

Report n. 132

**Proiezioni demografiche multiregionali
a due sessi, con immigrazione internazionale
e vincoli di consistenza**

Alessandro VALENTINI

Proiezioni demografiche multiregionali a due sessi, con immigrazione internazionale e vincoli di consistenza*

Alessandro Valentini
Dipartimento di Statistica e Matematica Applicata all'Economia
Facoltà di Economia. Università di Pisa
Via Ridolfi 10 – 56125 Pisa

Abstract

Il presente rapporto costituisce la parte preliminare di una ricerca avente ad oggetto l'impatto delle immigrazioni internazionali in una popolazione a bassa fecondità. In questa sede vengono infatti presentati gli strumenti di base necessari per ottenere realistiche previsioni demografiche con immigrazioni: un modello a due sessi con algoritmi di consistenza e un'indicazione in merito ai dati necessari per il funzionamento dello stesso (con particolare riferimento alla popolazione italiana).

1. Introduzione

Come indicato nell'abstract, il presente lavoro consiste nella presentazione di un sistema di previsioni demografiche basato su un realistico modello a due sessi con immigrazioni internazionali, di cui la componente immigratoria dall'estero costituisce un fattore rilevante.

Il modello presentato è predisposto specificamente per studiare l'evoluzione della popolazione italiana: il Paese con la più bassa fecondità del mondo che al momento sta cercando (forse inconsapevolmente) di contrastare la dinamica naturale negativa mediante il ricorso a massicci flussi immigratori.

Ciò non toglie che il modello stesso, vista la sua valenza generale, possa essere applicato in contesti territoriali diversi, specialmente in altri paesi industrializzati.

Il modello, oltre a porre una particolare enfasi sulla componente immigratoria, presenta contestualmente previsioni coerenti per i due sessi, ottenendo risultati senz'altro più realistici rispetto ai più evoluti modelli a dominanza femminile, quali quelli di Rogers (1995) o Willekens et al. (1984).

Il fatto che la consistenza tra componente maschile e componente femminile sia garantita in contesti territoriali più ristretti di quello nazionale, rende il sistema di previsioni in oggetto più sofisticato anche rispetto ai più realistici modelli a due sessi, specificamente progettati per popolazioni uni regionali, vedi per es. Keilman (1985), Long (1992) e Van Imhoff (1992). L'elevato numero di dimensioni implica l'utilizzo di una notevole mole di dati, non sempre disponibili per tutti i paesi.

* Il Presente lavoro fa parte di un progetto di ricerca avente ad oggetto l'impatto delle immigrazioni internazionali in una popolazione a bassa fecondità (con particolare riferimento all'evoluzione della popolazione italiana) svolta dall'autore presso il Dipartimento di Statistica e Matematica Applicata della Facoltà di Economia di Pisa in collaborazione e sotto la supervisione del Dottor Piero Manfredi. Tale studio è finanziato da una borsa di studio di Ateneo.

Nel corso del lavoro il modello in oggetto è presentato dettagliatamente in tutte le sue componenti, il che ne rende possibile l'implementazione diretta mediante un qualsiasi linguaggio di programmazione.

Tuttavia per le applicazioni è specificamente previsto l'utilizzo del programma LIPRO - Lifestyle PROjections, Van Imhoff and Keilman (1992) - nella versione 3.0 (Van Imhoff, 1994) -. Il ricorso al LIPRO da un lato elimina i problemi che consistono nella scrittura diretta del programma, ma dall'altro costituisce una limitazione per il "ridotto" numero di dimensioni possibili (60 nella versione attuale) che nel nostro caso impedisce la formulazione di un sistema con più di 5 regioni.

2. La definizione del problema

Come già accennato nel Paragrafo 1, l'Italia è al momento il Paese a più bassa fecondità del mondo e probabilmente quello che ha raggiunto il numero minore di figli per donna mai registrato nella storia dell'umanità, almeno per una popolazione di larghe dimensioni, Golini (1997).

Il paese è contestualmente sottoposto a notevoli flussi immigratori, provenienti in massima parte dai paesi in via di sviluppo, che nel breve periodo riescono (in qualche modo) a contrastare con il trend naturale negativo.

Le caratteristiche ora esaminate rendono di particolare interesse lo studio dell'impatto demografico delle immigrazioni all'interno della nostra popolazione, che da questo punto di vista costituisce una sorta di laboratorio pratico in cui studiare se ed in quale misura le immigrazioni sono in grado di contrastare con il declino e l'invecchiamento causati dalla bassa fecondità.

In due lavori preliminari, Valentini (1997) e Manfredi e Valentini (1997), è stata esaminata l'evoluzione della popolazione femminile italiana utilizzando un modello uniregionale. Rogers (1995), derivanti dalla mancata segmentazione della popolazione per macro-regione, ci hanno spinto a ripetere le proiezioni utilizzando un sistema multiregionale, Valentini (1998). I risultati ottenuti, pur significativi, sono però privi di valenza pratica per il fatto che non considerano l'altra componente della popolazione, quella maschile¹.

Ciò rende necessaria una ulteriore estensione del modello ai due sessi, in modo tale che sia garantita la coerenza tra numero di coniugati maschili e numero di coniugate femminili.

Come noto dalla letteratura, vedi per es. Shoen (1981), Keilman (1985) e Long (1992), la coerenza tra maschi e femmine viene realizzata esplicitando lo stato civile di ogni soggetto, ovvero distinguendo tra single e coniugati. Dato che nel nostro modello è prevista anche la differenziazione tra immigrati e nativi, la distinzione per stato civile deve essere mantenuta anche per questi gruppi. Ne conseguono diverse tipologie di coniugati: maschi nativi con femmine native, maschi nativi con femmine immigrate, maschi immigrati con femmine native ed infine maschi immigrati con femmine immigrate.

La fine del matrimonio, per ognuno dei due coniugi, può realizzarsi per decesso ma anche per altre due cause: il divorzio e la vedovanza. Divorziati e vedovi ritornano in gioco, nel senso che fanno parte della popolazione sottoposta al rischio di contrarre (nuovo) matrimonio. Per tale ragione dovrebbero costituire due gruppi a sé per ciascun sesso e distintamente per immigrati e nativi. Ciò comporterebbe nel concreto un aumento di 8 ($=2^3$) dimensioni per ogni regione, rendendo il sistema di fatto ingestibile.

¹ Valentini (1999) ha esteso i risultati delle proiezioni uniregionali ad un sistema a due sessi, utilizzando però la metodologia della dominanza femminile.

Allora si è ritenuto opportuno considerare i divorziati ed i vedovi come “single”, termine che nel nostro contesto deve essere riferito ai soggetti “sottoposti al rischio di contrarre matrimonio”, divorziati e vedovi compresi. Ne segue che i tassi di matrimonialità non faranno quindi riferimento al primo matrimonio, ma al generico evento del matrimonio (indipendentemente dal fatto che, per un determinato soggetto, sia il primo, od uno successivo).

Nonostante l’evidenziazione di quattro status di matrimonialità per ogni regione, il modello prevede che possano nascere figli anche al di fuori del matrimonio, la c.d. fecondità illegittima. Tale fenomeno, molto accentuato in alcuni Paesi del nord Europa, è in crescita in Italia, pur restando sempre piuttosto raro.

Le donne sottoposte al rischio di nascite illegittime non sottostanno ad alcun algoritmo di consistenza, dato che possono aver rapporti con qualsiasi partner (coniugato o non coniugato, nativo o immigrato).

3. Caratteristiche generali del modello

Abbiamo visto nel paragrafo precedente che il modello in oggetto è multiregionale, ma che presenta anche caratteristiche multistato in quanto la popolazione all’interno di ogni regione è distinta per sesso, per status di cittadinanza (nativi e immigrati) e per stato civile (single, coppie di nativi, coppie di immigrati, coppie di cui un coniuge sia nativo e l’altro immigrato).

Combinando i diversi casi si ottengono 12 gruppi (o strati o stadi, S) per ogni regione:

- Nativi, Single, Maschi (Gruppo S1)
- Nativi, Coniugati con native, Maschi (Gruppo S2)
- Nativi, Coniugati con non native, Maschi (Gruppo S3)
- Native, Single, Femmine (Gruppo S4)
- Native, Coniugate con nativi, Femmine (Gruppo S5)
- Native, Coniugate con non nativi, Femmine (Gruppo S6)
- Immigrati, Single, Maschi (Gruppo S7)
- Immigrati, Coniugati con native, Maschi (Gruppo S8)
- Immigrati, Coniugati con immigrate, Maschi (Gruppo S9)
- Immigrate, single, Femmine (Gruppo S10)
- Immigrate, Coniugate con nativi, Femmine (Gruppo S11)
- Immigrate, Coniugate con immigrati, Femmine (Gruppo S12)

Notare che il termine “nativo” viene utilizzato come sinonimo di “cittadino”, ovvero di soggetto con la cittadinanza italiana (anche se nato all’estero, come nel caso della naturalizzazione di immigrati).

Il termine “non nativo” è invece utilizzato come sinonimo di “immigrato”, o “non cittadino”. Vedremo successivamente che i figli di immigrati nati in Italia sono a tutti gli effetti immigrati, e possono essere naturalizzati solo in seguito (su loro richiesta).

Per quanto riguarda specificamente i cittadini coniugati, notare che esistono 4 possibili tipologie di gruppi di sposi, che potremo definire come “gruppi di matrimoniabilità”. Essi sono:

- marito e moglie entrambe nativi (Gruppo M1) (S2 = S5)
- marito nativo, moglie immigrata (Gruppo M2) (S3 = S11)
- marito immigrato, moglie nativa (Gruppo M3) (S8 = S6)
- marito e moglie entrambe immigrati (Gruppo M4) (S9 = S12)

e verranno esaminati più nel dettaglio nel corso della trattazione. Nel seguito faremo l'ipotesi che per ciascuna regione, nell'ambito di ogni gruppo, il numero di coniugati maschi coincida esattamente con il numero di coniugate femmine. Questo implica che le coppie di sposi siano residenti nella stessa regione.

I 12 stock per ogni regione vengono alimentati o decrementati dai seguenti eventi, che costituiscono i passaggi di stadio (con riferimento ad entrambe i sessi e sia ai nativi che agli immigrati, salvo i casi in cui sia diversamente specificato):

- nascite da donne coniugate (che danno luogo a figli "nativi" se almeno uno dei genitori è nativo ed "immigrati" se entrambe i genitori sono immigrati) o anche da single
- matrimoni secondo uno dei vari stati di matrimonialità
- decessi
- vedovanza (di coniugati)
- divorzio (di coniugati)
- emigrazione da una regione all'altra
- emigrazione all'estero
- immigrazione dall'estero
- naturalizzazione di immigrati (sia single che coniugati)

La naturalizzazione di un immigrato, acquisita a seguito della permanenza continuativa per più anni all'interno del territorio nazionale, fa acquistare al soggetto lo status di nativo. Se egli è coniugato l'evento naturalizzazione colpisce anche il coniuge, perché la coppia cambia la tipologia.

Riferendoci allo schema del Paragrafo 2, consideriamo per esempio il caso di un maschio immigrato coniugato con una nativa. Prima della naturalizzazione il suo status è S8, quello della moglie è S6. Dopo la naturalizzazione il soggetto diventa nativo (S2) e la moglie immigrata con nativo (e non più coniugata con immigrato) spostandosi quindi da S6 a S5.

Il sistema dei possibili cambiamenti di stato è piuttosto complesso, quindi per semplicità in Tabella 1 viene riportata una matrice di sintesi. In essa sono indicati, per una determinata regione, tutti i possibili "spostamenti" che si verificano nell'unità temporale di riferimento per la popolazione di una qualsiasi fascia per età.

E' prevista una matrice simile a quella della Tabella 1 per ogni classe di età (con classi quinquennali ne servono quindi 18, ammettendo che la classe aperta sia 85+).

La matrice per ogni classe per età nel sistema uniregionale della Tabella 1 ha dimensione [12, 12], dato che esistono 12 tipi di stato (chiaramente in un sistema con n regioni essa avrebbe dimensione [12*n, 12*n]).

Prima di procedere al commento è utile ricordare che:

S1, S2, S3	sono	Nativi	Maschi
S4, S5, S6	sono	Native	Femmine
S7, S8, S9	sono	Immigrati	Maschi
S10, S11, S12	sono	Immigrate	Femmine

Tabella 1. Matrice dei possibili passaggi di stato (all'interno di una regione)

AIDA	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
S1	X	X	X				X	X	X			
S2	X	X	X				X	X	X			
S3	X		X				X		X			
S4				X	X	X				X	X	X
S5				X	X	X				X	X	X
S6				X		X				X		X
S7							X	X	X			
S8							X	X	X			
S9							X		X			
S10										X	X	X
S11										X	X	X
S12										X		X

Dall'esame della matrice di relazioni si evince che da un quinquennio all'altro è sempre possibile il mantenimento dello stato di origine, mentre ci si può spostare solo su alcuni stati di destinazione.

Esaminiamo i passaggi di stato ammessi caso per caso:

- Da S1: passaggio da nativi single a nativi coniugati con moglie nativa o immigrata
- Da S2: passaggio da nativi coniugati con native a single (per vedovanza e divorzio)
- Da S3: passaggio da nativi coniugati con immigrate a single (per vedovanza e divorzio) od a nativi coniugati con native (per naturalizzazione della moglie)
- Da S4: passaggio da native single a native coniugate con marito nativo o immigrato
- Da S5: passaggio da native coniugate con nativi a single (per vedovanza e divorzio)
- Da S6: passaggio da native coniugate con immigrati a single (per vedovanza e divorzio) od a native coniugate con nativi (per naturalizzazione del marito)
- Da S7: passaggio da immigrati single a nativi single (per naturalizzazione) o a nativi coniugati con native o con immigrate (per naturalizzazione e matrimonio) od a immigrati coniugati con native o con immigrate (per solo matrimonio)
- Da S8: passaggio da immigrati coniugati con native a nativi single (per naturalizzazione e contestuale vedovanza o divorzio) o a immigrati single (per vedovanza o divorzio)
- Da S9: passaggio da immigrati coniugati con immigrate a nativi single (per naturalizzazione e contestuale vedovanza o divorzio) o coniugati con immigrate (per naturalizzazione) o coniugati con native (per naturalizzazione di entrambe i coniugi). E' anche possibile il passaggio ad immigrati single (per vedovanza o divorzio) o coniugati con native (per naturalizzazione del coniuge)
- Da S10: passaggio da immigrate single a native single (per naturalizzazione) o coniugate con nativi o con immigrati (per naturalizzazione e matrimonio). E' anche possibile il passaggio a immigrate coniugate con nativi o con immigrati (per matrimonio)
- Da S11: passaggio da immigrate coniugate con nativi a native single (per naturalizzazione e contestuale vedovanza o divorzio) o coniugate (per naturalizzazione). E' anche possibile il passaggio a immigrate single (per vedovanza o divorzio)
- Da S12: passaggio da immigrate coniugate con immigrati a native single (per naturalizzazione e contestuale vedovanza o divorzio) o coniugate con nativi (per naturalizzazione di entrambe i coniugi) o coniugate con immigrati (per

naturalizzazione). E' anche possibile il passaggio ad immigrate single (per vedovanza o divorzio) o coniugate con nativi (per naturalizzazione del coniuge)

Chiaramente la popolazione che raggiunge lo stato di destinazione (ovvero che rimane in quello di partenza) subisce (anche) una trasformazione di tipo diverso: invecchia per il numero di anni corrispondenti all'unità temporale di riferimento (nel nostro caso di 5 anni).

Così dicendo ci rendiamo anche conto che esistono due stati di destinazione non indicati direttamente dalla matrice della Tabella 1 in quanto comportano l'uscita definitiva dal sistema multiregionale – multistato dei soggetti che ne sono colpiti. Si tratta del decesso e dell'emigrazione all'estero (anch'essa fenomeno "definitivo" in quanto nel sistema non resta memoria del soggetto emigrato).

Per quanto riguarda la classe per età 0-4 bisogna aggiungere agli eventi della Tabella 1 le nascite, come vedremo nel Paragrafo 6.3.

Nonostante la sua complessità, la matrice di transizione non è in grado di rappresentare, da sola, l'evoluzione temporale della popolazione. Infatti non riesce a spiegare il fenomeno delle immigrazioni dall'estero, evento esogeno e quindi inserito nel sistema di proiezione mediante l'aggiunta di uno specifico vettore.

Tralasciando momentaneamente le migrazioni, ricordiamo anche che la matrice dei passaggi di stato della Tabella 1 deve essere estesa ad un sistema multiregionale, ove per "passaggio" va inteso più in generale lo spostamento dallo stato k in regione i al tempo t allo stato s in regione j a $t+5$ (ove k ed s e i e j possono coincidere o essere contemporaneamente o alternativamente diversi).

Ciò significa in altri termini che saranno possibili tutti i casi di cui alla matrice della Tabella 1, riferiti però ad ognuna delle n regioni del sistema.

Per esempio si consideri la categoria S1 (Single, Nativi, Maschi) della regione i :

- E' possibile il mantenimento dello status di maschio single nativo all'interno della regione i
- E' possibile il mantenimento dello status di maschio single nativo in altra regione (per effetto dell'emigrazione)
- E' possibile lo spostamento verso lo status di coniugato con nativa o con immigrata all'interno della stessa regione (per effetto del matrimonio)
- E' possibile lo spostamento verso lo status di coniugato con nativa o con immigrata verso un'altra regione (per effetto dell'emigrazione e del contestuale matrimonio)

Come già più volte rilevato in precedenza, le immigrazioni dall'estero costituiscono una componente molto rilevante dell'evoluzione della popolazione. Dato che, per ogni classe per età, la popolazione di ogni regione è ripartita in 12 strati, una disaggregazione simile deve essere utilizzata anche per il vettore dei flussi immigratori per unità di tempo, pur ricordando che non è previsto l'ingresso in tutti gli strati che in precedenza abbiamo definito come S1 - S12. Escludendo infatti per ovvie ragioni le categorie da S1 a S6 (tutte relative ai nativi), e vista l'assunzione che le coppie di sposi devono essere contemporaneamente residenti in una certa regione, sarà ammesso solo l'ingresso di immigrati single (categorie S7 e S10) e di coppie di sposi (categorie S9 e S12). Per questi ultimi due gruppi è necessario imporre il vincolo, come vedremo, che il numero di ingressi maschili coniugati corrisponda esattamente al numero di ingressi femminili coniugati ($I9=I12$).

Si consideri in questa sede il problema pratico degli ingressi di immigrati coniugati senza il coniuge a seguito (fenomeno molto frequente in Italia, specie con riferimento alla componente maschile della popolazione), ed il successivo ricongiungimento familiare.

Entrambe gli aspetti vengono previsti, seppure indirettamente, dal modello. Prima di tutto l'ingresso di un immigrato (o immigrata fa lo stesso) coniugato senza il coniuge a seguito è rilevato come ingresso di single perché tale soggetto, pur nel suo status di potenziale genitore, non è in grado di "produrre" figli. Il successivo arrivo del coniuge è ancora rilevato come ingresso di single, ma a questo punto la coppia è in grado di partecipare a tutti gli effetti al processo riproduttivo. Per ovviare a questo "gap", possono essere considerate probabilità di matrimonio tra immigrati ed immigrate single più alte rispetto ai valori che è ragionevole attendersi in base a prudenziali stime.

4. Vincoli di consistenza del modello

Per l'unità temporale di riferimento i diversi eventi dipendono dagli stock e dai tassi di passaggio da uno stadio all'altro, altrimenti esprimibili sotto forma di probabilità.

In base all'assunzione che le coppie di coniugati vivano nella stessa regione, è necessario che lo stock di matrimoni maschili (per ogni stato di matrimonialità) coincida esattamente con quello di matrimoni femminili. Ciò si realizza a monte imponendo la condizione che, per ogni intervallo temporale, il numero di nuovi matrimoni maschili (per ogni status coniugale) sia uguale al numero di nuovi matrimoni femminili e che il numero di matrimoni maschili disciolti (per decesso, divorzio, vedovanza o emigrazione) corrisponda al numero di matrimoni femminili disciolti.

Nessuna correzione è prevista per i matrimoni già in essere al tempo di inizio della proiezione, in quanto si tratta di dati esogeni. Una eventuale differenza tra componente maschile e componente femminile può essere ascrivibile o ad errori nella rilevazione dei dati od al fatto che vi sono coppie in cui i coniugi risiedono in regioni diverse (i due punti non sono però ammessi nel nostro modello).

Vediamo ora nel dettaglio i vincoli relativi ai flussi che condizionano il numero complessivo di matrimoni nell'unità di tempo. A tal proposito per una qualsivoglia regione si indichino con $+M1_m$, $+M2_m$, $+M3_m$ e $+M4_m$ i nuovi matrimoni maschili rispettivamente negli status di matrimonialità 1, 2, 3 e 4 (vedi paragrafo 1) che derivano da single già residenti all'interno della popolazione al tempo t , si definiscano come $cI4_m$ gli ingressi di coniugati maschili dall'estero nel periodo t , $t+5$. Si definiscano inoltre $-M1_m$ $-M2_m$ $-M3_m$ $-M4_m$ gli scioglimenti di matrimoni maschili nei 4 status di coniugalità, ed al solito si estenda la stessa notazione utilizzata per i maschi anche alle femmine (sostituendo evidentemente l'apice m con f).

I vincoli di consistenza per la regione in oggetto diventano allora, schematicamente:

Tabella 2: Vincoli di consistenza per ogni regione

1)	$-M1_m$	=	$-M1_f$
2)	$-M2_m$	=	$-M2_f$
3)	$-M3_m$	=	$-M3_f$
4)	$-M4_m$	=	$-M4_f$
5)	$+M1_m$	=	$+M1_f$
6)	$+M2_m$	=	$+M2_f$
7)	$+M3_m$	=	$+M3_f$
8)	$+M4_m + cI4_m$	=	$+M4_f + cI4_f$

Notare che la relazione di consistenza è semplificata per gli scioglimenti (casi da 1 a 4) e per i matrimoni misti (in cui un coniuge è nativo e l'altro è immigrato, casi da 5 a 7). E' invece necessario commentare la numero 8), che riguarda i matrimoni tra immigrati. Essa deve essere ripartito in due parti:

$$8a) \quad cI4_m = cI4_f$$

$$8b) \quad +M4_m = +M4_f$$

Il vincolo 8a) riguarda il fatto che è ammesso l'ingresso di immigrati coniugati purché si tratti di coppie. Esso si realizza agendo direttamente sui dati esogeni delle immigrazioni, imponendo lo stato di single all'eccesso di ingressi coniugati di un sesso (generalmente maschile).

In altri termini, ritornando alla solita notazione per i 12 strati di popolazione indicati per esempio nella tabella 1 ed antepoendo la sigla I per immigrazioni (ove per es $I9 = cI4_m$ e $I12 = cI4_f$), si ha che:

$$\text{se } I9 > I12 \text{ allora } I9 = I12 \text{ e } I7 = I7 + (I9 - I12)$$

$$\text{se } I12 > I9 \text{ allora } I12 = I9 \text{ e } I10 = I10 + (I12 - I9)$$

Il vincolo 8b) invece ripropone semplicemente il fatto che le nuove coppie di immigrati (con coniuge immigrato) maschili debbano corrispondere alle nuove coppie dello stesso tipo femminili, e si risolve nel modo consueto (casi da 1 a 7). Il ricongiungimento familiare verrà trattato alla stregua di (nuovo) matrimonio tra immigrati. A tale scopo il modello prevede specificamente una probabilità per un single immigrato di sposarsi con partner dello stesso sesso (anche nella stessa unità temporale di ingresso) più alta rispetto alla situazione normale.

Gli algoritmi di consistenza per le relazioni da 1) a 7) e per la 8b) si basano sul principio della c.d. media armonica, ripreso per esempio in Shoen (1981), Van Imhoff and Keilman (1992) e in Keilman and Ekamper (1993). Dopo la correzione degli eventi, i tassi di passaggio da uno stadio ad un altro vengono modificati rispetto ai valori originali in modo tale da riproporre, a parità dei volumi, il numero consistente sia per i maschi che per le femmine.

I dettagli tecnici relativi alle correzioni da apportare al sistema di previsione sono riportati estensivamente nel Paragrafo 6.

5. I dati richiesti dal modello

Il modello in oggetto necessita di una grande quantità di dati per poter essere applicato correttamente ad un qualsiasi Paese caratterizzato da peculiarità regionali. Per non cadere in errori di interpretazione, esaminiamo ora gli input necessari per utilizzare il modello con riferimento alla popolazione italiana (anche se le considerazioni che faremo hanno validità generale).

L'ottica multiregionale di indagine può essere quella di considerare un'aggregazione territoriale con cinque macroregioni ($n=5$): Nord Est, Nord Ovest, Centro, Meridione e Isole o, in mancanza di dati, una più ampia con tre ($n=3$): Nord, Centro e Sud (ove il Nord congloba il Nord Est ed il Nord Ovest ed il Sud il Meridione e le Isole).

Notare che l'implementazione del modello mediante il Lipro impedisce di considerare una ripartizione geografica più ampia di 5 regioni. Infatti il programma prevede la possibilità di utilizzare al massimo 60 strati, ma ogni regione del nostro modello ne necessita 12.

5.1. Gli eventi

Dato che abbiamo utilizzato l'approccio periodo-coorte, per la generica regione i , per ogni sesso e distintamente per immigrati e nativi, nonché per classi di età quinquennali, è necessario disporre dei seguenti dati:

- Tassi di mortalità per status di cittadinanza, stato civile (single e coniugati per stato di matrimonialità), sesso e regione
- Tassi di matrimonialità che trasformino i single nelle varie regioni j al tempo t in sposati sopravvissuti nella regione i al tempo $t+5$ (in ciascuno dei quattro stati di matrimonialità). Considerare che non si tratta necessariamente di primi matrimoni, ma anche di matrimoni di divorziati/e e di vedovi/e.
- Tassi di divorzialità e di vedovanza che trasformino gli sposati (in ciascuno dei quattro stadi di matrimonialità) al tempo t in una qualsiasi regione j in *single* sopravvissuti al tempo $t+5$ nella regione i .
- Tassi di emigrazione all'estero per i single e coniugati secondo i quattro stati di matrimonialità al tempo t nella regione i
- Tassi di emigrazione che trasformino i single nelle varie regioni j al tempo t in single sopravvissuti nella regione i al tempo $t+5$
- Tassi di emigrazione che trasformino i coniugati nelle varie regioni j al tempo t in coniugati sopravvissuti nella regione i al tempo $t+5$ (secondo il precedente stadio di matrimonialità)
- Tassi di naturalizzazione di immigrati dall'estero single (e di figli nati in Italia da coppie coniugate) che dalla generica regione j al tempo t vivono come nativi (cittadini italiani) in i al tempo $t+5$
- Tassi di naturalizzazione di emigrati dall'estero coniugati che dalla generica regione j al tempo t (nello stadio di matrimonialità di origine) vivono nella regione i al tempo $t+5$ come nativi (cittadini italiani) nello stato di matrimonialità di destinazione
- Numero di immigrati dall'estero nel periodo $t, t+5$ per regione di destinazione e sesso, distinti in single e coniugati
- Tassi di fecondità del numero medio di donne in età feconda (per status matrimoniale) in ciascuna delle n regioni
- Tassi di fecondità illegittima per regione
- Popolazione residente al 01/01/1996 (o data più recente) per sesso, regione, stato civile (single, coniugati per stato di matrimonialità) e status immigratorio

Se non è possibile disporre direttamente dei tassi di cui sopra, essi possono essere stimati mediante la conoscenza del livello della popolazione disaggregata per sesso, status di cittadinanza, stato civile a due successivi istanti temporali, t e $t+k$ (ponendo $k=5$) e dei seguenti eventi, accaduti nell'intervallo temporale $t, t+k$, separatamente per ciascun sesso (per classi di età quinquennali) e sia per gli immigrati che per i nativi

- Numero di decessi per status di cittadinanza, sesso e stato civile (single e coniugati secondo i 4 stati di matrimonialità)
- Numero di matrimoni (distinti per stato di matrimonialità) di single che, residenti in j al tempo t , sopravvivono e sono sposati in i al tempo $t+k$
- Numero di divorziati e vedovi che, sposati e residenti nella regione j al tempo t , si ritrovano a vivere come single nella regione i al tempo $t+k$
- Numero di single e coniugati (secondo i quattro stati di matrimonialità) che, residenti in t nella regione i , emigrano all'estero nell'intervallo temporale di riferimento
- Numero di single residenti al tempo t nella regione j che al tempo $t+k$ vivono in i

- Numero di coniugati residenti al tempo t nella regione j che al tempo $t+k$ vivono nella regione i (separatamente per ciascuno stato di matrimonialità)
- Numero di immigrati dall'estero – in precedenti periodi - single e coniugati che, viventi nella regione j al tempo t , sono naturalizzati (divenendo cittadini italiani) e sopravvivono nella regione i al tempo $t+k$
- Numero degli ingressi dall'estero nel periodo $t, t+k$ per regione di destinazione e per sesso, distintamente per single e coniugati
- Numero di figli nati da donne coniugate residenti nella regione i al tempo t (separatamente per stato di matrimonialità)
- Numero di figli illegittimi nati da single per regione (separatamente per native e per immigrate)

In questo caso i tassi vengono chiaramente calcolati rapportando gli eventi di ciascun tipo alla popolazione media “a rischio”, sempre in coerenza con l’approccio periodo-coorte e l’assunzione circa il vettore di input dall'esterno.

In Figura 1 per esempio, ammettendo che le date t e $t+5$ siano retrospettive, gli eventi relativi ad una certa classe di età sono racchiusi nel parallelogramma $EV = ABCD$. Ipotizzando che essi si distribuiscano uniformemente nell’unità temporale di osservazione, la popolazione

media di riferimento è costituita da: $P_M = \frac{AB + CD}{2}$. Il tasso medio per l’evento diventa

$$\text{quindi } m = \frac{EV}{P_M}.$$

La determinazione dei tassi di transizione da un stadio all’altro (per classi di età quinquennali) ci consente di utilizzare il modello al fine di formulare previsioni demografiche. Nessuna considerazione di merito circa la bontà delle stime può essere effettuata, dato che esse dipendono dal futuro trend di fenomeni (individuali e collettivi) che, al momento, possono solo essere stimati.

Tuttavia, nonostante gli elementi di aleatorietà che caratterizzano ogni sistema di previsione, il modello in oggetto (come già indicato nell’introduzione) presenta elementi di novità rispetto alle versioni multiregionali a dominanza femminile (vedi Rogers 1995 e Willekens et al. 1984), in quanto utilizza ragionevoli algoritmi di consistenza. Costituisce anche un’evoluzione rispetto a Van Imhoff (1992) per il fatto che estende il sistema multistato ad una popolazione multiregionale.

Dal punto di vista pratico le previsioni si sostanziano nell’applicazione dei tassi (attesi) ai contingenti di popolazione stimata per intervalli temporali quinquennali. Si tratta di previsioni “a tassi costanti” se si assume che gli stessi non cambino nel corso del tempo. Si parla invece di previsioni “a tassi variabili” quando si ipotizza che i tassi di transizione si modifichino da un periodo all’altro secondo una qualsivoglia legge di evoluzione temporale.

5.2 Applicazione del modello all’Italia: utilizzo delle fonti statistiche esistenti

L’applicazione diretta del modello in oggetto non è al momento possibile per l’Italia, in quanto nessuna fonte statistica ufficiale riporta il vettore della popolazione e gli eventi con il livello di dettaglio richiesto dal Paragrafo 5.1. Questi potrebbero però essere reperiti (in massima parte) dai nastri Istat individuali, mediante l’applicazione agli stessi di apposite procedure di estrazione.

Al fine di testare la validità del modello, è tuttavia utile (quanto meno in una prima fase) ricavare gli eventi richiesti basandosi in parte sulle fonti già pubblicate e per il resto facendo una serie di (ragionevoli) ipotesi circa la loro ripartizione nel livello di dettaglio richiesto.

Dato il relativamente elevato numero di operazioni necessarie, non è possibile riportare in questa sede il procedimento seguito per convertire le fonti ufficiali in input per il modello: questo sarà l'oggetto di un lavoro successivo. Basti sin da ora ricordare che le fonti ufficiali utilizzate riguardano alcune delle statistiche demografiche Istat relative agli anni dal 1991 al 1998. In particolare le seguenti: alcuni dati censuari, il *Compendio Statistico*, le pubblicazioni su *Nascite-caratteristiche demografiche e sociali*, *Decessi-caratteristiche demografiche e sociali*, *Matrimoni separazioni e divorzi*, *Tavole di mortalità della popolazione italiana per regione*, *Movimento migratorio della popolazione residente*, *Popolazione residente per sesso, età e regione (anni 1992-95)*. In relazione agli immigrati esiste la specifica pubblicazione del 1998 *La presenza straniera in Italia*, che costituisce una fonte molto ricca, dove è possibile astrarre molte informazioni.

Secondo un procedimento generale, gli eventi per gli autoctoni sono ricavati in base alla differenza tra quelli per i residenti e quelli per gli immigrati (residenti).

6. Il funzionamento del modello multiregionale-multistato

Per chiarire il funzionamento del sistema di proiezione qui descritto, presentiamo le caratteristiche tecniche del modello oggetto del presente lavoro. Per tale finalità è necessario unire gli insegnamenti multiregionali di Willekens e Drewe (1984) e Rogers (1995) con quelli multistato di Van Imhoff e Keilman (1992).

Nonostante la complessità delle relazioni interagenti, resta l'indipendenza tra le diverse fasce per età. Ciò attribuisce la possibilità di seguire l'evoluzione della popolazione per passi, secondo un algoritmo standard valido per tutte le classi meno che per quella più anziana (85 anni e più) e per il calcolo e la proiezione delle nascite, aspetti che verranno trattati separatamente.

6.1. Evoluzione temporale delle classi intermedie

Si consideri in primo luogo la proiezione della popolazione in età da x a $x+5$ al tempo t fino all'età $x+5$, $x+10$ a $t+5$ ove $x=5, 10, \dots, 80$ assumendo implicitamente che l'ultima classe contenga individui in età non inferiore ad 85 anni (che indicheremo come 85+). Vedere lo schema di Lexis della Figura 1.

Il sistema di proiezione viene realizzato in virtù della seguente relazione:

$$P_{t+5}^{x+5} = P_t^x - M_x \left[\frac{P_t^x + P_{t+5}^{x+5}}{2} \right] + I_t^x \quad [1]$$

Ove P_t^x, P_{t+5}^{x+5} (che d'ora in avanti indicheremo con P_t e P_{t+5}) rappresentano rispettivamente i vettori della popolazione in età da x a $x+5$ e da $x+5$ a $x+10$ disaggregata per regione e, nell'ambito di ogni regione, nei 12 strati in precedenza specificati.

I_t^x (I_t d'ora in poi) indica invece il vettore degli immigrati che entrano in età da x ad $x+5$ (distinti per sesso e per stato civile). La matrice M è un'estensione (multistato) della matrice di Leslie. Essa contiene i tassi che, applicati alla popolazione media di ciascuno stato nell'ambito di ogni regione, generano gli eventi.

A fini puramente didattici – illustrativi facciamo per un attimo riferimento ad un sistema con due regioni e quattro stati:

- S1,i = Nativi maschi in regione i (ed età x)
- S2, i = Native femmine in regione i (ed età x)
- S3, i = Immigrati maschi in regione i (ed età x)
- S4, i = Immigrate femmine in regione i (ed età x)

I possibili ingressi dall'estero sono:

- I3,i = Ingresso di immigrati maschi in regione i (ed età x)
- I4,i = Ingresso di immigrate femmine in regione i (ed età x)

I vettori della popolazione e degli immigrati assumono la seguente formulazione

$$P_t^x = \begin{bmatrix} s_{1,1} \\ s_{2,1} \\ s_{3,1} \\ s_{4,1} \\ s_{1,2} \\ s_{2,2} \\ s_{3,2} \\ s_{4,2} \end{bmatrix} \quad I_t^x = \begin{bmatrix} 0 \\ 0 \\ I_{3,1} \\ I_{4,1} \\ 0 \\ 0 \\ I_{3,2} \\ I_{4,2} \end{bmatrix} \quad [2]$$

La matrice M assume invece la seguente configurazione (ove, per semplicità di rappresentazione, l'esponente x relativo all'età è stato eliminato):

$$M_t^x = \begin{bmatrix} m_{11,11} & 0 & -m_{31,11} & 0 & -m_{11,21} & 0 & -m_{31,21} & 0 \\ 0 & m_{22,11} & 0 & -m_{42,11} & 0 & -m_{22,21} & 0 & -m_{42,21} \\ 0 & 0 & m_{33,11} & 0 & 0 & 0 & -m_{33,21} & 0 \\ 0 & 0 & 0 & m_{44,11} & 0 & 0 & 0 & -m_{44,21} \\ -m_{11,12} & 0 & -m_{31,12} & 0 & m_{11,22} & 0 & -m_{31,22} & 0 \\ 0 & -m_{22,12} & 0 & -m_{42,12} & 0 & m_{22,22} & 0 & -m_{42,22} \\ 0 & 0 & -m_{33,12} & 0 & 0 & 0 & m_{33,22} & 0 \\ 0 & 0 & 0 & -m_{44,12} & 0 & 0 & 0 & m_{44,22} \end{bmatrix} \quad [3]$$

$m_{sk,ij}$ (per $s \neq k$ e per $i \neq j$) indica invece il tasso di uscita dallo stato s in regione i per entrare nello stato k in regione j . Quando invece $s = k$ e $i = j$ si ha:

$$m_{ss,ii} = m_{sd,id} + m_{se,ie} + \sum_k \sum_j m_{sk,ij} \quad [4]$$

ove m_{sd} e m_{se} indicano i tassi di uscita dallo stato s in regione i rispettivamente per decesso e per emigrazione.

Per esempio $m_{33,11} = m_{3d,1d} + m_{3e,1e} + m_{31,11} + m_{31,12} + m_{33,12}$ indica il tasso di uscita dallo stato 3 della regione 1. Una parte delle uscite è causata dai decessi e dall'emigrazione all'estero ($m_{3d,1d} + m_{3e,1e}$), mentre la parte residua dipende dai passaggi di stato all'interno della stessa regione ($m_{31,11}$), dall'emigrazione in altra regione mantenendo lo stesso stato ($m_{33,12}$) e dall'emigrazione in altra regione e contestuale passaggio di stato ($m_{31,12}$).

I tassi di passaggio da uno stato-regione all'altro costituiscono decrementi per la popolazione di origine (per tale ragione si trovano con segno positivo lungo la diagonale principale di M, matrice che è preceduta dal segno negativo).

Tali tassi rappresentano invece incrementi per la popolazione dello stato-regione di destinazione (per tale ragione si trovano all'esterno della diagonale principale e sono preceduti dal segno negativo, che annulla il meno anteposto ad M).

Chiaramente la forma di P_t , I_t e di M_t risulta più complessa nel sistema generale in oggetto, composto da n regioni (nel caso italiano $n=3$ o $n=5$) e da 12 diversi strati di popolazione per ogni regione. Infatti la matrice M per ogni classe per età (nel nostro caso di ampiezza quinquennale) ha dimensione $(12 \cdot n, 12 \cdot n)$. I vettori P_t , P_{t+5} e I_t hanno invece dimensione $(12 \cdot n, 1)$.

Per semplicità di rappresentazione possiamo indicare gli elementi della matrice M con la consueta notazione $M[a,b]$ ove a è il numero di riga e b quello di colonna. Di conseguenza alla riga (o colonna) $(i-1) \cdot 12 + s$ corrisponde lo stato s ($1 \leq s \leq 12$) e la regione i ($1 \leq i \leq n$) ed inoltre:

$$M(v,v) = m_{vo} + m_{vd} + \sum_{j \neq v} m_{jv} \quad [5]$$

$$M(v,s) = -m_{sv}$$

In particolare $M[(i-1) \cdot 12 + k, (i-1) \cdot 12 + z]$ (con $k \neq z$) rappresenta il passaggio dallo stato z allo stato k nell'ambito della stessa regione i

$M[(i-1) \cdot 12 + k, (j-1) \cdot 12 + k]$ (con $i \neq j$) rappresenta invece il passaggio dalla regione j alla regione i nell'ambito dello stesso stato

$M[(i-1) \cdot 12 + k, (j-1) \cdot 12 + z]$ (con $i \neq j$ e $k \neq z$) indica infine il passaggio dallo stato z nella regione j allo stato k nella regione i

Chiaramente, come abbiamo visto sopra (vedi per esempio Tabella 1 per avere un'idea nell'ambito di una sola regione), alcune celle possono essere vuote, in quel caso $M[a,b]=0$.

Per esempio $M[4,1]=0$: Non è possibile il passaggio da maschi nativi a femmine native nell'ambito della regione 1.

$M[16,1]=0$ Non è nemmeno possibile il passaggio da maschi nativi a femmine native nell'ambito della regione 2, o di qualsiasi altra regione, da cui $M[(i-1) \cdot 12 + 1, (j-1) \cdot 12 + 4]=0$ per ogni i e j .

Ritorniamo ora al problema principale e risolviamo per P_{t+5} l'equazione [1], che è poi quella basilare del sistema di proiezione, ottenendo (come in Willekens et al., 1984):

$$P^{x+5}_{t+5} = \left[I + \frac{M_x}{2} \right]^{-1} \left[I - \frac{M_x}{2} \right] P^x_t + \left[I + \frac{M_x}{2} \right]^{-1} I^x_t \quad [6]$$

$$P^{x+5}_{t+5} = S^x \cdot P^x_t + H^x \cdot I^x_t$$

ove I è la matrice identità (da non confondersi con quella degli ingressi I_t), mentre S rappresenta la probabilità di sopravvivenza al tempo $t+5$ di soggetti che in t si trovano all'interno del paese ed H , invece, la probabilità di sopravvivenza a $t+5$ di immigrati che entrano nel periodo $t, t+5$.

E' importante notare che con le operazioni di inversione e di prodotto matriciale, è possibile proiettare al tempo $t+5$ la popolazione che a t si trova nella classe di età $x - x+5$. Nella Figura 1 ciò significa il passaggio dal segmento AB al segmento CD.

Tuttavia la proiezione così ottenuta verrà modificata in sede di applicazione di uno qualsiasi degli algoritmi di consistenza (una volta determinati gli eventi per tutte le classi di età).

6.2. Evoluzione temporale della classe per età più avanzata (85 anni ed oltre)

La proiezione per l'ultima classe d'età (85+) si ottiene tenendo in considerazione sia il numero di sopravvissuti dopo 5 anni degli individui in età da 80 ad 85 in t , sia la quota di cittadini che, sopravvissuti ed in età superiore ad 85 anni al tempo t , sono ancora in vita (con un'età di oltre 90 anni) a $t+5$.

Come indicato nella Figura 2, al tempo $t+5$ il segmento di popolazione in età superiore a 85 anni può essere suddiviso in due blocchi: coloro che hanno un'età compresa tra 85 e 90 anni (EF), che sono i sopravvissuti del gruppo 80-85 di 5 anni prima (AB), e coloro che hanno un'età superiore ai 90 anni (ED), che i sopravvissuti del gruppo 85 e più di 5 anni prima (BC).

Conoscendo le matrici M di mortalità nei 5 anni successivi per le classi 80-85 e 85 e più, si ha che:

$$\begin{aligned}
 P_{t+5}^{85} &= \left[I + \frac{M_{80}}{2} \right]^{-1} \left[I - \frac{M_{80}}{2} \right] P_t^{80} + \left[I + \frac{M_{89}}{2} \right]^{-1} I_t^{80} \\
 P_{t+5}^{90+} &= \left[I + \frac{M_{85+}}{2} \right]^{-1} \left[I - \frac{M_{85+}}{2} \right] P_t^{85+} + \left[I + \frac{M_{85+}}{2} \right]^{-1} I_t^{85+} \\
 P_{t+5}^{85*} &= P_{t+5}^{85} + P_{t+5}^{90+}
 \end{aligned} \quad [7]$$

Di conseguenza la popolazione nell'ultima classe di età si determina, nel concreto, sommando le previsioni relative a due distinti gruppi per età. Per le proiezioni della popolazione per il successivo intervallo temporale $t+10$ non è però possibile utilizzare le informazioni circa i soggetti che a $t+5$ sono in età da 85 a 90 anni.

In altri termini l'orizzonte temporale della Figura 2 viene shiftato di 5 anni ed il segmento DF diviene BC, ma non è più disaggregato in 2 blocchi.

6.3 La fascia di età 0-4 anni al tempo $t+5$

La popolazione in età da 0 a 4 anni al tempo $t+5$ dipende da due elementi: il numero di nascite nel periodo $t, t+5$ e il numero di ingressi dall'estero in tale fascia di età nello stesso periodo.

$$P_{t+5}^{0-4} = B_t - M^0 \left[\frac{P_{t+5}^{0-4}}{2} \right] + I_t^0 \quad [8]$$

Il vettore delle immigrazioni (I_0) costituisce un fattore esogeno, mentre quello delle nascite (B_t) è funzione della popolazione femminile media nelle varie classi di età feconda, coniugata (nei vari stati di matrimonialità) per le nascite legittime e single per quelle illegittime.

La popolazione femminile media (c.d. "a rischio di far figli") si ottiene (per ogni classe per età) mediante la semi somma tra il contingente iniziale (noto) e quello finale (stimato secondo il procedimento di evoluzione temporale specificato sopra, opportunamente corretto sulla base di specifici algoritmi di consistenza, come vedremo in dettaglio nel Paragrafo 6). L'applicazione degli algoritmi di consistenza rende utile calcolare (anticipatamente rispetto alle nascite) la popolazione in età $t+5$ per ogni stato e regione per tutte le età al di sopra di 5 anni.

Una volta conosciuto il vettore della popolazione a rischio, le nascite si calcolano applicando allo stesso la matrice dei tassi specifici di fecondità M_b (per ogni classe di età feconda):

$$B^x_t = M_b^x \left[\frac{P_t^x + P_{t+5}^{x+5}}{2} \right] \quad [9]$$

$$B_t = \sum_x B_t^x$$

Dato che il modello di previsione in oggetto non è a dominanza femminile e che presenta nel contempo caratteristiche multistato, la matrice dei tassi di fecondità M_b non è diagonale, come invece si verifica in Rogers (1995) e Willekens et al. (1984). Esaminiamo quindi la composizione di M_b , ricordando che, in virtù della Legge sulla Cittadinanza al momento vigente in Italia (Legge n.91 del 5 Febbraio 1992; D.P.R. n. 572/93; D.P.R. n. 362/94):

- i figli (legittimi) di cui almeno un genitore sia nativo sono considerati cittadini italiani (ed entrano a far parte del gruppo S1 se maschi e del gruppo S4 se femmine)
- i figli (legittimi) i cui genitori siano entrambe immigrati sono immigrati (ed entrano a far parte del gruppo S7 se maschi e del gruppo S10 se femmine)
- i figli (illegittimi) di native single sono nativi
- i figli (illegittimi) di immigrate single sono immigrati

Riportiamo in particolare la matrice dei tassi di fecondità di un sistema uniregionale per una qualsiasi classe per età feconda (che definiremo genericamente con x) basandoci, per semplicità espositive e sulla falsariga della Tabella 1, su un sistema con un solo paese. In quel caso:

Equazione [10]

$$M_b^x = \begin{bmatrix} 0 & 0 & 0 & mb_{4,1} & mb_{5,1} & mb_{6,1} & 0 & 0 & 0 & 0 & mb_{11,1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & mb_{4,4} & mb_{5,4} & mb_{6,4} & 0 & 0 & 0 & 0 & mb_{11,4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & mb_{10,7} & 0 & mb_{2,7} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & mb_{10,10} & 0 & mb_{2,10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Ove $mb_{5,1}$ e $mb_{5,4}$ rappresentano rispettivamente le nascite (legittime) maschili e femminili da donne nello stato coniugale M1; $mb_{6,1}$ e $mb_{6,4}$ le nascite da donne nello stato coniugale M2; $mb_{11,1}$ e $mb_{11,4}$ le nascite da donne nello stato coniugale M2 ed infine $mb_{12,7}$ e $mb_{12,10}$ le nascite da donne nello stato coniugale M4 (l'unico che prevede la nascita di figli immigrati legittimi).

$mb_{4,1}$ e $mb_{4,4}$ rappresentano invece le nascite illegittime (rispettivamente maschili e femminili) da donne single native, mentre $mb_{10,7}$ e $mb_{10,10}$ sono le nascite illegittime da donne single immigrate.

Notare che in generale si verifica che:

$$\frac{mb_{4,1}}{mb_{4,4}} = \frac{mb_{5,1}}{mb_{5,4}} = \frac{mb_{6,1}}{mb_{6,4}} = \frac{mb_{10,7}}{mb_{10,10}} = \frac{mb_{11,1}}{mb_{11,4}} = \frac{mb_{12,7}}{mb_{12,10}} = 1,05$$

per ogni classe di età e per ogni regione. Tuttavia tale assunzione può essere rimossa senza perdita di generalità.

Applicata la matrice di fecondità per tutte le età e determinata successivamente la popolazione in età da 0 a 4 anni al tempo $t+5$, il processo di previsione risulta essere concluso per il tempo $t+5$. E' a questo punto possibile ripetere il procedimento per le proiezioni del quinquennio successivo, dopo aver avuto l'accortezza di modificare i parametri di input se le previsioni sono a tassi variabili.

In quest'ultimo caso deve essere specificata la legge di evoluzione degli elementi della matrice M_x e del vettore I_x . Un'ipotesi può essere quella della variazione nelle stesse proporzioni di tutti gli elementi di un vettore o di una matrice, seguendo per esempio il seguente schema:

$$\begin{aligned} M_x^{t+5} &= \lambda_{M_x} M_x^t \\ Mb_x^{t+5} &= \lambda_b Mb_x^t \\ I_x^{t+5} &= \lambda_I I_x^t \end{aligned} \quad [11]$$

Ove i vari scalari λ assumono valori maggiori di uno se il fenomeno ha carattere incrementativo e minori di uno nel caso contrario.

7. Il funzionamento dell'algoritmo di consistenza

Discutiamo ora più in dettaglio il problema di consistenza, brevemente introdotto nel Paragrafo 4.

Come ormai noto, la procedura di consistenza risulta necessaria in un modello di previsione a due sessi in quanto il numero di matrimoni maschili celebrati (o dissolti) per unità di tempo nell'ambito di ogni regione per status di matrimonialità deve corrispondere esattamente al numero di matrimoni femminili celebrati (o dissolti) nello stesso intervallo temporale, regione e status di matrimonialità, vedi per esempio Keilman (1985).

Il problema della mancata corrispondenza degli eventi appena ricordati si crea perché la popolazione a rischio di matrimonio maschile e femminile non coincide, e inoltre per il fatto che i tassi di matrimonialità per status differiscono per i due sessi. Anche i tassi di mortalità, divorzialità e vedovanza per status di matrimonialità sono diversi per i maschi e per le femmine.

L'algoritmo di consistenza è definibile come una procedura di adattamento che modifica il numero di eventi maschili e femminili stimati secondo i tassi correnti in modo da realizzare lo stesso contingente di matrimoni per status coniugale per entrambe i sessi. Esistono varie procedure di adattamento, tra cui particolarmente usata è quella c.d. della media armonica, vedi per es. Shoen (1981), Keilman (1985), Manfredi e Salinelli (1991), Van Imhoff (1992), Ekamper and Keilman (1993).

La consistenza è relativa contestualmente ai nuovi matrimoni ed allo scioglimento di matrimoni già in essere. Per quanto riguarda gli scioglimenti per ogni stato di matrimonialità e regione, essa può essere intesa in senso più o meno restrittivo. E' restrittivo quando il numero di scioglimenti di un sesso (A) per causa (decesso, morte, vedovanza) uguaglia il

numero di scioglimenti dell'altro sesso (B) per la causa corrispondente. Per esempio Decessi A = Vedovanza B; Divorzi A = Divorzi B; Vedovanza A = Decessi B.

Affinché tutti questi vincoli vengano rispettati, è necessario ripartire la popolazione definita in precedenza come "single" in 3 sottogruppi: "soggetti mai sposati", "divorziati" e "vedovi". Ciò comporterebbe una ulteriore moltiplicazione del numero di strati per ogni regione, che passerebbero da 12 a 20, con l'aggiunta di divorziati e vedovi maschi e femmine nativi e non nativi. Nel Paragrafo 1 abbiamo escluso tale eventualità, che renderebbe il modello ingestibile nel concreto.

E' legato al precedente concetto di consistenza il problema del cambiamento della tipologia per quei matrimoni coinvolgenti gli immigrati che vengono naturalizzati. In quel caso dovrebbe infatti realizzarsi che il numero di sposi del sesso A che escono da un certo status per effetto della naturalizzazione (e che quindi entrano nello status di destinazione) corrisponda al numero di sposi dell'altro sesso che escono da quello status (per entrare a sua volta in quello di destinazione). Tener conto separatamente di questo ulteriore vincolo comporterebbe chiaramente un ulteriore appesantimento del modello.

E' allora utile, in pratica, far riferimento ad un concetto di consistenza meno rigido, imponendo semplicemente che il numero di nuovi matrimoni ed il numero di scioglimenti maschili corrispondano rispettivamente al numero di nuovi matrimoni e di scioglimenti femminili per status di matrimonialità e per unità di tempo, indipendentemente dalla causa.

Vediamo ora i dettagli tecnici relativi al calcolo delle consistenze nell'ambito del modello.

L'operazione preliminare è quella del calcolo del numero di eventi relativi ai nuovi matrimoni ed allo scioglimento di matrimoni in essere per ciascuna regione, sesso e status matrimoniale.

Definiremo a tal proposito i vettori degli eventi in ingresso ed in uscita, in età x , rispettivamente come $+EV_x$ e $-EV_x$. Per tale scopo viene opportunamente applicata la matrice dei tassi di transizione (M_x) alla popolazione media ($PM_x = \frac{1}{2} [P_x + P_{x+5}]$), ovvero agli esposti al rischio di ogni classe per età.

Ricordiamo in questa sede che (come indicato in Paragrafo 5.1 e come visualmente esposto nella Tabella 1 per una esemplificazione) gli elementi posti sulla diagonale principale di M_x ($Diag(M_x)$, vedi per esempio la cella $M_x[i,i]$) costituiscono i tassi di uscita dallo status i verso l'esterno del sistema (decessi, emigrazioni all'estero) e verso altri strati ($j \neq i$). Il vettore degli eventi in uscita in età x è quindi calcolabile come:

$$-EV_x = Diag(M_x) \left[\frac{P_x + P_{x+5}}{2} \right] \quad [12]$$

Gli elementi posti sulla generica riga i e colonna j ($i \neq j$) costituiscono invece i tassi di passaggio (preceduti dal segno negativo) dallo strato j allo strato i . Ne consegue che il numero di ingressi in ogni status i di soggetti già all'interno della popolazione al tempo t è pari alla somma in j dei prodotti tra la cella $M_x[i,j]$ e la relativa popolazione a rischio (riga j di PM_x). A ciò bisogna aggiungere il numero di ingressi dall'estero nell'unità temporale di riferimento opportunamente corretti, per quanto riguarda l'ingresso di coppie coniugate, secondo le istanze del Paragrafo 3, vincolo 8a. Ne consegue che il vettore degli eventi in ingresso è determinato nel modo seguente:

$$\begin{aligned}
+EV_x &= -[M_x - \text{Diag}(M_x)] \left[\frac{P_x + P_{x+5}}{2} \right] + I_x \\
+EV_x &= -[\text{Diag}(M_x) - M_x] \left[\frac{P_x + P_{x+5}}{2} \right] + I_x
\end{aligned}
\tag{13}$$

In accordo con quanto indicato nel Paragrafo 3, e secondo un algoritmo non eccessivamente restrittivo, la consistenza degli eventi è richiesta non per fasce d'età, ma con riferimento al numero complessivo espresso dai seguenti vettori:

$$\begin{aligned}
- EV &= \sum_x -EV_x \\
+ EV &= \sum_x +EV_x
\end{aligned}
\tag{14}$$

Dei vettori di cui all'equazione [14] interessano solo le righe: $12^*(r-1) + i$ con r indicante le varie regioni ($r=1, 2, \dots, n$) ed i ($i = 2, 3, 5, 6, 8, 9, 11, 12$) gli status di matrimonialità, vedi Paragrafo 2, in cui gli eventi (sia in ingresso che in uscita) devono coincidere per le seguenti coppie di strati (adattando la Tabella 2 al contesto generalizzato):

Tabella 3: I vincoli di consistenza in un sistema generalizzato

Per $r = 1$ fino a n

+EV [12*(r-1)+2]	=	+EV [12*(r-1)+5]
-EV [12*(r-1)+2]	=	-EV [12*(r-1)+5]
+EV [12*(r-1)+3]	=	+EV [12*(r-1)+11]
-EV [12*(r-1)+3]	=	-EV [12*(r-1)+11]
+EV [12*(r-1)+8]	=	+EV [12*(r-1)+6]
-EV [12*(r-1)+8]	=	-EV [12*(r-1)+6]
+EV [12*(r-1)+9] - I[12*(r-1)+9]	=	+EV [12*(r-1)+12] - I[12*(r-1)+12]
-EV [12*(r-1)+9]	=	-EV [12*(r-1)+12]

Successivo r

Notare che la Tabella 3 è semplicemente l'adattamento della Tabella 2 del Paragrafo 3 ad un sistema con n regioni. Considerare inoltre che, in virtù di quanto detto in quella sede:

$I[12^*(r-1)+9] = I[12^*(r-1)+12]$ ovvero che il fenomeno di ingresso dall'estero di immigrati coniugati è limitato alle coppie (che quindi entrano a far parte dello status di matrimonialità 4). Tale vincolo è imposto già in origine, in sede di controllo degli input esogeni. Come noto, l'eccesso di coniugati di un sesso è trattato come single.

La realizzazione di ciascuna delle coppie di uguaglianze di cui alla Tabella 3 si ottiene sostituendo ai valori non corretti dei membri di sinistra e di destra una opportuna media calcolata con i due termini stessi. In particolare il metodo della media armonica introdotto da Shoen (1981) soddisfa una serie di "buone proprietà" (Keilman, 1985) e quindi può essere utilizzato come base per un sistema coerente di proiezioni per i due sessi. Per un'introduzione vedi Manfredi e Salinelli (1991).

Definiamo allora come A l'evento non corretto di ciascuno dei membri di sinistra e come B l'evento non corretto di ciascuno dei membri di destra. Gli eventi corretti saranno allora A^* e B^* , ovvero la media armonica dei due termini A e B calcolata come segue:

$$A^* = B^* = \frac{2}{\frac{1}{A} + \frac{1}{B}} = \frac{2 \cdot A \cdot B}{A + B}$$

$$\lambda_A = \frac{A^*}{A} = \frac{2 \cdot B}{A + B} \quad [15]$$

$$\lambda_B = \frac{B^*}{B} = \frac{2 \cdot A}{A + B}$$

λ_A e λ_B rappresentano rispettivamente i fattori di correzione da apportare agli eventi del tipo A e del tipo B: $A^* = \lambda_A \cdot A$ e $B^* = \lambda_B \cdot B$. Indichiamo come λ_i il fattore di correzione per la generica riga i del vettore degli eventi. Questi vengono riportati in due specifiche matrici di correzione per gli eventi in ingresso ed in uscita, che definiremo rispettivamente come $+\lambda$ e $-\lambda$. Le matrici hanno dimensione $[n \times 12, n \times 12]$ ed il generico elemento è così costituito:

$$+\lambda[i,j]=0 \text{ per } i \neq j$$

$$-\lambda[i,j]=0 \text{ per } i \neq j$$

ovvero gli elementi al di fuori della diagonale principale assumono valori nulli

$$+\lambda[(r-1) \cdot 12 + i, (r-1) \cdot 12 + i] = 1 \text{ per } r = 1, 2, \dots, n \text{ e per } i = 1, 4, 7, 10$$

$$-\lambda[(r-1) \cdot 12 + i, (r-1) \cdot 12 + i] = 1 \text{ per } r = 1, 2, \dots, n \text{ e per } i = 1, 4, 7, 10$$

ovvero gli elementi sulla diagonale principale relativi però agli eventi che coinvolgono i single passano indenni l'algoritmo di consistenza.

Considerando che le matrici di correzione $+\lambda$ e $-\lambda$ sono uniche per tutte le classi di età, il vettore degli eventi corretti per la classe x - $x+5$ diviene:

$$+EV_x^* = +\lambda \cdot +EV_x$$

$$-EV_x^* = -\lambda \cdot -EV_x \quad [15]$$

A questo punto la popolazione corretta in età $x+5$ sarà determinata quindi come:

$$P_{x+5}^* = P_{x+5} + [+EV_x^* - +EV_x] - [-EV_x^* - -EV_x]$$

Da cui è calcolabile la popolazione media corretta:

$$PM_x^* = \frac{1}{2} [P_x + P_{x+5}^*]$$

L'algoritmo di consistenza si conclude con la correzione della matrice dei tassi di transizione (determinazione di M_x^*), in modo tale che, applicata alla popolazione media corretta, dia luogo al numero di eventi consistenti sia in uscita che in entrata (dopo aver sommato, in quest'ultimo caso, il vettore degli input immigratori dall'estero).

Questo è calcolabile considerando due diverse matrici contestualmente:

$$Diag[M_x^*] = -\lambda \cdot Diag[M_x] \quad [16]$$

$$[M_x^* - Diag[M_x^*]] = +\lambda \cdot [M_x - Diag[M_x]] \quad [17]$$

La somma delle due matrici di cui alle equazioni [16] e [17] dà luogo alle matrici dei tassi di transizione corretti per ogni classe per età:

$$M^*_x = \text{Diag}[M^*_x] + [M^*_x - \text{Diag}[M^*_x]] \quad [18]$$

La procedura è resa più trasparente dall'esame dettagliato dell'algoritmo di calcolo, indicato nella Tabella 4.

Tabella 4. Algoritmo di correzione della matrice dei tassi di transizione

Per r = 1 fino a n

 Per i = 2, 3, 5, 6, 8, 9, 11, 12

$$M_x^*[i,i] = -\lambda_i \cdot M_x[i,i]$$

 Per j = 1 fino a 12

 Se i ≠ j allora

$$M_x^*[i,j] = +\lambda_i \cdot M_x[i,j]$$

 Fine se

 Successivo j

Successivo i

Successivo r

In questo modo risulta garantito che:

$$P_{x+5}^* = P_x - M_x \cdot P M_x^* + I_x$$

per ogni fascia di età, ad eccezione di quella 0-4, non ancora determinata, vedi Paragrafo 5.3. La mancanza della consistenza per la classe per età 0 – 4 non costituisce alcuna limitazione in quanto nessun soggetto in tali età risulta spostato, o in grado di far figli! Da questo punto di vista è sovrabbondante anche l'applicazione dell'algoritmo per le classi di età al di sopra di 50 anni, che non condizionano più la fecondità e quindi il rapporto tra le nuove leve maschili e quelle femminili. Tuttavia viene comunque applicato perché la logica dell'algoritmo stesso rimane invariata anche per le età più avanzate.

8. Strutturazione logica del processo di proiezione

Nei paragrafi precedenti abbiamo esaminato in dettaglio le procedure di calcolo necessarie per le proiezioni nel futuro della popolazione di un determinato paese. Dato che la presentazione non ha seguito un esatto ordine logico, è opportuno in questa sede richiamare i principali passi di cui il complessivo sistema di stima consta.

Il primo punto di ogni previsione è quello della conoscenza del vettore della popolazione alla data più aggiornata possibile. Tale vettore sarà disaggregato per regione e, nell'ambito di ogni regione, per strato di popolazione ed ovviamente per classi di età di ampiezza quinquennale, vedi Paragrafo 3.

Il secondo punto riguarda la raccolta dei dati relativi ai tassi di transizione da uno strato all'altro (componenti della matrice M), secondo quanto indicato in Paragrafo 5, e dei dati relativi al vettore di ingressi (I). Gli elementi del vettore devono essere corretti in modo tale che il numero di ingressi di coniugati maschili corrisponda con il numero di ingressi di coniugati femminili. L'eventuale eccedenza di una componente verrà considerata come ingresso di single.

Il terzo punto del processo di previsione è relativo alla definizione dell'orizzonte temporale della proiezione ed alla formulazione di coerenti ipotesi circa l'evoluzione dei tassi

demografici implicati, che è poi il punto sostanzialmente più importante di tutti. A partire da questo punto le previsioni demografiche consistono semplicemente nella mera applicazione degli algoritmi di calcolo del modello presentato e sono, da questo punto di vista, esatte.

Per ogni intervallo temporale di dimensione quinquennale si procede in primo luogo al calcolo del numero di sopravvissuti alla fine del periodo dei soggetti che all'inizio si trovano all'interno della popolazione e dei soggetti che vi entrano nel periodo stesso, vedi Paragrafo 6.1. Il calcolo dei sopravvissuti dall'ultima classe per età (85 +) richiede una procedura più complessa, indicata in Paragrafo 6.2.

Notare che la determinazione della popolazione che al generico tempo $t+5$ si trova in età da 5 anni in poi secondo lo schema generale visto sopra non è soddisfacente. Infatti dà luogo ad incongruenze nel numero di sposi maschi e femmine nell'ambito dei vari strati di matrimonialità per ogni regione, come indicato in Paragrafo 4.

Per rendere coerente il numero di matrimoni dei due sessi risulta allora necessario adottare un algoritmo di consistenza, vedi Paragrafo 7. Solo dopo l'applicazione dell'algoritmo per le classi di età da 5 anni in poi risulta possibile procedere al calcolo delle nascite (legittime) nell'orizzonte temporale di riferimento. Al fine di calcolare le nascite illegittime non è previsto alcun algoritmo di consistenza.

Una volta noti i vettori delle nascite (e degli ingressi in età 0 – 4) si procede alla stima della popolazione in età 0 – 4 al tempo $t+5$. Vedi Paragrafo 6.3 per i dettagli. A questo punto la popolazione dal tempo t al tempo $t+5$ risulta concluso.

La ripetizione dell'iter seguito fino a questo punto consente di ottenere le stime anche per i periodi successivi. Notare tuttavia che, se le proiezioni sono a tassi variabili, il primo passo per la stima relativa al quinquennio $t+5$, $t+10$ consiste nel calcolare i nuovi valori della matrice M e del vettore I (sempre rispettando il vincolo di parità tra la componente maschile e quella femminile dei coniugati) secondo la funzione di previsione attesa.

La procedura descritta nel presente Paragrafo è riassunta nella Tabella 5.

Tabella 5. Principali componenti del processo di proiezione

- Inserimento dati di partenza
- Correzione input relativi ai matrimoni
- Definizione orizzonte temporale di riferimento (T)

Per $i = 5$ fino a T (con passo 5)

- Inserimento dati matrice di proiezione $M(t)$
- Calcolo popolazione in età da 5 in poi al tempo i
- Algoritmo di consistenza nel numero di matrimoni
- Correzione della matrice di proiezione $M(t)$
- Calcolo popolazione in età da 0 a 5 anni al tempo i

Successivo i

9. Conclusioni

Il presente lavoro è servito per introdurre un modello multiregionale di previsioni demografiche per i due sessi. Il modello, una volta fornito degli opportuni dati, serve per determinare l'evoluzione temporale della popolazione di un Paese industrializzato sottoposto

a consistenti flussi migratori provenienti dall'estero, in cui le caratteristiche degli immigrati siano notevolmente diverse rispetto a quelle dei nativi.

Uno specifico e molto avanzato modello multiregionale con separazione della componente immigratoria è stato presentato da Rogers (1995), ma non è soddisfacente (ai nostri fini) in quanto riferito ad un contesto con dominanza femminile. Con sistema di previsione da noi specificato si rimane nel contesto multiregionale, e nel contempo si superano i problemi di inconsistenza tra i due sessi mediante l'applicazione di specifici algoritmi di coerenza.

Il nostro modello ha validità generale, ma è pensato specificamente per studiare l'evoluzione della popolazione Italiana, in quanto laboratorio pratico in cui studiare entro quali limiti le immigrazioni sono in grado di contrastare la tendenza al declino ed all'invecchiamento derivanti dai veramente bassi coefficienti di fecondità del nostro paese. Per tali ragioni la presentazione della parte teorica è affiancata, nel Paragrafo 4, dall'indicazione dei dati necessari per una corretta applicazione del modello alla realtà.

La formulazione di plausibili scenari per il futuro e quindi il calcolo delle proiezioni demografiche corrispondenti a tali ipotesi, derivanti dall'applicazione del modello in oggetto, saranno l'argomento di un successivo lavoro.

10. Bibliografia

- Ekamper P. and Keilman N. (1993), Sensitivity analysis in a Multidimensional Demographic Projection Model with a Two-sex Algorithm, *Mathematical Population Studies*, Vol. 4(1), pp.21-36
- Gill R. and Keilman N. (1990), *On the estimation of Multidimensional Demographic Models with population registration data*, *Mathematical Population Studies*, Vol. 2 (2), pp. 119 – 143.
- Golini (1997), Demographic trends and ageing in Europe. Prospects, problems and policies, *Genus*, Vol LIII, n. 3-4, pp. 33 - 74
- Keifitz N. (1985), *Applied Mathematical Demography*, WILEY
- Keilman N, *Nuptiality models and the two-sex problem in national population forecasts*, *European Journal of Population* 1, pp. 207-235
- Long L. (1992), *Changing Residence: Comparative Perspectives on its Relationship to Age, Sex, and Marital Status*, *Population Studies*, 46, pp. 141-158
- Manfredi P. e Salinelli E. (1991), *About an interactive model for sexual Populations*, W.P. 35, Università di Pisa
- Manfredi P. e Valentini A. (1997), *Populations with below replacement fertility and immigrations: theoretical considerations and scenarios from the italian laboratory*, W.P. 127, Università di Pisa
- Rogers A. (1995), *Multiregional Demography. Principles, methods and extensions*, John Wiley & Soons, Chichester
- Shoen R. (1981), *The harmonic mean as the basis of a realistic two-sex marriage model*, *Demography* 18 (2), pp.201-216
- Valentini A. (1997), *Calo della fecondità ed immigrazioni: scenari e considerazioni sul caso italiano*, W.P. 125, Università di Pisa
- Valentini A. (1998), *Impact of international migrations on internal regional dynamics: theoretical facts and the italian case*, paper presentato alla International Conference on Immigrations, Bari 25-27 giugno 1998

- Valentini A. (1999), *Impatto delle immigrazioni sulla popolazione italiana: confronto tra scenari alternativi*, paper presentato alle Giornate di studio sulla popolazione, Firenze 7-9 gennaio 1999.
- Van Imhoff E. (1992), A general Characterization of Consistency Algorithms in Multidimensional Demographic Projection Models, *Population Studies*, Vol. 46, pp. 159-169.
- Van Imhoff E. and Keilman N. (1992), *Lipro 2.0: An Application of a Dynamic Demographic projection model to household structure in the Neitherlands*. NIDI CBGS Publications 23, Swets & Zeitlinger, Amsterdam/Lisse
- Van Imhoff E. (1994), *Lipro 3.0 user's guide and tutorial*. NIDI Working Paper 1994/1A and 1994/1B. The Hague: NIDI, Amsterdam
- Willekens F.J., Sham I., Sham J.M. and Ramachandran P. (1982), *Multi-state Analysis of Marital Status Life Tables: Theory and Application*, *Population Studies*, 36 (1), pp. 129-144
- Willekens F.J. and Drewe P. (1984), *A Multiregional Model for Regional Demographic Projection*, extracted from Ch. 15, *Demographic Research and Spatial Policy*, pp. 309-331, Academic Press, London

Figura 1.
L'Approccio periodo coorte per la generica età $x - x+5$

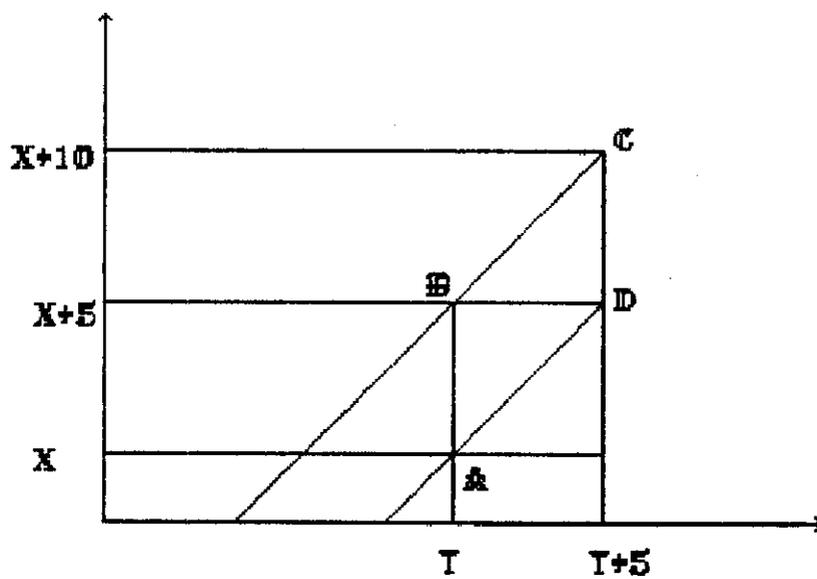


Figura 2.
Proiezioni al tempo $t+5$ per la popolazione con età superiore a 85 anni

