

Nota metodologica

La stima della speranza di vita per livello di istruzione è stata ottenuta attraverso l'integrazione tra l'archivio dei decessi¹ ed il censimento della popolazione e delle abitazioni del 2011 implementando una procedura di *record-linkage* individuale tra questi due archivi.

Tale procedura ha permesso di associare l'informazione del titolo di studio rilevata al censimento al 94,5 per cento dei deceduti nel 2012. L'analisi della performance della procedura di *record-linkage* per le principali variabili demografiche ha evidenziato in generale una omogeneità per genere (maschi 94,8 per cento e femmine 94,1 per cento), titolo di studio (superiore al 94 per cento per tutti i livelli di istruzione) ed età (tra l'88 per cento per la classe 25-34 anni ed il 95,5 per cento per la classe di età 75-84 anni).

Trattamento dei dati non linkati

Per quei decessi che non si sono accoppiati con l'archivio del censimento (d^{nl}) è disponibile solo l'informazione del livello di istruzione (m) già presente sulla scheda di morte.

Per tener conto della quota dei decessi non accoppiati sulla stima dei livelli di mortalità per titolo di studio (k al censimento), il numero dei decessi non linkati (d^{nl}) è stato redistribuito tenendo conto della duplice informazione sul titolo di studio presente nei decessi linkati (d^l) (Prospetto 1).

Prospetto 1. Decessi linkati al censimento (d^l). Distribuzione $d^l_{m,k}$ dei decessi per livello di istruzione presente nell'archivio dei decessi (m) e nel censimento (k) ($d^l_{m,k}$). (es: m e k con 4 livelli di istruzione).

Fonte	Censimento k=1,..4					Totale
	Variabile Livello di istruzione	Nessun titolo o licenza elementare (1)	Licenza media inferiore (2)	Licenza media superiore (3)	Laurea o titolo di studio superiore (4)	
Archivio dei decessi m=1,..4	Nessun titolo o licenza elementare (1)	$d^l_{1,1}$	$d^l_{1,2}$	$d^l_{1,3}$	$d^l_{1,4}$	$d^l_{1,0}$
	Licenza media inferiore (2)	$d^l_{2,1}$	$d^l_{2,2}$	$d^l_{2,3}$	$d^l_{2,4}$	$d^l_{2,0}$
	Licenza media superiore (3)	$d^l_{3,1}$	$d^l_{3,2}$	$d^l_{3,3}$	$d^l_{3,4}$	$d^l_{3,0}$
	Laurea o titolo di studio superiore (4)	$d^l_{4,1}$	$d^l_{4,2}$	$d^l_{4,3}$	$d^l_{4,4}$	$d^l_{4,0}$

A partire dal Prospetto 1 è stato possibile costruire la tabella dei 'fattori di ponderazione', ovvero delle frequenze relative di riga (Prospetto 2), dove il generico fattore di ponderazione per m e k è stato calcolato nel seguente modo:

$$[a] p_{m,k} = \frac{a^l_{m,k}}{a^l_{m,0}}$$

¹Indagine sui decessi e cause di morte, Istat.

Prospetto 2. Tabella dei coefficienti di ponderazione $p_{m,k}$

Fonte	Censimento $k=1,..4$					
	Variabile Livello di istruzione	Nessun titolo o licenza elementare (1)	Licenza media inferiore (2)	Licenza media superiore (3)	Laurea o titolo di studio superiore (4)	
Scheda di morte $m=1,..4$	Nessun titolo o licenza elementare (1)	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$	p_{10}
	Licenza media inferiore (2)	$p_{2,1}$	$p_{2,2}$	$p_{2,3}$	$p_{2,4}$	$p_{2,0}$
	Licenza media superiore (3)	$p_{3,1}$	$p_{3,2}$	$p_{3,3}$	$p_{3,4}$	$p_{3,0}$
	Laurea o titolo di studio superiore (4)	$p_{4,1}$	$p_{4,2}$	$p_{4,3}$	$p_{4,4}$	$p_{4,0}$

Il numero di decessi stimato $\hat{d}_{m,k}^{nl}$ è stato ottenuto moltiplicando la frequenza dei decessi non linkati con livello di istruzione m (d_m^{nl}) per il fattore di ponderazione $p_{m,k}$:

$$[b] \hat{d}_{m,k}^{nl} = d_m^{nl} p_{m,k}$$

Fissato il livello di istruzione k , sommando i decessi stimati $\hat{d}_{m,k}^{nl}$ per i livelli di istruzione m si ottiene una stima dei decessi per livello di istruzione al censimento k dovuti al mancato *linkage*:

$$[c] \hat{d}_k^{nl} = \sum_{m=1}^n \hat{d}_{m,k}^{nl}$$

In questo modo la somma per k (dei livelli di istruzione al censimento) dei decessi stimati è uguale al numero totale dei decessi non linkati $d^{nl} = \sum_{k=1}^n \hat{d}_k^{nl}$.

Il numero complessivo dei decessi con livello di istruzione studio k (d'_k) si ottiene dalla somma della frequenza dei decessi linkati (d_k^l) e della frequenza dei decessi stimati \hat{d}_k^{nl} :

$$[d] d'_k = d_k^l + \hat{d}_k^{nl}$$

La stima dei decessi per l'anno 2012 è stata fatta per ogni livello di istruzione k , genere e per singola età x (compresa tra i 25 ed i 89 anni).

Le probabilità di morte per titolo di studio

La serie delle probabilità di morte prospettive per livello di istruzione è stata ottenuta applicando i rischi relativi di morte per titolo di studio ($RR_{x,k}$) alla tavola di mortalità della popolazione italiana del 2012⁽¹⁾. In particolare le $q_{x,k}^{def}$ per età e livello di istruzione k [e] sono state ottenute dal prodotto delle q_x della popolazione italiana (q_x^{demo})⁽¹⁾ per gli $RR_{x,k}$ determinati dal *record-linkage* [cfr. paragrafo successivo].

$$[e] q_{x,k}^{def} = q_x^{demo} \cdot RR_{x,k} \text{ per } x = 25, 26 \dots 89$$

Per le età inferiori ai 25 anni e superiori o uguali ai 90 anni sono state utilizzate le probabilità di morte della popolazione italiana già pubblicate per il 2012. Ciò in quanto, prima dei 25 anni molti individui devono ancora completare gli studi e dai 90 anni l'esiguo numero di decessi nelle categorie più elevate di istruzione scolastica non consente l'analisi della mortalità per titolo di studio.

I rischi relativi di morte per età e titolo di studio

Per la costruzione dei rischi relativi per titolo di studio si è partiti dalle probabilità di morte prospettive, ottenute dai dati del *record-linkage*, nell'intervallo di età compresa tra 25 e 89 anni.

Le probabilità di morte ad età x per titolo di studio k sono state ottenute utilizzando l'espressione:

$$[f] \ q_{x,k} = \frac{d'_{x,k}}{pop_{x,k}} \quad \text{per } x = 25, 26 \dots 89$$

dove $d'_{x,k}$ rappresentano i deceduti nel 2012, residenti in Italia e che al 1° gennaio 2012 avevano un'età in anni compiuti uguale ad x e un titolo di studio k e $pop_{x,k}$ è il numero di individui rilevati dal censimento 2011 che al 1° gennaio del 2012 avevano la medesima età e titolo di studio, decurtato dei deceduti tra la data del censimento e il 31 dicembre 2011.

Le $q_{x,k}$ così ottenute presentano tuttavia un'elevata oscillazione che riguarda soprattutto le età più giovani dove la stima si basa su un numero esiguo di eventi. Per ovviare a questa problematica è stata applicata una procedura perequativa alle $q_{x,k}$ grezze [g] in analogia alla metodologia utilizzata per le tavole di mortalità per la popolazione italiana prodotte dall'Istat⁽²⁾.

$$[g] \ q'_{x,k} = \frac{1}{21} \cdot [7q_x + 6(q_{x-1} + q_{x+1}) + 3(q_{x-2} + q_{x+2}) - 2(q_{x-3} + q_{x+3})]$$

Questa procedura potrebbe aver allontanato le distribuzioni di probabilità per titolo di studio dal rispettivo ammontare assoluto di decessi. È stato quindi necessario fare in modo che ciascuna distribuzione esprimesse l'ammontare dei decessi osservati, attraverso una tecnica di posizionamento che prevede l'applicazione, alle probabilità perequate, di un fattore di correzione, dato dal rapporto tra il totale osservato di decessi (M_{oss}) e quello atteso (M_{att}) per ciascun titolo di studio [h]⁽²⁾.

$$[h] \ q''_{x,k} = q'_{x,k} \cdot \frac{M_{oss}}{M_{att}}$$

Queste tecniche hanno ridotto notevolmente l'instabilità delle stime delle q_x , riscontrata in particolare nelle età più giovani dove il numero di decessi per titolo di studio è ridotto.

Le $q''_{x,k}$ sono state utilizzate per il calcolo dei rischi relativi di morte specifici per titolo di studio ($RR_{x,k}$) [i].

$$[i] \ RR_{x,k} = \frac{q''_{x,k}}{q''_x}$$

dove $q''_{x,k}$ è la probabilità di morire nel 2012 di un individuo con età x al 1° gennaio 2012 e titolo di studio al censimento k , e q''_x è la probabilità di morte dell'intera popolazione della stessa età ottenuta anch'essa con i dati del *record-linkage*.

Derivazione delle funzioni biometriche

Le tavole di mortalità per titolo di studio e genere sono quindi costruite a partire dalle probabilità di morte per età calcolate come sopra illustrato. Semplificando la simbologia, è quindi possibile calcolare dalla serie delle q_x anche quella dei sopravvissuti, dei decessi, degli anni vissuti e delle speranze di vita.

Sopravvissuti (l_x): Individui, provenienti dalla generazione iniziale fittizia di 100 mila nati, che sopravvivono ai vari compleanni:

$$l_{x+1} = l_x \cdot (1 - q_x)$$

Decessi (d_x): Individui che muoiono tra l'età x e l'età $x+1$:

$$d_x = l_x \cdot q_x$$

Anni vissuti (L_x): Anni vissuti tra l'età x e $x+1$ da parte dei soggetti che hanno raggiunto l'età x . Il coefficiente utilizzato per la determinazione degli anni vissuti dai deceduti è 0.5, assumendo che i decessi avvengono a metà anno:

$$L_x = l_{x+1} + 0.5 d_x$$

ovvero si ottengono sommando i sopravvissuti al compleanno $x+1$ con il numero di anni vissuti dai soggetti d_x deceduti in età x . Solitamente si suppone che questi ultimi abbiano vissuto in media mezzo anno quindi:

Speranza di vita: numero medio di anni che un soggetto può ancora attendersi di vivere al compimento dell'età x

$$E_x = \frac{T_x}{l_x}$$

Bibliografia

1. Tavole di mortalità della popolazione italiana, Istituto nazionale di statistica, anno 2012, <http://demo.istat.it/unitav2012/note.html>
2. Tavole di mortalità della popolazione italiana per provincia e regione di residenza, Istituto nazionale di statistica, anno 1998, http://www3.istat.it/dati/catalogo/20020731_00/volumemortalita.pdf

Avvertenze

Tavole di mortalità della popolazione residente in Italia - Anno 2012

La serie delle probabilità di morte per livello di istruzione è stata ottenuta applicando rischi relativi di morte per titolo di studio alla tavola di mortalità della popolazione italiana del 2012 (<http://demo.istat.it/unitav2012/note.html>)

Funzioni biometriche:

q_x , probabilità che un individuo di età precisa x muoia prima di compiere l'età precisa $x+1$

l_x , funzione di sopravvivenza, rappresenta il numero di individui, provenienti da un ipotetico contingente iniziale di 100.000 nati vivi, che sopravvivono all'età x

d_x , numero di decessi tra le età x e $x+1$ (valori arrotondati)

L_x , numero di individui di età x espressa in anni compiuti (valori arrotondati)

e_x , speranza di vita, esprime il numero medio di anni che restano da vivere a quanti sopravvivono all'età precisa x