Probabilità e Processi Stocastici (455AA)

Lezione 18

Dario Trevisan

23/11/2022

Section 1

Stima dei parametri di una catena di Markov dalle osservazioni

Stima dei parametri

Consideriamo il problema di stimare i *parametri* di un processo di Markov omogeneo, sulla base di osservazioni di una traiettoria, ossia la matrice

- delle probabilità di transizione Q per una catena di Markov $(X_n)_n$

Stima dei parametri

Consideriamo il problema di stimare i *parametri* di un processo di Markov omogeneo, sulla base di osservazioni di una traiettoria, ossia la matrice

- delle probabilità di transizione Q per una catena di Markov $(X_n)_n$
- o delle intensità di salto L per un processo di Markov a salti $(X_t)_t$,

Due approcci

Si osserva che la catena X segue un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$, ossia $X_0 = x_0$, $X_1 = x_1$, ..., $X_n = x_n$ (scriviamo $X = \gamma$). Consideriamo i due approcci:

- Massima verosimiglianza: massimizzare

$$Q \mapsto L(\mathcal{Q} = Q; X = \gamma) = P(X = \gamma | \mathcal{Q} = Q) = P(X_0 = x_0) Q_\gamma$$

Due approcci

Si osserva che la catena X segue un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$, ossia $X_0 = x_0$, $X_1 = x_1$, ..., $X_n = x_n$ (scriviamo $X = \gamma$). Consideriamo i due approcci:

- Massima verosimiglianza: massimizzare

$$Q \mapsto L(\mathcal{Q} = Q; X = \gamma) = P(X = \gamma | \mathcal{Q} = Q) = P(X_0 = x_0) Q_\gamma$$

- Bayesiano: si introduce una densità a priori per $\mathcal{Q}$ (vista come variabile aleatoria) e si stima tramite Bayes la densità a posteriori, noto $X = \gamma$,

$$p(\mathcal{Q} = Q | X = \gamma) \propto p(\mathcal{Q} = Q) L(\mathcal{Q} = Q; X = \gamma),$$

Stima di massima verosimiglianza

Supponiamo per semplificare che sia $P(X_0 = x_0) = 1$, così

$$L(Q; X = \gamma) = Q_\gamma = \prod_{k=1}^{n-1} Q_{x_{k-1} \to x_k}.$$

- Raccogliendo i fattori ripetuti,

$$Q_\gamma = \prod_{k=1}^{n} Q_{x_{k-1} \to x_k} = \prod_{i,j \in E} Q_{i \to j}^{\gamma_{i \to j}},$$

  dove $\gamma_{i \to j}$ è il numero di transizioni dallo stato $i \in E$ a $j \in E$ che avvengono in γ.

Stima di massima verosimiglianza

Supponiamo per semplificare che sia $P(X_0 = x_0) = 1$, così

$$L(Q; X = \gamma) = Q_\gamma = \prod_{k=1}^{n-1} Q_{x_{k-1} \to x_k}.$$

- Raccogliendo i fattori ripetuti,

  $$Q_\gamma = \prod_{k=1}^{n} Q_{x_{k-1} \to x_k} = \prod_{i,j \in E} Q_{i \to j}^{\gamma_{i \to j}},$$

  dove $\gamma_{i \to j}$ è il numero di transizioni dallo stato $i \in E$ a $j \in E$ che avvengono in γ.

- In particolare,

  $$n = \sum_{i,j \in E} \gamma_{i \to j}.$$

Un massimo vincolato

Per calcolare il punto di massimo, di

$$Q \mapsto \prod_{i,j \in E} Q_{i\to j}^{\gamma_{i\to j}}$$

dobbiamo tenere conto del vincolo che la somma delle righe della matrice Q sia 1.

- Per determinare massimi o minimi di funzioni vincolate in generale si usano i moltiplicatori di Lagrange: nei punti critici il gradiente della funzione sia ortogonale al vincolo.

Una sostituzione

Nel nostro caso evitiamo i moltiplicatori esprimendo la diagonale di $\mathcal{Q}$ in termini delle altre entrate sulla riga:

$$Q_{i\to i} = 1 - \sum_{j\neq i} Q_{i\to j} \quad \text{per ogni } i \in E.$$

- Riscriviamo la verosimiglianza

$$L(\mathcal{Q} = Q; X = \gamma) = \prod_{i\in E} (1 - \sum_{j\neq i} Q_{i\to j})^{\gamma_{i\to i}} \prod_{j\neq i} Q_{i\to j}^{\gamma_{i\to j}}.$$

Una sostituzione

Nel nostro caso evitiamo i moltiplicatori esprimendo la diagonale di $\mathcal{Q}$ in termini delle altre entrate sulla riga:

$$Q_{i\to i} = 1 - \sum_{j\neq i} Q_{i\to j} \quad \text{per ogni } i \in E.$$

- Riscriviamo la verosimiglianza

$$L(\mathcal{Q} = Q; X = \gamma) = \prod_{i\in E} (1 - \sum_{j\neq i} Q_{i\to j})^{\gamma_{i\to i}} \prod_{j\neq i} Q_{i\to j}^{\gamma_{i\to j}}.$$

- Possiamo ragionare separatamente per ciascuna riga i, e massimizzare

$$(Q_{i\to j})_{j\neq i} \mapsto (1 - \sum_{j\neq i} Q_{i\to j})^{\gamma_{i\to i}} \prod_{j\neq i} Q_{i\to j}^{\gamma_{i\to j}},$$

Conclusione

Il metodo bayesiano

Se la densità a priori per $\mathcal{Q}$ è della forma

$$p(\mathcal{Q} = Q) \propto \prod_{i,j \in E} Q_{i \to j}^{\alpha_{ij}},$$

per opportuni parametri $\alpha_{ij} \geq 0$, la moda della densità sopra è data dalla matrice

$$Q_{i \to j} = \frac{\alpha_{ij}}{\sum_{k \in E} \alpha_{ik}} \propto \alpha_{ij}.$$

- La formula di Bayes darebbe quindi come densità a posteriori

$$p(\mathcal{Q} = Q | X = \gamma) \propto \prod_{i,j \in E} Q_{i \to j}^{\alpha_{ij} + \gamma_{i \to j}},$$

  con stima di massima verosimiglianza

$$Q_{i \to j} = \frac{\alpha_{ij} + \gamma_{i \to j}}{\sum_{k \in E} \alpha_{ik} + \gamma_{i \to k}} \propto \alpha_{ij} + \gamma_{ij}.$$

Processi di Markov a salti

Nel caso di processi di Markov a salti, l'argomento è analogo ma si basa sulla formula per la "densità" di probabilità di un cammino.

- Presentiamo solo la stima di massima verosimiglianza. Si consideri un cammino $\gamma = (x_0 \to x_1 \ldots x_n)$ che rimane per un tempo t_1 nello stato x_0, t_2 nello stato x_1 ecc., e si supponga di osservare $X = \gamma$, ossia tutta la traiettoria da $X_0 = x_0$ fino a $X_{t_1+\ldots+t_n} = x_n$.

Processi di Markov a salti

Nel caso di processi di Markov a salti, l'argomento è analogo ma si basa sulla formula per la "densità" di probabilità di un cammino.

- Presentiamo solo la stima di massima verosimiglianza. Si consideri un cammino $\gamma = (x_0 \to x_1 \dots x_n)$ che rimane per un tempo t_1 nello stato x_0, t_2 nello stato x_1 ecc., e si supponga di osservare $X = \gamma$, ossia tutta la traiettoria da $X_0 = x_0$ fino a $X_{t_1+\dots+t_n} = x_n$.
- La verosimiglianza per L è

$$\prod_{k=1}^{n} \exp\left(t_k L_{x_{k-1}\to x_{k-1}}\right) L_{x_{k-1}\to x_k} = \prod_{i\in E} \exp\left(\gamma_{i\to i} L_{i\to i}\right) \prod_{i\neq j\in E} L_{i\to j}^{\gamma_{i\to j}}$$

  dove $\gamma_{i\to j}$ per $i \neq j$ è come prima ma $\gamma_{i\to i}$ è il tempo totale trascorso dal cammino nello stato $i \in E$.

Massimizzazione

Eliminiamo il vincolo che la somma sulle righe di L è nulla:

$$\exp\left(\gamma_{i\to i} L_{i\to i}\right) = \exp\left(-\gamma_{i\to i} \sum_{j\neq i} L_{i\to j}\right).$$

- Passando ai logaritmi e derivando si ottiene che $\mathcal{L}_{MLE}$ è data dall'espressione, per $i \neq j$,

$$L_{i\to j} = \frac{\gamma_{i\to j}}{\gamma_{i\to i}}.$$

Modelli nascosti

Abbiamo supposto di osservare completamente la catena X in un intervallo (discreto o continuo) di tempi.

- Cosa accade se mancano le osservazioni delle variabili X_k in alcuni tempi?

Modelli nascosti

Abbiamo supposto di osservare completamente la catena X in un intervallo (discreto o continuo) di tempi.

- Cosa accade se mancano le osservazioni delle variabili X_k in alcuni tempi?
- oppure se si osserva solamente una funzione $g(X_k)$ della catena invece, di X_k, o più in generale una funzione $g(X_k, Z_k)$ dove Z è un processo indipendente da X?

Modelli nascosti

Abbiamo supposto di osservare completamente la catena X in un intervallo (discreto o continuo) di tempi.

- Cosa accade se mancano le osservazioni delle variabili X_k in alcuni tempi?
- oppure se si osserva solamente una funzione $g(X_k)$ della catena invece, di X_k, o più in generale una funzione $g(X_k, Z_k)$ dove Z è un processo indipendente da X?
- In questa situazione si parla di modelli di Markov nascosti (in inglese *Hidden Markov Models*, HMM) e la ricostruzione di X_k è il problema del filtraggio.

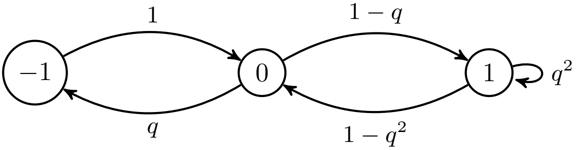
Modelli nascosti

Abbiamo supposto di osservare completamente la catena X in un intervallo (discreto o continuo) di tempi.

- Cosa accade se mancano le osservazioni delle variabili X_k in alcuni tempi?
- oppure se si osserva solamente una funzione $g(X_k)$ della catena invece, di X_k, o più in generale una funzione $g(X_k, Z_k)$ dove Z è un processo indipendente da X?
- In questa situazione si parla di modelli di Markov nascosti (in inglese *Hidden Markov Models*, HMM) e la ricostruzione di X_k è il problema del filtraggio.
- Opportuni algoritmi (EM) permettono di stimare i parametri di un HMM.

Un esempio/esercizio

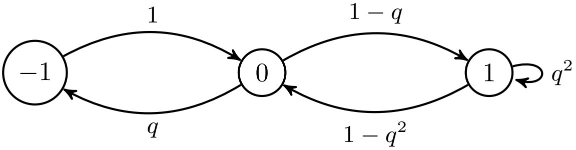
Si consideri una catena di Markov con probabilità di transizione rappresentate in figura, dove $q \in (0,1)$ è un parametro (non aleatorio).



1. Supponendo che X sia stazionaria, dire al variare di $q \in (0,1)$ se è più probabile che sia $X_0 = 0$ oppure $X_0 \neq 0$.

Un esempio/esercizio

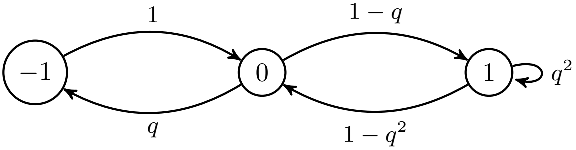
Si consideri una catena di Markov con probabilità di transizione rappresentate in figura, dove $q \in (0,1)$ è un parametro (non aleatorio).



1. Supponendo che X sia stazionaria, dire al variare di $q \in (0,1)$ se è più probabile che sia $X_0 = 0$ oppure $X_0 \neq 0$.
2. Si supponga noto a priori che $X_0 = 0$. Si osserva il cammino $0 \to -1 \to 0 \to -1 \to 0 \to 1 \to 1$. Determinare la stima di massima verosimiglianza q_{MLE}.

Un esempio/esercizio

Si consideri una catena di Markov con probabilità di transizione rappresentate in figura, dove $q \in (0,1)$ è un parametro (non aleatorio).



1. Supponendo che X sia stazionaria, dire al variare di $q \in (0,1)$ se è più probabile che sia $X_0 = 0$ oppure $X_0 \neq 0$.
2. Si supponga noto a priori che $X_0 = 0$. Si osserva il cammino $0 \to -1 \to 0 \to -1 \to 0 \to 1 \to 1$. Determinare la stima di massima verosimiglianza q_{MLE}.
3. Si supponga noto a priori che X sia stazionaria. Si osserva lo stesso cammino della domanda di prima. Determinare q_{MLE}.

1.

2

3. Stavolta

$$L(q; X=\gamma) = P(X_0 = 0 | X \text{ stazionaria}, q) Q_\gamma = \frac{1+q}{2+2q-q^2} q^2 (1-q) q^2.$$

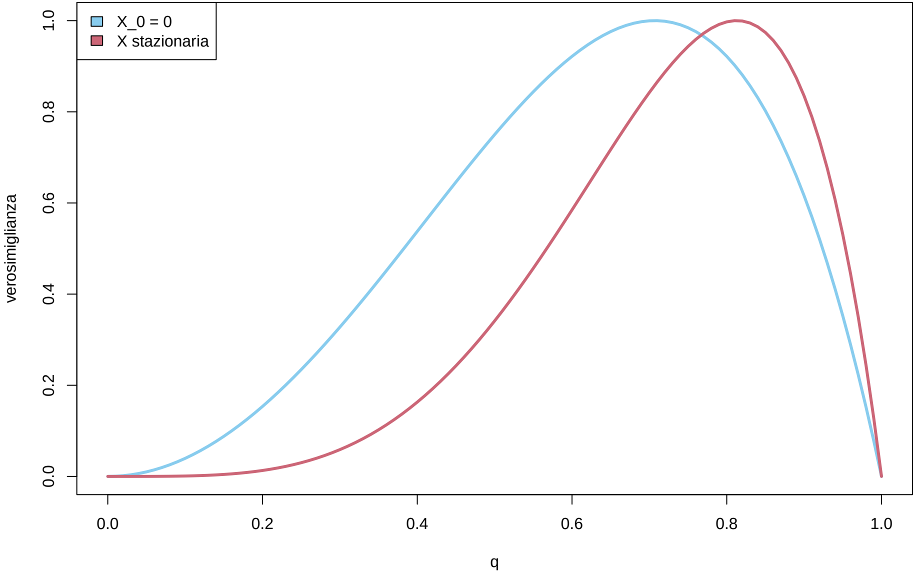
Stimiamo numericamente q_{MLE}.

```
likelihood_stazionaria = function(q){
  -(1-q^2)*q^4/(2+2*q-q^2)
}

x_mle = nlm(likelihood_stazionaria, 1/2)

x_mle$estimate
```

```
## [1] 0.8129379
```


X_0 = 0
X stazionaria
verosimiglianza
q
0.0
0.2
0.4
0.6
0.8
1.0

Section 2

Cenni alla teoria delle code

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale
- istanze di calcolo che devono essere eseguite da una o più processori in un computer...

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale
- istanze di calcolo che devono essere eseguite da una o più processori in un computer. . .
- L'obiettivo è individuare strategie per migliorare l'esperienza di chi è in attesa (ridurre i tempi) rendendone più efficiente il servizio.

Introduzione ai modelli di code

La *teoria delle code* studia le linee d'attesa che si possono formare in tante situazioni, ad esempio:

- persone che vogliono accedere ad un servizio (entrare in un negozio, o pagare alla cassa)
- veicoli che si presentano ad un casello autostradale
- istanze di calcolo che devono essere eseguite da una o più processori in un computer...
- L'obiettivo è individuare strategie per migliorare l'esperienza di chi è in attesa (ridurre i tempi) rendendone più efficiente il servizio.
- La teoria delle code è un campo molto esteso, presentiamo i modelli più semplici come esempi di processi di Markov a salti.

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,
- il tempo d'attesa che un servente prenda in carico il compito richiesto,

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,
- il tempo d'attesa che un servente prenda in carico il compito richiesto,
- e infine l'uscita dalla coda quando il compito è svolto

Clienti e serventi

Per indicare le persone, le auto, le istanze ecc. da servire usiamo il termine **clienti** (in inglese si usa a volte *jobs*)

Usiamo il termine **serventi** (in inglese *servers*) per chi eroga il servizio richiesto dei clienti.

Aspetti da modellizzare:

- l'ingresso di uno o più clienti,
- il tempo d'attesa che un servente prenda in carico il compito richiesto,
- e infine l'uscita dalla coda quando il compito è svolto
- Una volta introdotto un modello, è di interesse calcolare il tempo medio di attesa, il numero medio di clienti in coda e stimare i parametri di un modello sulla base di quantità osservate nella realtà.

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,

Notazione di Kendall

Kendall propose una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.
- Noi studiamo solo i modelli $M/M/c$: arrivi e tempi di servizio sono Markoviani (a tempi continui).

Notazione di Kendall

Kendall propose una notazione abbreviata $A/S/c$:

- A indica un "processo" di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.
- Noi studiamo solo i modelli $M/M/c$: arrivi e tempi di servizio sono Markoviani (a tempi continui).
- Formalmente i modelli sono processi di Markov a salti negli stati $E = \mathbb{N}$.

Notazione di Kendall

Kendall propone una notazione abbreviata $A/S/c$:

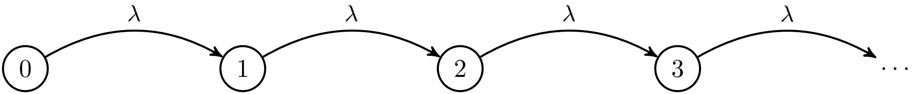
- A indica un “processo” di arrivo dei clienti,
- S la legge del tempo di servizio per ciascun cliente,
- c il numero dei serventi.
- Noi studiamo solo i modelli $M/M/c$: arrivi e tempi di servizio sono Markoviani (a tempi continui).
- Formalmente i modelli sono processi di Markov a salti negli stati $E = \mathbb{N}$.
- Lo stato n indica il numero di clienti in attesa o in corso di servizio.

Caso $M/M/0$: il processo di Poisson

Il modello più semplice è il caso in cui non vi siano serventi (oppure si è interessati solo al processo di arrivo dei clienti): il processo è detto *processo di Poisson* di intensità $\lambda > 0$.

Le intensità di salto sono

$$L_{n\to n+1} = \lambda, \quad L_{n\to n} = -\lambda \quad \text{e} \quad L_{n\to k} = 0 \qquad k \neq n,\, k \neq n+1.$$



- Ogni stato è transitorio, non vi sono distribuzioni invarianti.

Legame con la densità Poisson

Se $X_0 = 0$, allora la densità marginale di X_t è Poisson di intensità λt, ossia

$$\mu_n^t \propto \frac{(t\lambda)^n}{n!} = \frac{(t\lambda)^n}{n!} \exp(-t\lambda)\,.$$

- Basta verificare che valga la *master equation*, per ogni $n \in \mathbb{N}$, $t \geq 0$,

$$\frac{d}{dt}\mu_n^t = (\mu^t L)_n = \begin{cases} -\lambda \mu_0^t & \text{se } n = 0, \\ \lambda(\mu_{n-1}^t - \mu_n^t) & \text{se } n \geq 1. \end{cases}$$

Legame con la densità Poisson

Se $X_0 = 0$, allora la densità marginale di X_t è Poisson di intensità λt, ossia

$$\mu_n^t \propto \frac{(t\lambda)^n}{n!} = \frac{(t\lambda)^n}{n!} \exp(-t\lambda).$$

- Basta verificare che valga la *master equation*, per ogni $n \in \mathbb{N}$, $t \geq 0$,

$$\frac{d}{dt}\mu_n^t = (\mu^t L)_n = \begin{cases} -\lambda \mu_0^t & \text{se } n = 0, \\ \lambda(\mu_{n-1}^t - \mu_n^t) & \text{se } n \geq 1. \end{cases}$$

- Calcoliamo quindi

$$\frac{d}{dt} \exp(-t\lambda) \frac{(t\lambda)^n}{n!} = \begin{cases} -\lambda \exp(-t\lambda) & \text{se } n = 0, \\ \frac{nt^{n-1}\lambda^n}{n!} \exp(-t\lambda) - \lambda \frac{(t\lambda)^n}{n!} \exp(-t\lambda) & \text{se } n \geq 1 \end{cases}$$

Legame con la densità Poisson

Se $X_0 = 0$, allora la densità marginale di X_t è Poisson di intensità λt, ossia

$$\mu_n^t \propto \frac{(t\lambda)^n}{n!} = \frac{(t\lambda)^n}{n!} \exp\left(-t\lambda\right).$$

- Basta verificare che valga la *master equation*, per ogni $n \in \mathbb{N}$, $t \geq 0$,

$$\frac{d}{dt}\mu_n^t = (\mu^t L)_n = \begin{cases} -\lambda \mu_0^t & \text{se } n = 0, \\ \lambda(\mu_{n-1}^t - \mu_n^t) & \text{se } n \geq 1. \end{cases}$$

- Calcoliamo quindi

$$\frac{d}{dt} \exp\left(-t\lambda\right) \frac{(t\lambda)^n}{n!} = \begin{cases} -\lambda \exp\left(-t\lambda\right) & \text{se } n = 0, \\ \frac{nt^{n-1}\lambda^n}{n!} \exp\left(-t\lambda\right) - \lambda \frac{(t\lambda)^n}{n!} \exp\left(-t\lambda\right) & \text{se } n \geq 1 \end{cases}$$

- Per concludere nel caso $n \geq 1$ basta notare che

$$\frac{nt^{n-1}\lambda^n}{n!} \exp\left(-t\lambda\right) = \lambda \frac{(t\lambda)^{n-1}}{(n-1)!} \exp\left(-t\lambda\right) = \lambda \mu_{n-1}^t.$$

Stima del parametro λ dalle osservazioni

Supponiamo di osservare un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con tempi di permanenza t_1 (nello stato x_0), t_2 (in x_1), ..., t_n.

- Poiché i salti avvengono solo tra n e $n+1$, deve essere $x_1 = x_0 + 1$, $x_2 = x_0 + 2$, ecc.

Stima del parametro λ dalle osservazioni

Supponiamo di osservare un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con tempi di permanenza t_1 (nello stato x_0), t_2 (in x_1), ..., t_n.

- Poiché i salti avvengono solo tra n e $n+1$, deve essere $x_1 = x_0 + 1$, $x_2 = x_0 + 2$, ecc.
- La verosimiglianza è

$$L(\Lambda = \lambda; X = \gamma) = \prod_{k=1}^{n} \exp\left(-\lambda t_k\right) \lambda = \lambda^n \exp\left(-\lambda T\right).$$

dove supponiamo noto a priori che $X_0 = x_0$ e $T = \sum_{k=1}^{n} t_k$.

Stima del parametro λ dalle osservazioni

Supponiamo di osservare un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con tempi di permanenza t_1 (nello stato x_0), t_2 (in x_1), ..., t_n.

- Poiché i salti avvengono solo tra n e $n+1$, deve essere $x_1 = x_0 + 1$, $x_2 = x_0 + 2$, ecc.
- La verosimiglianza è

$$L(\Lambda = \lambda; X = \gamma) = \prod_{k=1}^{n} \exp\left(-\lambda t_k\right) \lambda = \lambda^n \exp\left(-\lambda T\right).$$

  dove supponiamo noto a priori che $X_0 = x_0$ e $T = \sum_{k=1}^{n} t_k$.

- La stima di massima verosimiglianza si trova annullando la derivata rispetto a λ e vale

$$\frac{n}{\lambda_{MLE}} - T = 0 \quad \text{quindi} \quad \lambda_{MLE} = \frac{n}{T}.$$

Code $M/M/1$

Supponiamo vi sia un solo servente e che il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (ogni cliente sia indipendente dagli altri).

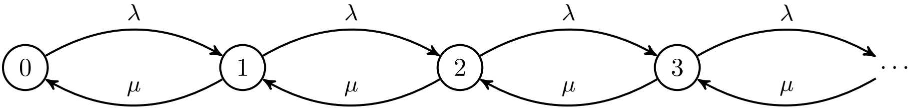
- Per arrivare al modello come un processo di Markov a salti, supponiamo non vi siano arrivi: si salta solo da n verso $n-1$ (se $n \geq 1$) con dei tempi di permanenza esponenziali di parametro μ. Pertanto, se $n \geq 1$,

$$L_{n \to n-1} = \mu.$$

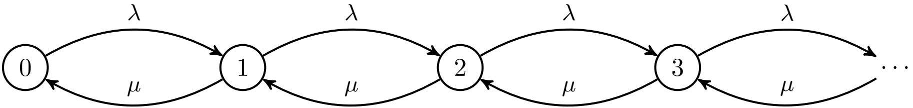
Code $M/M/1$

Supponiamo vi sia un solo servente e che il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (ogni cliente sia indipendente dagli altri).

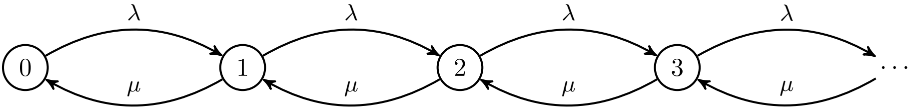
- Per arrivare al modello come un processo di Markov a salti, supponiamo non vi siano arrivi: si salta solo da n verso $n-1$ (se $n \geq 1$) con dei tempi di permanenza esponenziali di parametro μ. Pertanto, se $n \geq 1$,
$$L_{n \to n-1} = \mu.$$
- Aggiungiamo gli arrivi come un processo di Poisson di intensità λ: per $n \geq 0$,
$$L_{n \to n+1} = \lambda,$$
e di conseguenza
$$L_{n \to n} = \begin{cases} -\lambda & \text{se } n = 0 \\ -(\lambda + \mu) & \text{se } n \geq 1, \end{cases}$$



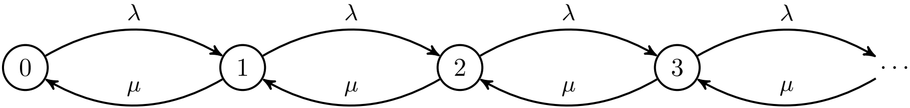
- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.



- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.
- Vi è una competizione tra il tasso di arrivo λ e di uscita μ.



- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.
- Vi è una competizione tra il tasso di arrivo λ e di uscita μ.
- Risolvendo l'equazione $\mu L = 0$ (o imponendo il bilancio di flusso) si trova
$$\mu_n \propto \left(\frac{\lambda}{\mu}\right)^n .$$



- Ogni stato è ricorrente, ma essendo infiniti stati non è ovvio che esista una distribuzione invariante.
- Vi è una competizione tra il tasso di arrivo λ e di uscita μ.
- Risolvendo l'equazione $\mu L = 0$ (o imponendo il bilancio di flusso) si trova
$$\mu_n \propto \left(\frac{\lambda}{\mu}\right)^n.$$
- Per garantire che μ sia una densità di probabilità, bisogna che
$$\sum_n \left(\frac{\lambda}{\mu}\right)^n < \infty,$$
ossia che $\lambda < \mu$.

- La distribuzione invariante è *geometrica* di parametro $1 - \lambda/\mu$, con valor medio
$$\mathbb{E}[N] = \sum_n n\mu_n = \frac{\lambda}{\mu - \lambda},$$

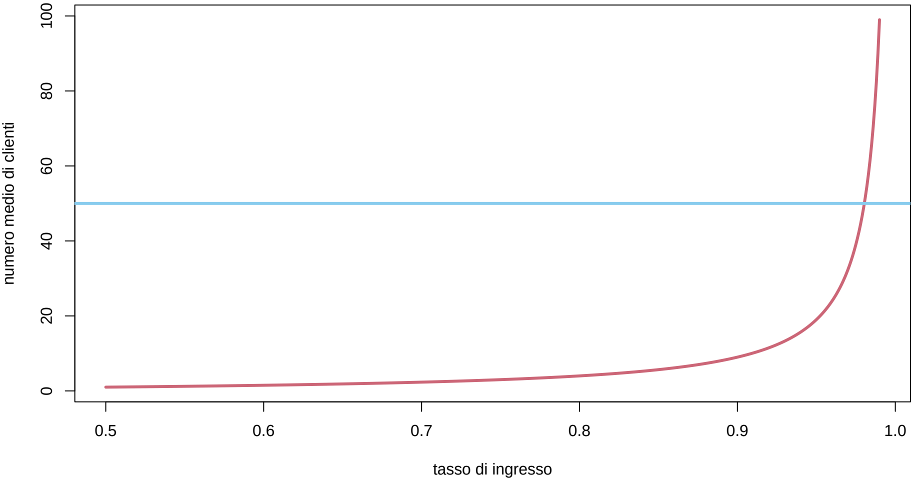


Figure 1: grafico di $\mathbb{E}[N]$ per $\mu = 1$ in funzione di λ (in rosso) e una soglia massima di possibili persone in coda (in azzurro)

Stima dei parametri dalle osservazioni

Stimiamo i parametri (λ, μ) avendo osservato un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_{\ell-1})$ con i tempi di permanenza t_1, t_2, ..., t_ℓ.

- Poniamo $T = \sum_{k=1}^{\ell} t_i$ e supponiamo inoltre che il cammino osservato non passi mai per lo stato 0.

Stima dei parametri dalle osservazioni

Stimiamo i parametri (λ, μ) avendo osservato un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_{\ell-1})$ con i tempi di permanenza t_1, t_2, ..., t_ℓ.

- Poniamo $T = \sum_{k=1}^{\ell} t_i$ e supponiamo inoltre che il cammino osservato non passi mai per lo stato 0.
- La verosimiglianza è

$$L(\lambda, \mu; X = \gamma) = \exp\left(-(\lambda + \mu) T\right) \lambda^{\gamma_+} \mu^{\gamma_-},$$

dove γ_+ indica il numero di arrivi osservati in γ (ossia transizioni da uno stato n a $n+1$), mentre γ_- il numero di uscite.

Stima dei parametri dalle osservazioni

Stimiamo i parametri (λ, μ) avendo osservato un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_{\ell-1})$ con i tempi di permanenza $t_1, t_2, \ldots, t_\ell$.

- Poniamo $T = \sum_{k=1}^{\ell} t_i$ e supponiamo inoltre che il cammino osservato non passi mai per lo stato 0.
- La verosimiglianza è

$$L(\lambda, \mu; X = \gamma) = \exp\left(-(\lambda + \mu) T\right) \lambda^{\gamma_+} \mu^{\gamma_-},$$

  dove γ_+ indica il numero di arrivi osservati in γ (ossia transizioni da uno stato n a $n+1$), mentre γ_- il numero di uscite.
- La stima di massima verosimiglianza è

$$\lambda_{MLE} = \frac{\gamma_+}{T}, \quad \mu_{MLE} = \frac{\gamma_-}{T}.$$

Se il cammino osservato trascorre un tempo T_0 nello stato 0, l'espressione per la verosimiglianza cambia: al posto di $-(\lambda+\mu)T$ si trova $-\lambda T - \mu(T - T_0)$

- λ_{MLE} non cambia, invece

$$\mu_{MLE} = \frac{\gamma_-}{T - T_0}.$$

Se il cammino osservato trascorre un tempo T_0 nello stato 0, l'espressione per la verosimiglianza cambia: al posto di $-(\lambda+\mu)T$ si trova $-\lambda T - \mu(T - T_0)$

- λ_{MLE} non cambia, invece

$$\mu_{MLE} = \frac{\gamma_-}{T - T_0}.$$

- Intepretazione: il tempo in cui la coda è vuota non si può usare per stimare il tasso di uscita.

Se il cammino osservato trascorre un tempo T_0 nello stato 0, l'espressione per la verosimiglianza cambia: al posto di $-(\lambda + \mu)T$ si trova $-\lambda T - \mu(T - T_0)$

- λ_{MLE} non cambia, invece

$$\mu_{MLE} = \frac{\gamma_-}{T - T_0}.$$

- Intepretazione: il tempo in cui la coda è vuota non si può usare per stimare il tasso di uscita.
- *Esempio*: In un intervallo di 10 minuti si osservano 5 persone arrivare alla cassa di un supermercato e 3 persone uscirne. Supponendo che la cassa non sia mai senza lavoro si stimano i parametri $\lambda = 1/2$ persone al minuto, $\mu = 3/10$ persone al minuto. Se invece la cassa è rimasta priva di persone in coda per 4 minuti, si stima $\mu = 3/6 = 1/2$ persone al minuto.

Code $M/M/\infty$

Consideriamo la situazione con un numero arbitrariamente grande, idealmente infinito, di serventi ($M/M/\infty$).

- Il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (e ogni cliente sia indipendente dagli altri).

Code $M/M/\infty$

Consideriamo la situazione con un numero arbitrariamente grande, idealmente infinito, di serventi ($M/M/\infty$).

- Il tempo di servizio per ciascun cliente sia una variabile esponenziale di parametro μ (e ogni cliente sia indipendente dagli altri).
- Per arrivare al modello , consideriamo il caso in cui non vi siano arrivi: si osservano salti da n a $n-1$ con dei tempi di permanenza dati dal minimo di n variabili aleatorie esponenziali indipendenti tra loro (la transizione avviene appena il cliente che impega meno tempo tra gli n in servizo lascia la coda).

Esercizio: il minimo di n variabili esponenziali indipendenti, tutte di parametro μ, ha densità esponenziale di parametro $n\mu$.

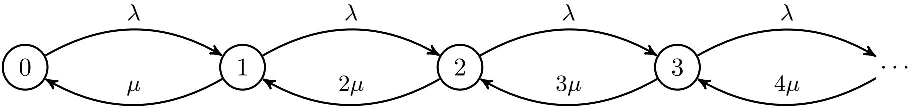
Pertanto, si avrà (se $n \geq 1$)

$$L_{n \to n-1} = n\mu.$$

Nel caso in cui vi siano arrivi con un processo di Poisson di intensità λ, poniamo, per $n \geq 0$,

$$L_{n \to n+1} = \lambda,$$

e di conseguenza



Distribuzione invariante

Come nel caso $M/M/1$, ogni stato è ricorrente.

- La competizione tra il tasso di arrivo λ e di uscita μ, è "smorzata" dal fatto che per n abbastanza grande si ha comunque $\lambda < n\mu$.

Distribuzione invariante

Come nel caso $M/M/1$, ogni stato è ricorrente.

- La competizione tra il tasso di arrivo λ e di uscita μ, è "smorzata" dal fatto che per n abbastanza grande si ha comunque $\lambda < n\mu$.
- Infatti una distribuzione invariante esiste sempre: risolvendo il sistema $\mu L = 0$ si trova una densità di Poisson di parametro λ/μ:

$$\mu_n \propto \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n .$$

Stima dei parametri dalle osservazioni

Come nel caso $M/M/1$, per stimare (λ, μ) sulla base dell'osservazione di un cammino $\gamma = (n_0 \to n_1 \to \ldots \to n_\ell)$ con i tempi di permanenza t_1, t_2, ..., t_ℓ scriviamo la verosimiglianza:

$$L(\lambda, \mu; X = \gamma) \propto \exp\left(-\lambda T - \mu T_\gamma\right) \lambda^{\gamma_+} \mu^{\gamma_-},$$

dove $T = \sum_{k=1}^{\ell} t_i$, γ_+ e γ_- sono come nel caso $M/M/1$.

Il termine nuovo è

$$T_\gamma = \sum_{k=1}^{\ell} t_i n_i,$$

(il tempo totale trascorso da tutti i clienti osservati nella coda). La stima di massima verosimiglianza è è

$$\lambda_{MLE} = \frac{\gamma_-}{T}, \quad \mu_{MLE} = \frac{\gamma_-}{T_\gamma}.$$

Il caso $M/M/c$

Il caso $M/M/c$ con $2 \leq c < \infty$ è intermedio tra i gli estremi che abbiamo considerato.

- Una distribuzione invariante esiste se e solo se $\lambda < c\mu$, ma le formule sono meno eleganti.