)) + \text{Var}(X(s) - X(0)) \\ &= g(t) + g(s).\end{aligned}$$

L'unica soluzione dell'equazione funzionale  $g(t+s) = g(t) + g(s)$  è  $g(t) = ct$ . Per  $t = 1$  si ha  $c = g(1)$ . Pertanto

$$\text{Var}(X(t)) = \sigma_1^2 t, \quad \sigma_1^2 = \text{Var}(X(1)).$$

Inoltre, per  $t > s$

$$\begin{aligned}\text{Var}(X(t)) &= \text{Var}(X(t) - X(s) + X(s) - X(0)) \\ \text{Var}(X(t) - X(s)) &= \text{Var}(X(t)) - \text{Var}(X(s)) = \sigma_1^2(t-s).\end{aligned}$$

Calcoliamo la funzione di autocovarianza di un processo a incrementi indipendenti e stazionari. Supponiamo  $t > s$ , si ha

$$\begin{aligned}\text{Cov}(X(t), X(s)) &= \frac{1}{2}\{\text{Var}(X(t) + X(s)) - [\text{Var}(X(t) - X(s))]\} = \\ &= \frac{1}{2}\{\sigma_1^2(t+s) + \sigma_1^2(t-s)\} = \sigma_1^2 s.\end{aligned}$$

Quindi,

$$\text{Cov}(X(t), X(s)) = \sigma_1^2 \min\{t, s\}. \quad (7.3)$$

## 7.2 Funzione caratteristica congiunta

La funzione caratteristica di un vettore aleatorio  $X = (X(1), \dots, X(n))$  è definita come

$$\psi_X(\theta_1, \dots, \theta_n) = \mathbb{E}(e^{i(\theta_1 X_1 + \dots + \theta_n X_n)}).$$

Osserviamo che per  $\theta_1 = \dots = \theta_n = \theta$

$$\psi_X(\theta_1, \dots, \theta_n) = \psi_{X_1 + \dots + X_n}(\theta).$$

$X_1, \dots, X_n$  sono stocasticamente indipendenti se e solo se

$$\psi_X(\theta_1, \dots, \theta_n) = \psi_{X_1}(\theta_1)\psi_{X_2}(\theta_2)\cdots\psi_{X_n}(\theta_n)$$

## 7.3 Processi Normali: cenni

Si dice che  $\{X(t), t \geq 0\}$  è un processo *Normale* o *Gaussiano* se, per qualsiasi numero intero  $n$  e qualsiasi sottoinsieme  $\{t_1, t_2, \dots, t_n\}$  di  $T$ , il vettore aleatorio variabili aleatorie  $(X(t_1), \dots, X(t_n))$  ha una distribuzione normale  $n$ -dimensionale. In tal caso si ha

$$\begin{aligned} \psi_X(\theta_1, \dots, \theta_n) &= \mathbb{E}(e^{i(\theta_1 X_1 + \dots + \theta_n X_n)}) = \dots \\ &\exp\left(i \sum_{j=1}^n \theta_j \mathbb{E}(X(t_j)) - \frac{1}{2} \sum_i \sum_j \theta_i \theta_j \text{Cov}(X(t_j), X(t_k))\right) \end{aligned} \quad (7.4)$$

L'equazione (7.4) in forma matriciale diviene

$$\begin{aligned} \psi_X(\theta_1, \dots, \theta_n) &= \\ &\exp\left(i(\underline{\theta}^t \cdot \underline{\mu}) - \frac{1}{2} \underline{\theta}^t K \underline{\theta}\right) \end{aligned} \quad (7.5)$$

con  $K$  matrice delle varianze-covarianze,  $K = [(\text{Cov}(X(t_j), X(t_k)))]$  e  $\underline{\mu} = (\mathbb{E}(X(t_1)), \dots, \mathbb{E}(X(t_n)))$ . Dalla (7.5) si evince che un processo Normale è caratterizzato dalle distribuzioni del secondo ordine. Inoltre essa è completamente determinata quando si conoscono i valori attesi e le covarianze, pertanto un processo Normale se è stazionario (debole) di ordine 2 allora sarà anche stazionario di ordine stretto.

Per l'importanza che ha il processo Normale, spesso si parla di processi stazionari, sottintendendo del secondo ordine.

### 7.3.1 Processo di Wiener-Moto Browniano

Si dice che un processo stocastico  $\{X(t), t \geq 0\}$  è un processo di Wiener se

- $\{X(t), t \geq 0\}$  ha incrementi ind. e staz.
- l'incremento  $X(t) - X(s)$ ,  $t > s$ , è distribuito normalmente
- $E(X(t)) = 0$
- $X(0) = 0$

In tal caso si ha (vedi (7.3))

$$\text{Var}(X(t)) = \sigma^2 t$$

dove  $\sigma^2 = \text{Var}(X(1))$  è l'unico parametro del processo. Inoltre si ha,  $t, s \geq 0$

$$R(t, s) = K(t, s) = \sigma^2 \min\{t, s\}.$$

Se  $\sigma = 1$  si ha un processo di Wiener standard.

Osserviamo che il vettore  $(X(t_1), \dots, X(t_n))$  si può esprimere come combinazione lineare delle variabili aleatorie normali e indipendenti

$$X(t_1), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1}).$$

Infatti, se poniamo  $Y = (X(t_1), \dots, X(t_n))$  e  $X = (X(t_1), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1}))$  si ha

$$Y^t = AX^t$$

con

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}$$

Pertanto, per ogni scelta  $t_1 < t_2 < \dots < t_n$ , si ha  $(X(t_1), \dots, X(t_n))$  ha una distribuzione normale multivariata. Quindi, un processo di Wiener è un processo Normale (a incrementi indipendenti e stazionari).

### 7.3.2 Moto Browniano e passeggiata aleatoria

Consideriamo una passeggiata aleatoria semplice con  $p = q = \frac{1}{2}$ , dove  $p$  (risp.  $q$ ) è la probabilità di fare un passo *unitario* a sinistra (risp. a destra). Adesso, supponiamo di raffinare il processo considerando passi sempre più piccoli in istanti di tempo più piccoli. Al limite di questo raffinamento (sotto certe condizioni) si ottiene un moto Browniano ([10] pag. 627, cap 10). Precisamente, supponiamo che in un unità di tempo  $\Delta t$  con la stessa probabilità si faccia un passo di ampiezza  $\Delta x$  a destra o a sinistra. Se  $X(t)$  denota la posizione al tempo  $t$  allora si ha

$$X(t) = \Delta x (X_1 + X_2 + \dots + X_{[\frac{t}{\Delta t}]})$$

con  $X_i \in \{-1, 1\}$  indipendenti con  $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$ . La notazione tra parentesi quadre  $[\frac{t}{\Delta t}]$  rappresenta la funzione massimo intero contenuto in  $\frac{t}{\Delta t}$ . Dato un reale  $\alpha$  il massimo intero contenuto  $[\alpha]$  è quel numero intero che soddisfa la seguente relazione

$$[\alpha] \leq \alpha < [\alpha] + 1.$$

Ad esempio  $[2.4] = 2$ ,  $[2] = 2$ ,  $[-2] = -2$ ,  $[-2.5] = -3$ ,  $[-3.5] = -4$ ,  $[-0.1] = -1$ .

Poichè  $\mathbb{E}(X_i) = 0$ ,  $Var(X_i) = \mathbb{E}(X_i^2) = 1$  si ha ovviamente che  $\mathbb{E}(X(t)) = 0$   
e

$$Var(X(t)) = \left[ \frac{t}{\Delta t} \right] (\Delta x)^2$$

Se si passa al limite  $\Delta t \rightarrow 0$  e  $\Delta X \rightarrow 0$  con la condizione che

$$\frac{(\Delta x)^2}{\Delta t} \rightarrow \sigma^2 > 0,$$

cioè  $\Delta x$  e  $\sqrt{\Delta t}$  sono infinitesimi dello stesso ordine, allora

$$\mathbb{E}(X(t)) = 0, \quad e \quad Var(X(t)) \rightarrow \sigma^2 t.$$

Ponendo  $\Delta x = \sigma\sqrt{\Delta t}$ , per  $\Delta t \rightarrow 0$ ,

- per il teorema centrale del limite si ha che  $X(t)$  ha una distribuzione normale con  $\mathbb{E}(X(t)) = 0$   $Var(X(t)) \rightarrow \sigma^2 t$ .
- $X(t)$  ha incrementi indipendenti, perchè relativamente ad intervalli disgiunti di tempo le variazioni della passeggiata aleatoria sono indipendenti
- $\{X(t), t \geq 0\}$  ha incrementi stazionari, la distribuzione di  $X(t+s) - X(t)$  non dipende da  $s$ .

Con questo passaggio al limite è possibile dimostrare che la variabile  $X(t)$  diventa una variabile continua e quindi la passeggiata aleatoria diventa un moto browniano.

**Osservazione 12.** Poichè un moto Browniano, o processo di Wiener-Lévy, dipende da un solo parametro  $\sigma$ , possiamo limitarci a studiare il cosiddetto moto Browniano standard  $B(t) = X(t)/\sigma$ .

### 7.3.3 Moto Browniano standard

Sia  $\{X(t), t \geq 0\}$  un moto Browniano standard ( $\sigma = 1$ ), si ha che  $X(t)$  ha una densità normale di parametri  $\mu_{X(t)} = 0$  e  $\sigma_{X(t)} = \sqrt{t}$ . Inoltre se consideriamo il vettore  $(X(t_1), X(t_2), \dots, X(t_n))$  con  $0 < t_1 < t_2 < \dots < t_n$  si ha che le seguenti equazioni

$$X(t_1) = x_1, X(t_2) = x_2, \dots, X(t_n) = x_n,$$

sono equivalenti a

$$X(t_1) = x_1, X(t_2) - X(t_1) = x_2 - x_1, \dots, X(t_n) - X(t_{n-1}) = x_n - x_{n-1}.$$

Inoltre, gli incrementi sono indipendenti, quindi

$$X(t_1), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$$

sono indipendenti, e stazionari, quindi  $X(t_h) - X(t_{h-1})$  ha distribuzione con media 0 e varianza  $t_h - t_{h-1}$ . Quindi la densità congiunta di  $(X(t_1), X(t_2), \dots, X(t_n))$  è data da

$$\begin{aligned} f(x_1, \dots, x_n) &= f_{t_1}(x_1) f_{t_2-t_1}(x_2 - x_1) \cdots f_{t_n-t_{n-1}}(x_n - x_{n-1}) = \\ &= \frac{\exp\left(-\frac{1}{2}\left(\frac{x_1^2}{t_1} + \frac{(x_2-x_1)^2}{t_2-t_1} + \dots + \frac{(x_n-x_{n-1})^2}{t_n-t_{n-1}}\right)\right)}{(2\pi)^{n/2} \sqrt{t_1(t_2-t_1)\cdots(t_n-t_{n-1})}} \end{aligned}$$

A partire dalla precedente formula si dimostra che  $X(s)|X(t) = x$ , con  $t > s$  ha una distribuzione normale di parametri

$$\mathbb{E}(X(s)|X(t) = x) = \frac{s}{t}x, \quad \text{Var}(X(s)|X(t) = x) = \frac{s}{t}(t - s)$$

**Esempio 40.** Vedi esempio 10.1 in [10]. In a bicycle ...

# Bibliografia

- [1] K. Baclawski, M. Cerasoli, and G. C. Rota. *Introduzione alla probabilità*. Unione Matematica Italiana, 1984.
- [2] P. Baldi. *Calcolo delle probabilità*. McGraw-Hill, 2007.
- [3] P. Billingsley. *Probability and measure*.
- [4] D.R. Cox and H.D. Miller. *The theory of Stochastic Processes*. Methuen & Co LTD, 1965.
- [5] L. Daboni. *Calcolo delle probabilità ed elementi di statistica*. UTET.
- [6] G. Dall'Aglio. *Calcolo delle probabilità*. Zanichelli.
- [7] W. Feller. *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons. varie edizioni.
- [8] S. Karlin and H. M. Taylor. *A first course in Stochastic Processes*. Academic Press, 1975.
- [9] V. G. Kulkarni. *Modeling and Analysis of Stochastic Systems*. Chapman & Hall, 1995.
- [10] S. M. Ross. *Introduction to Probability Models*. Academic Press.
- [11] S. M. Ross. *Stochastic Processes*. John Wiley and Sons, 1996.
- [12] R. Scozzafava. *Incertezza e Probabilità*. Zanichelli, 2001.