

ALCUNE INDIPENDENZE CONDIZIONALI NELLE SERIE  
STORICHE CATEGORIALI BIVARIATE

R. Colombi, S. Giordano

## 1. INTRODUZIONE

Per una serie storica categoriale bivariata, l'osservazione al tempo  $t$  è una coppia di modalità che indica come si manifestano due mutabili  $A_t$  e  $B_t$ , aventi rispettivamente  $a$  e  $b$  categorie. E' lecito in questo contesto domandarsi se sapere che nell'istante d'osservazione una mutabile ha assunto una certa modalità può suggerire quale categoria sarà più probabile per l'altra risposta nell'istante seguente; oppure chiedersi se la realizzazione di una mutabile dipenda o meno da come si manifesta l'altra nel medesimo istante dato il passato di entrambe; o ancora se conoscere la modalità di una mutabile nell'istante corrente renda del tutto ininfluenza il suo passato quando l'interesse è rivolto al presente dell'altra risposta. Tali interrogativi mostrano le potenzialità dell'ottica dinamica in cui la generica realizzazione di ciascuna delle due serie categoriali, è considerata come risultante dell'effetto del passato della medesima mutabile, e del passato e presente dell'altra. Per stabilire se la storia passata effettivamente influisca sul presente, e se ci sia dipendenza tra realizzazioni contemporanee bisognerà verificare delle condizioni descritte in seguito utilizzando il concetto di indipendenza condizionale.

Le condizioni d'indipendenza (introdotte nel paragrafo 2) includono tutte le possibili realizzazioni dei due processi fino all'istante attuale di osservazione. Nel paragrafo 3, sotto l'ipotesi che le due serie formino una catena di Markov bivariata, per semplicità del primo ordine, si mostra che le precedenti condizioni si semplificano in modo da coinvolgere soltanto le realizzazioni più recenti delle due mutabili. Nel quarto paragrafo, le indipendenze proposte sono ricondotte ad ipotesi lineari sui parametri di un'opportuna parametrizzazione con cui si specifica come le probabilità di transizione di una catena di Markov bivariata dipendano da covariate e dalla storia passata delle due risposte.

Grafi catena (nel paragrafo 5) illustreranno come l'assenza di connessioni tra nodi descriva le indipendenze condizionali introdotte ed un esempio nel paragrafo 6 servirà ad evidenziare l'utilità della metodologia proposta.

## 2. DEFINIZIONI ED IPOTESI DI INDIPENDENZA CONDIZIONALE

Conviene introdurre una notazione adeguata e richiamare alcuni utili teoremi. Per indicare che due insiemi di variabili  $X$  e  $Y$  sono indipendenti condizionatamente ad un terzo insieme di variabili  $Z$ , si userà la simbologia  $X \perp Y | Z$ , ad evidenziare che la conoscenza di uno dei due insiemi di variabili indipendenti non aggiunge informazioni sull'altro, noto l'insieme delle variabili condizionanti.

Richiamiamo ora una proprietà (Lauritzen, 1996) che sarà più volte usata in seguito.

*Teorema 1.* (PFIC) Proprietà fondamentale dell'indipendenza condizionale: dati gli insiemi di variabili  $X, Y, W, Z$ , le tre condizioni che seguono sono equivalenti

- i)  $X \perp (Y, Z) | W$ ;
- ii)  $X \perp Y | W$  e  $X \perp Z | (Y, W)$ ;
- iii)  $X \perp Z | W$  e  $X \perp Y | (W, Z)$ .

Inoltre, per funzioni di probabilità positive ciascuna delle tre precedenti equivale alla

- iv)  $X \perp Z | (Y, W)$  e  $X \perp Y | (W, Z)$ .

Sebbene sia una banale conseguenza del teorema 1, è utile ricordare per le dimostrazioni che seguiranno che la condizione i) implica ciascuna delle quattro indipendenze condizionali riportate in ii), iii) prese singolarmente.

Alla luce delle equivalenze appena citate possiamo ricavare un risultato più generale che servirà in seguito.

*Teorema 2.* Dati gli insiemi di variabili  $X, Y, Y', W, Z$ , le seguenti condizioni sono equivalenti per funzioni di probabilità positive

- 1)  $X \perp (Y, Y') | (W, Z)$  e  $X \perp (Y', Z) | (Y, W)$ ;
- 2)  $X \perp (Y, Y', Z) | W$ .

*Dimostrazione.* Considerando l'equivalenza tra la i) e la ii) nella PFIC, le due indipendenze condizionali riportate nella 1) possono essere riscritte in forma equivalente mediante le tre condizioni seguenti, considerate simultaneamente:

$$X \perp Y' | (Y, W, Z); \quad X \perp Y | (W, Z); \quad X \perp Z | (Y, W).$$

Le quali, sotto l'ulteriore restrizione di probabilità positive (come nella iv) della PFIC), risultano equivalenti alle due riportate di seguito:

$$X \perp Y' | (Y, W, Z); \quad X \perp (Y, Z) | W.$$

Queste ultime, infine, sempre in virtù della corrispondenza tra le condizioni *i*) e *ii*) della PFIC, valgono congiuntamente se e soltanto se è vera la condizione  $X \perp (Y, Y', Z) | W$  che ritroviamo nella 2).

Con  $A_t$  e  $B_t$  indichiamo la coppia di mutabili nell'istante attuale, il processo congiunto  $\{(A_t, B_t) \ t = \dots, -1, 0, 1, \dots\}$  verrà indicato per semplicità con la notazione  $\{A_t, B_t\}$ , ed in modo analogo i due processi marginali saranno  $\{A_t\}$  e  $\{B_t\}$ , la storia passata dei due processi, invece, costituirà gli insiemi di mutabili  $\mathbf{A}_{-\infty}^{t-1} = \{A_{t-j}, j = 1, \dots, \infty\}$  e  $\mathbf{B}_{-\infty}^{t-1} = \{B_{t-j}, j = 1, \dots, \infty\}$ . Inoltre, si dirà che una catena di Markov bivariata, di ordine uno, descrive l'evoluzione temporale di  $\{A_t, B_t\}$  se la probabilità che si manifesti una coppia di modalità in  $t$  non dipende da tutto il cammino passato delle due componenti, ma solo dalle modalità dei due caratteri osservate nell'istante immediatamente precedente. In termini di indipendenza condizionale, tale definizione di catena bivariata si specifica nel modo seguente

- CMB<sub>1</sub>:  $\{A_t, B_t\}$ , è una catena di Markov bivariata di ordine 1 se e solo se vale:  
 $A_t, B_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid A_{t-1}, B_{t-1}$  per ogni  $t$ ,  
 ovvero, qualunque sia l'istante di osservazione  $t$ , la realizzazione corrente dipende solo dal passato più recente e non da come il processo congiunto si sia manifestato fino ad allora<sup>1</sup>.

In riferimento alla sola serie  $\{A_t\}$  si definisce:

- CMU<sub>1</sub>:  $\{A_t\}$  è una catena di Markov univariata di ordine 1 se e solo se vale:  
 $A_t \perp \mathbf{A}_{-\infty}^{t-2} \mid A_{t-1}$  per ogni  $t$ ,  
 ovvero, conoscere l'ultima realizzazione rende ininfluenti tutte le precedenti se l'interesse è rivolto al valore corrente del processo.

In questo lavoro, l'obiettivo è studiare simultaneamente le due serie categoriali, e verificare se per prevedere la categoria più probabile in un certo istante per una mutabile sia utile conoscere solo il suo passato oppure il passato di entrambe e se conviene utilizzare anche l'informazione sul valore contemporaneo dell'altra risposta.

Questi diversi aspetti sono formulati più rigorosamente attraverso le definizioni seguenti.

- (G) Noncausalità secondo Granger<sup>2</sup> (Granger, 1969, Chamberlain<sup>3</sup>, 1982): il processo  $\{B_t\}$  non Granger causa il processo  $\{A_t\}$  se e solo se per ogni  $t$  si ha:

<sup>1</sup> Una definizione più generale ammette che possa differire l'ordine per le due componenti della catena CMB<sub>q</sub>:  $\{A_t, B_t\}$  è una catena di Markov bivariata di ordine  $q$  se e solo se per ogni  $t$  vale:  
 $A_t, B_t \perp \mathbf{A}_{-\infty}^{t-m-1}, \mathbf{B}_{-\infty}^{t-n-1} \mid \mathbf{A}_{t-m}^{t-1}, \mathbf{B}_{t-n}^{t-1}$ ; con  $q = \max(m, n)$ .

<sup>2</sup> Di seguito ci si riferirà alla medesima condizione (G) utilizzando anche le espressioni G-noncausalità e Granger noncausalità.

<sup>3</sup> Chamberlain riscrive la definizione di Granger noncausalità utilizzando la condizione di indipendenza condizionale invece del predittore lineare dell'originale versione di Granger. Per altri riferimenti si veda (Florens e Mouchart, 1982 e 1985, Hosoya, 1977, Pearl, 2000).

$A_t \perp \mathbf{B}_{-\infty}^{t-1} \mid \mathbf{A}_{-\infty}^{t-1}$ ; ovvero la conoscenza di  $\mathbf{B}_{-\infty}^{t-1} = \{B_{t-j}, j=1, \dots, \infty\}$ , oltre  $\mathbf{A}_{-\infty}^{t-1} = \{A_{t-j}, j=1, \dots, \infty\}$ , non migliora la previsione per la realizzazione di  $A_t$ .

– (IC) Indipendenza contemporanea dato il passato:

$$A_t \perp B_t \mid \mathbf{A}_{-\infty}^{t-1}, \mathbf{B}_{-\infty}^{t-1} \text{ per ogni } t,$$

vale a dire che se si condiziona a tutta la storia passata, il valore di una mutabile non dipende da come si manifesta l'altra nello stesso istante d'osservazione.

– (CC) Indipendenza dal passato noto il valore contemporaneo:

$$A_t \perp \mathbf{B}_{-\infty}^{t-1} \mid \mathbf{A}_{-\infty}^{t-1}, B_t \text{ per ogni } t,$$

cioè, nota la manifestazione corrente di un processo, il suo passato è ininfluenza sul contemporaneo valore dell'altro di cui si conoscono tutte le realizzazioni precedenti.

Queste tre condizioni coinvolgono un numero infinito di variabili casuali condizionanti. Dal momento che in realtà è difficile osservare l'evoluzione di mutabili su tutto l'orizzonte temporale, le ipotesi appena introdotte non sono sempre facilmente verificabili. D'altronde, non poter conoscere tutte le realizzazioni del passato non le rende ininfluenti poiché l'indipendenza condizionale ad un numero infinito di osservazioni non necessariamente risulta valida anche su un numero finito di osservazioni. Nel prossimo paragrafo si mostrerà che l'ipotesi di markovianità congiunta che introduciamo serve essenzialmente a coinvolgere un numero finito di variabili condizionanti nelle restrizioni di indipendenza che stiamo considerando.

### 3. INDIPENDENZE CONDIZIONALI PER CATENE DI MARKOV BIVARIATE

Di seguito esamineremo in dettaglio come le tre precedenti condizioni, che interpretano i nostri interrogativi sull'evoluzione congiunta di due mutabili, e che coinvolgono tutta la storia passata dei due processi fino al tempo  $t$ , vengano semplificate quando i due processi  $\{A_t, B_t\}$ , formano una catena di Markov bivariata, per semplicità del primo ordine. L'estensione ad un ordine della catena più elevato verrà trattata in un lavoro successivo.

#### 3.1. Granger noncausalità

Studiando il comportamento dinamico delle due risposte categoriali  $\{A_t, B_t\}$ , la verifica della noncausalità secondo Granger equivale a domandarsi se la storia

passata del processo  $\{B_t\}$  non fornisca informazioni utili, in aggiunta al passato di  $\{A_t\}$ , per prevedere il presente  $A_t$ .

I teoremi che seguono mostrano come la condizione (G) di G-noncausalità, sia sostituita da una più semplice condizione sotto la nostra ipotesi di markovianità bivariata del primo ordine CMB<sub>1</sub>.

*Teorema 3.* Se l'evoluzione temporale delle due mutabili è descritta da una catena di Markov bivariata del primo ordine, cioè se è vera l'ipotesi CMB<sub>1</sub>, e se vale la seguente condizione:

$$A_t \perp B_{t-1} \mid A_{t-1} \quad (g)$$

per ogni istante  $t$ , allora “ $\{B_t\}$  non causa  $\{A_t\}$ ” secondo la definizione di Granger (G).

*Dimostrazione.* Il risultato (g') discende direttamente dalla markovianità bivariata del primo ordine, mentre la (g) è un'assunzione del teorema

$$\begin{cases} A_t \perp B_{t-1} \mid A_{t-1} & (g) \\ A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid A_{t-1}, B_{t-1}. & (g') \end{cases}$$

Se consideriamo  $X = A_t$ ,  $Y = \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2}$ ,  $Z = B_{t-1}$  e  $W = A_{t-1}$ , la (g) e la (g') risultano analoghe alle condizioni in iii) della PFIC e quindi simultaneamente equivalenti alla condizione di indipendenza i) che risulta in questo caso

$$A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-1} \mid A_{t-1}. \quad (1)$$

Infine, la (1) implica direttamente la condizione (G), tesi del teorema.

Sotto l'ipotesi di markovianità bivariata del primo ordine delle due serie categoriali, il teorema 3 presentato ha il vantaggio di sostituire la condizione  $A_t \perp B_{t-1} \mid A_{t-1}$ , che coinvolge le mutabili solo negli ultimi due istanti consecutivi, alla condizione di Granger riferita all'intero cammino passato di entrambe le componenti della catena:  $A_t \perp \mathbf{B}_{-\infty}^{t-1} \mid \mathbf{A}_{-\infty}^{t-1}$ .

E' istruttivo considerare da un altro punto di vista il teorema 3 per stabilire sotto quale condizione un processo componente la catena di Markov bivariata sia markoviano marginalmente. Dalla (1) del teorema 3, infatti, in virtù della PFIC oltre la (G) discende anche la condizione CMU<sub>1</sub>:  $A_t \perp \mathbf{A}_{-\infty}^{t-2} \mid A_{t-1}$ , che assicura che il singolo processo  $\{A_t\}$  sia una catena di Markov del primo ordine. In particolare, (G) e CMU<sub>1</sub> sono congiuntamente equivalenti alla (1) e lo si dimostra facilmente riferendoci alle condizioni i) e ii) della PFIC in cui  $X = A_t$ ,  $Y = \mathbf{A}_{-\infty}^{t-2}$ ,  $Z = \mathbf{B}_{-\infty}^{t-1}$  e  $W = A_{t-1}$ .

Che la  $CMB_1$  e la (g) siano condizioni sufficienti per la noncausalità (G) è una tesi del teorema 3, ma in generale non è detto che, se le due serie sono congiuntamente markoviane e  $\{B_t\}$  non Granger causa  $\{A_t\}$ , allora valga la semplice condizione d'indipendenza (g). Questa implicazione vale solo per probabilità di transizione positive, come si evince dal teorema che segue, perchè si fa riferimento all'equivalenza dimostrata nel teorema 2 valida solo sotto l'ipotesi di probabilità positive.

*Teorema 4.* La markovianità bivariata  $CMB_1$  e la Granger noncausalità di  $\{B_t\}$  su  $\{A_t\}$  sono sufficienti per la condizione (g), nel caso di probabilità di transizione positive.

*Dimostrazione.* Dal teorema 2 deriva la seguente equivalenza, per probabilità positive:

$$A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \Big| A_{t-1}, B_{t-1} \text{ e } A_t \perp \mathbf{B}_{-\infty}^{t-2}, B_{t-1} \Big| \mathbf{A}_{-\infty}^{t-2}, A_{t-1} \Leftrightarrow A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-1} \Big| A_{t-1}.$$

La prima delle due condizioni a sinistra è la (g<sup>2</sup>) implicata dalla  $CMB_1$ , mentre la seconda è la condizione generale di Granger noncausalità (G). La condizione a destra, invece, è la (1) del teorema 3 per cui si dimostra facilmente è vera l'implicazione seguente che dà validità alla (g)

$$A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-1} \Big| A_{t-1} \Rightarrow A_t \perp B_{t-1} \Big| A_{t-1}.$$

In tal modo dalla  $CMB_1$  e dalla (G) abbiamo ricavato la condizione (g) nella classe delle catene bivariate con probabilità di transizione positive.

Unendo i due risultati dei teoremi 3 e 4 siamo in grado di affermare che se le due serie di mutabili sono descritte da una catena di Markov doppia, con probabilità di transizione non nulle, possiamo dedurre che l'intero processo  $\{B_t\}$  non Granger causa  $\{A_t\}$  se e soltanto se si verifica l'indipendenza  $A_t \perp B_{t-1} \Big| A_{t-1}$ .

Può essere interessante considerare quando due processi, di cui almeno uno markoviano, formano congiuntamente una catena di Markov bivariata. Per tale conclusione, ritornano utili la condizione di G-noncausalità ed un'ipotesi legata alle realizzazioni contemporanee delle due mutabili. A tal proposito, si consideri il seguente teorema.

*Teorema 5.* Se la serie categoriale  $\{A_t\}$  è markoviana di ordine uno, non è Granger causata dall'altro processo  $\{B_t\}$  ed è vera la seguente ipotesi:

$$B_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \Big| A_t, A_{t-1}, B_{t-1} \tag{I}$$

allora vale la condizione  $CMB_1$  ovvero i processi  $\{A_t, B_t\}$  compongono una catena di Markov bivariata del primo ordine.

*Dimostrazione.* Supponiamo che  $\{A_t\}$  sia una catena di Markov di ordine uno, quindi valga la  $CMU_1$ , e non sia causato dal processo  $\{B_t\}$  secondo la definizione (G). Come accennato a proposito del teorema 3, queste due condizioni (G,  $CMU_1$ ) sono equivalenti alla (1) che qui ritorna utile poichè implica la condizione (g').

Se a questo punto alla (g') aggiungiamo l'altra ipotesi (Γ), avvalendoci della PFIC otteniamo la condizione che caratterizza la markovianità bivariata:

$$\begin{cases} B_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid A_t, A_{t-1}, B_{t-1} & (\Gamma) \\ A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid A_{t-1}, B_{t-1} & (g') \end{cases} \Leftrightarrow A_t, B_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid A_{t-1}, B_{t-1}. \quad (CMB_1)$$

### 3.2. Indipendenza contemporanea

Quanto detto finora a proposito della condizione di causalità risponde solo a domande riguardanti il legame tra il cammino passato di una componente della catena ed il presente dell'altra, ma non viene affatto considerato il contemporaneo realizzarsi delle due mutabili.

L'idea di G-causalità nasce nell'ottica predittiva e quindi nel prevedere la realizzazione di una variabile al tempo  $t$  si ritengono già note le osservazioni fino all'istante  $t-1$  delle due risposte. La possibilità di disporre del valore dell'altra variabile in  $t$  merita una valutazione separata (si veda ad esempio Gourieroux e Monfort, 1997) e in quest'ottica s'introduce il concetto di indipendenza istantanea delle mutabili  $A_t$  e  $B_t$ , secondo cui includere tra le informazioni note il valore corrente  $B_t$  oltre tutto il passato  $\mathbf{A}_{-\infty}^{t-1}, \mathbf{B}_{-\infty}^{t-1}$  non consente una migliore previsione della realizzazione di  $A_t$ .

Tale considerazione viene qui formulata come condizione di indipendenza contemporanea condizionata alle osservazioni precedenti,

$$A_t \perp B_t \mid \mathbf{A}_{-\infty}^{t-1}, \mathbf{B}_{-\infty}^{t-1}. \quad (\text{IC})$$

Anche in questa condizione viene considerata l'intera storia passata delle due mutabili, ma la nostra ipotesi di markovianità bivariata per le due serie categoriali permette di verificare il concetto generale a partire da una condizione più semplice, come mostra il seguente teorema.

*Teorema 6.* Data una catena di Markov bivariata di ordine uno per il processo congiunto  $\{A_t, B_t\}$ , esiste indipendenza contemporanea (IC) tra le mutabili se e soltanto se, per ogni  $t$ , vale la condizione:

$$A_t \perp B_t \mid A_{t-1}, B_{t-1}. \quad (\text{ic})$$

*Dimostrazione.* Dall'equivalenza delle condizioni i) e iii) della PFIC deriva la seguente bimplicazione:

$$\left\{ \begin{array}{l} B_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid \mathcal{A}_{t-1}, B_{t-1}, \mathcal{A}_t \\ \mathcal{A}_t \perp B_t \mid \mathcal{A}_{t-1}, B_{t-1} \end{array} \right. \Leftrightarrow B_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2}, \mathcal{A}_t \mid \mathcal{A}_{t-1}, B_{t-1}.$$

La prima delle due precedenti condizioni a sinistra, è la (Γ) riferita ora al processo  $\{\mathcal{A}_t\}$ , che si ricava facilmente dalla  $\text{CMB}_1$ , la seconda è la (ic) vera per ipotesi. Inoltre si può osservare che, la condizione d'indipendenza a destra risulta equivalente alle due seguenti, sfruttando l'equivalenza tra le indipendenze i) e ii) nella PFIC:

$$\left\{ \begin{array}{l} B_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid \mathcal{A}_{t-1}, B_{t-1} \\ \mathcal{A}_t \perp B_t \mid \mathbf{A}_{-\infty}^{t-1}, \mathbf{B}_{-\infty}^{t-1}. \end{array} \right.$$

In queste ultime si riconosce l'indipendenza contemporanea (IC) che si voleva ricavare e l'altra è la condizione (g') che discende immediatamente dalla  $\text{CMB}_1$  che abbiamo ipotizzato.

L'implicazione inversa del teorema si dimostra ripercorrendo a ritroso le implicazioni della dimostrazione appena conclusa.

### 3.3. *Indipendenza dal passato noto il valore contemporaneo*

Quando l'interesse è rivolto a prevedere la realizzazione di  $\mathcal{A}_t$  si può considerare in alternativa alle due condizioni (G) e (IC), l'indipendenza condizionale (CC):  $\mathcal{A}_t \perp \mathbf{B}_{-\infty}^{t-1} \mid \mathbf{A}_{-\infty}^{t-1}, B_t$ , dove il valore contemporaneo  $B_t$  entra tra le condizionanti.

Questa condizione si semplifica sotto la consueta  $\text{CMB}_1$  e lo si può dimostrare in modo analogo ai precedenti teoremi 3 e 6.

*Teorema 7.* Se  $\{\mathcal{A}_t, B_t\}$  è una catena di Markov bivariata di ordine uno e se è valida la condizione:

$$\mathcal{A}_t \perp B_{t-1} \mid \mathcal{A}_{t-1}, B_t \tag{cc}$$

per ogni  $t$ , allora è vera la conclusione più generale (CC). L'implicazione inversa sotto l'ipotesi  $\text{CMB}_1$  vale solo per probabilità positive.

*Dimostrazione.* Alla semplice indipendenza ipotizzata  $\mathcal{A}_t \perp B_{t-1} \mid \mathcal{A}_{t-1}, B_t$  aggiungiamo la condizione  $\mathcal{A}_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid \mathcal{A}_{t-1}, B_{t-1}, B_t$ , che vale come semplice conseguenza della markovianità bivariata e non è altro che la (Γ) in cui è invertito il ruolo delle due mutabili.

Queste due condizioni per la PFIC sono equivalenti a quella che segue:  $\mathcal{A}_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-1} \mid \mathcal{A}_{t-1}, B_t$  da cui si deduce la tesi  $\mathcal{A}_t \perp \mathbf{B}_{-\infty}^{t-1} \mid \mathbf{A}_{-\infty}^{t-1}, B_t$ .



L'implicazione inversa richiede che dalle ipotesi  $\text{CMB}_1$  e (CC) si ricavi la condizione più semplice (cc). Queste due ipotesi danno validità alle condizioni seguenti

$$\begin{cases} A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-2} \mid A_{t-1}, B_{t-1}, B_t \\ A_t \perp \mathbf{B}_{-\infty}^{t-1} \mid \mathbf{A}_{-\infty}^{t-2}, A_{t-1}, B_t \end{cases}$$

ed il ricorso al teorema 2 per probabilità positive, mostra che queste ultime sono equivalenti all'indipendenza:  $A_t \perp \mathbf{A}_{-\infty}^{t-2}, \mathbf{B}_{-\infty}^{t-1} \mid A_{t-1}, B_t$ , la quale a sua volta implica la condizione richiesta (cc).

Un modo per trattare insieme gli aspetti descritti finora appare da alcune considerazioni.

Basta osservare le equivalenze tra le due condizioni (g) e (ic), ipotesi rispettivamente dei teoremi 3 e 6, e la seguente

$$A_t \perp B_{t-1}, B_t \mid A_{t-1} \tag{2}$$

da cui si deduce la condizione in esame (cc).

Inoltre, la (2) implica sicuramente le indipendenze (cc) e (ic) ed è da esse implicata in virtù della iv) della proprietà PFIC per funzioni di probabilità positive. Da ciò e dai precedenti commenti scaturisce che sotto l'ipotesi di indipendenza contemporanea dato il passato (ic), le condizioni (g) e (cc) risultano equivalenti, nell'ambito delle catene di Markov bivariate con probabilità di transizione positive.

### 3.4. Markovianità bivariata

Una considerazione finale serve a giustificare la scelta della nostra assunzione di markovianità bivariata, che nei teoremi 3, 4, 6, e 7 è stata utilizzata solo in parte. Nei teoremi 3 e 4 è stata considerata la condizione (g') per ricavare l'equivalenza tra la (g) e la (G) nel caso di probabilità di transizione positive. Nel teorema 6, sotto la condizione (Γ) è stata dimostrata l'equivalenza tra la (ic) e la (IC). Inoltre, la (Γ) riscritta con il ruolo delle due mutabili invertito, è stata utilizzata per ricavare l'equivalenza tra la condizione (cc) e la più generale (CC) nel teorema 7. Quindi la nostra assunzione di markovianità bivariata  $\text{CMB}_1$  è più forte di quel che basterebbe in ogni singolo teorema, ma ha il vantaggio di implicare tutte le condizioni necessarie nel caso in cui il ruolo di  $A_t$  e  $B_t$  sia invertito nel definire le condizioni (G) e (CC).

Si noti che Bouissou *et al.* (1986), hanno utilizzato la (1) come condizione equivalente alle due ricavate dalla definizione (G) e dall'assunzione  $\text{CMU}_1$  che il processo  $\{A_t\}$  sia descritto da una catena di Markov di ordine uno, in un contesto in cui l'interesse era però limitato alla sola Granger noncausalità di  $\{B_t\}$  su  $\{A_t\}$ . La nostra ipotesi di markovianità bivariata consente invece di trattare l'assenza di Granger causalità simmetricamente, per ciascun processo rispetto all'altro.

#### 4. COME VERIFICARE LE IPOTESI

In questa parte del lavoro proporremo un'ulteriore semplificazione nella verifica delle indipendenze condizionali (g), (ic) e (cc). Vedremo infatti come tali indipendenze siano riconducibili a vincoli di uguaglianza a zero di alcuni parametri di una delle tre parametrizzazioni, che introdurremo di seguito, di cui ci si può avvalere per modellare le probabilità di transizione di una catena di Markov bivariata.

##### 4.1. Modelli parametrici per catene di Markov bivariate

Il nostro obiettivo è quello di modellare il processo congiunto  $\{A_t, B_t\}$  mediante una catena di Markov bivariata. L'insieme delle  $a \times b$  coppie di categorie con cui le due mutabili si manifestano definisce lo spazio degli stati della catena e la probabilità di transizione congiunta  $p_{ij|z_t}^{A_t B_t} = P[A_t = i, B_t = j | z_t]$ , con  $i=1, \dots, a$  e  $j=1, \dots, b$ , è la probabilità che la catena visiti lo stato  $(i, j)$  nell'istante  $t$  nota la storia passata descritta dal vettore  $z_t$  delle variabili predeterminate. Per l'ipotesi di markovianità di ordine uno,  $z_t$  non dipende da tutto il cammino passato della catena, ma soltanto dalla coppia di categorie che identifica lo stato della catena nell'istante precedente  $t-1$  e da qualche covariata.

Più precisamente, gli elementi di  $z_t$  sono funzioni delle variabili dummy

$$I_{\{A_{t-1}=i\}} = \begin{cases} 1 & \text{se } A_{t-1} = i \\ 0 & \text{altrimenti} \end{cases} \quad I_{\{B_{t-1}=j\}} = \begin{cases} 1 & \text{se } B_{t-1} = j \\ 0 & \text{altrimenti} \end{cases}$$

( $i=1, \dots, a-1$  e  $j=1, \dots, b-1$ ) che indicano le realizzazioni delle due mutabili nell'istante  $t-1$  e di un vettore  $x_t = [x_{t1}, \dots, x_{tK}]'$  di covariate, assunte come costanti note.

La generica probabilità di transizione  $p_{ij|z_t}^{A_t B_t}$  può essere assimilata alla probabilità congiunta nella  $i$ -esima riga e  $j$ -esima colonna di una tabella di contingenza in cui le categorie della risposta  $A_t$ , sono disposte in riga e le categorie di  $B_t$  in colonna. Così, l'idea di parametrizzare le probabilità di transizione al variare delle variabili predeterminate si concretizza adattando ad un contesto dinamico la modellistica utilizzata per le tabelle di contingenza (Fahrmeir e Kaufmann, 1987). A tal fine, si definisce una trasformazione biunivoca  $\eta_t = g(\pi_t)$  tra il vettore  $\pi_t$  delle probabilità di transizione congiunte e un vettore di parametri  $\eta_t$  che siano di maggiore interesse o di interpretazione più diretta. Il vettore  $\eta_t$  è poi uguagliato ad un predittore lineare  $Z_t' \beta$  dove gli elementi della matrice  $Z_t$  sono funzioni delle variabili predeterminate e  $\beta$  è un vettore colonna di parametri incogniti. Pertanto, l'eventuale influenza delle covariate e dei valori passati delle due mutabili viene considerata in relazione a tali parametri.

Non sempre conviene modellare direttamente funzioni delle probabilità congiunte  $p_{ij|z_t}^{A_t B_t}$  in termini delle variabili predeterminate. Infatti, a seconda degli

obiettivi reali dello studio, si può ottenere un modello più facile da interpretare parametrizzando dapprima le probabilità di transizione condizionate  $p_{i|j,z_t}^{A_t|B_t} = P[A_t = i | B_t = j, z_t]$  e  $p_{j|i,z_t}^{B_t|A_t} = P[B_t = j | A_t = i, z_t]$ , oppure le probabilità di transizione marginali  $p_{i|z_t}^{A_t} = P[A_t = i | z_t]$  e  $p_{j|z_t}^{B_t} = P[B_t = j | z_t]$  e successivamente ricostruendo le probabilità di transizione congiunte.

A tal proposito, considereremo tre modi di modellare le probabilità di tabelle di contingenza che possono essere adattati per parametrizzare le probabilità di transizione di catene di Markov bivariate. In questi tre approcci gli elementi che compongono  $\eta_t$  sono tutti definiti come contrasti di logaritmi o delle singole probabilità congiunte  $\pi_t$  o di somme di tali probabilità. I tre approcci sono i seguenti:

- Approccio Log-lineare, in cui gli elementi in  $\eta_t$  sono gli effetti principali e le interazioni di ordine 1 di tipo log-lineare. Un esempio di queste interazioni è costituito dai logit (*local* o *baseline* con  $a$  e  $b$  come categorie di riferimento)  $\eta_{jt}^{B|A}$  definiti sulle probabilità di transizione condizionate  $p_{j|a,z_t}^{B_t|A_t}$ , dai logit (*local* o *baseline*)  $\eta_{it}^{A|B}$  definiti sulle probabilità di transizione condizionate  $p_{i|b,z_t}^{A_t|B_t}$ , e dai log-odds ratio (*local* o *baseline*)  $\eta_{ijt}^{AB}$ .
- Approccio Marginale (Bergsma, 1997, Colombi e Forcina, 2001) che ricava le probabilità di transizione congiunte  $p_{ij|z_t}^{A_t B_t}$  dalle marginali  $p_{i|z_t}^{A_t}$ ,  $p_{j|z_t}^{B_t}$  e da misure di associazione, che in genere sono log-odds ratio. In tale approccio il vettore  $\eta_t$  contiene logit delle probabilità di transizione marginali delle due risposte  $\eta_{it}^A$ ,  $i=1, \dots, a-1$ , e  $\eta_{jt}^B$ ,  $j=1, \dots, b-1$ , ed i log-odds ratio  $\eta_{ijt}^{AB}$  definiti sulle congiunte. In tale ambito la natura dei logit e log-odds ratio non è vincolata ad essere di tipo *local* o *baseline*. I logit possono essere scelti fra cinque differenti tipi (*baseline*, *local*, *continuation*, *reverse-continuation*, *global*) e il tipo di logit per le due marginali non necessariamente deve essere lo stesso. Anche gli odds ratio possono essere scelti tra tipi diversi, si vedano ad esempio (Barnhart e Sampson, 1994; Agresti, 1990; McCullagh e Nelder, 1989; Sen e Silvapulle, 2005).
- Approccio Condizionato-Marginale, è un approccio asimmetrico che presuppone che ci sia una direzione nella dipendenza contemporanea delle due componenti della catena. Infatti, le probabilità di transizione congiunte si determinano secondo la regola del prodotto della distribuzione condizionata per la marginale, ad esempio  $p_{ij|z_t}^{A_t B_t} = p_{j|i,z_t}^{B_t|A_t} \cdot p_{i|z_t}^{A_t}$ . In tal caso, il vettore dei parametri  $\eta_t$  è composto da logit marginali  $\eta_{it}^A$ , dai logit definiti sulle probabilità condizionate  $\eta_{jt}^{B|A}$ , ponendo a livello di riferimento la realizzazione di  $A_t$ , ed i log-odds ratio  $\eta_{ijt}^{AB}$ . Logit e log-odds ratio possono essere scelti tra gli stessi tipi utilizzati nell'approccio marginale.

#### 4.2. Vincoli sui parametri

Nella modellistica marginale, l'interesse è rivolto a descrivere la dipendenza di una mutabile dal passato di entrambe le mutabili, ma marginalmente rispetto al presente dell'altra mutabile. Ad esempio per l'*i*-esimo logit di tipo *baseline* delle probabilità di transizione marginali di  $A_t$  si può considerare il predittore lineare con effetti additivi:

$$\eta_{it}^A = \ln \frac{p_{i|z_t}^{A_t}}{p_{a|z_t}^{A_t}} = \beta_{i0}^A + \sum_{m=1}^{a-1} \beta_{im}^{AA} I_{\{A_{t-1}=m\}} + \sum_{n=1}^{b-1} \beta_{in}^{AB} I_{\{B_{t-1}=n\}} + \sum_{s=1}^K \beta_{is}^{Ax} x_{ts} \quad (3)$$

in cui i parametri  $\beta_{in}^{AB}$  che esplicitano l'influenza del passato  $B_{t-1}$  sul presente  $A_t$  hanno il significato di *log-odds ratio* (le categorie  $a$  e  $b$  sono i livelli di riferimento):

$$\beta_{in}^{AB} = \frac{P[A_t = i | A_{t-1} = a, B_{t-1} = n] P[A_t = a | A_{t-1} = a, B_{t-1} = b]}{P[A_t = a | A_{t-1} = a, B_{t-1} = n] P[A_t = i | A_{t-1} = a, B_{t-1} = b]}.$$

Così un valore nullo di tali coefficienti  $\beta_{in}^{AB}$  per ogni  $i, n$ , in presenza dei soli effetti additivi nel modello s'interpreta come indipendenza del presente  $A_t$  dall'immediato passato  $B_{t-1}$  condizionatamente soltanto alla categoria osservata per  $A_{t-1}$ :

$$\beta_{in}^{AB} = 0 \text{ per ogni } i, n \Leftrightarrow A_t \perp B_{t-1} | A_{t-1} \quad (g).$$

Quest'ultima è l'ormai nota condizione (g) che è equivalente all'assenza di Granger causalità del processo  $\{B_t\}$  su  $\{A_t\}$ , nella classe delle catene di Markov bivariate di ordine uno con probabilità di transizione positive (teorema 3).

L'interrogativo sulla Granger noncausalità si riduce quindi nei modelli marginali ad un problema di verifica d'ipotesi di nullità su alcuni parametri.

Nella modellistica log-lineare l'interesse invece riguarda la dipendenza di una mutabile dal passato di entrambe e dal presente dell'altra. Ad esempio il predittore lineare ad effetti additivi per l'*i*-esimo logit *baseline* di  $A_t$  sarà:

$$\ln \frac{p_{i|j, z_t}^{A_t|B_t}}{p_{a|j, z_t}^{A_t|B_t}} = (\beta_{i0}^A + \eta_{ij}^{AB}) + \sum_{m=1}^{a-1} \beta_{im}^{AA} I_{\{A_{t-1}=m\}} + \sum_{n=1}^{b-1} \beta_{in}^{AB} I_{\{B_{t-1}=n\}} + \sum_{s=1}^K \beta_{is}^{Ax} x_{ts} \quad (4)$$

se  $\eta_{ij}^{AB} = \eta_{ij}^{AB}$  cioè se gli *odds ratio baseline* calcolati sulle probabilità di transizione congiunte sono costanti nel tempo. Cosicchè vincoli di uguaglianza a zero dei coefficienti  $\beta_{in}^{AB}$  per ogni  $i, n$  nel precedente predittore che ora sono definiti come segue:

$$\beta_{in}^{AB} = \frac{P[A_t = i | B_t = b, A_{t-1} = a, B_{t-1} = n]P[A_t = a | B_t = b, A_{t-1} = a, B_{t-1} = b]}{P[A_t = a | B_t = b, A_{t-1} = a, B_{t-1} = n]P[A_t = i | B_t = b, A_{t-1} = a, B_{t-1} = b]}.$$

comporteranno ancora l'indipendenza tra  $A_t$  e  $B_{t-1}$ , ma, a differenza di quanto detto a proposito della (3), condizionatamente anche alla realizzazione corrente di  $B_t$  oltre a quella passata di  $A_{t-1}$ :

$$\beta_{in}^{AB} = 0 \text{ per ogni } i, n \Leftrightarrow A_t \perp B_{t-1} | A_{t-1}, B_t \quad (\text{cc}).$$

In quest'ultima ipotesi riconosciamo l'indipendenza indicata finora come (cc), e quindi non otteniamo la stessa conclusione cui si è pervenuti con gli stessi vincoli nel modello marginale, ovvero che il processo  $\{B_t\}$  non Granger causa  $\{A_t\}$ . Quando però introduciamo l'ulteriore ipotesi di indipendenza contemporanea dato il passato, nota come (ic), si annullano le interazioni  $\eta_{ij}^{AB}$  nel modello (4) e in questo particolare caso valori nulli dei coefficienti  $\beta_{in}^{AB} \forall i, n$  nella (4) equivalgono alla condizione (g) di Granger noncausalità, in accordo con quanto detto nel commento finale del paragrafo 3.3.

La considerazione importante che si trae da quanto detto è che nel modello log-lineare, soltanto unitamente alla condizione di indipendenza contemporanea, l'ipotesi della Granger noncausalità si esprime attraverso vincoli di uguaglianza a zero di alcuni parametri. In effetti, la condizione (g) della Granger noncausalità per catene di Markov bivariate riguarda una singola mutabile al tempo  $t$  ed è noto in letteratura che esprimere ipotesi riferite alle marginali partendo da un modello log-lineare comporta la costruzione di vincoli non lineari sui parametri. Nell'approccio marginale questo problema non si pone, e la condizione (g) oggetto d'interesse si traduce direttamente in vincoli lineari sui coefficienti. Inoltre soltanto nel contesto marginale le due forme della (g), che si ottengono invertendo il ruolo delle mutabili dando luogo alla noncausalità di ciascun processo rispetto all'altro, vengono simultaneamente testate come ipotesi di nullità su coefficienti.

Se l'approccio scelto è condizionato-marginale, si deve stabilire una precisa direzione di noncausalità, e trattandosi di un'ipotesi che riguarda il comportamento di una risposta rispetto al passato dell'altra, le considerazioni finora presentate a proposito della verifica della condizione (g) varranno per la componente marginale e le considerazioni fatte a proposito della (cc) per la componente condizionata.

Possiamo concludere che la possibilità di verificare la Granger noncausalità nelle catene di Markov bivariate, in modo semplice e per entrambe le mutabili, può essere annoverata dai sostenitori dell'approccio marginale a discapito delle altre modellistiche in cui poter dimostrare l'assenza del legame tra passato di una mutabile e presente dell'altra equivale a verificare vincoli non lineari sui parametri o ad aggiungere ulteriori ipotesi. Infine notiamo che l'ipotesi (ic) si traduce per tutti i tre approcci nella condizione di nullità di tutti i log-odds ratio  $\eta_{ji}^{AB}$ .

## 5. GRAFI

Le indipendenze condizionali (g), (ic) e (cc) costruite ed illustrate in dettaglio nei paragrafi iniziali, sono state in seguito interpretate in termini di vincoli lineari sui coefficienti nei modelli parametrici proposti per le probabilità di una catena di Markov doppia, ed ora saranno tradotte nell'assenza di precisi archi orientati in una struttura nota come grafo catena. Quindi la stessa indipendenza condizionale accompagnerà l'annullarsi di alcuni coefficienti di una precisa parametrizzazione alla mancanza di archi nei grafi. In un grafo catena (ad es. Lauritzen e Wermuth, 1990) l'insieme dei nodi che rappresentano le variabili è partizionato in sottoinsiemi detti elementi catena, e tali che i nodi contenuti all'interno dello stesso elemento catena sono collegati al più da archi non orientati, mentre i nodi contenuti in elementi diversi sono connessi mediante archi orientati. Nel nostro caso è la sequenza temporale che ne dà l'orientamento. Ciò che è utile ai nostri fini è che ad ogni arco mancante corrisponde una precisa indipendenza condizionale. Ogni arco mancante indica che la coppia di variabili coinvolte è indipendente condizionatamente alle variabili concorrenti, rappresentate dai nodi, nello stesso elemento catena o in elementi precedenti, che siano adiacenti ad almeno un nodo della coppia.

Sotto l'ipotesi di markovianità bivariata del primo ordine, nei grafi in figura 1, verranno omessi uno o più archi come conseguenza delle tre ipotesi d'indipendenza condizionale oggetto di questo lavoro.

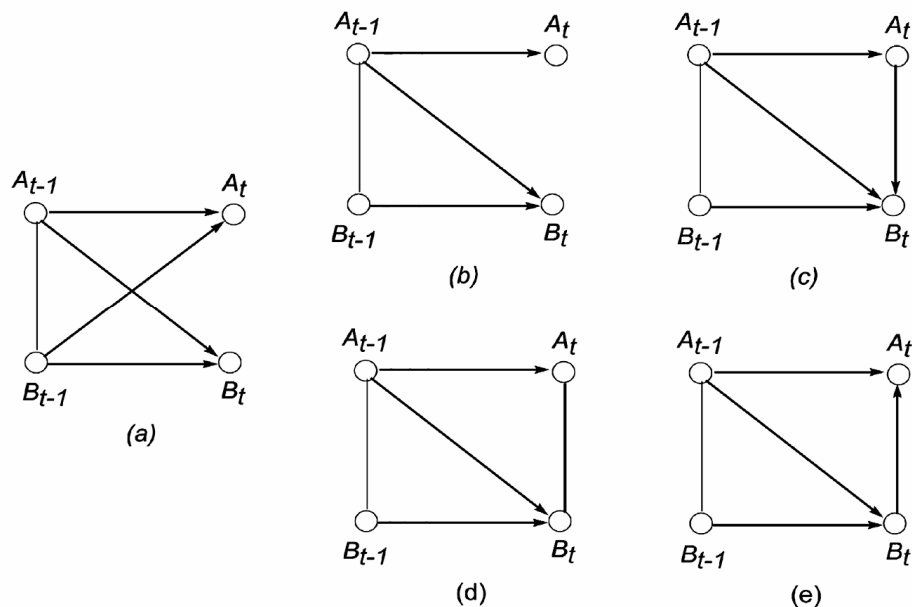


Figura 1 – Il grafo (a) illustra la condizione (ic), il grafo (b) corrisponde alle condizioni (g) e (ic), il grafo (c) evidenzia l'indipendenza (g), i grafi (d) ed (e) descrivono la condizione (cc).

Nel grafo (a), l'assenza dell'arco verticale che unisce le coppie di nodi delle realizzazioni contemporanee al tempo  $t$  rispecchia l'ipotesi (ic) di indipendenza contemporanea dato il passato. La mancanza del legame tra i punti  $A_t$  e  $B_t$  infatti, si

interpreta come indipendenza condizionale della coppia delle relative mutabili  $A_t$  e  $B_t$  e le condizionanti corrispondono ai nodi dell'elemento catena in  $t-1$  che con archi orientati vengono congiunti alle risposte in  $t$ .

I grafi (b) e (c) rappresentano due strutture in cui l'assenza dell'arco orientato  $(B_{t-1}, A_t)$  permette di dedurre la Granger noncausalità del processo  $\{B_t\}$  su  $\{A_t\}$ . Infatti, nel grafo (b), che rispecchia anche una situazione di indipendenza contemporanea (ic), la mancanza dell'arco tra i punti  $B_{t-1}$  e  $A_t$  si legge in termini di indipendenza tra le mutabili  $B_{t-1}$  e  $A_t$  dato l'unico nodo concorrente  $A_{t-1}$  che è contenuto nell'elemento catena che precede secondo l'ordine di dipendenza scelto. Analogo commento spetta alla mancanza dello stesso arco nel grafo (c) in cui però, a differenza del precedente, esiste una diretta dipendenza tra i due valori contemporanei delle mutabili, che viene rappresentata con archi orientati verso il basso da  $A_t$  a  $B_t$ . Anche in questo caso, se manca l'arco  $(B_{t-1}, A_t)$  si deduce direttamente la condizione (g).

L'assenza del medesimo arco  $(B_{t-1}, A_t)$  negli ultimi due grafi (d) ed (e) va interpretata in modo differente secondo l'indipendenza condizionale (cc):  $A_t \perp B_{t-1} \mid A_{t-1}, B_t$ , poichè  $B_t$  rientra tra i nodi concorrenti.

## 6. ESEMPIO

Nel proporre un esempio utilizziamo l'approccio marginale sia perchè consente di modellare la dinamica delle due componenti  $\{A_t\}$ ,  $\{B_t\}$  separatamente dalla loro associazione contemporanea, sia perchè permette di formulare l'ipotesi di Granger noncausalità come ipotesi di nullità di alcuni parametri.

L'esempio è relativo ai dati rilevati da una compagnia di telefonia mobile e servizi internet che deve monitorare l'utilizzo di una macchina (server) che eroga connessioni per servizi wap e web, in una zona del nord Italia. La compagnia controlla l'uso del server rilevando giornalmente i flussi di richieste soddisfatte di connessioni per entrambi i servizi e riportandoli alla capacità di massima attività della macchina stessa. Da ciò si ricavano due serie storiche dei valori (in percentuale) che indicano quanta parte della potenzialità complessiva della macchina risulta effettivamente sfruttata da ciascun servizio. Tale impegno quotidianamente richiesto al macchinario per soddisfare sia gli utenti wap che quelli web viene categorizzato in quattro modalità a seconda che si registri un livello basso (0), medio (1), medio-alto (2) ed elevato (3) ovvero che il livello sia al di sotto, compreso o superiore a due soglie stabilite dall'azienda tenendo conto delle potenzialità della macchina e del piano di erogazione adottato. Le serie storiche categorizzate relative alle connessioni wap  $\{A_t\}$  e  $\{B_t\}$  web analizzate di seguito riguardano il semestre novembre 2004-aprile 2005.

Descriviamo la dinamica dei livelli di capacità di erogazione utilizzate per i servizi wap e web della macchina mediante una catena di Markov bivariata del primo ordine. Quindi modelliamo le probabilità di transizione congiunte  $p_{ij|z_t}^{A_t B_t}$  che al

tempo  $t$  si registri il livello  $i$ -esimo di utilizzo della macchina per connessioni wap ( $i=0,1,2,3$ ), e il livello  $j$ -esimo per connessioni web ( $j=0,1,2,3$ ), noto l'impiego della macchina per i due servizi al tempo  $t-1$ . Così si può fornire uno strumento predittivo all'azienda per conoscere quale livello di utilizzo della macchina per il wap ed il web sia più probabile giorno per giorno, noti i volumi di attività della macchina per ambo i servizi nel giorno precedente. Altro obiettivo è verificare l'ipotesi di Granger noncausalità che in questo contesto equivale, ad esempio, a chiedere se il volume di richieste di connessioni web sia ininfluenza nel determinare il più probabile livello di impegno della macchina per connessioni wap nel giorno seguente. A tal fine, proprio perchè si utilizza l'approccio marginale, bisogna semplicemente sottoporre a verifica l'ipotesi che siano nulli quei coefficienti che esprimono l'effetto delle realizzazioni passate della mutabile web sui logit delle probabilità di transizione marginali della mutabile wap.

Le probabilità di transizione sono state parametrizzate utilizzando logit *continuation* per le probabilità di transizione marginali  $p_{i|z_t}^{A_t}$  e  $p_{j|z_t}^{B_t}$  e *odds ratio* di tipo *continuation*.

Per i logit sono stati utilizzati predittori lineari con effetti additivi del tipo semplificato nella (3), aggiungendo in questo caso l'ipotesi di parallelismo dei logit. Così, ad esempio per i logit di riga, con  $i=0, 1, 2$ , si ha:

$$\eta_{it}^A = \ln \frac{p_{i|z_t}^{A_t}}{p_{a|z_t}^{A_t}} = \beta_{i0}^A + \sum_{m=0}^2 \beta_m^{AA} I_{\{A_{t-1}=m\}} + \sum_{n=0}^2 \beta_n^{AB} I_{\{B_{t-1}=n\}}.$$

Relativamente ai logaritmi degli *odds ratio* sono state assunte le seguenti ipotesi: indipendenza contemporanea dato il passato ( $\eta_{ijt}^{AB} = 0 \quad \forall i,j \text{ e } \forall t$  che si identifica con la nostra condizione (ic)), associazione indipendente dal tempo ( $\eta_{ijt}^{AB} = \eta_{ij}^{AB} \quad \forall i,j \text{ e } \forall t$ ), ed associazione indipendente dal tempo e uniforme ( $\eta_{ijt}^{AB} = \eta \quad \forall i,j \text{ e } \forall t$ ).

Nella tavola 1 vengono presentati i diversi modelli riportando la log-verosimiglianza, il numero di parametri e l'ipotesi relativa all'associazione contemporanea.

In ogni cella della prima colonna figurano quattro valori, i primi due relativi alla mutabile servizi wap, ed i secondi riferiti alla mutabile servizi web. Essi indicano quali effetti compaiono nei predittori lineari che modellano i logit delle probabilità di transizione marginali. Così nella quaterna (1 1, 1 1), la prima coppia di 1 corrisponde alla presenza degli effetti di  $A_{t-1}$  e  $B_{t-1}$  ed indica che i logit relativi alle probabilità di transizione marginali della mutabile  $A_t$  dipendono dalle realizzazioni di  $A_{t-1}$  e  $B_{t-1}$ , cioè dai livelli d'impiego della macchina per le connessioni web e wap nel giorno precedente.

Analogamente la seconda coppia di valori unitari, indica che anche i logit delle probabilità di transizione marginali di  $B_t$  sono funzioni lineari additive degli effetti di  $A_{t-1}$  e  $B_{t-1}$ .



Mentre (1 0, 1 1) esprime la Granger noncausalità del processo web sul processo wap, che si ha quando i logit delle probabilità di transizione marginali di  $A_t$ , non dipendono dalla realizzazione di  $B_{t-1}$ . Infatti lo zero nella prima coppia va interpretato come assenza degli effetti di  $B_{t-1}$  tra le predeterminate che specificano i logit delle probabilità di transizione marginali di  $A_t$ . Questo equivale a ritenere che la conoscenza della capacità della macchina sfruttata per accordare connessioni web non incide nel determinare quale sarà il livello più probabile di utilizzo dello stesso macchinario per il servizio wap del giorno seguente.

TAVOLA 1

*Modelli per la G-noncausalità con differenti ipotesi di associazione*

Modello	N. parametri	Log-veros.	Associazione
(1 1, 1 1)	18	-339.7381	indip. cont. ( $\eta_{ijt}^{AB} = 0$ )
(1 1, 0 1)	15	-341.3716	indip. cont.
(1 0, 1 1)	15	-344.6896	indip. cont.
(1 0, 0 1)	12	-346.3230	indip. cont.
(1 1, 1 1)	19	-324.1459	unif. e cost. ( $\eta_{ijt}^{AB} = \eta$ )
(1 1, 0 1)	16	-325.8770	unif. e cost.
(1 0, 1 1)	16	-329.2831	unif. e cost.
(1 0, 0 1)	13	-331.7912	unif. e cost.
(1 1, 1 1)	27	-313.2495	cost. non vincolata ( $\eta_{ijt}^{AB} = \eta_{ij}^{AB}$ )
(1 1, 0 1)	24	-315.5416	cost. non vincolata
(1 0, 1 1)	24	-316.9356	cost. non vincolata
(1 0, 0 1)	21	-319.3879	cost. non vincolata

Analogamente (1 1, 0 1) rappresenta il modello con assenza di Granger causalità della componente wap della catena sulla componente web. Si esclude cioè l'influenza di  $A_{t-1}$  sui logit delle probabilità di transizione marginali della mutabile  $B_t$ .

Testare l'ipotesi di assenza di Granger noncausalità di  $\{B_t\}$  su  $\{A_t\}$  si riduce a confrontare il modello più parsimonioso (1 0, 1 1) con il modello di riferimento (1 1, 1 1) il che equivale a testare che siano nulli tutti i coefficienti indicanti l'influenza di  $B_{t-1}$  sui logit della componente  $A_t$ .

Il modello marginale consente di verificare anche la Granger noncausalità, di ogni mutabile rispetto all'altra e questo accade testando il modello (1 0, 0 1) in cui non compare il passato di ciascuna risposta nella combinazione lineare che modella i logit dell'altra contro il modello di riferimento (1 1, 1 1).

Infatti, nel modello (1 0, 0 1) sono nulli gli effetti delle realizzazioni di  $B_{t-1}$  nei logit delle probabilità di transizione marginali di  $A_t$  e simmetricamente sono uguali a zero anche gli effetti della mutabile  $A_{t-1}$  sui logit delle marginali di  $B_t$ .

Nella Tavola 2 si riporta l'esito dei test del rapporto di verosimiglianza e per evitare ridondanze inseriamo solo i risultati relativi ai modelli con associazione

costante nel tempo  $\eta_{ijt}^{AB} = \eta_{ij}^{AB}$  che come mostrato nella tavola 3 è da preferire alle altre ipotesi di associazione.

Viene accettata l'ipotesi di assenza di Granger causalità di ciascun processo rispetto all'altro, ed anche la noncausalità di entrambi, web su wap e wap su web, contro il modello completo di riferimento.

TAVOLA 2

*Test sulla Granger noncausalità con associazione costante*

Modello H <sub>1</sub>	Modello H <sub>0</sub>	G <sup>2</sup>	gdl
(1 1, 1 1)	(1 0, 1 1)	4.5842	3
(1 1, 1 1)	(1 1, 0 1)	7.3722	3
(1 1, 1 1)	(1 0, 0 1)	12.2768	6

Infine riportiamo nella tavola 3 gli esiti del test del rapporto delle verosimiglianze da cui si evince che l'ipotesi di associazione costante appare la più accreditata.

TAVOLA 3

*Test sulle ipotesi di associazione*

Modello	H <sub>0</sub>	H <sub>1</sub>	G <sup>2</sup>	gdl
(1 1, 1 1)	unif. e cost.	cost. non vincolata	21.7928	8
(1 1, 0 1)	unif. e cost.	cost. non vincolata	24.8066	8
(1 0, 1 1)	unif. e cost.	cost. non vincolata	20.6708	8
(1 0, 0 1)	unif. e cost.	cost. non vincolata	24.8066	8
(1 1, 1 1)	indip. cont.	cost. non vincolata	52.9772	9
(1 1, 0 1)	indip. cont.	cost. non vincolata	51.66	9
(1 0, 1 1)	indip. cont.	cost. non vincolata	55.508	9
(1 0, 0 1)	indip. cont.	cost. non vincolata	53.8702	9

## 7. CONCLUSIONI

Nel descrivere l'evoluzione temporale di due mutabili mediante catene di Markov ci siamo posti il problema di verificare alcune ipotesi di indipendenza condizionale che rispondono ad interrogativi sulla dipendenza tra passato di una mutabile e presente dell'altra (Granger causalità), e sul legame tra le realizzazioni contemporanee delle due mutabili.

La letteratura fornisce tre condizioni di indipendenza condizionale che coinvolgono tutta la storia passata dei due processi.

La nostra ipotesi di markovianità ci ha invece consentito di proporre delle condizioni di indipendenza condizionale che, sebbene riferite ad un numero finito di realizzazioni delle mutabili, sono equivalenti alle condizioni più generali.

Nella seconda parte del lavoro, dopo avere introdotto le parametrizzazioni log-lineare, marginale e condizionato-marginale per le probabilità di transizione congiunte, le precedenti condizioni di indipendenza sono state espresse attraverso vincoli sui parametri dei tre modelli. Si è evidenziato che solo nella parametrizza-

zione marginale l'ipotesi che ogni mutabile non Granger causi l'altra equivale a vincolare a zero alcuni parametri. Nel modello log-lineare la stessa conclusione in termini di G-noncausalità si ha solo se non c'è alcuna dipendenza tra le mutabili nello stesso istante, noto il passato di entrambe. Nell'approccio condizionato-marginale invece la scelta della mutabile marginale e di quella da condizionare impone una direzione all'eventuale causalità per cui è possibile esprimere attraverso vincoli lineari sui parametri solo l'ipotesi di Granger causalità di una mutabile sull'altra. Invece quando l'obiettivo è modellare l'influenza di una componente della catena sull'altra nel medesimo istante, sono preferibili le parametrizzazioni condizionato-marginale e log-lineare perchè per questi modelli le ipotesi di indipendenza, rilevanti in questo contesto, si traducono in vincoli lineari sui parametri. Rispetto all'approccio log-lineare e marginale nei modelli condizionato-marginali, è possibile descrivere attraverso vincoli lineari sui parametri sia la condizione di Granger noncausalità per la componente marginale, sia la dipendenza della componente condizionata dal valore contemporaneo della componente marginale.

Tutte e tre le parametrizzazioni introdotte sono infine adeguate a descrivere con vincoli di uguaglianza a zero dei parametri d'associazione, l'ipotesi di indipendenza contemporanea dato il passato.

Infine la teoria dei grafi ci ha permesso di descrivere in modo alternativo le indipendenze condizionali facendole corrispondere all'assenza di precisi archi nei grafi catena. Nell'ultimo paragrafo proponiamo un esempio che illustra la flessibilità dell'approccio marginale nel formulare ipotesi di Granger noncausalità sotto varie ipotesi di dipendenza contemporanea.

*Dipartimento di Ingegneria dell'informazione e metodi matematici  
Università di Bergamo*

ROBERTO COLOMBI

*Dipartimento di Economia e statistica  
Università della Calabria*

SABRINA GIORDANO

#### RIFERIMENTI BIBLIOGRAFICI

- A. AGRESTI, (1990), *Categorical data analysis*, Wiley & Son, New York.
- H. BARNHART, A.R. SAMPSON, (1994), *Overview of multinomial models for ordinal data*, "Communications in Statistics, Theory and Methods", 23, 12, pp. 3395-3416.
- W.P. BERGSMA, (1997), *Marginal models for categorical data*, Tilburg University Press, Tilburg.
- M.B. BOUISSOU, J.J. LAFFONT, Q.H. VUONG, (1986), *Tests of non causality under Markov assumptions for qualitative panel data*, "Econometrica", 54, pp. 395-414.
- G. CHAMBERLAIN, (1982), *The general equivalence of Granger and Sims causality*, "Econometrica", 50, pp. 569-581.
- R. COLOMBI, A. FORCINA, (2001), *Marginal regression models for the analysis of positive association of ordinal response variables*, "Biometrika", pp. 1007-1019.
- L. FAHRMEIR, H. KAUFMANN, (1987), *Regression models for nonstationary categorical time series*, "Journal of Time Series Analysis", 8, 2, pp. 147-160.
- J.P. FLORENS, M. MOUCHART, (1982), *A note on non causality*, "Econometrica", 50, 3, pp. 583-592.

- J.P. FLORENS, M. MOUCHART, (1985), *A linear theory for non causality*, "Econometrica", 53, 1, pp. 157-176.
- C. GOURIEROUX, A. MONFORT, (1997), *Time series and dynamic models*, Cambridge University Press.
- C.W.J. GRANGER, (1969), *Investigating causal relations by econometric models and cross-spectral methods*, "Econometrica", 37, 3, pp. 424-438.
- Y. HOSOYA, (1977), *On the Granger condition for non causality*, "Econometrica", 45, pp. 1735-1736.
- S.L. LAURITZEN, (1996), *Graphical models*, Clarendon Press, Oxford.
- S.L. LAURITZEN, N. WERMUTH, (1990), *On substantive research hypotheses, conditional independence graphs and graphical chain models*, "Journal of the Royal Statistical Society", B, 58, 1, pp. 21-50.
- P. McCULLAGH, J.A. NELDER, (1989), *Generalized linear models*, Chapman & Hall, Londra.
- J. PEARL, (2000), *Causality: models, reasoning and inference*, Cambridge University Press.
- P.K. SEN, M.J. SILVAPULLE, (2005), *Constrained statistical inference*, John Wiley & Son, New Jersey.

## RIASSUNTO

### *Alcune indipendenze condizionali nelle serie storiche categoriali bivariate*

In questo lavoro descriviamo congiuntamente due serie storiche di mutabili mediante una catena di Markov bivariata. L'assunzione di markovianità permette di semplificare le condizioni di indipendenza condizionale con cui si formalizza l'interrogativo sul legame tra presente di una mutabile e passato dell'altra (Granger causalità) e tra le realizzazioni contemporanee delle due mutabili. Successivamente, le tre condizioni proposte, che implicano quelle più generali definite su tutto l'orizzonte temporale, sono presentate in termini di vincoli lineari sui coefficienti nei modelli che parametrizzano le probabilità di transizione congiunte della catena. Infine, si mostra come le stesse condizioni corrispondano all'assenza di alcuni archi orientati nei grafi catena.

## SUMMARY

### *Some conditional independencies in bivariate categorical time series*

In this work we consider two time series of categorical data as a bivariate Markov chain. The markovianity assumption allows us to simplify some conditional independencies introduced in order to describe if the knowledge of past or present realizations of one of the two categorical variables can provide some additional information to forecast the current realization of the other. The three simple conditions introduced here, though referring only to the recent realizations of the two variables, imply the more general conditions defined by all of the past realizations. Moreover, we show that the proposed conditions are equivalent to the hypothesis of null coefficients in some parametric models for joint transition probabilities. Finally, we represent these conditions in terms of missing edges in chain graphs.