Un confronto tra diverse definizioni della scambiabilità parziale

L. Di Scala L. La Rocca

Tesina per il corso del Prof. Muliere, anno 2001

Ci proponiamo di confrontare la parziale scambiablità secondo de Finetti [1] con quella secondo Diaconis e Freedman [2], secondo la linea tracciata da Fortini, Ladelli, Petris e Regazzini in [3]. Sotto opportune ipotesi, vale un teorema di rappresentazione in termini di miscela di catene di Markov, che per esempio si applica ai processi d'urna rinforzati. Presenteremo brevemente tali processi, il cui studio si deve a Muliere, Secchi e Walker [4].

Scambiabilità parziale secondo de Finetti

La nozione di scambiabilità parziale è stata introdotta per la prima volta nel 1938 da Bruno de Finetti [1] con l'obiettivo di studiare il ragionamento induttivo nel caso di osservazioni eterogenee. Rimanendo per semplicità nell'ambito della teoria assiomatica di Kolmogorov, vale a dire nell'ambito delle probabilità numerabilmente additive, modelliamo le osservazioni con una famiglia $\xi = (\xi_{i,n})_{i\in I,n\in N}$ di variabili aleatorie definite su un comune spazio di probabilità $(\Omega, \mathcal{A}, \mathbf{P})$. Per lo studio della scambiabilità parziale nell'ambito più generale delle probabilità solo finitamente additive, cioè formulando (come de Finetti) la sola ipotesi di coerenza, si rimanda a Regazzini [5].

Def. 1 (Scambiabilità parziale alla d-F) *Diremo parzialmente scambiabile alla de Finetti la famiglia ξ (o parzialmente scambiabili le $\xi_{i,n}$ che la compongono) quando, comunque si fissi, per ogni $i \in I$, una permutazione finita π^i di N, risultino uguali in legge le famiglie $(\xi_{i,n})_{i\in I,n\in N}$ e $(\xi_{i,\pi^i_n})_{i\in I,n\in N}$.*

Ciò corrisponde a richiedere che le osservazioni siano divise in classi, una per ogni $i \in I$, all'interno delle quali esse possano ritenersi omogenee, ovvero pensarsi provenienti dalla ripetizione di un medesimo esperimento in fissate condizioni ambientali. Per esempio le $\xi_{i,n}$ potrebbero modellare la misura della lunghezza di una sbarra di alluminio e a ogni $i \in I$ potrebbe essere associata la temperatura alla quale si effettuano le misure $(\xi_{i,n})_{n\in N}$. Si richiede dunque, in particolare, che per $i \in I$ fissato le $(\xi_{i,n})_{n\in N}$ siano scambiabili e si richiede in più che osservazioni corrispondenti a temperature diverse possano essere permutate diversamente senza alterare la legge globale. Affinché la famiglia ξ sia parzialmente scambiabile (occorre e) basta che lo sia ogni sua sottofamiglia finita, come si verifica con un argomento monotono.

Per una famiglia parzialmente scambiabile ξ vale un teorema di rappresentazione analogo a quello per famiglie (completamente) scambiabili. Cioè esiste un elemento aleatorio $\tilde{\theta}$ condizionatamente al quale le variabili $\xi_{i,n}$ sono indipendenti e, per $i \in I$ fissato, identicamente distribuite. In altre parole, la legge di ξ è mistura di leggi prodotto opportune. Dimostriamo questo risultato, per semplicità nell'ipotesi che le $\xi_{i,n}$ prendano valori in uno spazio $(E, \mathcal{E})$ numerabile discreto.

Teo. 1 *Se ξ è una famiglia di variabili aleatorie parzialmente scambiabili a valori in uno spazio $(E, \mathcal{E})$ numerabile discreto, allora esiste un elemento aleatorio $\tilde{\theta}$ a valori in $[0,1]^{I \times N}$ condizionatamente al quale le $\xi_{i,n}$ sono indipendenti e isonome per $i \in I$ fissato.*

DIM.

Innanzi tutto la legge forte dei grandi numeri per successioni scambiabili permette di scrivere

$$\lim_m \frac{1}{m} \sum_{n=1}^{m} \mathbf{1}_{\{\xi_{i,n}=x\}} \stackrel{q.c.}{=} \liminf_m \frac{1}{m} \sum_{n=1}^{m} \mathbf{1}_{\{\xi_{i,n}=x\}} =: \tilde{\theta}_{i,x}$$

per ogni $i \in I, x \in E$. Definiamo $\tilde{\theta} := (\tilde{\theta}_{i,x})_{i \in I, x \in E}$ e studiamo la legge di $\xi | \tilde{\theta}$. Sia $\mathcal{T}$ la σ-algebra terminale di ξ, definita da $\mathcal{T} := \bigcap_{m \in N} \bigvee_{n \geq m} \sigma(\xi_{i,n} : i \in I)$. Per misurabilità asintotica risulta $\tilde{\theta} \in \mathcal{T}$. Se poi $A \in \mathcal{T}$, allora per ogni $m \in N$ si ha $\mathbf{1}_A = f_m \circ (\xi_{i,n})_{i \in I, n \geq m}$ per una certa f_m boreliana limitata; in virtù della parziale scambiabilità e sfruttando la linearità del valore atteso se ne deduce

$$\mathbf{E} \prod_{i \in F} \prod_{n=1}^{m_i} \mathbf{1}_{\{\xi_{i,n}=x_{i,n}\}} \mathbf{1}_{\{\xi_{j,m}=x\}} \mathbf{1}_A = \mathbf{E} \prod_{i \in F} \prod_{n=1}^{m_i} \mathbf{1}_{\{\xi_{i,n}=x_{i,n}\}} \frac{1}{l} \sum_{k=1}^{l} \mathbf{1}_{\{\xi_{j,m+k}=x\}} \mathbf{1}_A$$

per ogni F parte finita di I, $m_i \in N$, $j \in I$, $m > m_j$ ($m_i = 0$ se $i \notin F$), $l \in N$ e $x, x_{i,n} \in E$. Per convergenza dominata il secondo membro passa al limite per $l \to \infty$ e procedendo induttivamente si ottiene

$$\mathbf{E} \left[\prod_{i \in F} \prod_{n=1}^{m_i} \mathbf{1}_{\{\xi_{i,n}=x_{i,n}\}} \mathbf{1}_A \right] = \mathbf{E} \left[\prod_{i \in F} \prod_{n=1}^{m_i} \tilde{\theta}_{i,x_{i,n}} \mathbf{1}_A \right]$$

per ogni F parte finita di I, $m_i \in N$, $A \in \mathcal{T}$ e $x_{i,n} \in E$. Tanto basta a concludere che $\tilde{\theta}_{i,x} \in \mathbf{P}(\xi_{i,n} = x | \mathcal{T})$ per ogni $n \in N$ e le $\xi_{i,n}$ sono indipendenti condizionatamente a $\mathcal{T}$. Se si prende $A = \mathbf{1}_{\{\tilde{\theta} \in B\}}$ e si osserva che $\tilde{\theta}_{i,x}$ si ottiene da $\tilde{\theta}$ tramite proiezione canonica, resta dimostrato che $\tilde{\theta}_{i,x} \in \mathbf{P}(\xi_{i,n} = x | \tilde{\theta})$ per ogni $n \in N$ e le $\xi_{i,n}$ sono indipendenti condizionatamente a $\tilde{\theta}$.

*

Con le notazioni della dimostrazione, prendendo $A = \Omega$ e sfruttando la formula di integrazione rispetto alla misura immagine si ottiene la legge di ξ come miscela di opportune leggi prodotto, ovvero

$$\mathbf{E} \left[\prod_{i \in F} \prod_{n=1}^{m_i} \mathbf{1}_{\{\xi_{i,n}=x_{i,n}\}} \right] = \int_{[0,1]^{I \times N}} \prod_{i \in F} \prod_{n=1}^{m_i} \theta_{i,x_{i,n}} \mathcal{L}_{\tilde{\theta}}(d\theta)$$

per ogni F parte finita di I, $m_i \in N$ e $x_{i,n} \in E$. Osserviamo che per $i \in I$ fissato arbitrario si ha $\sum_{x\in E} \tilde{\theta}_{i,x} \overset{q.c.}{=} 1$ perché esiste una versione regolare di $\mathbf{P}(\xi_{i,n} = \cdot|\tilde{\theta})$. In particolare, se I coincide con E, allora $\tilde{\theta}$ è una matrice aleatoria di transizione su I (doppiamente infinita in generale). Questo caso avrà particolare importanza nel seguito. Se invece $|I| = 1$ si ritrova il classico Teorema di Rappresentazione di de Finetti.

Scambiabilità parziale alla Diaconis e Freedman

Una seconda interpretazione della scambiabilità parziale è stata data nel 1980 da Diaconis e Freedman [2]. Essa è caratteristica delle misture di leggi di catene di Markov, come ci accingiamo a verificare.

Def. 2 (Scambiabilità parziale alla D-F) *Un processo $Z = (Z_n)_{n\geq 0}$, a valori in uno spazio degli stati numerabile discreto $(I, \mathcal{I})$, si definisce parzialmente scambiabile alla Diaconis e Freedman se per ogni coppia di stringhe finite di stati in I, σ e τ, tali che $\sigma_1 = \tau_1$ e $a^{\sigma}_{i,j} = a^{\tau}_{i,j}$ (dove $a^{\sigma}_{i,j}$ è il numero di transizioni da i a j nella stringa σ), si ha $\boldsymbol{P}(A_\sigma) = \boldsymbol{P}(A_\tau)$ (dove $A_\sigma = \{Z_0 = \sigma_0, \dots, Z_{card(\sigma)} = \sigma_{card(\sigma)}\}$). Per brevità, si scriverà $\sigma \sim \tau$.*

La scambiabilità parziale alla Diaconis e Freedman necessita della ricorrenza di Z per dare luogo a un teorema di rappresentazione, perciò richiamiamo la definizione di ricorrenza.

Def. 3 *Un processo $(Z_n)_{n\geq 0}$ si dice ricorrente se vale che*

$$\boldsymbol{P}\{Z_n = Z_0 \text{ frequentemente}\} = 1.$$

Il risultato fondamentale è che per processi stocastici ricorrenti parzialmente scambiabili alla Diaconis e Freedman vale un teorema di rappresentazione in termini di mistura di catene di Markov.

Teo. 2 *Sia $(Z_n)_{n\geq 0}$ un processo stocastico definito su $(\Omega, \mathcal{A}, \boldsymbol{P})$ a valori in uno spazio degli stati numerabile discreto $(I, \mathcal{I})$. Sia $(Z_n)_{n\geq 0}$ ricorrente secondo la definizione (3). Allora $(Z_n)_{n\geq 1}$ è parzialmente scambiabile alla D-F se e solo se è mistura di catene di Markov, ossia se*

$$\boldsymbol{P}\{Z_0 = z_0, \dots, Z_n = z_n\} = \int_{\mathcal{P}} \prod_{k=0}^{n-1} p(z_k, z_{k+1}) \mu(z_0, dp) \tag{1}$$

per ogni $z_0, \dots, z_n \in I$ e $n \in N$, dove $\mathcal{P} = \{\text{matrici stocastiche su I} \times \text{I}\}$ è dotato della σ-algebra prodotto $\mathcal{B}[0,1]^{I\times I}$ e μ è un nucleo da I a $\mathcal{P}$.

Dim.

Se vale la scrittura (1), allora è immediato dedurre la scambiabilità parziale di Z. Supponiamo, quindi, che Z sia una catena parzialmente scambiabile alla D-F. Non lede le generalità supporre, secondo la costruzione canonica, che $(\Omega, \mathcal{A}, \mathbf{P}) = (I^\infty, P(I)^\infty, \mathbf{P} = \text{legge di } [\mathrm{Z_n}]_{\mathrm{n}\geq 0})$ e che $Z_n(\omega) = \omega_n$. Sia inoltre $\mathbf{P}\{Z_0 = 1\} = 1$. Si definisca la classe degli "uno-blocchi"

$$B = \{\sigma \in I^n : n \in N, \sigma_0 = 1 \text{ e } \sigma_\mathrm{i} \neq 1 \ \forall \mathrm{i} > 0\}.$$

Si noti che B è numerabile e quindi può pensarsi munito della σ-algebra di tutte le sue parti. Sia $(Y_n)_{n\geq 0}$ il processo degli "uno-blocchi" associato a Z, che è ben defino per la ricorrenza. La tesi segue dal dimostrare i seguenti tre passi successivi:

1. La successione $(Y_n)_{n\geq 0}$ è scambiabile.
2. La catena $(Z_n)_{n\geq 0}$ è ancora parzialmente scambiabile sotto la probabilità condizionale regolare $\mu_\omega(A) = \mathbf{P}\{Y_0 \in A|\mathcal{T}\}(\omega)$, dove $\mathcal{T}$ è la σ-algebra terminale di $[Y_n]_{n\geq 0}$.
3. Se $(Z_n)_{n\geq 0}$ è ricorrente e parzialmente scambiabile e se le Y_n sono i. i. d. , allora Z è una catena di Markov.

Innanzitutto, la scambiabilità di Y segue immediatamente dalla parziale scambiabilità di Z. Infatti, se $\tau = \pi(\sigma)$, poiché $\sigma_0 = \tau_0 = 1$, si ha che $\sigma \sim \tau$. Quindi

$$\begin{aligned}&\mathbf{P}\{Y_0 = \sigma_0, Y_1 = \sigma_1, \ldots, Y_{|\sigma|} = \sigma_{|\sigma|}\} = \mathbf{P}(A_\sigma) = \mathbf{P}(A_\tau) = \\ &= \mathbf{P}\{Y_0 = \pi(\sigma_0), Y_1 = \pi(\sigma_1), \ldots, Y_{|\sigma|} = \pi(\sigma_{|\sigma|})\}.\end{aligned}$$

La precedente espressione sfrutta il fatto che se $\sigma \sim \tau$, allora $|\sigma| = |\tau|$.

Quindi, poiché I^∞ è polacco, per il Teorema di Rappresentazione di de Finetti, le $(Y_n)_{n\geq 0}$ sono indipendenti e identicamente distribuite condizionalmente alla σ-algebra $\mathcal{T} = \cap_{n\geq 1}\sigma(Y_n, Y_{n+1}, \ldots)$. Equivalentemente, esiste una probabilità condizionale regolare $\mu_\omega(A) = \mathbf{P}\{Y_0 \in A|\mathcal{T}\}(\omega)$, definita su $\Omega \times P(I)^\infty$ tale che

$$\mathbf{P}\{Y_0 \in A_0, Y_1 \in A_1, \ldots, Y_n \in A_n\} = \int_\Omega \mu_\omega(A_0)\cdots\mu_\omega(A_n)\mathbf{P}(d\omega).$$

Per provare il secondo punto, grazie a un argomento monotono è sufficiente dimostrare che, se $\sigma \sim \tau$, allora

$$\mathbf{P}\{A_\sigma|Y_n\cdots Y_{n+m}\} = \mathbf{P}\{A_\tau|Y_n\cdots Y_{n+m}\}$$

per $n \geq n_0$ e per ogni m, ossia che

$$\mathbf{P}\{A_\sigma, Y_n = y_n, \ldots, Y_{n+m} = y_{n+m}\} = \mathbf{P}\{A_\tau, Y_n = y_n, \ldots, Y_{n+m} = y_{n+m}\}$$

per ogni $y_n, \ldots, y_{n+m}$ stringhe finite, per $n \geq n_0$ e per ogni m. Sia

$$\begin{aligned}S_\sigma \quad = \quad &\{\delta \in I^k : k \in N \text{ e gli n + i "uno} - \text{blocchi" di } \sigma\delta \\ &\text{siano uguali a } \mathrm{y_{n+i}}, \forall \mathrm{i} = 1, \ldots, \mathrm{m}\}.\end{aligned}$$

Allora, se $\sigma \sim \tau$, $S_\sigma = S_\tau$ e $\sigma\delta = \tau\delta$, per ogni $\delta \in S_\sigma$. Quindi $\mathbf{P}(A_{\sigma\delta}) = \mathbf{P}(A_{\tau\delta})$, per ogni $\delta \in S_\sigma$. Sommando su δ ed essendo S_σ numerabile si ottiene:

$$\mathbf{P}\{A_\sigma, Y_n = y_n, \ldots, Y_{n+m} = y_{n+m}\} = \sum_{\delta\in S_\sigma} \mathbf{P}(A_{\sigma\delta}) =$$

$$= \sum_{\delta\in S_\sigma} \mathbf{P}(A_{\tau\delta}) = \mathbf{P}\{A_\tau, Y_n = y_n, \ldots, Y_{n+m} = y_{n+m}\}.$$

Si osservi, ora, che Z è una catena di Markov se e solo se

$$\mathbf{P}(A_{\sigma j}|A_{\sigma}) = \mathbf{P}(A_{\tau j}|A_{\tau})$$

dove σ e τ sono stringhe finite di stati di I che iniziano e finiscono nello stesso stato (1 e i rispettivamente) ma non sono necessariamente equivalenti rispetto a $\sim$. In altre parole, occorre verificare l'uguaglianza

$$\mathbf{P}(A_{\sigma j})\mathbf{P}(A_{\tau}) = \mathbf{P}(A_{\tau j})\mathbf{P}(A_{\sigma}).$$

Si noti che, per ogni σ, τ stringhe finite, sotto l'ipotesi di indipendenza delle Y_n, vale

$$\mathbf{P}(A_{1\sigma 1\tau}) = \mathbf{P}(A_{1\sigma})\mathbf{P}(A_{1\tau}).$$

Sia δ una qualunque stringa finita non contenente 1. Allora si ha $\mathbf{P}(A_{\sigma}) = \sum_{\delta} \mathbf{P}(A_{\sigma\delta 1})$. Inoltre, $\sigma\delta\tau j \sim \tau\delta\sigma j$. Pertanto $\mathbf{P}(A_{\sigma\delta\tau j}) = \mathbf{P}(A_{\tau\delta\sigma j})$. Per l'indipendenza degli "uno-blocchi", sommando su δ si ottiene la tesi del punto (3).

Si conclude pertanto che se le variabili Y_n sono indipendenti e identicamente distribuite rispetto alla probabilità $\mu_{\omega}(\cdot) = \mathbf{P}\{Y_0 \in \cdot|\mathcal{T}\}(\omega)$, allora

$$\mu_{\omega}(Z_{n+1} = z_{n+1}|Z_n = z_n, \ldots, Z_0 = z_0) = \mu_{\omega}(Z_{n+1} = z_{n+1}|Z_n = z_n)$$

$\mathbf{P}$-q. c. e infine, definite le probabilità di transizione

$$p_{\omega}(z_n, z_{n+1}) = \mu_{\omega}(Z_{n+1} = z_{n+1}|Z_n = z_n)$$

si ottiene

$$\mathbf{P}\{Z_0 = z_0, \ldots, Z_n = z_n\} = \int_{\Omega} \prod_{k=0}^{n-1} p_{\omega}(z_k, z_{k+1})\mathbf{P}(d\omega).$$

*

Il controesempio a seguire (tratto da Diaconis e Freedman [2]) mostra la necessità dell'ipotesi di ricorrenza. Supponiamo $I = \{0, 1\}$. Sia Z il processo deterministico dotato dell'unica traiettoria $\tau = (0, 0, 1, 1, \ldots)$. Chiaramente Z non è ricorrente, ma è parzialmente scambiabile. Infatti, per ogni σ stringa finita di stati in I che sia l'inizio di τ, l'unica stringa ad essa equivalente è σ stessa. Ma Z non è mistura di catene di Markov. Infatti, se per assurdo lo fosse, si avrebbe $\tilde{p}_{11} = 1$ q.c. e si potrebbe scrivere

$$\theta_k(Z) := \mathbf{P}\{Z_0 = 0, \ldots, Z_k = 0, Z_{k+1} = 1, \ldots\} = \int_{[0,1]} p_{00}^{k-1}(1 - p_{00})\mu(0, dp_{00})$$

per la probabilità di una traiettoria che consista di k zeri seguiti da tutti 1. Si noti che l'evento descritto ha probabilità decrescente in k. Al contrario, data la natura deterministica di Z, è immediato verificare che $\theta_1(Z) = 0$, $\theta_2(Z) = 1$ e $\theta_3(Z) = 0$.

Equivalenza tra le due nozioni di scambiabilità parziale

Sia $Z = (Z_n)_{n\geq 1}$ un processo stocastico a valori nello spazio degli stati numerabile discreto $(I, \mathcal{I})$. Supponiamo $Z_0 = z_0$ quasi certamente e definiamo per ricorrenza i tempi opzionali

$$\begin{aligned} T_{i,1} &:= \inf\{n \geq 0 : Z_n = i\} \\ T_{i,m} &:= \inf\{n > T_{i,m-1} : Z_n = i\} \quad m > 1 \end{aligned}$$

per ogni $i \in I$. A parole $T_{i,m}$ è il tempo (aleatorio) della m-esima visita di Z in i. Se aggiungiamo a I un punto fittizio Δ, ottenendo la sua compattificazione di Alexandrov e quindi $(\bar{I}, \bar{\mathcal{I}})$ ancora numerabile discreto (almeno come spazio misurabile) possiamo poi definire la famiglia degli stati successivi $V = (V_{i,n})_{i\in\bar{I}, n\in N}$ associata al processo Z, ponendo

$$\begin{aligned} V_{i,n} &:= Z_{T_{i,n}+1} \quad \text{se } \mathrm{T_{i,n}} < \infty \\ V_{i,n} &:= \Delta \quad \text{se } \mathrm{T_{i,n}} = \infty. \end{aligned}$$

In buona sostanza la famiglia V descrive le transizioni del processo Z: infatti "$V_{i,n} = v$" significa "Z passa da i a v dopo l'n-esima visita in i". Peraltro, essendo fissato lo stato iniziale, le transizioni descrivono completamente Z.

Teo. 3 *La famiglia V è parzialmente scambiabile alla d-F se e solo se il processo Z è parzialmente scambiabile alla D-F e ricorrente.*

Seguendo la linea di Fortini, Ladelli, Petris e Regazzini [3], dimostreremo il teorema 3 mediante quattro lemmi: dai primi due seguirà il "solo se", dai secondi due il "se". Preliminarmente, introduciamo la nozione, in questo contesto ausiliaria, di ricorrenza forte.

Def. 4 *Un processo $(Z_n)_{n\geq 0}$ si dice fortemente ricorrente se per ogni $m \in N$ e $z_0, \ldots, z_m \in I$ vale che*

$$\boldsymbol{P}\left(\bigcap_{n=1}^{m}\{Z_n = z_n\}\right) = \boldsymbol{P}\left(\bigcap_{n=1}^{m}(\{Z_n = z_n\} \cap \Lambda_{z_n})\right)$$

dove $\Lambda_z := \{Z_n = z \text{ frequentemente}\}$.

In altre parole, assegnato uno stato $z \in I$, un processo Z fortemente ricorrente o non lo visita o lo visita infinite volte. Per riunione numerabile sui possibili valori iniziali e prendendo $m = 1$ si ritrova la nozione di ricorrenza data con la definizione (3).

È utile osservare che sia la scambiabilità parziale, sia la ricorrenza forte, sono condizioni sufficienti affinché le righe di V contengano o tutti o nessun Δ. Più formalmente, se V è parzialmente scambiabile o se Z è fortemente ricorrente, allora $\exists n : V_{i,n} = \Delta \overset{q.c.}{\Longrightarrow} \forall n : V_{i,n} = \Delta$ qualunque sia $i \in I$. Ciò è semplice conseguenza della definizione nel caso della ricorrenza forte, mentre segue dal teorema di rappresentazione nel caso della parziale scambiabilità: l'unico modo per una successione i cui elementi valgano definitivamente Δ di essere indipendente è che *tutti* i suoi elementi valgano Δ.

Lemma 1 *Se V è parzialmente scambiabile, Z è fortemente ricorrente.*

DIM.

Si consideri l'evento $A = \{Z_0 = z_0, Z_1 = z_1, \ldots, Z_n = z_n, Z_{n+1} = z_{n+1}\}$. Su A, quasi certamente, le righe di V corrispondenti a $z_0, \ldots, z_n$ non contengono alcun Δ. Allora $\mathbf{P}(A \cap \bigcap_{m=1}^{n} \Lambda_{z_m}) = \mathbf{P}(A)$. La tesi si ottiene per riunione numerabile su $z_{n+1} \in I$.

*

Lemma 2 *Se V è parzialmente scambiabile alla d-F, Z lo è alla D-F.*

DIM.

Sia $(z_0, \ldots, z_n) \sim (z'_0, \ldots, z'_n)$ e sia $F \subset I$ l'insieme degli stati distinti che compaiono in tali stringhe. Indicata con m_i la molteplicità di $i \in F$, si ha

$$\begin{aligned}
\{Z_0 = z_0, \ldots, Z_n = z_n\} &= \bigcap_{i \in F} \bigcap_{n=1}^{m_i} \{V_{i,n} = v_{i,n}\} \\
\{Z_0 = z'_0, \ldots, Z_n = z'_n\} &= \bigcap_{i \in F} \bigcap_{n=1}^{m_i} \{V_{i,n} = v_{i,\pi^i_n}\}
\end{aligned}$$

per opportune permutazioni π^i di $\{1, \ldots, m_i\}$. La tesi ne segue immediatamente.

*

Risulta così dimostrato che se V è parzialmente scambiabile alla d-F, allora Z è (fortemente) ricorrente e parzialmente scambiabile alla D-F. Otterremo il viceversa, come annunciato, mediante due ulteriori lemmi.

Lemma 3 *Se Z è fortemente ricorrente e parzialmente scambiabile alla D-F, allora V è parzialmente scambiabile alla d-F.*

DIM.

Per ogni $i \in I$ sia π^i una permutazione finita di N fissata ad arbitrio. Mostriamo l'uguaglianza in legge delle famiglie $(V_{i,n})_{i \in I, n \in N}$ e $(V_{i,\pi^i_n})_{i \in I, n \in N}$. A tal fine è sufficiente considerare insiemi del tipo

$$A = \bigcap_{i \in F} \bigcap_{n=1}^{m_i} \{V_{i,n} = v_{i,n}\} = \{Z_0 = z_0, \ldots, Z_n = z_n\}$$

con F parte finita di I, $m_i \in N$ tali che $\pi^i_{m_i} = m_i$ per ogni $i \in F$ e $v_{i,n} \in E$ tali che si trovino opportuni $z_0, \ldots, z_n \in E$. Insiemi che non ammettano una tale rappresentazione in termini di Z (o che non verifichino $\pi^i_{m_i} = m_i$ per ogni $i \in F$) si ottengono per riunione numerabile di insiemi A del tipo suddetto (in altre parole, per marginalizzazione). Si ha

$$A' = \bigcap_{i \in F} \bigcap_{n=1}^{m_i} \{V_{i,n} = v_{i,\pi^i_n}\} = \{Z_0 = z'_0, \ldots, Z_n = z'_n\}$$

con $(z_0, \ldots, z_n) \sim (z'_0, \ldots, z'_n)$, grazie all'ipotesi che gli m_i siano punti fissi delle π^i. Osserviamo l'importanza dell'ipotesi di ricorrenza forte, che rende banalmente vera la tesi per gli eventi in cui compaia Δ.

Lemma 4 *Se Z è ricorrente e parzialmente scambiabile alla D-F, allora Z è fortemente ricorrente.*

DIM.

Il processo Z esce quasi certamente da z_0 ed è ricorrente, pertanto $(V_{z_0,n})_{n\geq 1}$ è scambiabile (si veda al proposito la dimostrazione del lemma precedente). Allora le $V_{z_0,n}$ sono indipendenti e identicamente distribuite condizionatamente all'elemento aleatorio

$$\tilde{\theta} = (\tilde{\theta}_i)_{i\in I} = \lim_{m\to\infty} \frac{1}{m} \sum_{n=1}^{m} \mathbf{1}_{\{V_{z_0,n}=i\}}$$

e anzi $\tilde{\theta}_i \in \mathbf{P}(V_{z_0,1} = i|\tilde{\theta})$. Sia ora $z \in I$ e $A_z^+ := \{\tilde{\theta}_z > 0\}$. Sfruttando il secondo lemma di Borel e Cantelli, possiamo calcolare

$$\begin{aligned}
\mathbf{P}\{Z_1 = z\} &= \mathbf{P}\{V_{z_0,1} = z\} \\
&= \mathbf{E}\left[\tilde{\theta}_z \mathbf{1}_{A_z^+}\right] \\
&= \mathbf{E}\left[\mathbf{P}\left(V_{z_0,1} = z, V_{z_0,n} = z\, freq.|\tilde{\theta}\right) \mathbf{1}_{A_z^+}\right] \\
&= \mathbf{P}\{V_{z_0,1} = z, V_{z_0,n} = z\, freq.\}
\end{aligned}$$

e pertanto, siccome $\{V_{z_0,n} = z\, freq.\} \subset \Lambda_z$, la condizione di ricorrenza forte è verificata con $m = 1$. Non resta che procedere per induzione, osservando che condizionatamente a $\{Z_0 = z_0, \ldots, Z_m = z_m\}$ il processo $(Z_{m+n})_{n\geq 0}$ esce da z_m, è parzialmente scambiabile alla D-F ed è anche ricorrente per l'ipotesi induttiva. Il ragionamento sopra svolto può allora iterarsi e la forte ricorrenza di Z risulta dimostrata.

*

Mostriamo infine la necessità dell'ipotesi di ricorrenza, mediante un opportuno controesempio. Siano $I = \{0, 1, 2\}$, $z_0 = 0$ e $\mathbf{P}$ tale che

$$\mathbf{P}\{Z_1 = 1, Z_2 = 0, Z_3 = 0\} = \mathbf{P}\{Z_1 = 0, Z_2 = 1, Z_3 = 0\} = \frac{1}{2}$$

e $\mathbf{P}\{Z_n = 2\} = 1$ per $n \geq 4$. Allora il processo Z è parzialmente scambiabile alla D-F, ma la relativa famiglia V non è parzialmente scambiabile alla d-F perché risulta

$$\frac{1}{2} = \mathbf{P}\{V_{0,1} = 1, V_{0,2} = 0, V_{0,3} = 2\} \neq \mathbf{P}\{V_{0,1} = 1, V_{0,2} = 2, V_{0,3} = 0\} = 0.$$

Ovviamente un siffatto Z non è ricorrente.

Riassumendo, per un processo con spazio degli stati numerabile discreto, la scambiabilità secondo de Finetti degli stati successivi alle visite è in generale una condizione più forte della scambiabilità secondo Diaconis e Freedman; tuttavia le due condizioni si equivalgono nell'ipotesi di ricorrenza.

Un'applicazione

Un esempio di processo parzialmente scambiabile alla D-F viene fornito da Muliere, Secchi e Walker in [4]. Il tipo di processo trattato dai suddetti autori viene definito "processo d'urna rinforzato" (Reinforced Urn Process).

Siano S uno spazio di stati numerabile discreto ed E un insieme finito di colori tale che $card(E) = k \geq 1$. Sia $U : S \to C$ la funzione rappresentante la composizione iniziale dell'urna associata allo stato x, $\forall x \in S$, dove $C = \{(r_1, \ldots, r_k) : r_i \in N, \forall i, \text{ e } \sum_{\mathrm{i=1}}^{\mathrm{k}} \mathrm{r_i} > 0\}$. Sia $q : S \times E \to S$ la legge evolutiva del processo. Il processo con i suddetti elementi costitutivi verrà indicato con $X \in RUP(S, E, U, q)$. A ogni stato x è dunque associata un'urna con composizione iniziale $U(x)$, da cui si estrae una pallina di colore c, presente inizialmente nell'urna in numero $n_x(c)$. L'estrazione avviene ogni volta che il processo viene a trovarsi nello stato x, con reimmissione della pallina estratta e di una seconda dello stesso colore. La legge evolutiva q determina lo stato $q(x, c)$ successivamente visitato dal processo. Si supponga che $\forall\ x, y \in S$ esista al più un colore $c = c(x, y) \in E$ tale che $q(x, c(x, y)) = y$.

Si consideri il processo che ha inizio quasi certo in $x_0 \in S$ fissato. Sia $\sigma = (x_0, \ldots, x_n)$ una sequenza finita di stati di S. Essa si dirà *ammissibile* se $\sigma_0 = x_0$ e se esiste una sequenza $(c_0, \ldots, c_{n-1})$ in E tale che $q(x_i, c_i) = x_{i+1}$ e $n_{x_i}(c_i) > 0$, per $i = 0, \ldots, n-1$. Sia $t(x, y)$ il numero di transizioni in σ da x a y, $l_x(c) := t(x, q(x, c))$ il numero di transizioni da x causate dal colore c e $t(x) := \sum_{y \in S} t(x, y)$ il numero totale di transizioni uscenti da x.

Il seguente risultato mostra che il processo X è parzialmente scambiabile alla D-F, grazie alla particolare forma delle distribuzioni finito-dimensionali che ne determinano la legge.

Teo. 4 *Sia $n \geq 1$. Allora, se $\sigma = (x_0, \ldots, x_n)$ è ammissibile, si ha*

$$\boldsymbol{P}\{X_0 = x_0, \ldots, X_n = x_n\} = \prod_{x \in S} \left[\frac{\prod_{c \in E} \prod_{i=0}^{l_x(c)-1} (n_x(c) + i)}{\prod_{j=0}^{t(x)-1} (j + \sum_{c \in E} n_x(c))} \right].$$

Altrimenti, $\boldsymbol{P}\{X_0 = x_0, \ldots, X_n = x_n\} = 0$.

DIM.

Si usi la convenzione $\prod_{i=0}^{-1} = 1$. Se σ non è ammissibile, la tesi è immediata. Altrimenti, si fattorizzi $\mathbf{P}\{X_0 = x_0, \ldots, X_n = x_n\}$ in

$$\begin{aligned} &\mathbf{P}\{X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = x_0\} \cdots \\ &\quad \cdots \mathbf{P}\{X_1 = x_1 | X_0 = x_0\} \mathbf{P}\{X_0 = x_0\} \end{aligned}$$

e si osservi che

$$\begin{aligned} &\mathbf{P}\{X_k = x_k | X_{k-1} = x_{k-1}, \ldots, X_0 = x_0\} = \\ &= \frac{n_{x_{k-1}}(c(x_{k-1}, x_k)) + m_{x_{k-1}}(c(x_{k-1}, x_k))}{\sum_{c \in E} \left(n_{x_{k-1}}(c) + m_{x_{k-1}}(c) \right)} \end{aligned}$$

dove $m_{x_0}(c) = 0$ e $m_{x_k}(c) = \sum_{j=0}^{k-1} \mathbf{1}_{\{(x_j, c(x_j, x_{j+1}))\}}(x_k, c)$. Quindi

$$\mathbf{P}\{X_0 = x_0, \ldots, X_n = x_n\} = \prod_{k=0}^{n-1} \left[\frac{n_{x_k}(c(x_k, x_{k+1})) + m_{x_k}(c(x_k, x_{k+1}))}{\sum_{c \in E} (n_{x_k}(c) + m_{x_k}(c))} \right] \quad (2)$$

da cui ci accingiamo a derivare la tesi. Per ogni stato x che figura in $\sigma = (x_0, \ldots, x_{n-1})$, sia $(y_1, \ldots, y_{t(x)})$ la sequenza ordinata degli stati visitati immediatamente dopo x in σ. Allora, i fattori associati a x in (2) possono essere riscritti nella forma

$$\frac{n_x(c(x,y_1))}{\sum_{c\in E} n_x(c)} \frac{n_x(c(x,y_2))+\mathbf{1}_{\{y_1\}}(y_2)}{1+\sum_{c\in E} n_x(c)} \cdots \frac{n_x(c(x,y_{t(x)}))+\sum_{i=1}^{t(x)-1} \mathbf{1}_{\{y_i\}}(y_{t(x)})}{t(x)-1+\sum_{c\in E} n_x(c)} =$$
$$= \frac{\prod_{c\in E} \prod_{i=0}^{l_x(c)-1} (n_x(c)+i)}{\prod_{j=0}^{t(x)-1} (j+\sum_{c\in E} n_x(c))} \ .$$

Ne segue la tesi.

*

Riferimenti bibliografici

[1] B. de Finetti. Sur la condition d'"équivalence partielle". In *VI colloque Genève: "Act. Sc. Ind.", n^o 739*. Herman, Paris, 1938.

[2] P. Diaconis and D. Freedman. De finetti's theorem for Markov chains. *The Annals of Probability*, 8:115–130, 1980.

[3] S. Fortini, L. Ladelli, G. Petris, and E. Regazzini. On mixtures of distributions of Markov chains. *Quaderno IAMI-CNR, n^o 9*, 1999.

[4] P. Muliere, P. Secchi, and S. G. Walker. Urn schemes and reinforced random walks. *Stochastic Processes and their Applications*, 88:59–78, 2000.

[5] E. Regazzini. Coherence, exchangeability and statistical models. In *Atti del convegno "Sviluppi metodologici nei diversi approcci all'inferenza statistica"*, volume 2, pages 101–137. Pitagora Editrice, Bologna, 1991.