Alma Mater Studiorum · Università di Bologna

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Matematica, Curriculum applicativo

UN'ANALISI PROBABILISTICA DELLA MOBILITA' SOCIALE

Tesi di Laurea in Meccanica Statistica

Relatore:
Chiar.mo Prof.
Pierluigi Contucci

Correlatore:
Dott.ssa
Alessandra Bianchi

Presentata da:
Sara Garzia

II Sessione
Anno Accademico 2011-2012

La mente è come un paracadute.

Funziona solo se si apre.

A. Einstein

Introduzione

Da oltre mezzo secolo, in molti Paesi europei e negli Stati Uniti, la mobilità sociale costituisce un argomento di studio di grande rilievo. Su questo tema numerosi sociologi in tutto il mondo hanno condotto ricerche vaste ed accurate, pur partendo da impostazioni teoriche molto diverse.
A raccogliere ed analizzare dati sulla mobilità sociale si sono dedicati infatti sia coloro che volevano capire come si formano, si trasformano ed agiscono le classi sociali, sia coloro che, mirando alla realizzazione dell'ugualianza delle opportunità di partenza, erano più interessati a misurare il grado di apertura o di fluidità della società.
Per quanto riguarda il nostro Paese, vi è una eccezione. Nei giornali, alla televisione, nelle conversazioni di tutti i giorni si fa spesso uso del concetto di mobilità sociale per analizzare fenomeni e tendenze diverse, per esempio l'espansione dell'istruzione superiore o la secolarizzazione religiosa o la diffusione di alcune forme di comportamento deviante. Ma i sociologi italiani non hanno dedicato finora molto tempo allo studio della mobilità sociale, sono pochissime le ricerche su questo tema condotte nel nostro paese.

Questa tesi nasce con l'intento di intraprendere un piccolo passo verso l'incontro delle scienze *dure* (matematica, fisica, etc.) con quelle *morbide* (scienze economiche e sociali). Un grande supporto ci è stato fornito dal Prof. Marzio Barbagli. Grazie a lui abbiamo potuto apprendere le nozioni e le caratteristiche proprie della mobilità sociale.
Il nostro scopo è quello di applicare i metodi della matematica allo studio

della mobilità sociale. Grazie agli strumenti di probabilità si cercherà di dare una previsione sulle frequenze delle varie classi sociali prese in esame nelle generazioni future. L'approccio matematico è questo: considerare la mobilità sociale come una catena di Markov a stati discreti, cioè come un processo stocastico avente come matrice di transizione la tavola di mobilità, ottenuta dall'incrocio delle frequenze dell'intervistato e del padre; calcoleremo il vettore della distribuzione di probabilità stazionaria relativa alla suddetta tavola di mobilità e infine calcoleremo il tempo di mescolamento.

Le catene di Markov costituiscono uno dei modelli probabilistici più conosciuti, grazie al loro vasto impiego in numerose discipline, dalla statistica alla fisica, dalla genetica all'informatica. Esse forniscono un modello matematico per descrivere sistemi soggetti a cambiamenti casuali di stato, la cui evoluzione dipende dallo stato momentaneo del sistema e non dalla sua storia passata. Il termine *catena di Markov* deriva dal matematico russo Andrei Andreevic Markov che per primo, all'inizio del 900, le studiò e ne definì formalmente la teoria.
La distribuzione di probabilità stazionaria ci assicura una condizione di equilibrio per una catena di Markov, sotto particolari ipotesi. Infine per tempo di mescolamento si intende il tempo che impiega una catena di Markov a raggiungere tale distribuzione.

I dati utilizzati in questa ricerca ci sono stati forniti, grazie al Prof. Marzio Barbagli, dall'Istituto Nazionale di Statistica ISTAT. Si tratta di un'indagine multiscopo sulle famiglie dal titolo 'Aspetti della vita quotidiana, anno 2001' in cui vengono rilevati gli aspetti fondamentali della vita quotidiana e i comportamenti relativi all'anno in corso. L'indagine ha raggiunto approssimativamente 19.920 famiglie per un totale di 53.113 individui.
In questa indagine ci basiamo su uno schema a otto classi sociali e i dati sono stati elaborati con il software di statistica SPSS.

La tesi è così organizzata. Nel primo capitolo viene trattato il tema della mobilità sociale sotto il punto di vista sociologico e verranno introdotte delle nozioni per poter comprendere al meglio il problema. Il secondo capitolo tratta della natura dei dati e in che modo si è arrivati a suddividere le professioni in classi sociali e la costruzione della tavola di mobilità. Nel terzo e nel quarto capitolo sono presenti gli aspetti matematici quali le catene di Markov a stati discreti, la distribuzione di probabilità stazionaria e il tempo di mescolamento. Infine, nel quinto capitolo, ci sono i risultati di questa ricerca, parte fondamentale di tutto il lavoro.

I risultati mostrano che con un tempo di mescolamento stimato pari a tre generazioni, la nostra tavola di mobilità converge ad una situazione di equilibrio, detta distribuzione di probabilità stazionaria.
Tale risultato è basato sull'ipotesi che la matrice di transizione rimanga invariata e pertanto potrebbe sembrare lontano dalla realtà fenomenologica in cui condizioni sociali ed economiche cambiano nel corso di tre generazioni. Va tuttavia osservato che sogni ed aspirazioni degli individui rimangono focalizzati sui propri obiettivi anche se le condizioni esterne cambiano a causa di una forma di inerzia legata alla propria cultura. Il risultato quindi ha ragioni per descrivere una buona approssimazione della realtà.

Indice

Capitolo 1

La mobilità sociale: aspetti sociologici

In questo capitolo trattiamo il tema della mobilità sociale, in particolare delle classi sociali, sotto il punto di visto sociologico. Daremo delle definizioni e nozioni che ci permetteranno di introdurre la ricerca svolta in questa tesi.

1.1 Tipi di mobilità

Definiamo *mobilità sociale* ogni passaggio di un individuo da uno strato, un ceto, una classe sociale ad un altro. I sociologi che nell'ultimo mezzo secolo si sono dedicati allo studio sistematico di questi movimenti distinguono fra mobilità orizzontale e verticale, ascendente e discendente, intergenerazionale ed intragenerazionale, di breve e di lungo raggio, assoluta e relativa, individuale e di gruppo.
Per *mobilità sociale orizzontale* si intende il passaggio di un individuo da una posizione sociale ad un'altra allo stesso livello. E' questo il caso di coloro che nascono in una famiglia di artigiani o di commercianti e che, quando raggiungono l'età adulta, fanno gli impiegati. Si parla invece di *mobilità sociale verticale* per indicare lo spostamento ad una posizione più alta o più bassa

nel sistema di stratificazione sociale. Nel primo caso è ascendente e nel secondo caso è discendente. Per comprendere meglio: il figlio dell'operaio di fabbrica che prende la laurea in ingegneria e va a fare il dirigente d'azienda è un caso di mobilità sociale ascendente. Invece, colui che viene da una famiglia di imprenditori ma non prosegue gli studi dopo la licenza media inferiore ed entra nel mercato del lavoro come commesso di un negozio è un caso di mobilità discendente. Si tratta sempre comunque di *mobilità di lungo raggio* avvenuta fra classi molto lontani; quando invece, come accade spesso nella realtà, questi strati o classi sono contigue, si parla di *mobilità di breve raggio.* I passaggi fra una classe sociale e l'altra possono essere esaminati mettendo a confronto la posizione della famiglia di origine di un individuo con quella che egli ha raggiunto in un determinato momento della sua vita. Si parla, in questo caso, di *mobilità sociale intergenerazionale.* Ma il confronto può essere anche fatto fra le posizioni che una persona ha occupato nel corso della sua esistenza, per esempio fra quando è entrato nel mercato del lavoro e dieci anni dopo. Questa mobilità viene chiamata *intragenerazionale* o *di carriera.* La *mobilità assoluta* è data dal numero complessivo di persone che si spostano da una classe all'altra (ad esempio, nel caso di quella intergenerazionale, che appartengono ad una classe diversa da quella dei genitori) e per *mobilità relativa* si intende il grado di eguaglianza delle possibilità di mobilità dei membri delle varie classi.
Le distinzioni fatte finora sono riferite alla *mobilità individuale* perché hanno come soggetto il singolo individuo. Del tutto diversa è la *mobilità sociale collettiva* che riguarda i movimenti verso l'alto o verso il basso di un intero gruppo (uno strato, una classe o una casta) di persone rispetto a tutti gli altri gruppi sociali.
Il sociologo russo Pitirim Sorokin ha presentato la differenza fra mobilità discendente individuale e di gruppo scrivendo che 'il primo caso di caduta può essere paragonato ad un individuo che cade da una nave, mentre il secondo ricorda l'affondamento della nave con tutte le persone a bordo, o una nave ridotta a un relitto che si sfascia' [Sorokin 1927, trad. it. 1965]. Un esempio

interessante di mobilità di gruppo ascendente è stato in Italia quello della professione di notaio, quando la legge di riforma del 1913 richiese l'obbligo della laurea in giurisprudenza per esercitarla (mentre prima bastavano due anni di università) e riconobbe lo status libero professionale a fianco di quello pubblico ufficiale [Santoro 1994].

1.2 Due tradizioni teoriche

Lo studio della mobilità sociale è stato intrapreso da due diversi punti di vista e per dare risposta a due diversi interrogativi teorici:

- Il primo ha a che fare con il concetto di apertura di una società o di fluidità sociale cioè con le opportunità che le persone con origini sociali diverse hanno di raggiungere le varie posizioni nel sistema di stratificazione. Gli studiosi che si richiamano a questa impostazione si chiedono di solito se l'apertura di una determinata società sia mutata nel corso del tempo e cercano di individuare i fattori che favoriscono o ostacolano la fluidità sociale.

- Il secondo filone ruota invece intorno al problema della formazione e dell'azione delle classi. Alcuni sociologi hanno sostenuto che una classe diventa una formazione stabile quando coloro che ne fanno parte condividono valori, idee, stili di vita e ritengono di avere interessi comuni. Per questo, essi si sono chiesti se la mobilità intergenerazionale, riducendo la componente permanente di una classe, cioè la quota di persone che restano per tutta la vita nella stessa posizione dei genitori, impedisca a questa di diventare una collettività sociale.

1.3 Le ricerche sulla mobilità sociale

Il primo studio sistematico sulla mobilità sociale è stato scritto nel 1927 da *Pitirim Sorokin*, un sociologo nato nel 1889 nella Russia nord-occidentale, che di questo tema aveva una conoscenza personale diretta. Come avrebbe raccontato molti anni dopo nella sua autobiografia, egli era figlio di un operaio e di una contadina e, nel corso della sua vita, era stato successivamente bracciante, artigiano ambulante, operaio di fabbrica, impiegato, insegnante, capo corista, rivoluzionario, prigioniero politico, giornalista, studente, direttore di un giornale, membro del governo Kerensky, esule, professore in università russe, ceche e americane, e studioso di fama internazionale.
Nel suo libro Sorokin analizzava la mobilità sociale in numerosi Paesi (non solo occidentali) per un lunghissimo arco temporale (dall'antica Roma alla fine del secolo scorso) basandosi su una ricchissima documentazione. Una parte di questa è costituita da dati statistici, ricavati non da inchieste sistematiche su campioni rappresentativi di tutta la popolazione di un Paese, ma dalle cosiddette indagini di élite, riguardanti cioè le origini sociali di alcuni gruppi particolari (i sovrani, gli uomini di genio, i santi, i dirigenti).
Il libro di Sorokin è considerato un classico. Alcune delle sue tesi di fondo sono ancora oggi valide. La documentazione empirica su cui queste si basano non è priva di interesse. Tuttavia, dal punto di vista della raccolta e dell'analisi dei dati, le ricerche di mobilità sociale hanno fatto molti passi avanti nell'ultimo periodo, soprattutto dopo la seconda guerra mondiale. E' da allora che i sociologi hanno iniziato a condurre indagini sistematiche su campioni rappresentativi della popolazione di un paese. La prima è stata realizzata nel 1949 da un gruppo della London School of Economics diretta da David Glass [1954] su 10 mila adulti residenti in Inghilterra, in Scozia e nel Galles. Numerose altre sono state compiute in seguito, specialmente nei paesi occidentali. Fra le più importanti vanno ricordate quelle condotte negli Stati Uniti da Blau e Duncan [1967], in Gran Bretagna da Goldthorpe [1980], in Francia da Thélot [1980], in Italia da Cobalti e Schizzerotto [1994]. Tutte queste ricerche hanno alcune caratteristiche in comune. Si servono di

campioni della popolazione molto grandi (da 5.000 a 25.000 casi) e dunque sono assai costose. Mirano a rilevare la posizione sociale degli individui che fanno parte di questi campioni e delle loro famiglie di origine. Per questo, di solito chiedono agli intervistati l'occupazione che svolgono in quel momento, quella che avevano prima (quando sono entrati nel mercato del lavoro e dieci anni dopo), l'occupazione del padre (quando essi avevano 12 o 14 anni) e, se sono sposati, quella del suocero. Nell'analisi dei dati fanno uso di tecniche avanzate e complesse. E' proprio nelle indagini di mobilità che i sociologi hanno per la prima volta sperimentato alcune di queste tecniche.
Tutte le più importanti indagini di questo tipo condotte negli Stati Uniti e in Europa (ad eccezione di quella svolta in Italia) si basano su dati riguardanti soltanto la popolazione maschile. Il motivo è che i sociologi che le hanno dirette si rifacevano alla concezione tradizionale della posizione delle donne nel sistema di stratificazione sociale. Tale concezione si basa su due assunti di fondo. Il primo è che l'unità base del sistema di stratificazione non è l'individuo ma la famiglia. Il secondo è che la posizione della famiglia in questo sistema è determinata interamente da quella del capofamiglia, cioè del marito o del padre. L'occupazione della donna non ha, secondo questa concezione, alcun peso per la collocazione della famiglia nel sistema di stratificazione sociale.
Ultimamente molti studiosi e studiose hanno criticato questa concezione, sostenendo che essa rende impossibile l'analisi di una delle più importanti forme di disuguaglianza sociale, quella appunto basata sulle differenze di genere. Essi hanno messo in rilievo che oggi la concezione della stratificazione sociale è meno accettabile di un tempo perché contrasta con alcune tendenze di fondo delle società avanzate e in particolare con due di queste. La prima è che una quota crescente di famiglie ha come 'capo' non un uomo ma una donna. La seconda tendenza è che aumentano il tasso di attività della popolazione femminile: le donne entrano sempre più spesso nel mercato del lavoro, occupano posizioni più elevate di un tempo, vi restano più a lungo di prima. Questa tendenza ha prodotto effetti di grande rilievo sulla famiglia:

ha fatto diminuire il grado di omogamia (percentuale di sposi che fanno lo stesso lavoro) e ha fatto aumentare il numero di coppie nella quali la moglie ha un'occupazione superiore a quella del marito.
Il dibattito su questo problema è ancora aperto nelle riviste scientifiche ma oggi nessuno studioso condurrebbe più una ricerca di mobilità sociale coinvolgendo soltanto la popolazione maschile.

1.4 L'importanza delle classi sociali

Alcuni sociologi ritengono che ormai il concetto di classe sociale sia una cosa passata, inutilizzabile da chi voglia capire la realtà delle società contemporanee. Terry Clark e Seymour Martin Lipset [1991] scrivono: 'Negli ultimi decenni man mano che le gerarchie tradizionali perdevano di importanza ed emergevano nuove differenze sociali, l'analisi di classe si è rivelata sempre più inadeguata [...] Il concetto di classe, per quanto utile per lo studio dei precedenti periodi storici, è sempre più superato'.
Si è visto che la classe sociale influisce meno di un tempo sul comportamento di voto. Mentre un tempo gli elettori di classe operaia tendevano a preferire i partiti di sinistra e quelli delle classi medio-alte i partiti di destra, oggi questo non si verifica più o per lo meno non nella stessa misura di prima.
Vi sono però altri sociologi che ritengono che il concetto di classe sociale sia utile anche per l'analisi delle società contemporanee. Sono anche convinti che la classe sociale resti un criterio significativo di strutturazione delle disuguaglianze e che tuttora l'appartenenza ad una classe influisca su molti aspetti della vita di un individuo.
Alla metà degli anni settanta il sociologo americano Randall Collins scrive: 'Nella ricerca empirica la classe sociale si è rivelata la variabile esplicativa di gran lunga più importante' [1975, trad. it. 1980]. I risultati delle ricerche condotte nei diversi paesi occidentali continuano a mostrare che la classe sociale esercita ancora una influenza rilevante su molte forme di comporta-

mento. Questo vale ad esempio per le scelte elettorali, perché le analisi più sofisticate dei dati esistenti hanno mostrato che, contrariamente a quanto sostengono Terry Clark e Seymour Martin Lipset, fra classe sociale e preferenza di voto vi è anche oggi una relazione significativa [Manza 1995; Goldthorpe 1996; Hout 1996].

1.5 Le tecniche di analisi

Nella letteratura sociologica contemporanea sono rintracciabili due principali strategie di trattamento statistico dei dati di mobilità. Si tratta, da un lato, delle tecniche di regressione e, dall'altro lato, delle procedure di analisi per variabili categoriali. Il ricorso alle prime presuppone che la struttura delle posizioni sociali formi un continuum gerarchicamente ordinato. Questa strategia analitica risulta, dunque, particolarmente congruente con la concezione dei processi di mobilità sottostante al modello meritocratico delle diseguaglianze sociali. Le tecniche di analisi per variabili categoriali sono, invece, preferibili qualora si assuma che le disparità sociali diano luogo a raggruppamenti discreti tra i quali intercorrono relazioni complesse, non sempre riducibili a una gerarchia lineare, di dominio e di subordinazione. Esse risultano, pertanto, meglio adatte allo studio dei processi di mobilità nell'ottica della formazione e dell'azione delle classi sociali.
In particolare, in questo lavoro si utilizzano le procedure per il trattamento di dati categoriali e lo strumento di base che si userà saranno le *tavole di mobilità*. Esse sono delle matrici quadrate che hanno sulle righe le classi di origine e sulle colonne le classi di arrivo. Gli elementi di queste matrici riportano la frequenza con cui i soggetti di una data classe di provenienza sono presenti in una determinata classe di destinazione.

1.6 Le conseguenze della mobilità sociale

Il metodo più rigoroso per individuare gli effetti della mobilità sociale è certamente quello che consiste nel seguire per un lungo periodo di tempo dei campioni di soggetti socialmente stabili, mobili ascendenti e mobili discendenti (rispetto a diverse classi sociali) e nel registrare delle informazioni su alcuni valori e comportamenti sia della famiglia di origine che di questi stessi soggetti, in vari momenti della loro vita.
Nella realtà, però, è molto difficile (se non quasi impossibile) percorrere questa strada. Per questo motivo, si fa di solito ricorso a indagini trasversali, utilizzando varie procedure di analisi. La procedura metodologicamente più corretta consiste nell'individuare innanzitutto i valori ed i comportamenti delle persone socialmente stabili nelle diverse classi, nel considerare tali valori e comportamenti come tipici delle varie classi e dunque come parametri validi di analisi. Mettendo a confronto con questi parametri i valori ed i comportamenti delle persone mobili si cerca di accertare gli effetti dei diversi passaggi fra le varie classi prese in considerazione.
E' proprio con questa procedura che con il Prof. Marzio Barbagli, in questo lavoro, si è voluto studiare la mobilità sociale con il campione di dati preso in esame. Purtroppo questa tecnica non ha portato risultati soddisfacenti perché si è riscontrato che non ci sono particolari differenze tra le varie classi sociali considerate. I comportamenti dei vari membri dei ceti non si ripetono in maniera sistematica come invece si può riscontrare nella ricerca effettuata dallo stesso Prof. Marzio Barbagli con i professori Vittorio Capecchi e Antonio Cobalti. Tale ricerca riguarda lo studio della mobilità sociale in Emilia Romagna ed è stata svolta in collaborazione con l'Unione Regionale delle Camere di Commercio dell'Emilia Romagna e con le singole Camere provinciali. Si tratta di una indagine svoltasi nel novembre 1983, su un campione di 4.493 unità familiari residenti nei 94 comuni dell'Emilia Romagna inclusi nel campione ISTAT per la rilevazione delle forze di lavoro. Con questo questionario sono state raccolte informazioni sulle 12.946 persone che facevano parte di queste famiglie.

Le diverse aree di comportamenti prese in considerazione sono: la politica, la religione, l'uso del tempo libero, il matrimonio e la riproduzione.
Esamimando la popolazione emiliana si arriva alla conclusione che la mobilità sociale ha determinato uno spostamento politico verso il centro-destra, piuttosto che verso sinistra, ha favorito l'instabilità elettorale ed ha scoraggiato l'iscrizione ai partiti. Invece sul comportamento politico delle classi urbane la mobilità sociale ha avuto effetti opposti. Essa ha infatti provocato un aumento, in ciascuna classe sociale, della quota di persone che votano per l'allora PCI, che sono elettoralmente stabili e che aderiscono ad un partito.
Per quanto riguarda l'uso del tempo libero, esaminando la popolazione emiliana nel suo complesso si è visto che la mobilità sociale favorisce l'abitudine ad andare in vacanza. E per quanto riguarda le singole classi urbane la mobilità sociale ha effetti opposti perché determina un aumento, in ciascuna di esse, della quota di coloro che non vanno in vacanza.
In particolare si è visto che in Emilia gli operai di origine agricola hanno stili di vita diversi da quelli degli operai di seconda generazione.
L'analisi dei dati mette in rilievo che a livello micro la mobilità sociale non determina lo sradicamento e l'isolamento delle persone che cambiano classe, ma mette in moto invece un processo di ridefinizione dell'identità sociale, di sostituizione dei modi di pensare e di agire della classe di partenza con quella di arrivo.

Capitolo 2

I dati

In questo capitolo verranno introdotti i dati usati nella ricerca e in particolare si elencheranno le classi sociali ottenute dalla classificazione dell'attività lavorativa degli intervistati e dei rispettivi padri.

2.1 La natura dei dati

I dati utilizzati in questa ricerca sono stati resi disponibili, grazie al Prof. Marzio Barbagli, dall'Istituto Nazionale di Statistica ISTAT.
Si tratta di un'indagine multiscopo sulle famiglie dal titolo '**Aspetti della vita quotidiana, anno 2001**' in cui vengono rilevati gli aspetti fondamentali della vita quotidiana e i comportamenti relativi all'anno in corso.
Sono presenti varie aree tematiche che permettono di cogliere come vive la popolazione, se è soddisfatta dei servizi di pubblica utilità che devono contribuire alla qualità della vita. I principali contenuti informativi dell'indagine sono: famiglia, abitazione, zona in cui si vive, istruzione e formazione, lavoro domestico ed extradomestico, spostamenti quotidiani, tempo libero e partecipazione sociale, stili di vita e condizione di salute, consumo di farmaci e utilizzo dei servizi sanitari, micro-criminalità, funzionamento dei servizi di pubblica utilità.

L'indagine ha raggiunto approssimativamente 19.920 famiglie per un totale di 53.113 individui, ma come si vedrà in seguito nella sezione dedicata ai risultati sperimentali, il numero effettivo delle persone è molto minore.
Le informazioni sono state raccolte con intervista diretta per una parte dei quesiti. Nei casi in cui l'individuo non fosse disponibile all'intervista per particolari motivi, le informazioni sono state fornite da un altro componente della famiglia. Per un'altra parte dei quesiti è stata prevista l'autocompilazione.

2.2 La suddivisione in classi sociali

In questa analisi ci basiamo su uno schema a otto classi. Questo schema raggruppa nel modo che segue le varie occupazioni:

1. borghesia: dirigenti, direttivi, quadri, imprenditori, liberi professionisti;

2. classe media impiegatizia: impiegati, intermedi, insegnanti;

3. dipendenti agricoltura: capi operaio, operai subalterni e assimilati, apprendisti e lavoranti a domicilio per conto di impresa nel ramo di agricoltura, caccia e pesca;

4. dipendenti industria: capi operaio, operai subalterni e assimilati, apprendisti e lavoranti a domicilio per conto di impresa nel ramo di industria, estrazione e costruzioni;

5. dipendenti terziario: capi operaio, operai subalterni e assimilati, apprendisti e lavoranti a domicilio per conto di impresa nel ramo di commercio, alberghi, ristoranti, trasporti, magazzini e comunicazioni, intermediazioni, noleggio, pubblica amministrazione e difesa, istruzione, sanità e altri servizi sociali;

6. autonomi agricoltura: lavoratori in proprio, soci cooperativa produzione beni e/o prestazioni di servizio, coadiuvanti nel ramo di agricoltura, caccia e pesca;

7. autonomi industria: lavoratori in proprio, soci cooperativa produzione beni e/o prestazioni di servizio, coadiuvanti nel ramo di industria, estrazione e costruzioni;

8. autonomi terziario: lavoratori in proprio, soci cooperativa produzione beni e/o prestazioni di servizio, coadiuvanti nel ramo di commercio, alberghi, ristoranti, trasporti, magazzini e comunicazioni, intermediazioni, noleggio, pubblica amministrazione e difesa, istruzione, sanità e altri servizi sociali.

I dati sono stati elaborati con il software di statistica SPSS (Statistical Package for Social Science); di seguito riportiamo le tabelle di frequenza degli intervistati e dei padri.

Intervistato

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	borghesia	3915	7,4	11,7	11,7
	classe media impiegatizia	9431	17,8	28,1	39,7
	dipendenti agricoltura	1664	3,1	5,0	44,7
	dipendenti industria	6658	12,5	19,8	64,5
	dipendenti terziario	5232	9,9	15,6	80,1
	autonomi agricoltura	1758	3,3	5,2	85,3
	autonomi industria	1144	2,2	3,4	88,7
	autonomi terziario	3799	7,2	11,3	100,0
	Totale	33601	63,3	100,0	
Mancanti	Mancante di sistema	19510	36,7		
Totale		53111	100,0		

Padre

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulata
Validi	borghesia	3704	7,0	10,5	10,5
	classe media impiegatizia	5099	9,6	14,5	25,0
	dipendenti agricoltura	4433	8,3	12,6	37,6
	dipendenti industria	9061	17,1	25,8	63,4
	dipendenti terziario	3811	7,2	10,8	74,2
	autonomi agricoltura	4522	8,5	12,9	87,1
	autonomi industria	1224	2,3	3,5	90,6
	autonomi terziario	3316	6,2	9,4	100,0
	Totale	35170	66,2	100,0	
Mancanti	non so	2024	3,8		
	Mancante di sistema	15917	30,0		
	Totale	17941	33,8		
Totale		53111	100,0		

Questo studio si basa sulle relazioni tra le variabili mediante tavole di contingenza (i cosiddetti ‘incroci’) percentualizzati per riga. Tali tavole sono delle matrici quadrate di dimensione 8×8 in cui sulle righe sono riportate le classi di origine (ossia quelle dei padri) e sulle colonne le classi di arrivo (quelle degli intervistati). In particolare, durante tutto questo lavoro si usa la suddetta matrice:

Tavola di contingenza Padre * Intervistato

			Intervistato								
			borghesia	classe media impiegatizia	dipendenti agricoltura	dipendenti industria	dipendenti terziario	autonomi agricoltura	autonomi industria	autonomi terziario	Totale
Padre	borghesia	Conteggio	956	996	33	205	228	75	67	279	2839
		% entro Padre	33,7%	35,1%	1,2%	7,2%	8,0%	2,6%	2,4%	9,8%	100,0%
	classe media impiegatizia	Conteggio	756	1918	27	287	370	38	46	267	3709
		% entro Padre	20,4%	51,7%	,7%	7,7%	10,0%	1,0%	1,2%	7,2%	100,0%
	dipendenti agricoltura	Conteggio	131	400	736	946	539	222	126	329	3429
		% entro Padre	3,8%	11,7%	21,5%	27,6%	15,7%	6,5%	3,7%	9,6%	100,0%
	dipendenti industria	Conteggio	518	2024	131	2191	1309	120	253	646	7192
		% entro Padre	7,2%	28,1%	1,8%	30,5%	18,2%	1,7%	3,5%	9,0%	100,0%
	dipendenti terziario	Conteggio	244	935	49	506	719	30	63	323	2869
		% entro Padre	8,5%	32,6%	1,7%	17,6%	25,1%	1,0%	2,2%	11,3%	100,0%
	autonomi agricoltura	Conteggio	297	596	249	685	478	819	155	397	3676
		% entro Padre	8,1%	16,2%	6,8%	18,6%	13,0%	22,3%	4,2%	10,8%	100,0%
	autonomi industria	Conteggio	111	284	16	180	129	26	120	128	994
		% entro Padre	11,2%	28,6%	1,6%	18,1%	13,0%	2,6%	12,1%	12,9%	100,0%
	autonomi terziario	Conteggio	337	828	50	249	337	66	62	678	2607
		% entro Padre	12,9%	31,8%	1,9%	9,6%	12,9%	2,5%	2,4%	26,0%	100,0%
Totale		Conteggio	3350	7981	1291	5249	4109	1396	892	3047	27315
		% entro Padre	12,3%	29,2%	4,7%	19,2%	15,0%	5,1%	3,3%	11,2%	100,0%

Nel capitolo 5 mostreremo l'utilizzo di questa matrice nell'ambito matematico. Nello specifico, esamineremo il problema del calcolo della distribuzione di probabilità stazionaria e del tempo di mescolamento per la matrice appena citata.

Capitolo 3

Le catene di Markov a stati discreti

Le catene di Markov a stati discreti costituiscono uno dei modelli probabilistici più conosciuti grazie al loro vasto impiego in numerosi settori: biologia, informatica, economia, fisica. Il loro studio richiede strumenti elementari (in confronto a quelli domandati da altri modelli stocastici) e inoltre sono caratterizzate da una immediata raffigurazione intuitiva dei fenomeni aleatori coinvolti.

3.1 Definizioni e prime proprietà

In tutto il seguito supponiamo fissati uno spazio probabilizzato $(\Omega, \mathcal{A}, P)$ e un insieme finito o numerabile S.

Definizione 3.1. Una *matrice di transizione* (o *nucleo di transizione*) relativa a S è una famiglia $\mathcal{P} = (p_{i,j})_{(i,j)\in S\times S}$ di numeri reali tale che:

- $0 \leq p_{i,j} \leq 1$ per ogni $(i, j) \in S \times S$

- $\displaystyle\sum_{j\in S} p_{i,j} = 1$ per ogni $i \in S$.

Definizione 3.2. Una *catena di Markov* con spazio degli stati S e matrice di transizione $\mathcal{P}$ è un processo stocastico $X = (X_n)_{n\geq 0}$ di variabili aleatorie, definite su $(\Omega, \mathcal{A}, P)$ e a valori in S, che verificano la seguente condizione:
$P(X_0 = i_0, X_1 = i_1, ..., X_n = i_n, X_{n+1} = i_{n+1}) =$
$P(X_0 = i_0)p_{i_0,i_1} \cdot p_{i_1,i_2} \cdot ... \cdot p_{i_n,i_{n+1}}$
per ogni $i_0, ..., i_{n+1} \in S$ e ogni $n \geq 0$.

Definizione 3.3. La distribuzione di probabilità ν su S definita da

$$\nu(i) = P(X_0 = i) \text{ per ogni } i \in S$$

è detta *distribuzione iniziale di X.*
Nel caso in cui ν coincida con la legge di Dirac δ_{i_0} concentrata in $i_0 \in S$, cioè nel caso in cui $\nu(i_0) = 1$ e $\nu(i) = 0$ per ogni $i \neq i_0$, diciamo che la catena di Markov X *parte* dallo stato i_0 oppure che è *uscente* dallo stato i_0.

Dalla definizione 3.2 è chiaro che la distribuzione di X, intesa come variabile aleatoria definita su $(\Omega, \mathcal{A}, P)$ e a valori in S^∞, dipende solo dalla legge iniziale ν e dalla matrice di transizione $\mathcal{P}$. Due catene di Markov con la stessa legge iniziale e la stessa matrice di transizione sono quindi identicamente distribuite. Inoltre, dati un insieme finito o numerabile S, una misura di probabilità ν, e una matrice di transizione $\mathcal{P}$ relativa a S, è possibile costruire uno spazio probabilizzato $(\Omega, \mathcal{A}, P)$ e su di esso una successione $(X_n)_{n\geq 0}$ di variabili aleatorie che sia una catena di Markov con spazio degli stati S, legge iniziale ν e matrice di transizione $\mathcal{P}$.

Una catena di Markov gode della seguente proprietà, detta *Proprietà di Markov.*

Proposizione 3.1.1. *Sia $X = (X_n)_{n\geq 0}$ una catena di Markov con spazio degli stati S e matrice di transizione $\mathcal{P}$. Allora, per ogni $i_0, ..., i_{n+1} \in S$ con*

$P(X_0 = i_0, ..., X_n = i_n) > 0$ *si ha*

$$P(X_{n+1} = i_{n+1} | X_n = i_n, ..., X_0 = i_0) = P(X_{n+1} = i_{n+1} | X_n = i_n) = p_{i_n, i_{n+1}}. \tag{3.1}$$

In particolare vale la relazione
$P(X_n = i_n, X_{n+1} = i_{n+1}) = P(X_n = i_n) p_{i_n, i_{n+1}}$ *per ogni* $i_n, i_{n+1} \in S$.

Dimostrazione. Verifichiamo che vale la proprietà (3.1) sviluppando i primi due termini della relazione. Per quanto riguarda il primo termine, abbiamo

$$P(X_{n+1} = i_{n+1} | X_n = i_n, ..., X_0 = i_0) = \frac{P(X_0 = i_0, ..., X_n = i_n, X_{n+1} = i_{n+1})}{P(X_0 = i_0, ..., X_n = i_n)}$$

che, grazie alla definizione 3.2, coincide con $p_{i_n, i_{n+1}}$.

Per il secondo termine, valgono le seguenti uguaglianze:

$$P(X_{n+1} = i_{n+1} | X_n = i_n) = \frac{P(X_n = i_n, X_{n+1} = i_{n+1})}{P(X_n = i_n)} =$$

$$\frac{\sum_{i_0, i_1, ..., i_{n-1} \in S} P(X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i_n, X_{n+1} = i_{n+1})}{\sum_{i_0, i_1, ..., i_{n-1} \in S} P(X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i_n)} =$$

$$\frac{\sum_{i_0, ..., i_{n-1} \in S} P(X_0 = i_0) p_{i_0, i_1} \cdot ... \cdot p_{i_{n-1}, i_n} \cdot p_{i_n, i_{n+1}}}{\sum_{i_0, ..., i_n \in S} P(X_0 = i_0) p_{i_0, i_1} \cdot ... \cdot p_{i_{n-1}, i_n}} = p_{i_n, i_{n+1}}.$$

Abbiamo così dimostrato che i due termini sono entrambi uguali a $p_{i_n, i_{n+1}}$.

□

Vale anche il viceversa della proposizione precedente.

Proposizione 3.1.2. *Se $X = (X_n)_{n\geq 0}$ è una successione di variabili aleatorie a valori in S tale che, per ogni $i_0, ..., i_{n+1} \in S$ con $P(X_0 = i_0, ..., X_n = i_n) > 0$, vale*

$$P(X_{n+1} = i_{n+1} | X_n = i_n, ..., X_0 = i_0) = p_{i_n, i_{n+1}}$$

allora X è una catena di Markov con matrice di transizione $\mathcal{P} = (p_{i,j})_{(i,j)\in S\times S}$.

Dimostrazione. Dall' ipotesi ricaviamo che per ogni $i_0, ..., i_{n+1} \in S$, vale
$P(X_0 = i_0, ..., X_n = i_n, X_{n+1} = i_{n+1}) = P(X_0 = i_0, ..., X_n = i_n)p_{i_n,i_{n+1}}$
e quindi, procedendo per induzione su n, otteniamo l'uguaglianza
$P(X_0 = i_0, ..., X_n = i_n, X_{n+1} = i_{n+1}) = P(X_0 = i_0)p_{i_0,i_1} \cdot ... \cdot p_{i_{n-1},i_n} \cdot p_{i_n,i_{n+1}}$. □

La proprietà di Markov ha un importante significato, infatti dice che la probabilità di occupare la posizione i_{n+1} nel 'futuro' (l'istante $n+1$), dipende solo dalla posizione occupata nel 'presente' (l'istante n) e non dalle posizioni occupate nel 'passato' (gli istanti $0, ..., n-1$). Il numero $p_{i,j}$ rappresenta quindi la probabilità di passare nello 'stato futuro' j, subordinatamente al fatto di trovarsi nello 'stato presente' i. Più generalmente vale:

Proposizione 3.1.3. *Sia $X = (X_n)_{n\geq 0}$ una catena di Markov con matrice di transizione $\mathcal{P}$ e legge iniziale ν. Allora si ha*
$P(X_{n+k} = i_{n+k} | X_n = i_n, X_{n+1} = i_{n+1}) = P(X_{n+k} = i_{n+k} | X_{n+1} = i_{n+1})$
per ogni $k \geq 1$ e ogni $i_n, i_{n+1}, i_{n+k} \in S$ con $P(X_n = i_n, X_{n+1} = i_{n+1}) > 0$.

Dimostrazione. Proviamo la proposizione sviluppando i primi due termini dell'uguaglianza. Per quanto riguarda il primo termine, abbiamo:

$P(X_{n+k} = i_{n+k} | X_{n+1} = i_{n+1}, X_n = i_n) =$

$$\frac{P(X_{n+k}=i_{n+k},X_{n+1}=i_{n+1},X_n=i_n)}{P(X_{n+1}=i_{n+1},X_n=i_n)}=$$

$$\frac{\displaystyle\sum_{i_0,\dots,i_{n-1},\dots,i_{n+2},\dots,i_{n+k-1}\in S} P(X_{n+k}=i_{n+k},X_{n+k-1}=i_{n+k-1},...,X_{n+1}=i_{n+1},X_n=i_n,...,X_0=i_0)}{P(X_n=i_n)p_{i_n,i_{n+1}}}.$$

Sfruttando la definizione 3.2, otteniamo che l'espressione precedente è uguale a

$$\frac{P(X_n=i_n)p_{i_n,i_{n+1}}\displaystyle\sum_{i_{n+2},\dots,i_{n+k-1}\in S} p_{i_{n+1},i_{n+2}}\cdot\ldots\cdot p_{i_{n+k-1},i_{n+k}}}{P(X_n=i_n)p_{i_n,i_{n+1}}},$$

ossia

$$\sum_{i_{n+2},\dots,i_{n+k-1}\in S} p_{i_{n+1},i_{n+2}}\cdot\ldots\cdot p_{i_{n+k-1},i_{n+k}}$$

Per il secondo termine, procediamo in modo analogo e valgono quindi le seguenti uguaglianze:

$$P(X_{n+k}=i_{n+k}|X_{n+1}=i_{n+1})=\frac{P(X_{n+k}=i_{n+k},X_{n+1}=i_{n+1})}{P(X_{n+1}=i_{n+1})}=$$

$$\frac{\displaystyle\sum_{i_0,\dots,i_{n-1},i_n,i_{n+2},\dots,i_{n+k-1}\in S} P(X_{n+k}=i_{n+k},X_{n+k-1}=i_{n+k-1},...,X_{n+1}=i_{n+1},X_n=i_n,...,X_0=i_0)}{P(X_{n+1}=i_{n+1})}=$$

$$\frac{P(X_{n+1}=i_{n+1})\displaystyle\sum_{i_{n+2},\dots,i_{n+k-1}\in S} p_{i_{n+1},i_{n+2}}\cdot\ldots\cdot p_{i_{n+k-1},i_{n+k}}}{P(X_{n+1}=i_{n+1})}=$$

$$\sum_{i_{n+2},\dots,i_{n+k-1}\in S} p_{i_{n+1},i_{n+2}}\cdot\ldots\cdot p_{i_{n+k-1},i_{n+k}}.$$

Abbiamo così dimostrato che i due termini sono uguali.

□

Un altro importante risultato è dato dalle *Equazioni di Chapman-Kolmogorov*, di seguito enunciate:

Proposizione 3.1.4. *Sia* $X = (X_n)_{n\geq 0}$ *una catena di Markov con matrice di transizione* $\mathcal{P}$ *e legge iniziale* ν *. Per ogni* $i, j \in S$ *con* $P(X_n = i) > 0$*, la probabilità di passaggio dallo stato* i *allo stato* j *in* k *passi è*
$P(X_{n+k} = j|X_n = i) = p_{i,j}^{(k)}$ *dove*

$$\begin{cases} p_{i,j}^{(0)} = \delta_i(j) \\ \\ p_{i,j}^{(k)} = \sum_{h\in S} p_{i,h}\, p_{h,j}^{(k-1)} \qquad \text{per ogni } k \geq 1 \end{cases} \tag{3.2}$$

In altri termini, abbiamo $p_{i,j}^{(k)} = \left[\mathcal{P}^k\right]_{i,j}$ *dove* $\mathcal{P}^k$ *è la potenza* k*-esima della matrice* $\mathcal{P}$ *secondo il prodotto righe per colonne.*

Dimostrazione. Dimostriamo la proposizione procedendo per induzione su k.
Per $k = 0$ e $n \geq 0$, la tesi è ovvia perché abbiamo la legge di Dirac $\delta_i(j)$ concentrata in i.
Per $k = 1$ e $n \geq 0$, la tesi equivale alla relazione $P(X_{n+1} = j|X_n = i) = p_{i,j}$ che è la proprietà di Markov, provata in precedenza.
Procediamo quindi alla dimostrazione della tesi per $k > 1$ e $n \geq 0$. Supponiamo che la tesi sia vera per $k-1$ e per ogni $n \geq 0$ e la proviamo per k e ogni $n \geq 0$. Valgono le seguenti uguaglianze:

$$P(X_{n+k} = j|X_n = i) = \sum_{h\in S} P(X_{n+k} = j, X_{n+1} = h|X_n = i) =$$

$$\sum_{h\in S} P(X_{n+k} = j|X_{n+1} = h, X_n = i)P(X_{n+1} = h|X_n = i)$$

dove la seconda somma è estesa a tutti gli stati h per cui
$P(X_{n+1} = h, X_n = i) > 0$.
Osservando che $P(X_{n+1} = h, X_n = i)$ è diverso da zero se e solo se $p_{i,h}$ è

diverso da zero e usando la proposizione 3.1.3, otteniamo

$$P(X_{n+k} = j | X_n = i) = \sum_{h \in S} p_{i,h} P(X_{n+k} = j | X_{n+1} = h).$$

Per ipotesi induttiva concludiamo che

$$P(X_{n+k} = j | X_n = i) = \sum_{h \in S} p_{i,h} p_{h,j}^{(k-1)} = p_{i,j}^{(k)}.$$

Procedendo per induzione su k, dalle relazioni (3.2) ricaviamo facilmente che $\quad p_{i,j}^{(k)} = \left[\mathcal{P}^k\right]_{i,j}$.

$\square$

Da questo risultato, possiamo ricavare la distribuzione della variabile aleatoria X_n. Vale infatti la seguente proposizione.

Proposizione 3.1.5. *Sia $X = (X_n)_{n \geq 0}$ una catena di Markov con matrice di transizione $\mathcal{P}$ e legge iniziale ν. La distribuzione della variabile aleatoria X_n è la misura di probabilità $\nu\mathcal{P}^n$ su S definita da*

$$\begin{cases} \nu\mathcal{P}^0 = \nu \\ \\ \nu\mathcal{P}^n(A) = \sum_{i \in A} \sum_{j \in S} \nu(j) p_{j,i}^{(n)} \qquad \text{per ogni } A \subseteq S \end{cases}$$

Dimostrazione. Tale affermazione discende immediatamente da quella precedente osservando che, grazie alla formula delle probabilità totali, abbiamo

$$P(X_n = i) = \sum_{j \in S, P(X_0=j)>0} P(X_0 = j)P(X_n = i|X_0 = j) = \sum_{j \in S} \nu(j)p_{j,i}^{(n)}.$$

Ne segue che, per ogni sottoinsieme A di S, vale

$$P(X_n \in A) = \sum_{i \in A} \sum_{j \in S} \nu(j)p_{j,i}^{(n)}.$$

□

3.2 Proprietà forte di Markov

Sia $X = (X_n)_{n \geq 0}$ una catena di Markov con spazio degli stati S e matrice di transizione $\mathcal{P}$ e denotiamo con $\mathcal{F}_n^X$ la $\sigma-$algebra generata dalle variabili aleatorie $X_0, X_1, ..., X_n$, ossia la $\sigma-$algebra formata dalle unioni (numerabili) disgiunte di eventi della forma $(X_0 = i_0, , ..., X_n = i_n)$ con $i_0, ..., i_n$ in S. La famiglia $\mathcal{F}^X = (\mathcal{F}_n^X)_{n \geq 0}$ è detta *filtrazione naturale* di X.

Proposizione 3.2.1. *Sia $X = (X_n)_{n \geq 0}$ una catena di Markov con spazio degli stati S e matrice di transizione $\mathcal{P}$. Allora per ogni $n \geq 0$, ogni $F \in \mathcal{F}_n^X$ e ogni funzione f limitata su S abbiamo*

$$E\left[f(X_{n+1})I_F\right] = E\left[(\mathcal{P}f \circ X_n)I_F\right]$$

dove $(\mathcal{P}f)(x) = \sum_{j \in S} p_{x,j} f(j)$ *per ogni x in S.*

Se f è la funzione indicatrice I_A di un sottoinsieme A di S, si usa scrivere $P(x, A)$ al posto di $\mathcal{P}I_A(x)$ e quindi abbiamo

$$P(x, A) = \sum_{j \in S} p_{x,j} I_A(j) = \sum_{j \in A} p_{x,j}.$$

Dimostrazione. Supponiamo F della forma $(X_0 = i_0, ..., X_n = i_n)$. Allora valgono le seguenti uguaglianze:

$$E\left[f(X_{n+1})I_F\right] = \sum_{j \in S} E\left[f(X_{n+1})I_{(X_{n+1}=j)}I_F\right] =$$
$$\sum_{j \in S} f(j)P(X_0 = i_0, ..., X_n = i_n, X_{n+1} = j)$$

Per la proprietà di Markov, otteniamo che l'ultimo termine coincide con

$$\sum_{j \in S} f(j)P(X_0 = i_0, ..., X_n = i_n)p_{i_n,j} = P(X_0 = i_0, ..., X_n = i_n)(\mathcal{P}f)(i_n) =$$
$$E\left[\mathcal{P}f(X_n)I_F\right].$$

Essendo il generico elemento F di $\mathcal{F}_n^X$ unione numerabile di eventi del tipo di quello considerato sopra, è facile ottenere la tesi nel caso generale.

□

Vale anche il viceversa della proposizione precedente.

Proposizione 3.2.2. *Se $X = (X_n)_{n \geq 0}$ è una successione di variabili aleatorie a valori in S tale che*

$$E\left[f(X_{n+1})I_F\right] = E\left[(\mathcal{P}f \circ X_n)I_F\right]$$

per ogni $n \geq 0$, ogni $F \in \mathcal{F}_n^X$ e ogni funzione f limitata su S, allora X è una catena di Markov con matrice di transizione $\mathcal{P}$.

Dimostrazione. Prendiamo $F = (X_0 = i_0, ..., X_n = i_n)$ e sia f la funzione indicatrice $I_{\{i_{n+1}\}}$. Sviluppando il primo termine della relazione, abbiamo
$E\left[f(X_{n+1})I_F\right] = P(X_0 = i_0, ..., X_{n+1} = i_{n+1})$.
Per quanto riguarda il secondo termine, otteniamo
$E\left[(\mathcal{P}f \circ X_n)I_F\right] = E\left[(\mathcal{P}f)(i_n)I_F\right] = P(i_n, i_{n+1})P(F) =$

$p_{i_n,i_{n+1}} P(X_0 = i_0, ..., X_n = i_n)$.
Otteniamo così la proprietà di Markov e dunque la tesi è provata.

□

Denotiamo con $\mathcal{F}_\infty^X$ la più piccola $\sigma-$algebra contenente $\bigcup_{n\geq 0} \mathcal{F}_n^X$.

Definizione 3.4. Un *tempo di arresto discreto* relativo alla filtrazione naturale $\mathcal{F}^X = (\mathcal{F}_n^X)_{n\geq 0}$ di $X = (X_n)_{n\geq 0}$ è una variabile aleatoria discreta

$$T: \ (\Omega, \mathcal{A}, P) \longrightarrow \mathbb{N} \cup \{+\infty\}$$

tale che per ogni $k \geq 0$ l'evento $(T = k)$ appartiene alla $\sigma-$algebra $\mathcal{F}_k^X$.
Nel seguito, il simbolo $\mathcal{F}_T^X$ denoterà la $\sigma-$algebra

$$\left\{F \in \mathcal{F}_\infty^X : F \cap (T = k) \in \mathcal{F}_k^X \text{ per ogni } k \geq 0\right\}.$$

Dato un tempo di arresto discreto T, relativo alla filtrazione $\mathcal{F}^X$, è possibile definire la seguente variabile aleatoria:

$$X_T = \sum_{k\in\mathbb{N}} X_k I_{\{T=k\}} + cI_{\{T=+\infty\}}$$

dove c è un'arbitraria costante reale. È facile verificare che si tratta di una variabile aleatoria misurabile rispetto alla $\sigma-$algebra $\mathcal{F}_T^X$.
Infine, poiché, per ogni $n \geq 0$, la variabile aleatoria $T + n$ è ancora un tempo di arresto relativo alla filtrazione $\mathcal{F}^X$, è possibile sostituire nelle definizioni precedenti T con $T + n$, ottenendo le definizioni di $\mathcal{F}_{T+n}^X$ e X_{T+n}.

Con queste premesse, possiamo enunciare il seguente teorema.

Teorema 3.2.3. (Proprietà forte di Markov)
Sia $X = (X_n)_{n \geq 0}$ *una catena di Markov con spazio degli stati* S *e matrice di transizione* $\mathcal{P}$. *Sia* T *un tempo di arresto discreto relativo alla filtrazione naturale* $\mathcal{F}^X$ *di* X *e tale che* $P(T < +\infty) > 0$. *Allora sussistono i seguenti fatti:*

1. *Secondo la misura di probabilità* $P(\,\cdot\,|T < +\infty)$, *la successione* $Y = (X_{T+n})_{n \geq 0}$ *è una catena di Markov di matrice di transizione* $\mathcal{P}$.

2. *Se sull'evento* $(T < +\infty)$ *la variabile aleatoria* X_T *è costante, allora secondo la misura di probabilità* $P(\,\cdot\,|T < +\infty)$, *la successione di variabili aleatorie* Y *è indipendente da* $\mathcal{F}_T^X$.

Dimostrazione. Poniamo $H = (T < +\infty)$, $P_H = P(\,\cdot\,|H)$ e $E_H = E\,[\,\cdot\,|H]$.

1. Grazie alla proposizione 3.2.2, basta verificare che per ogni $n \geq 0$, ogni funzione f limitata su S e ogni $F \in \mathcal{F}_n^Y$ abbiamo

$$E_H\,[f(X_{T+n+1})I_F] = E_H\,[(\mathcal{P}f \circ X_{T+n})I_F]$$

ossia

$$E\,[f(X_{T+n+1})I_{F\cap H}] = E\,[(\mathcal{P}f \circ X_{T+n})I_{F\cap H}]\,.$$

Essendo $T+n$ un tempo di arresto con $(T+n < +\infty) = (T < +\infty) = H$, è sufficiente verificare l'ultima uguaglianza solo per $n = 0$. A questo proposito, osserviamo che H coincide con $\bigcup_{k\in\mathbb{N}}(T = k)$ e che per $n = 0$ l'evento F è contenuto in $\mathcal{F}_T^X$. Ne segue, grazie alla proposizione 3.2.2, che sussistono le seguenti uguaglianze:

$$E\left[f(X_{T+1})I_{F\cap H}\right] = \sum_{k\in\mathbb{N}} E\left[f(X_{T+1})I_{F\cap(T=k)}\right] = \sum_{k\in\mathbb{N}} E\left[f(X_{k+1})I_{F\cap(T=k)}\right]$$

$$= \sum_{k\in\mathbb{N}} E\left[(\mathcal{P}f)(X_k)I_{F\cap(T=k)}\right] = \sum_{k\in\mathbb{N}} E\left[(\mathcal{P}f)(X_T)I_{F\cap(T=k)}\right]$$

$$= E\left[(\mathcal{P}f)(X_T)I_{F\cap H}\right].$$

2. Sia $X_T = i$ sull'evento $H = (T < +\infty)$.
   Vogliamo dimostrare che, secondo la misura di probabilità P_H, la successione $Y = (X_{T+n})_{n\geq 0}$ è indipendente da $\mathcal{F}_T^X$, cioè che per ogni A appartenente a $\mathcal{F}_T^X$ con $P_H(A) = \dfrac{P(A\cap H)}{P(H)} > 0$, abbiamo

   $P_H(X_T = i, X_{T+1} = i_1, ..., X_{T+n} = i_n | A) =$

   $P_H(X_T = i, X_{T+1} = i_1, ..., X_{T+n} = i_n)$
   per ogni $n \geq 1$ e ogni $i_1, ..., i_n$ in S.

   A questo proposito, definiamo il tempo di arresto U nel modo seguente:

$$U = \begin{cases} T & su\ A \\ +\infty & su\ A^c. \end{cases}$$

   Allora abbiamo

$$(U < +\infty) = A \cap (T < +\infty) = A \cap H$$

   e, per il punto 1 applicato ad X e al tempo di arresto U, abbiamo che $(X_{U+n})_{n\geq 0}$ è, secondo la misura di probabilità $P_{A\cap H}$, una catena di Markov con matrice di transizione $\mathcal{P}$. Inoltre, per il fatto che U coincide con T su A e che $X_T = i$ su H, si tratta di una catena di Markov uscente dallo stato i, ossia abbiamo

$$P(X_U = i | A\cap H) = P(X_T = i | A\cap H) = 1.$$

Valgono quindi le seguenti uguaglianze:
$P_H(X_T = i, X_{T+1} = i_1, ..., X_{T+n} = i_n|A) =$
$P_{A\cap H}(X_T = i, X_{T+1} = i_1, ..., X_{T+n} = i_n) =$
$P_{A\cap H}(X_U = i, X_{U+1} = i_1, ..., X_{U+n} = i_n) =$
$p_{i,i_1} \cdot p_{i_1,i_2} \cdot ... \cdot p_{i_{n-1},i_n}$.
Per concludere, basta osservare che, per il punto 1 e per il fatto che $X_T = i$ su H, abbiamo anche
$P_H(X_T = i, X_{T+1} = i_1, ..., X_{T+n} = i_n) = P_H(X_T = i)p_{i,i_1} \cdot ... \cdot p_{i_{n-1},i_n} =$
$p_{i,i_1} \cdot ... \cdot p_{i_{n-1},i_n}$.

$\square$

3.3 Relazione di equivalenza fra stati

In questo paragrafo verranno date delle definizioni riguardanti delle relazioni che possono intercorrere tra due stati appartenenti a S.

Sia $\mathcal{P}$ una matrice di transizione relativa a S.

Definizione 3.5. Dati due stati i e j, diciamo che i *comunica con* j oppure che j *è accessibile da* i (e scriviamo $i \hookrightarrow j$), se esiste $k \geq 0$ tale che $p_{i,j}^{(k)} \neq 0$.

Osserviamo che, per definizione, abbiamo sempre $i \hookrightarrow i$ perché $p_{i,i}^{(0)} = 1 > 0$ per ogni i in S.

Definizione 3.6. Diciamo che i *e* j *comunicano tra loro* oppure che i *e* j *sono equivalenti* (e scriviamo $i \longleftrightarrow j$) se $i \hookrightarrow j$ e $j \hookrightarrow i$, ossia se esistono $k_1, k_2 \geq 0$ tali che $p_{i,j}^{(k_1)} \neq 0$ e $p_{j,i}^{(k_2)} \neq 0$.

Il termine *equivalenti* è giustificato dal fatto che la relazione '$\longleftrightarrow$' appena introdotta è una relazione di equivalenza. Infatti gode delle proprietà seguenti.

- Proprietà *riflessiva*: i comunica con se stesso sempre.
- Proprietà *simmetrica*: segue immediatamente dalla definizione.
- Proprietà *transitiva*: se i comunica con j e j comunica con s allora i comunica con s. Infatti esistono k_1, k_2, h_1, h_2 tali che $p_{i,j}^{(k_1)} \neq 0$, $p_{j,i}^{(k_2)} \neq 0$, $p_{j,s}^{(h_1)} \neq 0$ e $p_{s,j}^{(h_2)} \neq 0$.

  Ne segue che

$$p_{i,s}^{(k_1+h_1)} = \sum_{z \in S} p_{i,z}^{(k_1)} p_{z,s}^{(k_2)} \geq p_{i,j}^{(k_1)} p_{j,s}^{(h_1)} \neq 0$$

$$p_{s,i}^{(h_2+k_2)} = \sum_{z \in S} p_{s,z}^{(h_2)} p_{z,i}^{(k_2)} \geq p_{s,j}^{(h_2)} p_{j,i}^{(k_2)} \neq 0$$

La *classe di equivalenza dello stato* s è quindi

$$[s] = \{i \in S \ : \ i \longleftrightarrow s\}\,.$$

Definizione 3.7. Una matrice di transizione $\mathcal{P}$ si dice *irriducibile* se esiste una sola classe di equivalenza, ossia se tutti gli stati comunicano fra loro.

3.4 Stati transienti e stati ricorrenti

In questo paragrafo e nel seguente, verrà affrontato il problema della classificazione degli stati.

Sia $\mathcal{P}$ una matrice di transizione relativa allo spazio degli stati S e sia x un elemento di S.

Denotiamo con *r(x)* la probabilità di ritorno in x relativa a $\mathcal{P}$, ossia la probabilità che una *qualsiasi* catena di Markov con matrice di transizione $\mathcal{P}$ e uscente da x ha di ritornare in x.

Osserviamo che, se $X = (X_n)_{n\geq 0}$ è una catena di Markov con matrice di transizione $\mathcal{P}$ e uscente da x, allora abbiamo

$$r(x) = P(\bigcup_{n\geq 1}(X_n = x)) = P(inf\,\{n \geq 1\ :\ X_n = x\} < +\infty) =$$
$$\sum_{k\geq 1} P(inf\,\{n \geq 1\ :\ X_n = x\} = k) = \sum_{k\geq 1} f_x^{(k)} \qquad \text{dove}$$

$$f_x^{(k)} = \begin{cases} P(X_1 = x) & se\ k = 1 \\ P(X_1 \neq x, ..., X_{k-1} \neq x, X_k = x) & se\ k > 1 \end{cases}$$

Poiché la distribuzione di $X = (X_n)_{n\geq 0}$ è univocamente determinata dalla sua legge iniziale e dalla sua matrice di transizione $\mathcal{P}$, le probabilità $f_x^{(k)}$ dipendono solo da x e da $\mathcal{P}$ e quindi $r(x)$ dipende effettivamente solo da x e da $\mathcal{P}$ (e non dalla particolare catena di Markov con matrice di transizione $\mathcal{P}$ e uscente da x considerata).

Definizione 3.8. Dato uno stato x, se

- $r(x) < 1$, si dice che lo stato x é *transiente* oppure *transitorio* (per la matrice di transizione $\mathcal{P}$)

- $r(x) = 1$, si dice che lo stato x é *ricorrente* oppure *persistente* (per la matrice di transizione $\mathcal{P}$)

Proposizione 3.4.1. *Sia $X = (X_n)_{n\geq 0}$ una catena di Markov con matrice di transizione $\mathcal{P}$ e legge iniziale ν. Sia $x \in S$ e consideriamo la successione $(T_k)_{k\geq 1}$ degli istanti di visita a x diversi da zero, cioè*

$$\begin{cases} T_1 = inf\,\{n \geq 1 \;:\; X_n = x\} \\ T_{k+1} = inf\,\{n > T_k \;:\; X_n = x\} \qquad \forall\, k \geq 1. \end{cases}$$

Per ogni $k \geq 1$, vale l'uguaglianza

$$P(T_{k+1} < +\infty) = r(x)P(T_k < +\infty).$$

Dimostrazione. Se $P(T_k < +\infty) = 0$, allora la relazione è banalmente verificata. Infatti, essendo $(T_{k+1} < +\infty) \subseteq (T_k < +\infty)$, abbiamo
$P(T_{k+1} < +\infty) = 0$.
Se $P(T_k < +\infty) > 0$, possiamo applicare la proprietà forte di Markov. Ponendo $H = (T_k < +\infty)$, abbiamo che, secondo la misura di probabilità P_H, la successione $(X_{T_k+n})_{n\geq 0}$ è una catena di Markov con matrice di transizione $\mathcal{P}$. Inoltre, per definizione di T_k, si tratta di una catena di Markov uscente da x. Otteniamo quindi

$$P_H(T_{k+1} < +\infty) = P_H(\,\bigcup_{n\geq 1}(X_{T_k+n} = x)\,) = r(x),$$

da cui, essendo $(T_{k+1} < +\infty) \subseteq H$, ricaviamo

$$P(T_{k+1} < +\infty) = r(x)P(T_k < +\infty).$$

□

Corollario 3.4.2. *Sia $X = (X_n)_{n\geq 0}$ una catena di Markov con matrice di transizione $\mathcal{P}$ e legge iniziale ν. Denotiamo con V_x^X il numero totale di visite* *che X fa allo stato x a partire dall'istante 1, ossia poniamo*

$$V_x^X = \sum_{n \geq 1} I_{\{X_n = x\}}.$$

Allora abbiamo

$$\sum_{n \geq 1} P(X_n = x) = E\left[V_x^X\right] = P(T_1 < +\infty)(\sum_{h=0}^{\infty} r(x)^h).$$

Dimostrazione. Per la linearità della speranza, abbiamo

$$E\left[V_x^X\right] = \sum_{n \geq 1} E\left[I_{\{X_n = x\}}\right] = \sum_{n \geq 1} P(X_n = x).$$

Inoltre abbiamo

$$E\left[V_x^X\right] = \sum_{k=1}^{\infty} P(V_x^X \geq k) = \sum_{k=1}^{\infty} P(T_k < +\infty).$$

Grazie alla proposizione precedente, otteniamo

$$E\left[V_x^X\right] = P(T_1 < +\infty) + \sum_{k=2}^{\infty} P(T_{k-1} < +\infty) r(x) = P(T_1 < +\infty) \sum_{h=0}^{\infty} (r(x))^h.$$

$\square$

Sia $y \in S$. Se X esce da y, ossia se $\nu = \delta_y$, allora abbiamo

$$E\left[V_x^X\right] = \sum_{n \geq 1} P(X_n = x) = \sum_{n \geq 1} P(X_n = x | X_0 = y) = \sum_{n \geq 1} p_{y,x}^{(n)}.$$

In particolare, se $y = x$, allora abbiamo $E\left[V_x^X\right] = \sum_{n \geq 1} p_{x,x}^{(n)}$.

Dai risultati precedenti, discendono le seguenti proposizioni.

Proposizione 3.4.3. *Data una catena di Markov $X = (X_n)_{n\geq 0}$ con matrice di transizione $\mathcal{P}$ e legge iniziale ν e fissato x in S, sono possibili solo i due seguenti casi:*

1. *x è transiente e V_x^X è integrabile (quindi $P(V_x^X < +\infty) = 1$).*

2. *x è ricorrente e V_x^X è quasi certamente a valori in $\{0, +\infty\}$, ossia abbiamo*

$$P(V_x^X = +\infty) = 1 - P(V_x^X = 0).$$

In particolare, nel secondo caso, se $\nu = \delta_x$, allora $P(V_x^X = +\infty) = 1$.

Dimostrazione. Dal corollario precedente ricaviamo che, se $r(x) < 1$, allora $E\left[V_x^X\right] < +\infty$, cioè V_x^X è integrabile e dunque $P(V_x^X < +\infty) = 1$.

Se invece lo stato x è ricorrente, grazie alla proposizione 3.4.1 abbiamo $P(T_k < +\infty) = P(T_1 < +\infty)$ per ogni $k \geq 1$ e quindi

$$P(V_x^X = +\infty) = P(\bigcap_{k\geq 1}(T_k < +\infty)\ ) = \lim_{k\to\infty} P(T_k = +\infty)$$

$$= P(T_1 < +\infty) = 1 - P(V_x^X = 0).$$

In particolare, se $\nu = \delta_x$, otteniamo $P(T_1 < +\infty) = r(x) = 1$ e quindi $P(V_x^X = +\infty) = 1$.

□

Proposizione 3.4.4. *Data una matrice di transizione $\mathcal{P}$, sono possibili solo i due seguenti casi:*

1. *x è transiente e $\sum_{n\geq 1} p_{y,x}^{(n)} < +\infty$ per ogni y in S.*

2. *x è ricorrente e*

$$\sum_{n\geq 1} p_{y,x}^{(n)} = \begin{cases} +\infty & \text{se } y \text{ comunica con } x \\ 0 & \text{altrimenti} \end{cases}$$

Dimostrazione. Sia y un elemento di S e $X = (X_n)_{n\geq 0}$ una catena di Markov con matrice di transizione $\mathcal{P}$ e legge iniziale $\nu = \delta_y$. Grazie a quanto osservato precedentemente, abbiamo

$$\sum_{n\geq 1} p_{y,x}^{(n)} = E\left[V_x^X\right].$$

Dalla Proposizione 3.4.3, discende quindi immediatamente che, se x è transiente, allora necessariamente si deve avere

$$\sum_{n\geq 1} p_{y,x}^{(n)} < +\infty.$$

Inoltre, se x è ricorrente, si deve avere

$$\sum_{n\geq 1} p_{x,x}^{(n)} = +\infty.$$

Per definizione, se y non comunica con x, allora $\sum_{n\geq 1} p_{y,x}^{(n)} = 0$.

Infine, se x è ricorrente e se y è uno stato diverso da x che comunica con x, allora, per definizione, esiste $n \geq 1$ tale che $p_{y,x}^{(n)} > 0$ e quindi, dalla proposizione precedente, otteniamo

$$E\left[V_x^X\right] = \sum_{n\geq 1} p_{y,x}^{(n)} > 0.$$

Ne segue che $P(V_x^X > 0) > 0$ e quindi, per la proposizione precedente, abbiamo $P(V_x^X = +\infty) > 0$, da cui $\sum_{n\geq 1} p_{y,x}^{(n)} = E\left[V_x^X\right] = +\infty$.

□

Dalle proposizioni precedenti si deducono immediatamente i seguenti corollari.

Corollario 3.4.5. *Lo stato x è ricorrente se e solo se si ha* $\sum_n p_{x,x}^{(n)} = +\infty$.

Corollario 3.4.6. *Per una catena di Markov $X = (X_n)_{n\geq 0}$ con spazio degli stati S, per ogni x in S, vale l'implicazione*

$$P(V_x^X < +\infty) = 1 \Longrightarrow E\left[V_x^X\right] < +\infty.$$

Dimostrazione. Sono possibili solo due casi: x è transiente oppure x è ricorrente con $P(V_x^X = 0) = 1 - P(V_x^X = +\infty) = 1$. In entrambi i casi otteniamo $E\left[V_x^X\right] < +\infty$.

□

Corollario 3.4.7. *Se lo spazio degli stati S è finito, allora deve esistere sicuramente almeno uno stato ricorrente per $\mathcal{P}$.*

Dimostrazione. Supponiamo per assurdo che ogni x appartenente a S sia transiente, allora abbiamo $P(V_x^X < +\infty) = 1$ per ogni $x \in S$. Ne segue che
$P(X_n \in S \; solo \; per \; un \; numero \; finito \; di \; indici \; n) =$
$P(\bigcap_{x\in S}(V_x^X < +\infty)) = 1$
Ma questo è assurdo!

□

Teorema 3.4.8. *Sia x uno stato ricorrente per $\mathcal{P}$ e sia y uno stato tale che x comunica con y. Allora abbiamo che lo stato y è ricorrente.*

Dimostrazione. Per ipotesi x comunica con y e quindi per definizione esiste $k \geq 0$ tale che $p_{x,y}^{(k)} > 0$. Inoltre, per quanto visto nel punto precedente, abbiamo che anche y comunica con x, ossia esiste $h \geq 0$ per cui $p_{y,x}^{(h)} > 0$. Dalla relazione

$$p_{y,y}^{(k+n+h)} = \sum_{z \in S} \sum_{i \in S} p_{y,z}^{(k)} \cdot p_{z,i}^{(n)} \cdot p_{i,y}^{(h)} \geq p_{y,x}^{(k)} \cdot p_{x,x}^{(n)} \cdot p_{x,y}^{(h)},$$

sommando su n, troviamo

$$\sum_n p_{y,y}^{(n)} \geq \sum_n p_{y,y}^{(k+n+h)} \ \geq \ p_{y,x}^{(k)} \ (\sum_n p_{x,x}^{(n)}) \ p_{x,y}^{(h)} \ = \ +\infty.$$

Ne segue che lo stato y è ricorrente.

$\square$

Da teorema appena enunciato, discendono i seguenti utili corollari.

Corollario 3.4.9. *Se x è ricorrente allora ogni stato equivalente a x è ricorrente. In particolare, se la matrice di transizione $\mathcal{P}$ è irriducibile, sono possibili solo i seguenti casi:*

- *$\mathcal{P}$ è irriducibile ricorrente (ossia tutti gli stati sono ricorrenti)*
- *$\mathcal{P}$ è irriducibile transiente (ossia tutti gli stati sono transienti).*

Se, per giunta, $\mathcal{P}$ è irriducibile con spazio degli stati S finito, allora $\mathcal{P}$ è sicuramente irriducibile ricorrente.

Corollario 3.4.10. *Se x comunica con y ma y non comunica con x, allora x è sicuramente transiente.*

Di seguito enunciamo un risultato che ci permetterà di dimostrare il teorema sulla convergenza all'equilibrio.

Teorema 3.4.11. *Sia $\mathcal{P}$ una matrice di transizione irriducibile e ricorrente. Allora per ogni $j \in S$ abbiamo che $P(T_j < +\infty) = 1$.*

Capitolo 4

Distribuzione di probabilità stazionaria e tempo di mescolamento

In questo capitolo tratteremo il tema della distribuzione di probabilità stazionaria e del tempo di mescolamento per una catena di Markov $X = (X_n)_{n\geq 0}$ a stati discreti S e con matrice di transizione $\mathcal{P} = (p_{i,j})_{(i,j)\in S\times S}$ relativa ad S.

4.1 Definizioni e proprietà

In questa sezione ci concentriamo sulle catene di Markov irriducibili, ossia quei processi in cui tutti gli stati della matrice di transizione comunicano tra di loro.
Elenchiamo una serie di definizioni importanti che ci serviranno per enunciare il teorema sulla convergenza all'equilibrio.

Definizione 4.1. Diciamo che lo stato i ha *periodo* d se $p_{i,i}^n = 0$ dove d è il più piccolo intero tale che n non è divisibile per d.

Uno stato con periodo 1 è detto *aperiodico*.

Può essere mostrato che la periodicità è una proprietà delle classi equivalenza. Cioè, se lo stato i ha periodo d, e gli stati i e j comunicano allora anche lo stato j ha periodo d.

Definizione 4.2. Se lo stato i è ricorrente, allora i è detto *positivo ricorrente* se, partendo da i, il tempo atteso fino a quando il processo ritorna allo stato i è finito.

Anche in questo caso, si può mostrare che la ricorrenza positiva è una proprietà delle classi di equivalenza. In particolare, in una catena di Markov a stati finiti tutti gli stati ricorrenti sono ricorrenti positivi.

Definizione 4.3. Gli stati ricorrenti positivi e aperiodici sono chiamati *ergodici*.

Diamo ora la definizione di probabilità invariante.

Definizione 4.4. Una misura di probabilità π si dice *stazionaria* o *invariante* se

$$\pi\mathcal{P} = \pi.$$

4.2 Convergenza all'equilibrio

Di seguito riportiamo una proposizione che ci servirà per dimostrare il teorema della convergenza all'equilibrio.

Proposizione 4.2.1. *Sia $\mathcal{P}$ una matrice di transizione irriducibile. Allora sono equivalenti le seguenti affermazioni:*

- *tutti gli stati sono positivi ricorrenti;*
- *qualche stato i è positivo ricorrente;*
- *$\mathcal{P}$ ha una distribuzione di probabilità stazionaria π.*

Vale il seguente teorema.

Teorema 4.2.2. *Sia S un insieme finito. Supponiamo che per tutti gli stati i di S si ha*

$$p_{i,j}^{(n)} \to \pi_j \quad per\ n \to +\infty \quad \forall j \in S.$$

Allora $\pi = (\pi_j : j \in S)$ è una distribuzione stazionaria.

Dimostrazione. Per come è stato costruito π_j abbiamo:

$$\sum_{j \in S} \pi_j = \sum_{j \in S} \lim_{n \to +\infty} p_{i,j}^{(n)} = \lim_{n \to +\infty} \sum_{j \in S} p_{i,j}^{(n)} = 1$$

e inoltre:

$$\pi_j = \lim_{n \to +\infty} p_{i,j}^{(n)} = \lim_{n \to +\infty} \sum_{k \in S} p_{i,k}^{(n)} p_{k,j} = \sum_{k \in S} \lim_{n \to +\infty} p_{i,k}^{(n)} p_{k,j} = \sum_{k \in S} \pi_k p_{k,j}$$

dove abbiamo usato il fatto che S è finito per giustificare lo scambio delle operazioni di somma e limite. Allora π è una distribuzione stazionaria.

□

Con queste premesse possiamo enunciare il seguente risultato.

Teorema 4.2.3. *(Convergenza all'equilibrio) Sia $\mathcal{P}$ una matrice di transizione irriducibile e aperiodica, e supponiamo che $\mathcal{P}$ abbia una distribuzione di probabilità stazionaria π. Sia λ un'altra qualsiasi distribuzione. Supponiamo che $X = (X_n)_{n\geq 0}$ è una catena di Markov con matrice di transizione $\mathcal{P}$ e legge iniziale λ. Allora vale*

$$P(X_n = j) \to \pi_j \quad \textit{per } n \to +\infty \quad \forall j.$$

In particolare abbiamo che

$$p_{i,j}^{(n)} \to \pi_j \quad \textit{per } n \to +\infty \quad \forall i, j.$$

Dimostrazione. Sia $(Y_n)_{n\geq 0}$ una catena di Markov con matrice di transizione $\mathcal{P}$ e distribuzione di probabilità π e supponiamo che sia indipendente da $(X_n)_{n\geq 0}$. Fissiamo uno stato di riferimento b e un insieme

$$T = \inf \{n \geq 1 : X_n = Y_n = b\}.$$

`I passo.` Dimostriamo che $P(T < +\infty) = 1$.
Il processo $W_n = (X_n, Y_n)$ è una catena di Markov sullo spazio degli stati $S \times S$ con matrice di transizione

$$\tilde{\mathcal{P}}_{(i,k)(j,l)} = p_{i,j} p_{k,l}$$

e distribuzione iniziale

$$\mu_{(i,k)} = \lambda_i \pi_k.$$

Se $\mathcal{P}$ è una matrice di transizione aperiodica, per tutti gli stati i, j, k, l si ha

$$\tilde{\mathcal{P}}^{(n)}_{(i,k)(j,l)} = p^{(n)}_{i,j} p^{(n)}_{k,l} > 0$$

per tutti gli n sufficientemente grandi; allora risulta che anche $\tilde{\mathcal{P}}$ è irriducibile. Inoltre, $\tilde{\mathcal{P}}$ ha una distribuzione di probabilità stazionaria data da

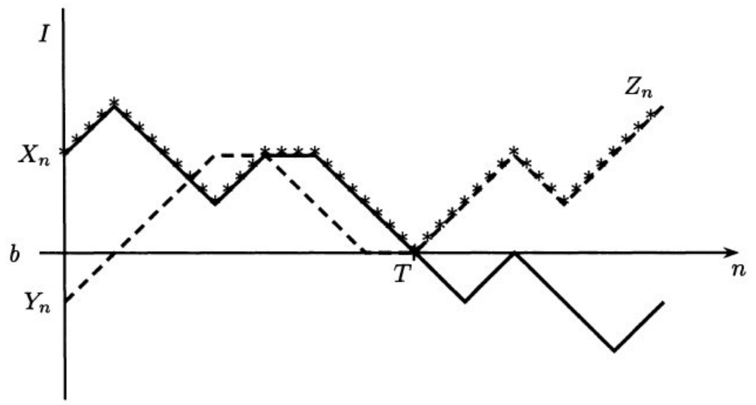
$$\tilde{\pi}_{(i,k)} = \pi_i \pi_k$$

e dalla proposizione 4.2.1 abbiamo che $\tilde{\mathcal{P}}$ è ricorrente positiva. Ma T è il primo tempo di passaggio in W_n in (b, b), allora dal teorema 3.4.11 segue che $P(T < +\infty) = 1$.

`II passo.` Consideriamo

$$Z_n = \begin{cases} X_n & se \;\; n < T \\ Y_n & se \;\; n \geq T. \end{cases}$$

Il diagramma seguente illustra l'idea. Dimostriamo che $(Z_n)_{n\geq 0}$ è una catena di Markov con matrice di transizione $\mathcal{P}$ e legge λ.



Viene applicata la proprietà forte di Markov a $(W_n)_{n\geq 0}$ al tempo T, in modo da avere che $(X_{T+n}, Y_{T+n})_{n\geq 0}$ è una catena di Markov con matrice di transizione $\tilde{\mathcal{P}}$ e legge $\delta_{(b,b)}$ ed è indipendente da $(X_0, Y_0), (X_1, Y_1), ..., (X_T, Y_T)$.

Dalla proprietà di simmetria, possiamo sostituire il processo $(X_{T+n}, Y_{T+n})_{n\geq 0}$ con $(Y_{T+n}, X_{T+n})_{n\geq 0}$ che è ancora una catena di Markov con matrice di transizione $\tilde{\mathcal{P}}$ e legge $\delta_{(b,b)}$ e rimane indipendente da $(X_0, Y_0), (X_1, Y_1), ..., (X_T, Y_T)$. Allora il processo $W'_n = (Z_n, Z'_n)$ è una catena di Markov con matrice di transizione $\tilde{\mathcal{P}}$ e legge μ in cui

$$Z'_n = \begin{cases} Y_n & se \;\; n < T \\ X_n & se \;\; n \geq T. \end{cases}$$

In particolare, $(Z_n)_{n\geq 0}$ è una catena di Markov avente matrice di transizione $\mathcal{P}$ e legge λ.

`III passo.` Vale la seguente uguaglianza

$$P(Z_n = j) = P(X_n = j \;\; e \;\; n < T) + P(Y_n = j \;\; e \;\; n \geq T).$$

A questo punto possiamo scrivere che

$$|P(X_n = j) - \pi_j| = |P(Z_n = j) - P(Y_n = j)| =$$

$$|P(X_n = j \;\; e \;\; n < T) - P(Y_n = j \;\; e \;\; n < T)| \leq P(n < T)$$

ma $P(n < T) \to 0$ per $n \to +\infty$ e questo dimostra l'asserto.

$\square$

4.3 Tempo di mescolamento

In questa parte ci occuperemo del tempo di mescolamento ossia il tempo che impiega una catena di Markov a raggiungere la distribuzione di probabilità stazionaria.
La convergenza alla distribuzione di equilibrio di una catena di Markov tende a trovarsi brutalmente in modo asintotico quando un certo parametro tende all'infinito, che di solito è la cardinalità dello spazio degli stati della catena. Se questo accade, il tempo in cui si è verificata la convergenza è detto *tempo*

di mescolamento.

Per cominciare, abbiamo bisogno di un modo per misurare quanto siamo lontani dalla distribuzione di probabilità stazionaria. Per fare questo useremo il concetto di distanza in variazione totale tra due misure di probabilità sullo spazio degli stati della catena di Markov.

Siano μ e ν due misure di probabilità sullo spazio degli stati S per una data catena di Markov $X = (X_n)_{n\geq 0}$.

Definizione 4.5. La *distanza in variazione totale* tra le misure di probabilità μ e ν è data da:

$$d_{TV}(\mu,\nu) = \|\mu-\nu\| = \sup_{A\subset S} |\mu(A)-\nu(A)|\,.$$

Così definita la distanza in variazione totale misura l'errore massimo che si fa quando si approssima μ a ν per predire la probabilità di un dato evento. Abbiamo il seguente risultato.

Proposizione 4.3.1. *Se lo spazio degli stati S è un insieme discreto abbiamo le seguenti identità:*

$$\|\mu-\nu\| = \frac{1}{2}\sum_{s\in S}(\mu(s)-\nu(s))^+ =$$

$$\frac{1}{2}\sum_{s\in S}|\mu(s)-\nu(s)| =$$

$$\frac{1}{2}\max_{\|f\|_\infty\leq 1}|E_\mu(f)-E_\nu(f)|\,.$$

Come conseguenza della terza identità, notiamo che vale

$$0 \leq d_{TV}(\mu, \nu) \leq 1.$$

Cioè, il valore massimo che la distanza in variazione totale può assumere è 1.

Definizione 4.6. Sia $X = (X_n)_{n \geq 0}$ una catena di Markov irriducibile e aperiodica con matrice di transizione $\mathcal{P}$ sullo spazio degli stati finito S. Sia π_x la distribuzione di probabilità stazionaria di X, definita da $\sum_x \pi_x p_{x,y} = \pi_y$. Definiamo la *funzione distanza* per tutti gli $n \geq 0$ in questo modo:

$$d(n) = \max_{x \in S} \left\| p^{(n)}_{x,\cdot} - \pi_\cdot \right\| .$$

La teoria classica delle catene di Markov ci dice che $d(n)$ tende a zero quando n tende all'infinito. Infatti, il Teorema di Perron Frobenius afferma che, asintoticamente per n che tende all'infinito, la distanza $d(n)$ decresce velocemente in modo esponenziale, con un tasso di decrescita dato dallo spazio spettrale della catena.

Proposizione 4.3.2. *Sia λ il massimo autovalore della matrice di transizione $\mathcal{P}$ che è strettamente minore di* 1*. Allora esiste una costante C tale che:*

$$d(n) \approx C\lambda^n, \quad n \to +\infty.$$

A priori questo sembra che ci dice tutto ciò che vogliamo riguardo i tempi di mescolamento. Invece, per rendere piccolo λ^n è sufficiente considerare n

più grande di $-\frac{1}{\log \lambda}$.
Se $d(n)$ converge a zero quando n tende all'infinito, ha sempre senso definire, per $0 < \epsilon < \frac{1}{2}$:

$$\tau(\epsilon) = \inf \{n \geq 0 : d(n) \leq \epsilon\}.$$

Definizione 4.7. Si definisce il *tempo di mescolamento* nel modo seguente

$$t_{mix} = \inf \left\{ n \geq 0 : d(n) \leq \frac{1}{e} \right\}$$

dove la costante $\frac{1}{e}$ è arbitraria.

Proposizione 4.3.3. *Valgono le seguenti proprietà:*

1. *$d(n)$ è non crescente rispetto al tempo;*

2. *sia ρ definita da*

$$\rho(n) = \max_{x,y \in S} \left\| p_{x,\cdot}^{(n)} - p_{y,\cdot}^{(n)} \right\|.$$

   Allora

$$d(n) \leq \rho(n) \leq 2d(n);$$

3. *ρ è submoltiplicativo, cioè per ogni $k, n \geq 0$ vale che*

$$\rho(n+k) \leq \rho(n)\rho(k).$$

Dimostrazione. Dimostriamo un punto alla volta.

1. Per $k \leq n$ e x, y, z in S sono vere le seguenti uguaglianze:

$$p_{x,z}^{(n)} = \sum_{y \in S} p_{x,y}^{(k)} p_{y,z}^{(n-k)}$$

$$\pi_z = \sum_{y \in S} \pi_y p_{y,z}^{(n-k)}.$$

   Quindi possiamo scrivere:

$$p_{x,z}^{(n)} - \pi_z = \sum_{y \in S} p_{y,z}^{(n-k)} (p_{x,y}^{(k)} - \pi_y).$$

Da cui:

$$\left\|p^{(n)}_{x,\cdot} - \pi_{\cdot}\right\| = \frac{1}{2}\sum_{z\in S}\left|p^{(n)}_{x,z} - \pi_z\right| =$$

$$\frac{1}{2}\sum_{z\in S}\left|\sum_{y\in S} p^{(n-k)}_{y,z}(p^{(k)}_{x,y} - \pi_y)\right| \leq$$

$$\frac{1}{2}\sum_{z\in S}\sum_{y\in S} p^{(n-k)}_{y,z}\left|p^{(k)}_{x,y} - \pi_y\right| =$$

$$\frac{1}{2}\sum_{y\in S}\left|p^{(k)}_{x,y} - \pi_y\right| = \left\|p^{(k)}_{x,\cdot} - \pi_{\cdot}\right\|.$$

Se consideriamo il massimo su x in S, otteniamo:

$$d(n) = \max_{x\in S}\left\|p^{(n)}_{x,\cdot} - \pi_{\cdot}\right\| \leq \max_{x\in S}\left\|p^{(k)}_{x,\cdot} - \pi_{\cdot}\right\| = d(k) \ \ \forall n \geq k.$$

2. La parte a destra della disuguaglianza è semplicemente la disuguaglianza triangolare. Per quanto riguarda la parte a sinistra, osserviamo che dalla proprietà di stazionarietà, se $A \subset S$, vale:

$$\pi_A = \sum_{y\in S}\pi_y p^{(n)}_{x,A}.$$

Allora, grazie alla disuguaglianza triangolare, abbiamo:

$$\left\|\pi - p^{(n)}_{x,\cdot}\right\| = \max_{A\subset S}\left|p^{(n)}_{x,A} - \pi_A\right| =$$

$$\max_{A\subset S}\left|\sum_{y\in S}\pi_y\left[p^{(n)}_{x,A} - p^{(n)}_{y,A}\right]\right| \leq$$

$$\max_{A\subset S}\sum_{y\in S}\pi_y\left|p^{(n)}_{x,A} - p^{(n)}_{y,A}\right| \leq$$

$$\rho(n)\sum_{y\in S}\pi_y = \rho(n).$$

3. Sia (X_k, Y_k) una coppia di variabili aleatorie ottenuta accoppiando $p^{(n)}_{x,\cdot}$ con $p^{(n)}_{y,\cdot}$. Allora

$$\left\|p^{(k)}_{x,\cdot} - p^{(k)}_{y,\cdot}\right\| = P(X_k \neq Y_k).$$

Dalle variabili X_k e Y_k possiamo costruire X_{k+n} e Y_{k+n} in modo da formare un particolare accoppiamento di $p_{x,\cdot}^{(k+n)}$ con $p_{y,\cdot}^{(k+n)}$. Chiediamo che sull'evento $A_z = \{X_k = Y_k = z\}$ valga:

$$X_{k+n} = Y_{k+n} \approx p_{z,\cdot}^{(n)}$$

e che se $A_{z,z'} = \{X_k = z, Y_k = z'\}$ con $z \neq z'$, allora

$$X_{k+n} \approx p_{z,\cdot}^{(k+n)} \;\; e \;\; Y_{k+n} \approx p_{z',\cdot}^{(k+n)}.$$

La scelta di X_{k+n} e Y_{k+n} è tale che, sull'evento $A_{z,z'}$, X_{k+n} e Y_{k+n} è una coppia di variabili aleatorie ottenute mettendo insieme $p_{z,\cdot}^{(n)}$ con $p_{z',\cdot}^{(n)}$: allora

$$P(X_{k+n} = Y_{k+n} | A_{z,z'}) = \rho_{z,z'}(n) \leq \rho(n).$$

Con queste definizioni abbiamo:

$$\rho(k+n) \leq P(X_{k+n} \neq Y_{k+n}) =$$

$$P(X_k \neq Y_k) P(X_{k+n} \neq Y_{k+n} | X_k \neq Y_k) =$$

$$\rho(k) P(X_{k+n} \neq Y_{k+n} | X_k \neq Y_k).$$

Sia $\mu_{(x,y)}$ la legge di (X_k, Y_k) dato $X_k \neq Y_k$. Dalla proprietà di Markov al tempo k, otteniamo l'asserto:

$$\rho(k+n) \leq \rho(k) \sum_{z \neq z'} \mu_{(z,z')} P(X_{k+n} \neq Y_{k+n} | A_{z,z'}) \leq$$

$$\rho(k)\rho(n) \sum_{z \neq z'} \mu_{(z,z')} = \rho(k)\rho(n).$$

$\square$

Capitolo 5

I risultati

Il presente capitolo racchiude i risultati di questa ricerca, la parte fondamentale di tutta la tesi.
Il nostro scopo è quello di considerare la mobilità sociale come un processo stocastico, o meglio come una catena di Markov $X = (X_n)_{n \geq 0}$ intesa come la distribuzione dei lavori nelle varie classi sociali per la generazione n-esima.

Consideriamo come spazio degli stati S il numero della suddivisione in classi sociali che abbiamo espresso nel capitolo 2, quindi si ha che
$S = \{1, 2, ..., 8\}$.

Per ogni elemento i e j che variano su S si può denotare con n_i il numero degli intervistati che hanno i padri appartenenti alla classe i e con $n_{i,j}$ il numero degli intervistati appartenenti alla classe j e che hanno i padri nella classe i. Risulta che

$$\sum_{j=1}^{8} n_{i,j} = n_i \quad \forall i \in S. \tag{5.1}$$

Per ogni i e j in S possiamo definire la seguente probabilità

$$p_{i,j} = \frac{n_{i,j}}{n_i}$$

e ovviamente vale che $0 \leq p_{i,j} \leq 1$. Inoltre, per la (5.1) si ha che

$$\sum_{j=1}^{8} p_{i,j} = 1.$$

In particolare la matrice $\mathcal{P} = (p_{i,j})_{(i,j)\in S\times S}$ è una matrice di transizione relativa a S della catena di Markov X.

Nel capitolo 2 abbiamo parlato di come siamo riusciti a suddividere le varie attività lavorative in classi sociali e inoltre abbiamo riportato la tavola di contingenza ottenuta mediante l'incrocio tra le informazioni riguardanti gli intervistati e i rispettivi padri.

Tavola di contingenza Padre * Intervistato

			Intervistato								
			borghesia	classe media impiegatizia	dipendenti agricoltura	dipendenti industria	dipendenti terziario	autonomi agricoltura	autonomi industria	autonomi terziario	Totale
Padre	borghesia	Conteggio	956	996	33	205	228	75	67	279	2839
		% entro Padre	33,7%	35,1%	1,2%	7,2%	8,0%	2,6%	2,4%	9,8%	100,0%
	classe media impiegatizia	Conteggio	756	1918	27	287	370	38	46	267	3709
		% entro Padre	20,4%	51,7%	,7%	7,7%	10,0%	1,0%	1,2%	7,2%	100,0%
	dipendenti agricoltura	Conteggio	131	400	736	946	539	222	126	329	3429
		% entro Padre	3,8%	11,7%	21,5%	27,6%	15,7%	6,5%	3,7%	9,6%	100,0%
	dipendenti industria	Conteggio	518	2024	131	2191	1309	120	253	646	7192
		% entro Padre	7,2%	28,1%	1,8%	30,5%	18,2%	1,7%	3,5%	9,0%	100,0%
	dipendenti terziario	Conteggio	244	935	49	506	719	30	63	323	2869
		% entro Padre	8,5%	32,6%	1,7%	17,6%	25,1%	1,0%	2,2%	11,3%	100,0%
	autonomi agricoltura	Conteggio	297	596	249	685	478	819	155	397	3676
		% entro Padre	8,1%	16,2%	6,8%	18,6%	13,0%	22,3%	4,2%	10,8%	100,0%
	autonomi industria	Conteggio	111	284	16	180	129	26	120	128	994
		% entro Padre	11,2%	28,6%	1,6%	18,1%	13,0%	2,6%	12,1%	12,9%	100,0%
	autonomi terziario	Conteggio	337	828	50	249	337	66	62	678	2607
		% entro Padre	12,9%	31,8%	1,9%	9,6%	12,9%	2,5%	2,4%	26,0%	100,0%
Totale		Conteggio	3350	7981	1291	5249	4109	1396	892	3047	27315
		% entro Padre	12,3%	29,2%	4,7%	19,2%	15,0%	5,1%	3,3%	11,2%	100,0%

Questa tavola di contingenza può essere interpretata come una matrice di transizione avente per elementi le percentuali normalizzate per riga. Si tratta di una matrice irriducibile perché ogni stato può essere raggiunto a partire da ogni altro stato ed è aperiodica.
Essendo S un insieme finito e grazie ai risultati mostrati nel capitolo 4, possiamo affermare che esiste un'unica distribuzione stazionaria tale che

$$\lim_{n\to+\infty} p_{i,j}^{(n)} = \pi_j \quad \forall j \in S.$$

Vale a dire che siamo in grado di calcolare la distribuzione limite, ossia la situazione futura della mobilità sociale per questa particolare tavola di contingenza, e in seguito il tempo necessario per ottenere l'equilibrio.

5.1 Convergenza all'equilibrio

Consideriamo la matrice di transizione avente per elementi le percentuali normalizzate per riga della tavola di contingenza mostrata in precedenza. Risulta:

$$\mathcal{P} = \begin{pmatrix} 0.337 & 0.351 & 0.012 & 0.072 & 0.08 & 0.026 & 0.024 & 0.098 \\ 0.204 & 0.517 & 0.007 & 0.077 & 0.1 & 0.01 & 0.013 & 0.072 \\ 0.038 & 0.117 & 0.214 & 0.276 & 0.157 & 0.065 & 0.037 & 0.096 \\ 0.072 & 0.281 & 0.018 & 0.305 & 0.182 & 0.017 & 0.035 & 0.09 \\ 0.085 & 0.326 & 0.017 & 0.176 & 0.251 & 0.01 & 0.022 & 0.113 \\ 0.081 & 0.162 & 0.068 & 0.186 & 0.13 & 0.223 & 0.042 & 0.108 \\ 0.111 & 0.286 & 0.016 & 0.181 & 0.13 & 0.026 & 0.121 & 0.129 \\ 0.129 & 0.318 & 0.019 & 0.096 & 0.129 & 0.025 & 0.024 & 0.26 \end{pmatrix}$$

Come abbiamo visto, la matrice $\mathcal{P}$ è una matrice di transizione perché soddisfa le seguenti condizioni:

- $0 \leq p_{i,j} \leq 1$ per ogni $(i, j) \in S \times S$

- $\displaystyle\sum_{j=1}^{8} p_{i,j} = 1$ per ogni $i \in S$.

Possiamo calcolare la distribuzione stazionaria in due modi.

`[I metodo.]` In generale vale che le matrici di transizione del tipo $\mathcal{P}$ hanno un autovalore $\lambda = 1$. Quando ci si trova nel nostro caso in cui si ha a che fare con una catena irriducibile (per cui esiste un intero k relativamente al quale tutti gli stati comunicano tra loro), abbiamo che tale autovalore è il massimo in modulo. Troviamo che, per ogni distribuzione iniziale, esiste un'unica distribuzione stazionaria ed è l'autovettore corrispondente all'autovalore $\lambda = 1$. La distribuzione stazionaria π è quindi la soluzione di:

$$\pi = \mathcal{P}\pi$$

con la condizione di normalizzazione:

$$\sum_{i=1}^{8} \pi_i = 1 \quad \forall i \in S.$$

Possiamo calcolare in Matlab l'autovettore grazie alla funzione *eigs*.

`[II metodo.]` Calcoliamo la distribuzione stazionaria π considerando che vale la condizione

$$\pi = \mathcal{P}\pi$$

se e solo se

$$(I - \mathcal{P})\pi = 0$$

è un sistema omogeneo con una soluzione non nulla (la distribuzione stazionaria), ossia con un'equazione ridondante. Nella formula il termine I denota la matrice identica di dimensione 8×8. Per calcolare tale soluzione sostituiamo ad un'equazione del sistema la condizione di normalizzazione che sappiamo essere $\displaystyle\sum_{i=1}^{8} \pi_i = 1 \quad \forall i \in S$.

In questo caso abbiamo usato in Matlab il metodo di eliminazione di Gauss che è un metodo diretto per la risoluzione dei sistemi lineari. Essendo un metodo diretto la soluzione viene calcolata in un numero finito di passi modificando la matrice del problema in modo da rendere più agevole il calcolo della soluzione.

La distribuzione stazionaria è la seguente:

$$\pi = \begin{pmatrix} 0.1795 \\ 0.3904 \\ 0.0169 \\ 0.1288 \\ 0.1324 \\ 0.0212 \\ 0.0238 \\ 0.1071 \end{pmatrix}$$

e possiamo esprimerla in termini di percentuali secondo le varie classi sociali in questo modo:

Borghesia	17.95
Classe media impiegatizia	39.04
Dipendenti agricoltura	1.69
Dipendenti industria	12.88
Dipendenti terziario	13.24
Autonomi agricoltura	2.12
Autonomi industria	2.38
Autonomi terziario	10.71

Si può notare che la classe media impiegatizia è la più numerosa mentre i dipendenti dell'agricoltura tendono a diminuire, come pure gli autonomi dello stesso settore. I rami dell'industria e del terziario hanno un leggero calo mentre la borghesia tende ad aumentare di numero.
Questo fenomeno potrebbe essere una conseguenza del fatto che il grado di istruzione si è alzato e la maggior parte delle persone preferiscono continuare gli studi andando all'università per ricoprire un ruolo di spicco piuttosto che entrare subito nel mondo del lavoro.

5.2 Tempo di mescolamento

Consideriamo la matrice di transizione $\mathcal{P} = (p_{i,j})_{(i,j)\in S\times S}$ relativa a S della catena di Markov X e la sua distribuzione stazionaria π appena calcolata.

Si vuole misurare il tempo $\tilde{n}$ necessario affinchè, per ogni j in S, $P(X_{\tilde{n}} = j)$ sia vicina a $\pi(j)$.
Sia ϵ fissato, generalmente si considera $\epsilon = e^{-1}$. Si definisce $\tilde{n}$ come il più piccolo n per cui vale la relazione

$$d(p^{(n)}_{i,\cdot}, \pi) \leq \epsilon \quad \forall i \in S.$$

In questo caso si è utilizzata la distanza nello spazio L^1 che risulta essere:

$$d_1(p^{(n)}_{i,\cdot}, \pi) = \sum_{j=1}^{8} |p^{(n)}_{i,j} - \pi_j| \quad \forall i \in S.$$

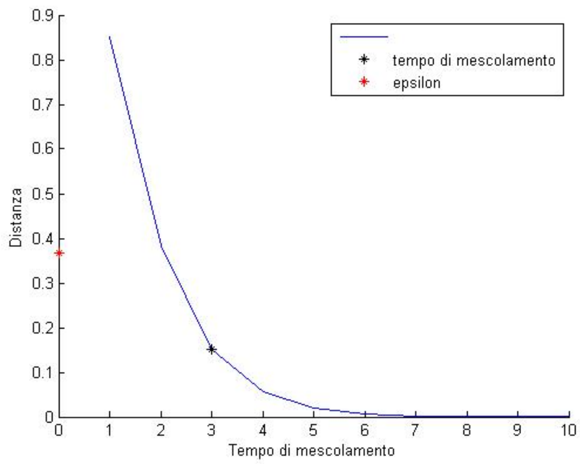
Possiamo calcolare il tempo di mescolamento costruendo un algoritmo in Matlab e imponendo che tale tempo $\tilde{n}$ sia il primo n per cui si verifica che

$$\max_{i\in S} \left\{ \sum_{j=1}^{8} |p^{(n)}_{i,j} - \pi_j| \right\} \leq \epsilon$$

con ϵ fissato e pari a e^{-1}.

Ne risulta che il tempo di mescolamento per la catena di Markov $X = (X_n)_{n\geq 0}$ con matrice di transizione $\mathcal{P} = (p_{i,j})_{(i,j)\in S\times S}$ e distribuzione stazionaria π è uguale a 3.

Possiamo rappresentare l'andamento della distanza $d(p_{i,\cdot}^{(n)}, \pi)$ che decresce all'aumentare delle iterazioni.



Abbiamo così dimostrato una importante proprietà del tempo di mescolamento, cioè che la funzione distanza è non crescente rispetto al tempo (punto numero 1 della Proposizione 4.3.3).

Bibliografia

[1] A. Bagnasco, M. Barbagli, A. Cavalli, *Sociologia II. Differenziazione e riproduzione sociale*, 1997, Il Mulino.

[2] M. Barbagli, V. Capecchi, A. Cobalti, *La mobilità sociale in Emilia Romagna*, 1988, Il Mulino.

[3] N. Berestycki, *Eight lectures on Mixing Times*, 2009.

[4] F. Biagini, M. Campanino, *Elementi di probabilità e statistica*, 2006, Springer.

[5] A. Cobalti, A. Schizzerotto, *La mobilità sociale in Italia*, 1994, Il Mulino.

[6] P. Contucci, S. Isola, *Probabilità elementare*, 2008, Zanichelli.

[7] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1, 1957, John Wiley and Sons, Inc.

[8] J. R. Norris, *Markov Chains*, 1997, Cambridge University Press.

[9] M. Pisati, *La mobilità sociale*, 2000, Il Mulino.

[10] S. M. Ross, *Introduction to Probability Models*, 2003, Academic press.

Ringraziamenti

Un grande ringraziamento va al Professore Pierluigi Contucci per l'aiuto, il sostegno e la tanta pazienza che ha avuto durante questo lavoro e per i preziosi insegnamenti. Il suo continuo incoraggiamento mi ha spronato soprattutto nei momenti in cui pensavo di non farcela e di voler abbandonare tutto.
Un ringraziamento sentito spetta al Professore Marzio Barbagli per avermi introdotta nel tema della mobilità sociale e per il tempo e l'attenzione che mi ha dedicato.
Ringrazio la Dottoressa Alessandra Bianchi per l'interesse e la disponibilità e il Professore Asher Daniel Colombo per avermi aiutata con il software SPSS.

Un grazie enorme va alla mia mamma e alla mia sorellina che hanno sempre creduto in me e mi hanno sostenuta nei momenti difficili di questo cammino. Ringrazio la mia vera famiglia: i miei nonni materni per l'aiuto e perché se sono qui è anche merito loro, i miei zii materni e i miei cugini.
Un grazie particolare ai miei amici di avventura matematica: Gianluca per il grande sostegno, le componenti dell'integrale triplo Lucia e Virginia per l'aiuto durante gli esami, nei momenti di sconforto oltre che per le serate e Chiara che mi ha aiutata molto in quest'ultimo anno.
Ringrazio la mia coinquilina Mary per avermi supportato e sopportato, per i tanti pranzi, cene e dolci di casa Lembo-Garzia e soprattutto per avermi fatto tornare il sorriso grazie ai suoi stupendi balletti artistici.
Vorrei ringraziare i miei migliori amici Stefania e Ivan per avermi aiutata

nell'ultimo periodo e infine tutti gli altri amici vicini e lontani, vecchi e nuovi...

Un grazie 'insolito' va a coloro che non hanno creduto in me, che pensavano fossi una fallita e che mi hanno voltato le spalle.