

# Introduzione alle Catene di Markov

**Giuseppe BIANCHI**

**Dipartimento di Ingegneria Elettronica  
Università di Roma, Tor Vergata  
giuseppe.bianchi@uniroma2.it**

Versione 1.1 – Giugno 2009

A.A. 2008-2009

*Questa breve dispensa ha lo scopo di introdurre lo studente al concetto di Catena di Markov, evitando complicazioni di carattere formale, ma cercando di derivare le caratteristiche di una catena di Markov da esempi e considerazioni di carattere intuitivo. La trattazione seguente è pertanto necessariamente incompleta: si rimanda a testi specialistici per una trattazione più formale.*

# INDICE

|            |  |           |
|------------|--|-----------|
| <b>I</b>   | <b>LA DISTRIBUZIONE ESPONENZIALE NEGATIVA.....</b>   | <b>3</b>  |
| I.1        | DEFINIZIONE.....   | 3         |
| I.1.1      | <i>Processo degli arrivi markoviano</i> .....  | 3         |
| I.1.2      | <i>Definizione: Processo di servizio markoviano</i> .....  | 3         |
| I.1.3      | <i>Esempio</i> .....   | 3         |
| I.2        | PROPRIETÀ FONDAMENTALI DELLA DISTRIBUZIONE ESPONENZIALE NEGATIVA ..  | 4         |
| I.2.1      | <i>Derivazione della distribuzione esponenziale negativa</i> .....   | 5         |
| I.2.2      | <i>Esempio 1</i> .....   | 6         |
| I.2.3      | <i>Esempio 2 – paradosso della vita residua</i> .....  | 6         |
| I.3        | DISTRIBUZIONE ESPONENZIALE E DISTRIBUZIONE DI POISSON .....  | 6         |
| I.3.1      | <i>Dimostrazione</i> .....   | 6         |
| <b>II</b>  | <b>ARRIVI ESPONENZIALI E PROCESSI DI PURA NASCITA .....</b>  | <b>8</b>  |
| II.1       | FREQUENZE DI TRANSIZIONE DI STATO .....  | 9         |
| II.2       | FLUSSI DI PROBABILITÀ .....  | 10        |
| <b>III</b> | <b>CATENE DI MARKOV E DISTRIBUZIONI STAZIONARIE.....</b>   | <b>11</b> |
| III.1      | DEFINIZIONE DI CATENA DI MARKOV .....  | 12        |
| III.1.1    | <i>Esempio: dimostriamo che un sistema M/M/1 è una catena di Markov</i> .....  | 12        |
| III.2      | DISTRIBUZIONE STAZIONARIA: ESEMPIO DI PROCESSO ON/OFF.....   | 14        |
| III.2.1    | <i>Calcolo della distribuzione stazionaria senza passare dalla soluzione del sistema di equazioni differenziali.</i> ..... | 17        |
| III.2.2    | <i>Teorema di conservazione dei flussi di probabilità</i> .....  | 18        |
| III.3      | DISTRIBUZIONE STAZIONARIA: CASO M/M/1 .....  | 20        |
| III.4      | DISTRIBUZIONE STAZIONARIA: CASO M/M/N/N .....  | 21        |

# I La distribuzione esponenziale negativa

## I.1 Definizione

Una variabile casuale  $T$  ha distribuzione esponenziale negativa, con parametro  $\lambda$ , quando la sua funzione di distribuzione è:

$$F_T(t) = P\{T \leq t\} = 1 - e^{-\lambda t}$$

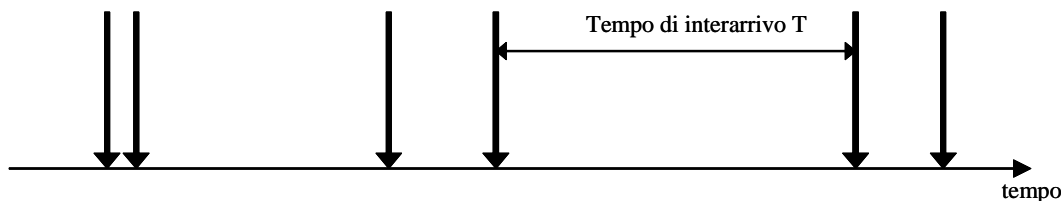
Il valor medio della variabile casuale  $T$  è legato al suo parametro  $\lambda$  dalla relazione<sup>1</sup>:

$$E[T] = \frac{1}{\lambda}$$

Quando la v.c.  $T$  rappresenta un “tempo” (per esempio tempo di interarrivo tra due chiamate, o tempo di servizio di un pacchetto, misurato ad esempio in secondi) allora il parametro  $\lambda$  ha le dimensioni di una frequenza (secondi<sup>-1</sup>).

### I.1.1 Processo degli arrivi markoviano

Un processo degli arrivi è detto **markoviano** se la distribuzione dei tempi di interarrivo è esponenziale negativa.



### I.1.2 Definizione: Processo di servizio markoviano

Un processo di servizio è detto **markoviano** se la distribuzione del tempo di servizio è esponenziale negativa.

### I.1.3 Esempio

Supponiamo che in un'ora arrivino mediamente 2834 chiamate ad una centrale telefonica. Assumiamo che la distribuzione del tempo di interarrivo tra due chiamate consecutive sia esponenziale negativa. Notando che il tempo medio di interarrivo tra due

<sup>1</sup> Infatti, applicando la ben nota formula per il calcolo del valor medio di una v.c. a partire dall'espressione della

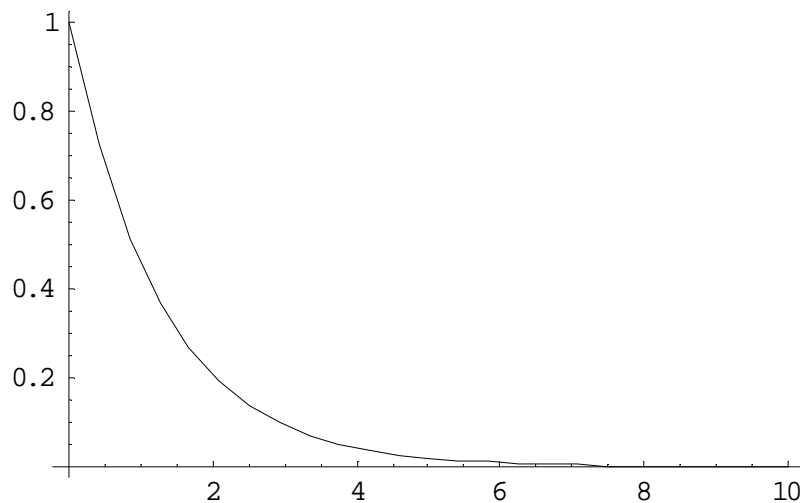
sua funzione di distribuzione,  $E[T] = \int_0^{\infty} (1 - F_T(t)) dt = \int_0^{\infty} (1 - [1 - e^{-\lambda t}]) dt = \int_0^{\infty} e^{-\lambda t} dt = \frac{1}{\lambda}$

chiamate consecutive è  $E[T] = 3600/2834 = 1.27$  secondi, la frequenza di arrivo risulta essere  $\lambda = 1/E[T] = 2834/3600 = 0.787$  chiamate/secondo.

Grazie alla conoscenza della distribuzione di probabilità del tempo di interarrivo, è semplice calcolare, ad esempio, la probabilità che due chiamate consecutive siano intervallate da più di due secondi:

$$P\{T > 2\} = 1 - P\{T \leq 2\} = 1 - (1 - e^{-0.787 \cdot 2}) = e^{-0.787 \cdot 2} = 0.207 = 20.7\%$$

A titolo di esempio si riporta, nel grafico seguente, l'andamento della probabilità che due chiamate siano intervallate da più di  $t$  secondi, per  $t$  variabile tra 0 e 10. Il parametro  $\lambda$  è fissato a 0.787 chiamate/secondo.



Come si vede graficamente, per  $t=10$  la probabilità è prossima allo zero. Ma non è ovviamente zero, ed infatti il valore numerico è

$$P\{T > 10\} = e^{-0.787 \cdot 10} = 0.00038 = 0.038\%$$

## I.2 Proprietà fondamentali della distribuzione esponenziale negativa

La distribuzione esponenziale negativa è l'unica distribuzione statistica che gode delle seguenti proprietà.

1. **assenza di memoria:** la probabilità di avere un evento (ad esempio un arrivo) a partire da un dato istante di tempo non dipende dalla storia precedente;
2. la probabilità di avere un evento (ad es. un arrivo) in un intervallo di tempo piccolo  $\Delta t$  è data da  $\lambda \Delta t$ . In altre parole, tale probabilità è, a meno di infinitesimi di ordine superiore, **proporzionale alla durata dell'intervallo di tempo considerato**, essendo  $\lambda$  (parametro della distribuzione esponenziale negativa) la costante di proporzionalità.

### I.2.1 Derivazione della distribuzione esponenziale negativa

Supponiamo di avere un generico processo degli arrivi, dove per “generico” intendiamo non nota la distribuzione di probabilità del tempo di interarrivo. Supponiamo però di sapere che

1. la probabilità di avere un arrivo in un intervallo di tempo  $\Delta t$  “piccolo” sia (a meno di infinitesimi di ordine superiore) proporzionale alla durata dell’intervallo stesso, ovvero  $P\{\text{arrivo in } \Delta t\} \propto \Delta t$ ; supponiamo in particolare che la costante di proporzionalità sia  $\lambda$ , ovvero  $P\{\text{arrivo in } \Delta t\} = \lambda \Delta t$ ;
2. che tale probabilità sia indipendente dalla scelta dell’intervallo di tempo  $\Delta t$ ; come caso particolare, tale probabilità è indipendente da quanto è successo precedentemente a questo intervallo (proprietà di assenza di memoria);
3. che la probabilità di avere più di un arrivo in  $\Delta t$  sia piccola rispetto a quella di avere un singolo arrivo - ovvero che la probabilità di avere arrivi multipli sia  $o(\Delta t)$ .

Ci chiediamo: quale è la probabilità di NON avere alcun arrivo in un intervallo di tempo di durata  $t$ ? Per rispondere a questa domanda, chiamiamo  $P_0(t)$  la probabilità di non avere alcun arrivo nell’intervallo temporale  $(0,t)$ . Notando che, a meno di infinitesimi di ordine superiore, la probabilità di non avere alcun arrivo in un intervallo di tempo  $t + \Delta t$  è data (per ipotesi di indipendenza) dal prodotto della probabilità di non aver avuto alcun arrivo nel tempo  $t$  (ovvero, per definizione,  $P_0(t)$ ) e di non avere alcun arrivo nell’intervallo  $\Delta t$  (ovvero  $1 - \lambda \Delta t$ ), possiamo scrivere:

$$P_0(t + \Delta t) = P_0(t) \cdot (1 - \lambda \Delta t) = P_0(t) - \lambda P_0(t) \Delta t \quad \Rightarrow \quad \frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t)$$

Passando al limite per  $\Delta t \rightarrow 0$  otteniamo una semplice equazione differenziale del primo ordine:

$$P_0'(t) = -\lambda P_0(t)$$

che possiamo risolvere imponendo la condizione al contorno  $P_0(0) = 1$ , condizione che vale in quanto  $P_0(t)$  è una probabilità, e che la probabilità di non avere alcun arrivo in un tempo nullo è ovviamente 1. Il risultato finale è:

$$P_0(t) = e^{-\lambda t}$$

Il risultato appena dimostrato dice la seguente cosa: **se valgono le ipotesi 1-3 precedenti, il tempo di interarrivo è distribuito esponenzialmente**. Infatti, per definizione di  $P_0(t)$ ,

$$P\{T \leq t\} = 1 - P_0(t) = 1 - e^{-\lambda t}$$

### I.2.2 Esempio 1

Supponiamo che una chiamata telefonica sia distribuita esponenzialmente. Supponiamo che la sua durata media sia di 120 secondi. Quale è la probabilità che la chiamata termini in un intervallo piccolo  $\Delta t$ , dato che la chiamata è durata già 90 secondi? Cambia qualcosa se la chiamata è durata 180 secondi? La risposta ovviamente è sempre  $\mu \Delta t$ , con  $\mu=1/120$ , a prescindere da quanto la chiamata sia durata in precedenza.

Si noti che in questo esempio abbiamo scelto una distribuzione esponenziale negativa in cui l'evento di riferimento è la fine della chiamata.

### I.2.3 Esempio 2 – paradosso della vita residua

Supponiamo che il tempo di interarrivo tra due autobus ad una fermata segua una legge esponenziale negativa. Supponiamo che mediamente arrivi un autobus ogni 10 minuti. Se un passeggero arriva ad un istante casuale, quanto dovrà mediamente aspettare?

Ora, intuitivamente potrebbe sembrare ragionevole rispondere dicendo “5 minuti”, ovvero la metà del tempo di interarrivo tra due autobus. In realtà, la risposta corretta è 10 minuti (!) in quanto il tempo di attesa del prossimo autobus non dipende, per le proprietà della distribuzione esponenziale negativa, da quanto tempo è già passato.

Si noti che questa proprietà vale solo ed esclusivamente perché abbiamo assunto un processo di interarrivo degli autobus con distribuzione esponenziale negativa: se avessimo assunto una distribuzione deterministica (ovvero un autobus esattamente ogni 10 minuti), la risposta corretta sarebbe stata 5 minuti.

Questo esempio è chiamato “paradosso della vita residua”, in quanto la risposta sembra essere paradossale. In realtà non c'è nessun paradosso; semplicemente il tempo di attesa del prossimo autobus è sicuramente maggiore della metà del valore medio (i 5 minuti di cui sopra) perché è più probabile che un passeggero arrivi in un intervallo “grande” tra due arrivi consecutivi di autobus.

## I.3 Distribuzione esponenziale e distribuzione di Poisson

Consideriamo un processo di arrivi markoviano (ovvero un processo per cui il tempo di interarrivo sia distribuito esponenzialmente). Supponiamo che il tempo medio di interarrivo sia  $1/\lambda$ . Consideriamo un generico intervallo  $t$ . Vale il seguente risultato: **il numero di arrivi in un intervallo  $t$  è una variabile casuale con distribuzione di Poisson:**

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

### I.3.1 Dimostrazione

Già sappiamo (v. sezione I.2.1) che la probabilità  $P_0(t)$  che non vi siano arrivi

nell'intervallo di tempo  $(0,t)$  è data dalla la soluzione dell'equazione differenziale  $P_0'(t) = -\lambda P_0(t)$ , ovvero:

$$P_0(t) = e^{-\lambda t}$$

Concentriamoci su un istante di tempo  $t + \Delta t$  con  $\Delta t$  piccolo. Possiamo calcolare la probabilità di avere un arrivo in  $t + \Delta t$  come: i) la probabilità di avere avuto 0 arrivi al tempo  $t$  e di avere 1 arrivo nell'intervallo di tempo  $\Delta t$  più la probabilità di avere avuto 1 arrivo al tempo  $t$  e di non avere nuovi arrivi nell'intervallo di tempo  $\Delta t$ . In formule:

$$P_1(t + \Delta t) = P_0(t) \cdot \lambda \Delta t + P_1(t)(1 - \lambda \Delta t) \Rightarrow \frac{P_1(t + \Delta t) - P_1(t)}{\Delta t} = -\lambda P_1(t) + \lambda P_0(t)$$

Analogamente, possiamo generalizzare e calcolare la probabilità di avere  $k$  arrivi nel tempo  $t + \Delta t$  come:

$$P_k(t + \Delta t) = P_{k-1}(t) \cdot \lambda \Delta t + P_k(t)(1 - \lambda \Delta t) \Rightarrow \frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = -\lambda P_k(t) + \lambda P_{k-1}(t)$$

dove abbiamo trascurando l'eventualità di arrivi multipli nell'intervallo di tempo  $\Delta t$  (la cui probabilità, come detto precedentemente, è  $o(\Delta t)$  per una distribuzione esponenziale). Passando al limite per  $\Delta t \rightarrow 0$ , otteniamo il seguente sistema di equazioni differenziali:

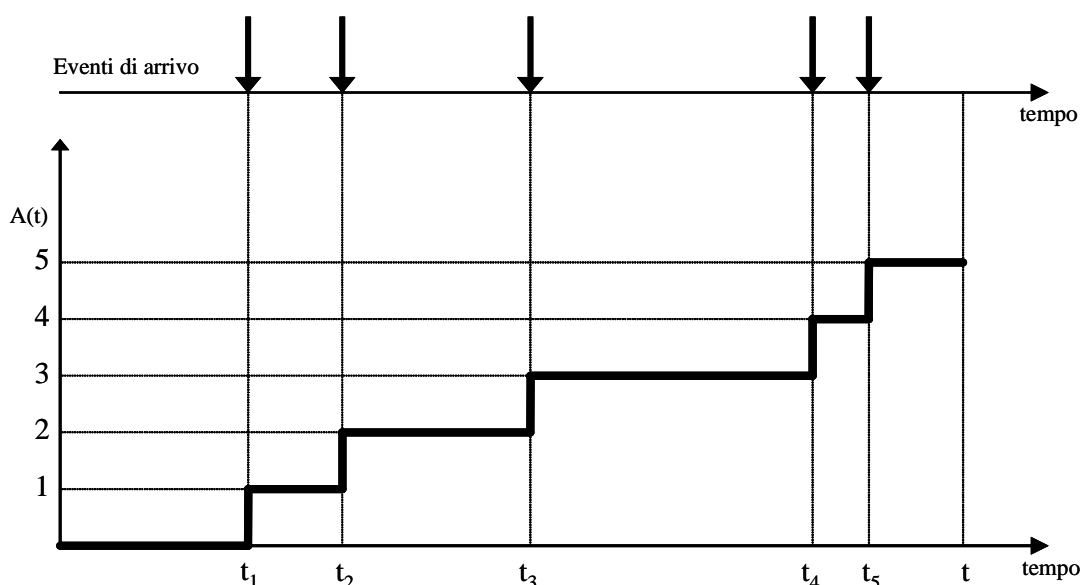
$$\begin{cases} P_0(t) = e^{-\lambda t} \\ P_1'(t) = -\lambda P_1(t) + \lambda P_0(t) \\ P_2'(t) = -\lambda P_2(t) + \lambda P_1(t) \\ \dots \\ P_k'(t) = -\lambda P_k(t) + \lambda P_{k-1}(t) \end{cases}$$

L'equazione relativa alla seconda riga è immediatamente risolta usando il termine  $P_0(t)$  noto dalla prima riga, ed applicando l'ovvia condizione al contorno  $P_1(0) = 0$  (la probabilità di avere un arrivo al tempo 0 è ovviamente nulla). Una volta calcolato  $P_1(t)$  è possibile calcolare  $P_2(t)$  risolvendo l'equazione differenziale indicata nella terza riga, e così via per qualunque valore  $k$  scelto. Il risultato è, come volevasi dimostrare, la distribuzione di Poisson:

$$\begin{cases} P_0(t) = e^{-\lambda t} \\ P_1(t) = \lambda t e^{-\lambda t} \\ P_2(t) = \frac{(\lambda t)^2}{2} e^{-\lambda t} \\ \dots \\ P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \end{cases}$$

## II Arrivi esponenziali e processi di pura nascita

Lo studio fatto fino ad ora ci ha permesso di mettere in relazione un processo degli arrivi, espresso in termini della *distribuzione di probabilità della variabile casuale continua tempo di interarrivo*, con un processo di “conteggio”, espresso in termini di *distribuzione di probabilità della variabile casuale discreta numero di arrivi al tempo  $t$* . Nell’importante caso di arrivi markoviani che stiamo considerando, le distribuzioni di probabilità in questione sono la distribuzione esponenziale negativa per il tempo di interarrivo, e la distribuzione di Poisson per il numero di arrivi al tempo  $t$ .



La Figura precedente riporta una illustrazione grafica della relazione tra eventi di arrivo e processo  $A(t)$  che rappresenta il numero di arrivi al tempo  $t$ . Il processo  $A(t)$  è un “processo di pura nascita”, in quanto non decresce mai all’aumentare del tempo, ma si incrementa sempre di una unità (nascita) ad ogni arrivo successivo. Nel caso particolare di arrivi markoviani, si noti che il processo  $A(t)$  non fa mai “salti” di più di una unità (abbiamo infatti detto che la probabilità di avere arrivi multipli in un intervallo  $\Delta t \rightarrow 0$  è un infinitesimo di ordine superiore rispetto alla probabilità di avere un arrivo singolo, data da  $\lambda \Delta t$ ).

La figura precedente evidenzia come il processo  $A(t)$  si evolva attraverso “stati” discreti<sup>2</sup>, dove uno stato rappresenta il numero di eventi di arrivo conteggiati fino all’istante  $t$  considerato. E’ utile rappresentare graficamente gli stati del processo con il diagramma illustrato nella figura seguente:

---

<sup>2</sup> Un processo stocastico a stati discreti è spesso chiamato “catena”. Nel caso specifico in questione il processo  $A(t)$  è una catena tempo-continua in quanto è dipendente dal parametro  $t$  continuo.





dove i cerchi rappresentano un possibile **stato** del sistema,  $A(t)=0, 1, 2, \dots$ , e le frecce rappresentano le possibili **transizioni di stato**, ovvero la presenza di una freccia tra due stati implica che è possibile avere un “salto” tra i due stati in questione. Per esempio, nel caso del processo di pura nascita  $A(t)$ , come si vede dal diagramma, le uniche transizioni di stato possibile sono tra lo stato  $k$  e lo stato  $k+1$ . Non sono infatti possibili salti multipli, e non è possibile – si noti che le transizioni di stato vanno nel verso della freccia - un ritorno allo stato  $k$  una volta raggiunto lo stato  $k+1$  (essendo il processo di pura nascita, una volta contati 10 arrivi non c’è modo di tornare allo stato 9, intendendo per stato il numero di arrivi contati!!).

## II.1 Frequenze di transizione di stato

Siamo arrivati al momento più importante di questa trattazione. Ci chiediamo: è possibile (e/o utile) assegnare un valore numerico alle frecce, disegnate nel diagramma a stati precedente, che rappresentano le transizioni di stato? La risposta è: per processi markoviani, sicuramente sì. In particolare diamo la seguente definizione:

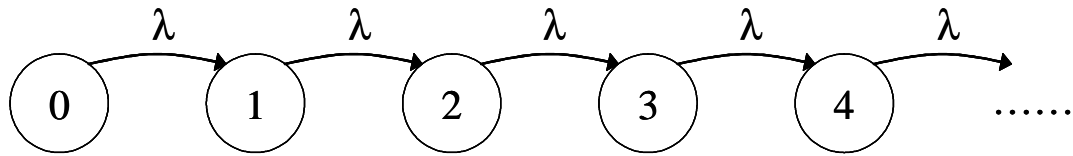
*Chiamiamo **frequenza di transizione di stato** tra gli stati  $k$  ed  $h$  il rapporto fra la probabilità che, assumendo il sistema in un dato stato  $k$ , avvenga una transizione di stato  $k \rightarrow h$  in un intervallo di tempo tendente a 0, e l’intervallo di tempo considerato. In altre parole, detto  $k$  uno stato origine e  $h$  uno stato destinazione, la frequenza di transizione di stato  $q(k \rightarrow h)$  è definita come:*

$$q(k \rightarrow h) = \lim_{\Delta t \rightarrow 0} \frac{P(k \rightarrow h \text{ in } \Delta t \mid \text{sistema nello stato } k)}{\Delta t}$$

L’unità di misura di una frequenza di transizione è *probabilità condizionata su tempo*, ovvero secondi<sup>-1</sup>. Nel caso particolare in questione (arrivi markoviani), è facile derivare le frequenze di transizione. Infatti, l’unico caso in cui la frequenza di transizione assume valore non nullo è la transizione da un generico stato  $k$  allo stato  $k+1$ . La probabilità di transizione dallo stato  $k$  allo stato  $k+1$  in un intervallo di tempo  $\Delta t$  coincide con la probabilità di avere un arrivo nell’intervallo di tempo  $\Delta t$ . Poiché i tempi di interarrivo hanno distribuzione esponenziale con parametro  $\lambda$ , la probabilità di un arrivo in  $\Delta t$  non dipende dalla storia passata, ed è banalmente data da  $\lambda \Delta t$ . Concludiamo quindi che, per qualunque stato  $k$  considerato,

$$q(k \rightarrow k + 1) = \lim_{\Delta t \rightarrow 0} \frac{\lambda \Delta t}{\Delta t} = \lambda$$

E' consuetudine riportare graficamente le frequenze di transizione nel diagramma di stato, associandole agli archi che rappresentano le transizioni di stato (vedi figura seguente).



Una importante osservazione è che la definizione di frequenza di transizione data precedentemente **non risulterebbe univoca per processi non markoviani**. Infatti, la probabilità di transizione da uno stato  $k$  ad uno stato  $k+1$  dipenderebbe dall'evoluzione del processo e dal tempo di permanenza nello stato considerato. Per esempio, se gli arrivi fossero deterministici (ad esempio 1 arrivo al secondo), la frequenza di transizione dallo stato  $k$  allo stato  $k+1$  sarebbe di 1 transizione al secondo, ma ovviamente la probabilità di avere o meno una transizione a partire dallo stato  $k$  considerato dipenderebbe da quanto tempo il processo ha passato nello stato stesso!

## II.2 Flussi di probabilità

L'analisi dell'esempio di processo di pura nascita considerato finora ci permette di trarre un'ulteriore importante informazione. Definiamo **flusso di probabilità** uscente da un dato stato  $k$  ed entrante in un altro stato  $h$ , al tempo  $t$ , come il prodotto tra la probabilità di essere nello stato considerato al tempo  $t$  e la frequenza di transizione tra gli stati considerati, ovvero:

$$\varphi(k \rightarrow h, t) = P_k(t)q(k \rightarrow h)$$

Ricordando la definizione di frequenza di transizione di stato data nella sezione II.1, e sostituendo nella formula precedente, otteniamo:

$$\begin{aligned} \varphi(k \rightarrow h, t) &= P_k(t) \lim_{\Delta t \rightarrow 0} \frac{P(k \rightarrow h \text{ in } \Delta t \mid \text{ sistema nello stato } k)}{\Delta t} = \\ &= \lim_{\Delta t \rightarrow 0} \frac{P(k \rightarrow h \text{ in } \Delta t, \text{ sistema nello stato } k)}{\Delta t} \end{aligned}$$

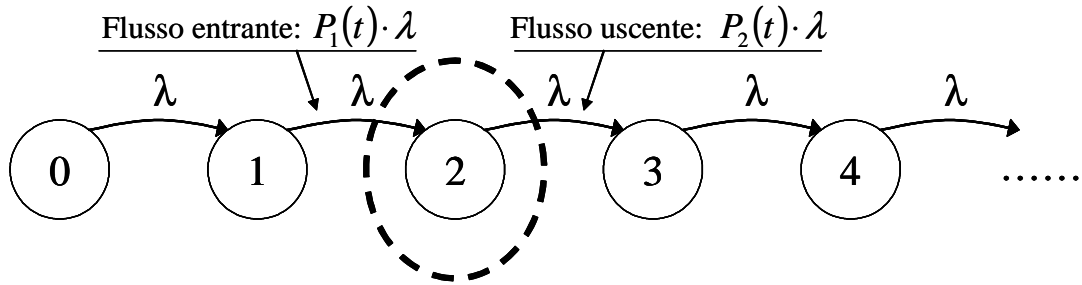
Tale formula ci dice che un flusso di probabilità rappresenta la frequenza (ovvero probabilità su tempo) assoluta (si noti la differenza con la corrispondente probabilità condizionata nel caso di transizione di stato) di passare, al tempo  $t$ , dallo stato  $k$  allo stato  $h$ . In generale in flusso di probabilità così definito dipende dal tempo  $t$  considerato. Si dimostra<sup>3</sup> che:

<sup>3</sup> La dimostrazione coincide con quella proposta nella sezione I.3.1. In particolare, per il caso particolare di un processo di pura nascita, basta dimostrare che, in un intervallo  $(t, t+\Delta t)$ , vale l'espressione

$$\lim_{\Delta t \rightarrow 0} \frac{P_k(t + \Delta t) - P_k(t)}{\Delta t} = P_k'(t) = -\lambda P_k(t) + \lambda P_{k-1}(t)$$

e rileggere tale espressione come somma algebrica dei flussi attraverso lo stato  $k$ .

la variazione di probabilità istantanea per un dato stato è uguale alla somma algebrica dei flussi in entrata ed in uscita dallo stato considerato.



Applicando questa considerazione, sarebbe stato immediato ottenere, con riferimento alla figura precedente che rappresenta il diagramma di transizione di stato per un processo di pura nascita, l'equazione differenziale:

$$P_2'(t) = -\lambda P_2(t) + \lambda P_1(t)$$

dove:

- La variazione istantanea di probabilità nello stato 2 è la derivata prima  $P_2'(t)$  ;
- Il flusso entrante nello stato 2 è  $\lambda P_1(t)$  : essendo entrante nello stato è preso con segno positivo;
- Il flusso uscente dallo stato 2 è  $\lambda P_2(t)$  : essendo uscente dallo stato è preso con segno negativo.

### III Catene di Markov e distribuzioni stazionarie

Il processo di pura nascita studiato nella sezione precedente rappresenta un primo esempio, il piu' elementare possibile, di una catena di Markov tempo-continua. Tale processo è infatti:

- Una catena, in quanto lo spazio degli stati è discreto
- E' una catena di Markov in quanto (come vedremo tra breve) la probabilità di transizione di stato dipende esclusivamente dallo stato di partenza e non dall'evoluzione precedente del processo e/o dal tempo speso nello stato considerato.

Tuttavia il processo di pura nascita è una catena un po' anomala, ed in particolare di scarso interesse pratico. Infatti, l'evoluzione del processo di pura nascita è particolare in quanto, all'aumentare del tempo  $t$ , la probabilità di essere in uno stato  $k$  comunque scelto tende a 0. Infatti, per  $t$  tendente ad infinito, è intuitivo comprendere come il numero di arrivi tenderà a diventare infinito, e quindi la probabilità di avere un numero finito di arrivi tende a 0. Si usa dire che il processo di pura nascita **non ammette una distribuzione a regime** (ovvero una **distribuzione stazionaria** ovvero una distribuzione in condizioni di **equilibrio statistico**). Si usa anche dire che la catena in questione è instabile.

In questa sezione introdurremo catene di Markov che, a differenza del processo di pura nascita, ammettono una distribuzione a regime, capiremo a cosa serve tale distribuzione ed impareremo a calcolarla.

### III.1 Definizione di catena di Markov

Nel prosieguo, come peraltro fatto finora, ci limitiamo all'analisi di processi tempo-continui. La definizione che stiamo per proporre non è tipicamente quella data nei testi di probabilità, ma ne è una diretta conseguenza (ovvero risulta dimostrabile a partire dalla definizione rigorosa<sup>4</sup> - rimandiamo il lettore interessato alla consultazione dei testi specialistici). Riteniamo però che la definizione seguente sia per noi preferibile in quanto, oltre ad essere decisamente più semplice, mette immediatamente in luce alcune proprietà chiave delle catene di Markov che poi verranno sfruttate nella nostra trattazione.

*Con il termine “Catena di Markov” definiamo un processo stocastico a stati discreti che gode delle seguenti due proprietà: 1) il tempo di permanenza in ogni stato è una variabile casuale con distribuzione esponenziale negativa, e 2) le frequenze di transizione di stato dipendono solo dallo stato di partenza e non dagli stati visitati precedentemente.*

#### III.1.1 Esempio: dimostriamo che un sistema M/M/1 è una catena di Markov

Un sistema M/M/1 è un sistema caratterizzato da un numero infinito di utenti, da un server e da numero infinito di posti in coda. Il processo degli arrivi al sistema è un processo di Poisson con frequenza  $\lambda$  (ovvero la distribuzione del tempo di interarrivo è esponenziale negativa con parametro  $\lambda$ ), ed il tempo di servizio per ogni cliente nel sistema è una variabile casuale con distribuzione esponenziale negativa con parametro  $\mu$  (ovvero il tempo medio di servizio è  $1/\mu$ ).

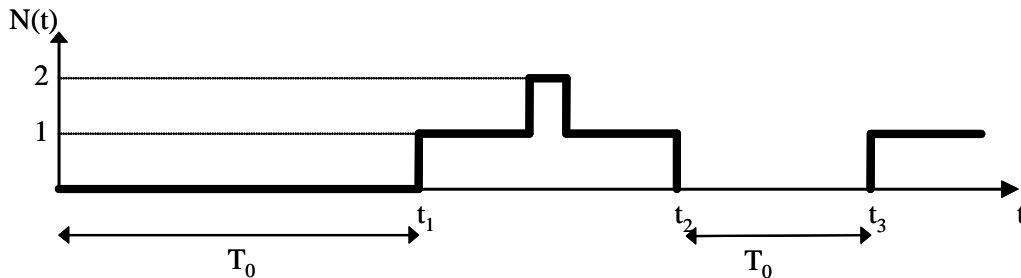
Per dimostrare che il processo  $N(t)$  definito come numero di utenti nel sistema rappresenta una catena di Markov, cominciamo ad analizzare lo stato 0, ovvero lo stato che caratterizza il sistema vuoto. Supponiamo in particolare che al tempo  $t=0$  il sistema si trovi nello stato 0. In tale condizioni, non essendoci alcun cliente in servizio, l'unico evento possibile è l'arrivo di un nuovo cliente al sistema. Poiché il tempo di interarrivo è distribuito esponenzialmente con parametro  $\lambda$ , sfruttando la proprietà di assenza di memoria, possiamo concludere che la variabile casuale  $T_0$  è a sua volta distribuita esponenzialmente con parametro  $\lambda$  e con valor medio  $E[T_0]=1/\lambda$ , ed è rappresentativa (vedi figura seguente):

- o sia del tempo che intercorre fra l'istante  $t=0$  e l'istante  $t_1$  di arrivo di un cliente

---

<sup>4</sup> Una catena di Markov è un processo stocastico a stati discreti  $X(t)$ , che gode della seguente proprietà: scelti arbitrariamente  $n+1$  istanti di tempo  $t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n$ , e per qualunque scelta arbitraria degli stati discreti  $s_0, s_1, \dots, s_{n-1}, s_n$  vale la seguente proprietà:  
 $P(X(t_n) = s_n | X(t_0) = s_0, X(t_1) = s_1, \dots, X(t_{n-1}) = s_{n-1}) = P(X(t_n) = s_n | X(t_{n-1}) = s_{n-1})$

- o che del tempo di permanenza del sistema nello stato 0 tra un istante di ritorno allo stato 0 (tempo  $t_2$  in figura) ed il successivo istante di arrivo di un nuovo cliente (tempo  $t_3$ )



Consideriamo a questo punto lo stato 1. Abbiamo una transizione di stato in due soli casi:

- o se un nuovo cliente arriva prima che il cliente nel sistema abbia finito il servizio (in gergo una “nascita”), avremo una transizione  $1 \rightarrow 2$ ;
- o se invece il cliente finisce il servizio prima che un nuovo utente entri nel sistema (in gergo una “morte”), avremo una transizione  $1 \rightarrow 0$ .

Si noti che abbiamo trascurato il caso in cui nello stesso istante di tempo si abbia contemporaneamente la fine del servizio e l’arrivo di un nuovo utente, in quanto questo evento ha probabilità infinitesima di ordine superiore rispetto ad un singolo arrivo o partenza.

Per valutare se il tempo di permanenza nello stato 1 è distribuito esponenzialmente, ed in questo caso con quale parametro, possiamo operare come segue. La probabilità di uscire dallo stato 1 in un intervallino di tempo  $\Delta t$  può essere scritta come segue:

$$P(\text{arrivo in } \Delta t) \times [1 - P(\text{partenza in } \Delta t)] + P(\text{partenza in } \Delta t) \times [1 - P(\text{arrivo in } \Delta t)]$$

dove

$$P(\text{arrivo in } \Delta t) = \lambda \Delta t$$

$$P(\text{partenza in } \Delta t) = \mu \Delta t$$

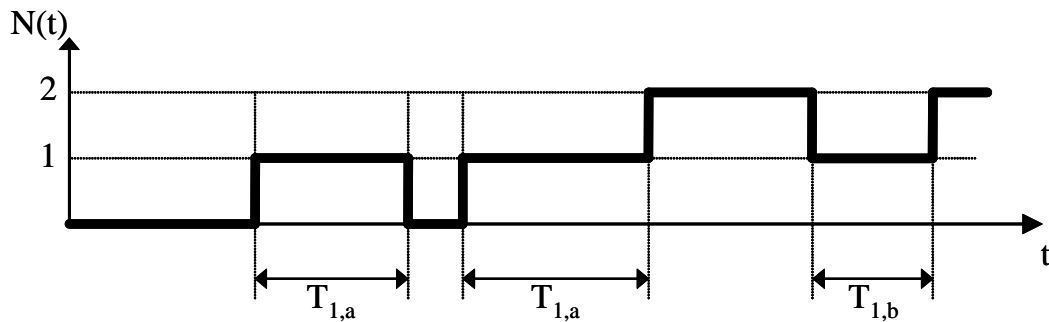
pertanto la probabilità di uscire dallo stato 1 in un intervallo di tempo  $\Delta t$  può essere espressa come:

$$\begin{aligned} & \lambda \Delta t \times [1 - \mu \Delta t] + \mu \Delta t \times [1 - \lambda \Delta t] = \\ & = \lambda \Delta t - \lambda \mu (\Delta t)^2 + \mu \Delta t - \lambda \mu (\Delta t)^2 = \\ & = (\lambda + \mu) \Delta t + o(\Delta t) \end{aligned}$$

Da cui concludiamo che, a meno di infinitesimi di ordine superiore, la probabilità di uscire dallo stato 1 è proporzionale all’intervallino di tempo  $\Delta t$  considerato (e che quindi la distribuzione è esponenziale negativa), e che la costante di proporzionalità è  $\lambda + \mu$  (e che quindi questo è il parametro della distribuzione esponenziale negativa in questione, ovvero

che, detta  $T_1$  la variabile casuale che rappresenta il tempo di permanenza del sistema nello stato 1, è  $P(T_1 \leq t) = 1 - e^{-(\lambda+\mu)t}$ .

Si noti infine che la dimostrazione precedente è valida sia per il caso in cui il sistema sia entrato nello stato 1 a partire dallo stato 0 (ovvero grazie ad un arrivo – caso  $T_{1,a}$  nella figura seguente), sia che il sistema sia entrato nello stato 1 a partire dallo stato 2 (ovvero grazie ad una partenza – caso  $T_{1,b}$  nella figura seguente).



Rimane da dimostrare che la probabilità di transizione di stato dipende esclusivamente dallo stato di partenza, e non dagli stati visitati precedentemente. A tale proposito è banale<sup>5</sup> dimostrare che, assumendo di avere una transizione di stato, la probabilità di passare dallo stato 1 allo stato 2 è data da

$$P(1 \rightarrow 2 | \text{transizione di stato}) = \frac{\lambda}{\lambda + \mu}$$

ed analogamente che

$$P(1 \rightarrow 0 | \text{transizione di stato}) = \frac{\mu}{\lambda + \mu}$$

da cui si evince che tali probabilità non dipendono dalla storia passata, ma solo dai tassi di nascita e morte.

Infine, l'analisi svolta per il caso di stato 1 può essere ripetuta per qualunque altro stato  $k$ . Da cui concludiamo che il processo  $N(t)$  che rappresenta gli utenti in un sistema M/M/1 è una catena di Markov.

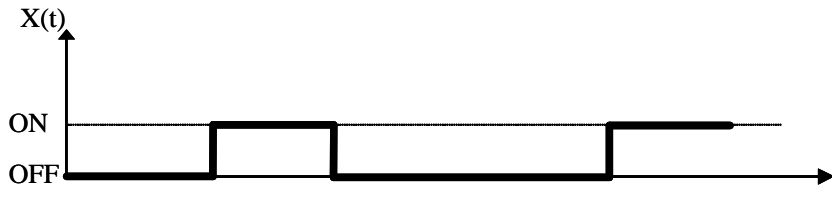
### III.2 Distribuzione stazionaria: esempio di processo ON/OFF

L'esempio più elementare di catena di Markov è un processo a due stati, ad esempio una risorsa che si alterna tra stato ON (occupato) ed OFF (libero). Il processo è markoviano se il

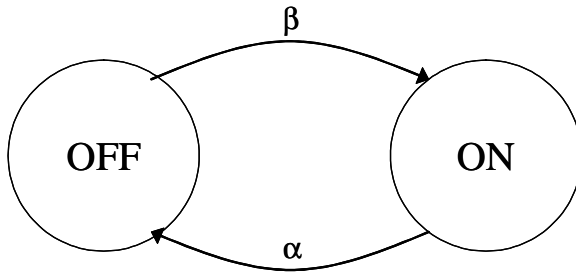
---

<sup>5</sup>  $P(1 \rightarrow 2 | tx\_stato) = \int_0^{\infty} P(\text{arrivo in } (t, t+dt))P(\text{servizio} > t)dt = \int_0^{\infty} \lambda e^{-\lambda t} e^{-\mu t} dt = \frac{\lambda}{\lambda + \mu}$

tempo di permanenza negli stati è distribuito esponenzialmente. Chiamiamo con  $T_{ON}$  e  $T_{OFF}$  i tempi medi di permanenza nei relativi stati.



Notiamo in primo luogo che la transizione di stato tra uno stato ON ed uno stato OFF avviene nel momento in cui termina il periodo di tempo in cui la risorsa è occupata, ovvero  $T_{ON}$ . Pertanto, la frequenza di transizione  $ON \rightarrow OFF$  è data dal reciproco del tempo di permanenza  $T_{ON}$  nello stato ON, ovvero  $\alpha = 1/ T_{ON}$  (in altre parole, per assunzione di esponenzialità, il tempo di permanenza nello stato ON è una variabile casuale esponenziale negativa con funzione di distribuzione  $1 - e^{-\alpha t}$ ). Analogamente, sia  $\beta = 1/ T_{OFF}$  la frequenza di transizione dallo stato OFF verso lo stato ON. La catena di Markov a due stati così ottenuta è illustrata nel diagramma di stato seguente.



Consideriamo l'evoluzione del processo a partire da un generico istante  $t$  fino ad un generico istante di tempo  $t + \Delta t$  con  $\Delta t$  piccolo. Allora, considerando che, in un tempo  $\Delta t$ , transizioni di stato multiple hanno una probabilità che risulta essere un infinitesimo di ordine superiore rispetto alle transizioni singole, e considerando che, grazie all'assunzione di esponenzialità dei tempi di permanenza negli stati considerati, la probabilità di avere una transizione dallo stato ON ad OFF è data da  $\alpha \Delta t$  e viceversa la probabilità di transizione da OFF ad ON è data da  $\beta \Delta t$ , possiamo scrivere:

$$P_{ON}(t + \Delta t) = P_{ON}(t)(1 - \alpha \Delta t) + P_{OFF}(t) \cdot \beta \Delta t$$

$$P_{OFF}(t + \Delta t) = P_{ON}(t) \cdot \alpha \Delta t + P_{OFF}(t)(1 - \beta \Delta t)$$

Riarrangiando i termini nelle equazioni e dividendo per  $\Delta t$ :

$$\frac{P_{ON}(t + \Delta t) - P_{ON}(t)}{\Delta t} = -\alpha \cdot P_{ON}(t) + \beta \cdot P_{OFF}(t)$$

$$\frac{P_{OFF}(t + \Delta t) - P_{OFF}(t)}{\Delta t} = \alpha \cdot P_{ON}(t) - \beta \cdot P_{OFF}(t)$$

Passando al limite per  $\Delta t \rightarrow 0$ :

$$P'_{ON}(t) = -\alpha \cdot P_{ON}(t) + \beta \cdot P_{OFF}(t)$$

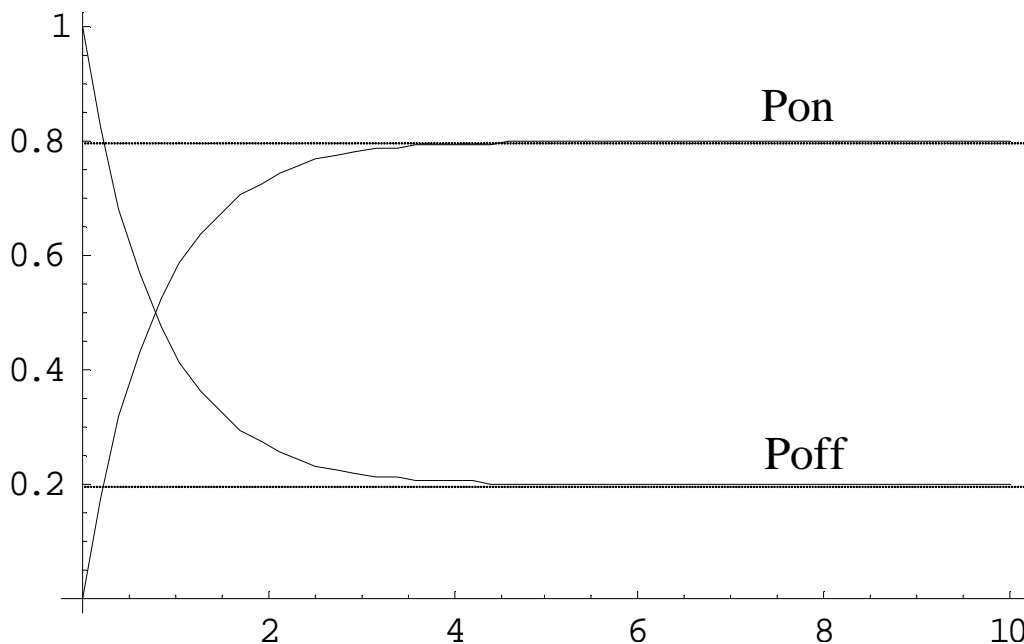
$$P'_{OFF}(t) = \alpha \cdot P_{ON}(t) - \beta \cdot P_{OFF}(t)$$

che rappresenta un sistema di equazioni differenziali lineari. Tali equazioni, nel caso generale di una catena di Markov arbitraria (ovvero non necessariamente limitata a due stati) prendono il nome di **equazioni di Chapman-Kolmogorov**. E' immediato verificare che, nell'assunzione in cui il sistema al tempo 0 si trovi nello stato OFF (ovvero  $P_{ON}(0) = 0, P_{OFF}(0) = 1$ ), il sistema ha la seguente soluzione:

$$P_{ON}(t) = \frac{\beta}{\alpha + \beta} - \frac{\beta}{\alpha + \beta} e^{-(\alpha + \beta)t}$$

$$P_{OFF}(t) = \frac{\alpha}{\alpha + \beta} + \frac{\beta}{\alpha + \beta} e^{-(\alpha + \beta)t}$$

E' istruttivo rappresentare graficamente i risultati ottenuti. In particolare supponiamo di considerare un sistema in cui risulti  $T_{ON}=4s$  e  $T_{OFF}=1s$ . Pertanto,  $\alpha=1/4$  e  $\beta=1$ .



Gli andamenti delle probabilità di essere nello stato ON ed OFF at tempo  $t$  permettono di trarre alcuni interessanti considerazioni.

- Essendo partiti dall'assunzione di sistema inizialmente nello stato OFF, per  $t$  piccolo la probabilità  $P_{ON}(t)$  risulta, come intuitivamente aspetato, essere piccola.
- Per  $t$  crescente fino a circa  $t=4$ , la probabilità  $P_{ON}(t)$  cresce.
- Per  $t$  elevato (da 4 in poi),  $P_{ON}(t)$  tende a convergere verso un asintoto orizzontale dato dal valore 0.8.

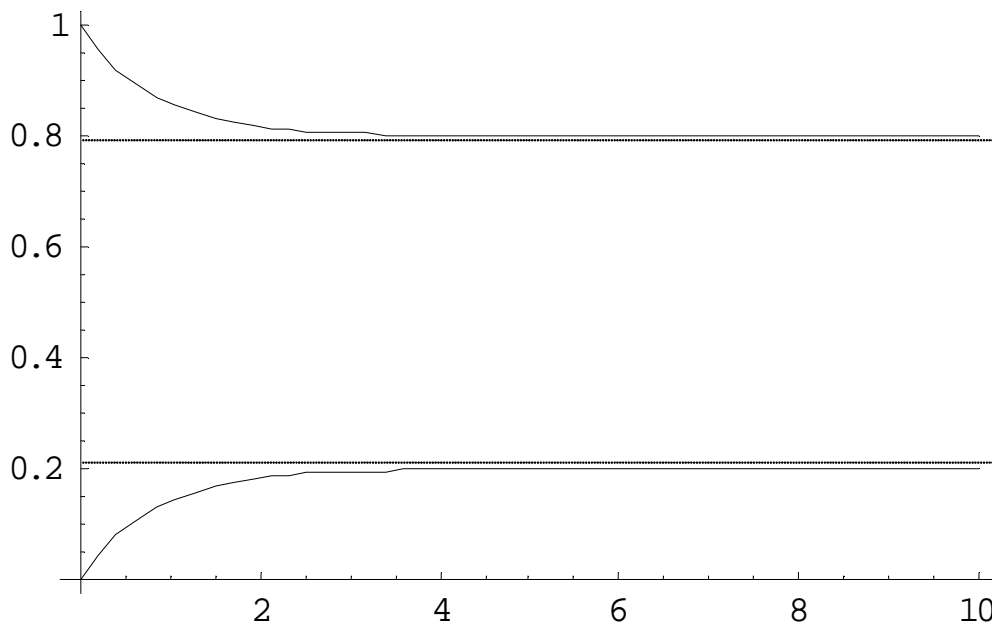


Proviamo a risolvere il sistema di equazioni differenziali, cambiando le condizioni iniziali, ed in particolare assumendo che il sistema inizialmente si trovi nello stato ON (ovvero  $P_{ON}(0) = 1, P_{OFF}(0) = 0$ ). I risultati numerici sono, in questo caso:

$$P_{ON}(t) = \frac{\beta}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta} e^{-(\alpha + \beta)t}$$

$$P_{OFF}(t) = \frac{\alpha}{\alpha + \beta} - \frac{\alpha}{\alpha + \beta} e^{-(\alpha + \beta)t}$$

ed il grafico risultante con i parametri numerici  $\alpha = 1/4$  e  $\beta = 1$  è il seguente:



*Dal confronto tra i due grafici, concludiamo che a prescindere dalle condizioni iniziali, il sistema evolve asintoticamente verso una distribuzione di probabilità fissata. Chiamiamo i valori asintotici: **probabilità limite di stato**. Chiamiamo infine la distribuzione limite: **distribuzione stazionaria** (o **distribuzione a regime** o **distribuzione all'equilibrio**).*

### III.2.1 Calcolo della distribuzione stazionaria senza passare dalla soluzione del sistema di equazioni differenziali.

Sebbene la risoluzione delle equazioni di Chapman-Kolmogorov permetta di ottenere l'evoluzione probabilistica del sistema al variare del tempo  $t$ , risulta evidente come, in gran parte dei casi, sia sufficiente avere la sola distribuzione stazionaria. Per esempio, la probabilità di trovare il sistema occupato in un istante di tempo scelto casualmente è bene approssimata dalla probabilità asintotica di trovare il sistema nello stato ON<sup>6</sup>.

<sup>6</sup> Si noti che la conoscenza della distribuzione stazionaria non permetterebbe di rispondere a domande di questo tipo: "dato che al tempo 0 una chiamata in arrivo al sistema lo trova occupato, con che probabilità il sistema è

Con riferimento al caso elementare di catena di Markov a due stati, il calcolo della distribuzione stazionaria sarebbe stato banale, partendo da considerazioni meramente intuitive. Infatti, essendo  $T_{ON}=4s$  e  $T_{OFF}=1s$ , la probabilità in un istante di ispezione casuale di trovare il sistema nello stato ON è di 4 volte quella di trovare il sistema nello stato OFF.

Ma supponiamo di non volerci avvalere di queste considerazioni intuitive. Possiamo notare che **il calcolo della distribuzione stazionaria comunque NON richiede la soluzione delle equazioni di Chapman-Kolmogorov, ma solo la loro formulazione!** Infatti, ricordiamo che nel caso in questione le equazioni di Chapman-Kolmogorov sono date da:

$$\begin{aligned} P'_{ON}(t) &= -\alpha \cdot P_{ON}(t) + \beta \cdot P_{OFF}(t) \\ P'_{OFF}(t) &= \alpha \cdot P_{ON}(t) - \beta \cdot P_{OFF}(t) \end{aligned}$$

Sapendo che a regime (ovvero per  $t \rightarrow \infty$ ) le probabilità  $P_{ON}$  e  $P_{OFF}$  convergono a valori costanti, possiamo concludere che a regime le corrispondenti derivate tendono a 0. Pertanto, per il calcolo della distribuzione stazionaria, sarebbe bastato scrivere il seguente sistema di equazioni lineari (N.B: non più equazioni differenziali):

$$\begin{aligned} 0 &= -\alpha \cdot P_{ON}(\infty) + \beta \cdot P_{OFF}(\infty) & \Rightarrow & P_{ON}(\infty) = \frac{\beta}{\alpha} P_{OFF}(\infty) \\ 0 &= \alpha \cdot P_{ON}(\infty) - \beta \cdot P_{OFF}(\infty) & \Rightarrow & P_{ON}(\infty) = \frac{\beta}{\alpha} P_{OFF}(\infty) \end{aligned}$$

Da cui si nota che le due equazioni sono linearmente dipendenti tra loro. Al fine di pervenire alla soluzione numerica, sarebbe bastato imporre, in luogo di una delle due equazioni, la condizione “di normalizzazione”:

$$\begin{aligned} 0 &= -\alpha \cdot P_{ON}(\infty) + \beta \cdot P_{OFF}(\infty) & \Rightarrow & P_{ON}(\infty) = \frac{\beta}{\alpha} P_{OFF}(\infty) \\ P_{ON}(\infty) + P_{OFF}(\infty) &= 1 & \Rightarrow & P_{ON}(\infty) = \frac{\beta}{\alpha + \beta}, \quad P_{OFF}(\infty) = \frac{\alpha}{\alpha + \beta} \end{aligned}$$

### III.2.2 Teorema di conservazione dei flussi di probabilità

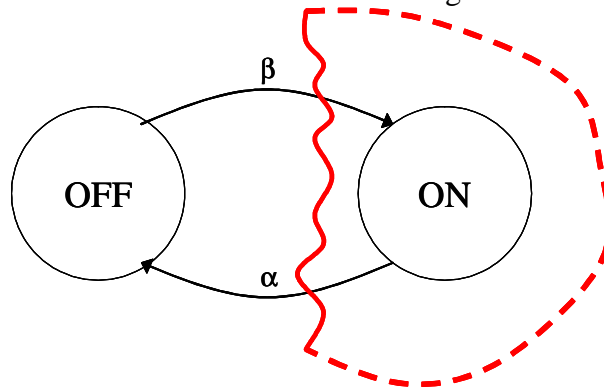
Chiarito che il calcolo della distribuzione stazionaria non richiede la soluzione esplicita delle equazioni di Chapman-Kolmogorov, ma la loro sola formulazione, in questa sezione enunciamo un teorema che permette di semplificare enormemente stesura del sistema di equazioni lineari la cui soluzione permette di calcolare la distribuzione stazionaria. Il teorema di conservazione dei flussi di probabilità si può enunciare come segue.

---

ancora occupato al tempo  $t=1$ ?”. Tuttavia, quello che interessa nella pratica è la probabilità che una generica chiamata trovi il sistema occupato a prescindere dalla conoscenza di cosa è successo precedentemente, e questa coincide con la probabilità a regime di trovare il sistema nello stato ON.

Chiamiamo **taglio** operato sull'insieme degli stati una suddivisione degli stati in due insiemi disgiunti e complementari. Per un sistema in equilibrio statistico (a regime), **la somma algebrica dei flussi di probabilità attraverso un qualunque taglio è nulla.**

Nel caso elementare di catena di Markov a due stati trattata finora, è evidente che esiste un solo possibile taglio, illustrato nella figura seguente. Nella figura, al fine di sottolineare graficamente come un taglio “isoli” un insieme di stati rispetto agli altri stati, il taglio è stato illustrato come una linea chiusa. Che non sia possibile operare un secondo taglio è evidente: un'eventuale taglio “intorno” allo stato OFF (in luogo di quello disegnato “intorno” allo stato ON) avrebbe l'esito di suddividere l'insieme dei due stati negli stessi identici due sottinsiemi!



Per convenienza di notazione, chiamiamo  $\pi_{ON}$  e  $\pi_{OFF}$  le probabilità limite di stato, in luogo della precedente notazione  $P_{ON}(\infty)$  e  $P_{OFF}(\infty)$ . Notiamo che:

- Il flusso entrante attraverso il taglio considerato è  $\pi_{OFF} \cdot \beta$ ;
- Il flusso uscente attraverso il taglio considerato è  $\pi_{ON} \cdot \alpha$ ;

Applicando il teorema enunciato precedentemente, possiamo pertanto scrivere la relazione:

$$0 = -\alpha \cdot \pi_{ON} + \beta \cdot \pi_{OFF} \quad \Rightarrow \quad \pi_{ON} = \frac{\beta}{\alpha} \pi_{OFF}$$

Imponendo la condizione di normalizzazione (la somma delle probabilità  $\pi_x$  deve essere 1, altrimenti non avremmo una distribuzione di probabilità) otteniamo le probabilità limite di stato:

$$\pi_{ON} = \frac{\beta}{\alpha + \beta}$$

$$\pi_{OFF} = \frac{\alpha}{\alpha + \beta}$$

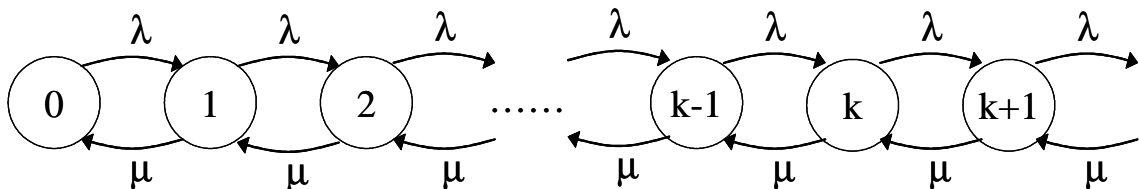
che, a parte la differente notazione, ovviamente coincidono con quelle calcolate nella sezione III.2.1.

### III.3 Distribuzione stazionaria: caso M/M/1

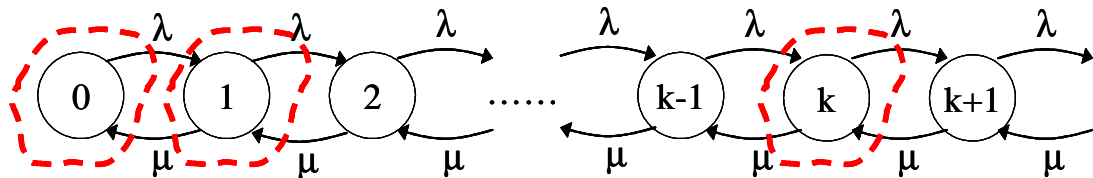
Riprendiamo il caso importante del sistema M/M/1. Nella sezione III.1.1 abbiamo verificato che il processo  $N(t)$  che rappresenta il numero di clienti in un sistema M/M/1 è una catena di Markov. Il processo  $N(t)$  è un processo di nascita e morte, in quanto prevede salti di al più una unità (partenze e/o arrivi singoli). Pertanto, le uniche frequenze di transizione di stato non nulle sono tra stati adiacenti. In particolare:

- Si ha una “nascita” in presenza di un arrivo di un utente al sistema. Visto che in un intervallo di tempo  $\Delta t$  si ha un arrivo con probabilità  $\lambda \Delta t$ , la frequenza di transizione da un generico stato  $k$  ( $k \geq 0$ ) allo stato  $k+1$  è  $\lambda$ .
- Si ha una “morte” quando un utente completa il servizio. Per qualunque stato  $k > 0$ , in un intervallo di tempo  $\Delta t$ , ciò avviene con probabilità  $\mu \Delta t$ . Pertanto, la frequenza di transizione da un generico stato  $k$  ( $k > 0$ ) allo stato  $k-1$  è  $\mu$ . Si noti che non è possibile avere “morti” nello stato 0.

Ciò porta al diagramma di stato illustrato nella figura seguente.



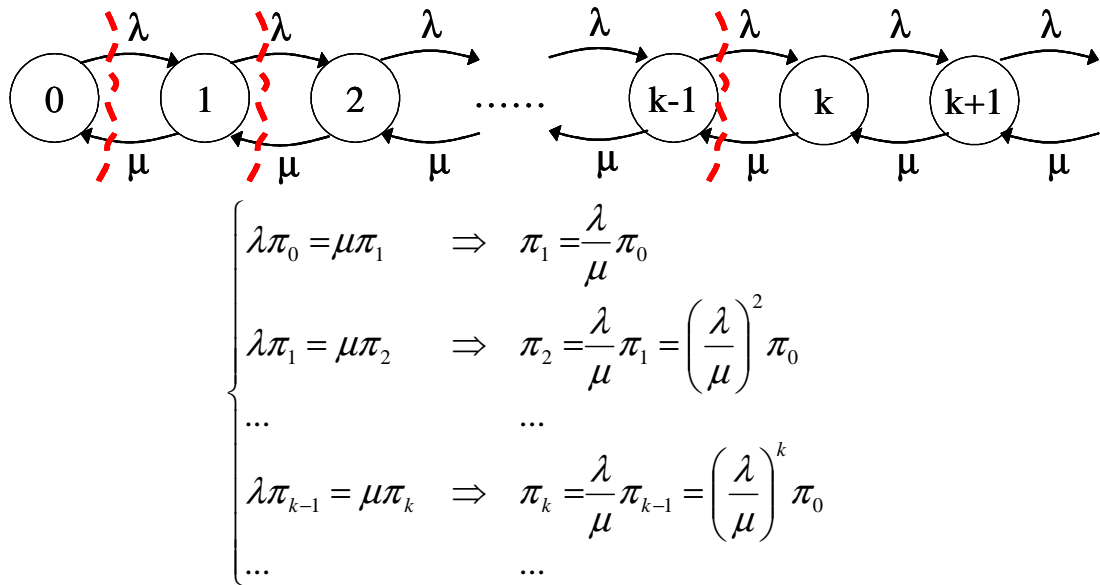
Al fine di scrivere il sistema lineare risolutivo, si può procedere applicando il teorema di conservazione dei flussi a tagli effettuati intorno ad ogni singolo stato, come illustrato in figura:



Dai tagli intorno agli stati 0, 1, ..., k, si otterrebbe il sistema lineare ad infinite equazioni

$$\begin{cases} \lambda\pi_0 = \mu\pi_1 \\ \lambda\pi_0 + \mu\pi_2 = (\lambda + \mu)\pi_1 \\ \dots \\ \lambda\pi_{k-1} + \mu\pi_{k+1} = (\lambda + \mu)\pi_k \\ \dots \end{cases}$$

Che può essere risolto ricorsivamente, esprimendo nella prima equazione  $\pi_1$  in funzione di  $\pi_0$ , nella seconda equazione  $\pi_2$  in funzione di  $\pi_0$  (essendo  $\pi_1$  oramai noto), etc. Tuttavia è didatticamente istruttivo mostrare come una scelta più “furba” dei tagli permette di semplificare la stesura del sistema (ovviamente il risultato finale non cambia):



Si noti che, al fine di completare il calcolo della distribuzione stazionaria, è necessario trovare  $\pi_0$ . Ciò è fatto imponendo la condizione di normalizzazione:

$$1 = \sum_{k=0}^{\infty} \pi_k = \pi_0 \sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^k = \pi_0 \frac{1}{1 - \frac{\lambda}{\mu}} \Rightarrow \pi_0 = 1 - \frac{\lambda}{\mu}$$

dove la sommatoria converge solo se  $\lambda/\mu < 1$ .

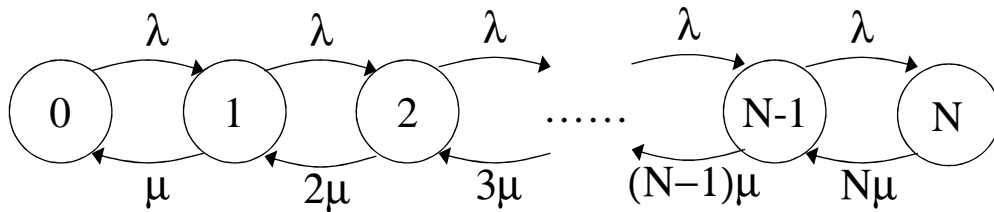
Riassumendo, detta  $\rho = \lambda/\mu$  l'utilizzazione del sistema, la distribuzione stazionaria della catena considerata (ovvero la probabilità di trovare, a regime,  $k$  utenti nel sistema) è data da  $\pi_k = (1 - \rho)\rho^k$ .

### III.4 Distribuzione stazionaria: caso M/M/N/N

Siamo ora in grado di derivare formalmente la distribuzione stazionaria per il caso particolarmente importante di chiamate offerte ad un fascio di circuiti ed il suo caso speciale rappresentato dalla formula B di Erlang per il calcolo della relativa probabilità di blocco. Possiamo infatti modellare questo sistema come un sistema a coda composto da  $N$  circuiti (serventi) che operano in parallelo. Il numero di chiamate (utenti) che può essere ammesso al sistema considerato è al più uguale ad  $N$ . Diciamo che il sistema è a pura perdita, in quanto non può accomodare (in una eventuale fila di attesa) utenti in soprannumero rispetto al numero di serventi.

Consideriamo il processo stocastico  $N(t)$  che rappresenta il numero di circuiti (serventi) occupati al tempo  $t$ . E' banale dimostrare che tale processo stocastico è una catena di markov nell'assunzione di processo di arrivi markoviano (ovvero tempo di interarrivo tra chiamate distribuito esponenzialmente) e tempo di servizio esponenziale. Detto  $\lambda$  il tasso di arrivo delle chiamate al sistema (chiamate/secondo), e detto  $1/\mu$  il tempo medio di servizio di una

chiamata, possiamo rappresentare il sistema mediante il seguente diagramma di transizione di stato.



Il diagramma si può spiegare immediatamente come segue, studiando come caso particolare ciò che succede nello stato 2 e poi generalizzando. Il sistema è nello stato 2 quando due chiamate, diciamo la chiamata A e la chiamata B, sono attive e tutti gli altri circuiti sono liberi. In tale stato consideriamo un piccolo intervallo di tempo  $\Delta t$ . In tale intervallo, per assunzione di markovianità del processo degli arrivi e distribuzione esponenziale negativa del tempo di servizio possono indipendentemente avvenire i seguenti eventi con le seguenti probabilità:

$$P(\text{arrivo in } \Delta t) = \lambda \Delta t$$

$$P(\text{fine chiamata A in } \Delta t) = \mu \Delta t$$

$$P(\text{fine chiamata B in } \Delta t) = \mu \Delta t$$

La frequenza di transizione dallo stato 2 allo stato 3 è quantificabile come il limite del rapporto tra la probabilità che nell'intervallo  $\Delta t$  considerato si abbia un arrivo e nessuna partenza (tutti gli eventi multipli sono trascurabili in quanto infinitesimi di ordine superiore a  $\Delta t$ ) e l'intervallo stesso  $\Delta t$ , ovvero:

$$q(2 \rightarrow 3) = \lim_{\Delta t \rightarrow 0} \frac{\lambda \Delta t (1 - \mu \Delta t)(1 - \mu \Delta t)}{\Delta t} = \lambda$$

La frequenza di transizione dallo stato 2 allo stato 1 si può invece calcolare considerando che tale transizione avviene, trascurando gli eventi multipli, nel caso in cui o la chiamata A o la chiamata B terminano, ovvero:

$$q(2 \rightarrow 1) = \lim_{\Delta t \rightarrow 0} \frac{(1 - \lambda \Delta t) \mu \Delta t (1 - \mu \Delta t) + (1 - \lambda \Delta t) (1 - \mu \Delta t) \mu \Delta t}{\Delta t} = 2\mu$$

Le altre probabilità di transizione (verso stati non adiacenti, ad esempio  $2 \rightarrow 0$  o  $2 \rightarrow 4$ ) sono nulle in quanto la probabilità di occorrenza di tali transizioni in  $\Delta t$  risulta essere  $o(\Delta t)$ .

Una volta giustificato il diagramma sopra illustrato, la distribuzione stazionaria può essere immediatamente calcolata applicando il metodo del bilanciamento dei flussi di probabilità, che porta a scrivere le seguenti equazioni:

$$\begin{aligned}
\lambda\pi_0 = \mu\pi_1 &\Rightarrow \pi_1 = \frac{\lambda}{\mu}\pi_0 \\
\lambda\pi_1 = 2\mu\pi_2 &\Rightarrow \pi_2 = \frac{\lambda}{2\mu}\pi_1 = \frac{1}{2}\left(\frac{\lambda}{\mu}\right)^2\pi_0 \\
&\dots \dots \dots \\
\lambda\pi_{k-1} = k\mu\pi_k &\Rightarrow \pi_k = \frac{\lambda}{k\mu}\pi_{k-1} = \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\pi_0 \\
&\dots \dots \dots \\
\lambda\pi_{N-1} = N\mu\pi_N &\Rightarrow \pi_N = \frac{\lambda}{N\mu}\pi_{N-1} = \frac{1}{N!}\left(\frac{\lambda}{\mu}\right)^N\pi_0
\end{aligned}$$

Da cui concludiamo, una volta imposta la condizione di normalizzazione, e chiamato con  $A=\lambda/\mu$  il rapporto tra tasso di arrivo e tasso di servizio,

$$\pi_k = \frac{A^k}{k!}\pi_0 = \frac{A^k / k!}{\sum_{i=0}^N A^i / i!}$$

Pertanto ritrovando, per il caso particolare di probabilità di blocco  $\pi_N$  la formula B di Erlang<sup>7</sup> a suo tempo data.

---

<sup>7</sup> Si noti che nella derivazione qui proposta, è stata fatta l'assunzione di tempi di servizio distribuiti esponenzialmente. In realtà è possibile dimostrare (ma la dimostrazione è decisamente più complessa e ben oltre gli obiettivi di queste dispense) che la formula B di Erlang è valida anche nel caso in cui le chiamate hanno distribuzione di tipo generale, ovvero anche per sistemi M/G/N/N. Tale proprietà è chiamata "insensibilità" della probabilità di blocco alla distribuzione del tempo di servizio. In pratica, la probabilità di blocco in un sistema a pura perdita NON dipende dalla distribuzione statistica della durata delle chiamate, ma dipende SOLO dal suo valor medio.