La modellizzazione della mortalità

Fabio Bellini

Università di Milano-Bicocca

fabio.bellini@unimib.it

7 ottobre 2021

La modellizzazione probabilistica della durata della vita umana è fondamentale nella valutazioni delle prestazioni attuariali e nel risk management delle compagnie di assicurazione. Si tratta di un tema classico, a cui hanno lavorato a partire dalla fine del ‘600 grandi scienziati come Sir Edmund Halley, Abraham De Moivre, Daniel Bernoulli. I modelli tradizionali, detti a mortalità deterministica, assumono una popolazione in condizioni stazionarie, in cui le nascite equilibrano le morti, e un'unica distribuzione di probabilità consente di calcolare le probabilità di vita e di morte per tutte le età, senza variazioni nel tempo, come vedremo tra poco. Questi modelli sono spessi descritti attraverso una *tavola di mortalità*, che come vedremo esprime il numero di sopravvissuti in funzione del tempo per una coorte di partenza fissata.

Il modello probabilistico

Ipotizziamo una popolazione in condizioni stazionarie. Immaginiamo che la durata della vita di un individuo dalla nascita sia espressa da una variabile casuale nonnegativa T_0, che è la stessa per tutti gli individui e non cambia nel tempo. Indichiamo con F_0 la sua funzione di ripartizione

$$F_0(x) = \mathbb{P}(T_0 \leq x)$$

e con $S_0(x)$ la sua funzione di sopravvivenza

$$S_0(x) = \mathbb{P}(T_0 > x) = 1 - F_0(x).$$

Dato che T_0 è nonnegativa, si ha $S_0(0) = 1$. Dalla Statistica sappiamo che $S_0(x)$ è noncrescente e $\lim_{x \to +\infty} S_0(x) = 0$. Nei modelli attuariali può esistere una massima età raggiungibile che si indica con la lettera greca ω, per cui si ha $S_0(\omega) = 0$; comunemente si pone $\omega = 120$.

La funzione di sopravvivenza condizionata

Nella ipotesi di popolazione in condizioni stazionarie, la conoscenza della variabile casuale T_0 consente di ricavare la distribuzione della lunghezza della vita residua di un individuo di età x, che indichiamo con la variabile casuale T_x. La funzione di ripartizione di T_x si calcola nel modo seguente:

$$F_x(t) = \mathbb{P}(T_x \leq t) = \mathbb{P}(T_0 \leq x + t \mid T_0 > x).$$

Si tratta cioè della *probabilità condizionata* dell'evento $A = \{T_0 \leq x + t\}$ rispetto all'evento $B = \{T_0 > t\}$. Dato che come noto

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)},$$

possiamo calcolare

La funzione di sopravvivenza condizionata

$$F_x(t) = \mathbb{P}(T_0 \leq x + t \mid T_0 > x) = \frac{\mathbb{P}(\{T_0 \leq x + t\} \cap \{T_0 > x\})}{\mathbb{P}(T_0 > x)} =$$

$$= \frac{\mathbb{P}(x < T_0 \leq x + t)}{\mathbb{P}(T_0 > x)} = \frac{F_0(x + t) - F_0(x)}{S_0(x)} =$$

$$= \frac{S_0(x) - S_0(x + t)}{S_0(x)} = 1 - \frac{S_0(x + t)}{S_0(x)}.$$

Di conseguenza, la funzione di sopravvivenza condizionata è data da

$$S_x(t) = 1 - F_x(t) = \frac{S_0(x + t)}{S_0(x)}.$$

Un esempio

Consideriamo T_0 con funzione di ripartizione

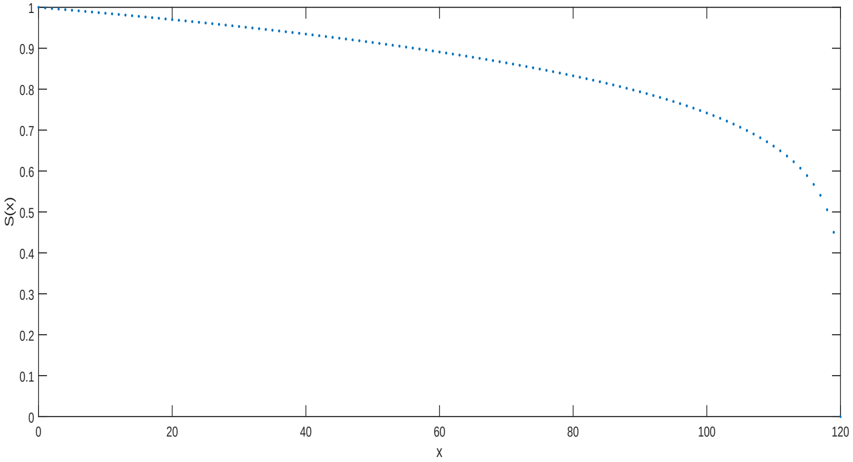
$$F_0(x) = 1 - \left(1 - \frac{x}{120}\right)^{1/6}, \quad 0 \leq x \leq 120,$$

da cui otteniamo

$$S_0(x) = 1 - F_0(x) = \left(1 - \frac{x}{120}\right)^{1/6}, \quad 0 \leq x \leq 120.$$

Osserviamo subito che l'età massima è $\omega = 120$ in quanto $S_0(120) = 0$.
Rappresentiamo graficamente l'andamento di $S_0(x)$ nella slide sucessiva.

Un esempio



La funzione di sopravvivenza S_0 dell'esempio.

Un esempio

Utilizzando questo modello probabilistico possiamo calcolare alcune probabilità di vita e di morte. La probabilità che un nuovo nato superi i 30 anni è data da

$$S_0(30) = \left(1 - \frac{30}{120}\right)^{1/6} \simeq 95,32\%.$$

La probabilità che un 40enne sopravviva oltre i 65 anni può essere calcolata usando la formula ricavata precedentemente:

$$S_{40}(25) = \frac{S_0(40+25)}{S_0(40)} \simeq 93,95\%.$$

La probabilità che un 30enne muoia entro 20 anni può essere calcolata come il complementare a uno della probabilità che sopravviva per 30 anni:

$$F_{30}(20) = 1 - S_{30}(20) = 1 - \frac{S_0(30+20)}{S_0(30)} = \frac{S_0(30) - S_0(50)}{S_0(30)} \simeq 4,10\%.$$

Probabilità di vita e di morte

Introduciamo alcune notazioni attuariali standard per x e k interi.
La probabilità che un individuo di età x sopravviva per k anni si indica con

$$ {}_{\mathbf{k}}\mathbf{p}_{\mathbf{x}} := \mathbb{P}(T_x > k) = \frac{S_0(x+k)}{S_0(x)}. $$

La probabilità che un individuo di età x muoia entro k anni è data da

$$ {}_{\mathbf{k}}\mathbf{q}_{\mathbf{x}} := 1 - {}_{\mathbf{k}}\mathbf{p}_{\mathbf{x}} = 1 - \frac{S_0(x+k)}{S_0(x)} = \frac{S_0(x) - S_0(x+k)}{S_0(x)}. $$

Per definizione si ha ${}_{\mathbf{k}}\mathbf{p}_{\mathbf{x}} + {}_{\mathbf{k}}\mathbf{q}_{\mathbf{x}} = 1$. Se $k = 0$, si ha ${}_{\mathbf{0}}\mathbf{p}_{\mathbf{x}} = 1$ e ${}_{\mathbf{0}}\mathbf{q}_{\mathbf{x}} = 0$.
Nel caso $k = 1$, si introduce la notazione abbreviata

$$ \mathbf{p}_{\mathbf{x}} := {}_{\mathbf{1}}\mathbf{p}_{\mathbf{x}}, \quad \mathbf{q}_{\mathbf{x}} := {}_{\mathbf{1}}\mathbf{q}_{\mathbf{x}}. $$

$\mathbf{p}_{\mathbf{x}}$ rappresenta quindi la probabilità che l'individuo di età x sopravviva per un anno, $\mathbf{q}_{\mathbf{x}}$ la probabilità che muoia entro l'anno.

Un individuo di età x sopravvive per k anni se e solo se sopravvive per un anno, e poi sopravvive per altri $k-1$ anni; ne segue la relazione

$$\mathbf{{}_k p_x = p_x \cdot {}_{k-1}p_{x+1}}$$

che possiamo verificare attraverso il calcolo

$$\frac{S_0(x+k)}{S_0(x)} = \frac{S_0(x+1)}{S_0(x)} \frac{S_0(x+1+k-1)}{S_0(x+1)}.$$

Ripetendo lo stesso ragionamento al periodo successivo otteniamo

$$\mathbf{{}_k p_x = p_x \cdot p_{x+1} \cdot {}_{k-2}p_{x+2}}$$

e ripetendolo su tutti i periodi si ottiene

$$\mathbf{{}_k p_x = p_x \cdot p_{x+1} \cdot \ \ldots \ \cdot p_{x+k-1}}.$$

Probabilità di morte differite

La probabilità che un individuo di età x muoia in un intervallo di tempo compreso tra h e $h+k$ anni a partire da oggi prende il nome di probabilità di morte differita e si indica con

$$\mathbf{{}_{h|k}q_x} := \mathbb{P}(h < T_x \leq h+k)$$

Notare che il primo indice h indica il numero di anni di differimento, il secondo indice k indica la lunghezza dell'intervallo di tempo in cui avviene il decesso. Anche le probabilità di morte differite possono essere facilmente a partire dalla funzione di sopravvivenza alla nascita S_0:

$$\mathbb{P}(h < T_x \leq h+k) = S_x(h) - S_x(h+k) = \frac{S_0(x+h)}{S_0(x)} - \frac{S_0(x+h+k)}{S_0(x)} =$$
$$= \frac{S_0(x+h) - S_0(x+h+k)}{S_0(x)}.$$

Probabilità di morte differite

Un individuo di età x muore in un intervallo compreso tra h e $h+k$ anni a partire da oggi se e solo se sopravvive per h anni, diventando un individuo di età $x+h$, e poi muore nei successivi k anni. Si ha quindi

$$\mathbf{{}_{h|k}q_x} = \mathbf{{}_hp_x} \cdot \mathbf{{}_kq_{x+h}},$$

come possiamo anche verificare calcolando

$$\frac{S_0(x+h) - S_0(x+h+k)}{S_0(x)} = \frac{S_0(x+h)}{S_0(x)} \cdot \frac{S_0(x+h) - S_0(x+h+k)}{S_0(x+h)}.$$

Possiamo anche osservare che

$$\mathbf{{}_{h|k}q_x} = \mathbf{{}_hp_x} - \mathbf{{}_{h+k}p_x},$$

$$\mathbf{{}_{h|k}q_x} = \mathbf{{}_{h+k}q_x} - \mathbf{{}_hq_x}.$$

Probabilità di morte differite

La prima uguaglianza deriva dal fatto che un individuo di età x muore in un intervallo compreso tra h e $h+k$ anni a partire da oggi se e solo se sopravvive per h anni (con probabilità ${}_{\mathbf{h}}\mathbf{p}_{\mathbf{x}}$) e non sopravvive poi oltre $h+k$ anni, quindi dobbiamo togliere da ${}_{\mathbf{h}}\mathbf{p}_{\mathbf{x}}$ la probabilità che sopravviva oltre $h+k$ anni, data da ${}_{\mathbf{h+k}}\mathbf{p}_{\mathbf{x}}$.
Analogamente, la seconda uguaglianza deriva dal fatto che un individuo di età x muore in un intervallo compreso tra h e $h+k$ anni a partire da oggi se e solo se muore entro $h+k$ anni (con probabilità ${}_{\mathbf{h+k}}\mathbf{q}_{\mathbf{x}}$) ma non è morto nei primi oltre h anni, quindi dobbiamo togliere da ${}_{\mathbf{h+k}}\mathbf{q}_{\mathbf{x}}$ la probabilità di morte entro i primi k anni, data da ${}_{\mathbf{h}}\mathbf{q}_{\mathbf{x}}$.
Infine, osserviamo che un individuo di età x muore entro k anni se e solo se muore in uno dei k intervalli di tempo $(h, h+1]$, per $h = 0, \ldots, k-1$. Si ha quindi

$$ {}_{\mathbf{k}}\mathbf{q}_{\mathbf{x}} = \sum_{h=0}^{k-1} {}_{\mathbf{h|1}}\mathbf{q}_{\mathbf{x}}. $$

Le tavole di mortalità

Le informazioni storiche relative alla mortalità si presentano usualmente attraverso le *tavole di mortalità*. Queste tavole fanno riferimento a un ipotetico gruppo di individui (detto coorte) di numerosità pari alla *radice* della tavola (usualmente pari a 100.000), che viene seguito nel tempo. Si indica convenzionalmente con ℓ_x il numero di sopravviventi al tempo x e con d_x il numero dei decessi nell'intervallo $(x, x+1]$, che è dato da

$$d_x = \ell_x - \ell_{x+1}.$$

Le probabilità di sopravvivenza sono calcolate come

$$\mathbf{{}_k p_x} = \frac{\ell_{x+k}}{\ell_x}.$$

Il tasso di mortalità è dato da

$$\mathbf{q_x} = \frac{d_x}{\ell_x}.$$

Le tavole ISTAT 2013

Età	ℓ_x	Età	ℓ_x	Età	ℓ_x	Età	ℓ_x	Età	ℓ_x
0	100000	11	99543	22	99239	33	98651	44	97667
1	99662	12	99534	23	99191	34	98589	45	97521
2	99636	13	99524	24	99141	35	98521	46	97362
3	99618	14	99512	25	99089	36	98450	47	97184
4	99605	15	99496	26	99036	37	98375	48	96990
5	99596	16	99475	27	98983	38	98296	49	96778
6	99587	17	99448	28	98930	39	98210	50	96545
7	99578	18	99415	29	98877	40	98120	51	96286
8	99569	19	99376	30	98824	41	98021	52	96000
9	99560	20	99332	31	98769	42	97913	53	95686
10	99552	21	99286	32	98711	43	97795	54	95341

Tavola ISTAT della mortalità italiana, anno 2013, maschi

Le tavole ISTAT 2013

Età	ℓ_x	**Età**	ℓ_x	**Età**	ℓ_x	**Età**	ℓ_x	**Età**	ℓ_x
56	94543	67	86244	78	65835	89	24431	99	1419
57	94073	68	85045	79	62943	90	20537	100	894
58	93564	69	83712	80	59796	91	16898	101	535
59	93008	70	82269	81	56384	92	13687	102	293
60	92410	71	80649	82	52763	93	10972	103	151
61	91746	72	78915	83	49015	94	8609	104	72
62	91010	73	77080	84	45096	95	6590	105	32
63	90189	74	75125	85	41073	96	4783	106	13
64	89313	75	73079	86	36918	97	3289	107	5
65	88346	76	70882	87	32676	98	2188	108	2

Tavola ISTAT della mortalità italiana, anno 2013, maschi

Le tavole ISTAT 2013

Età	ℓ_x	Età	ℓ_x	Età	ℓ_x	Età	ℓ_x	Età	ℓ_x
0	100000	11	99626	22	99493	33	99262	44	98704
1	99722	12	99619	23	99475	34	99231	45	98613
2	99702	13	99612	24	99457	35	99196	46	98513
3	99688	14	99604	25	99439	36	99159	47	98404
4	99677	15	99594	26	99420	37	99118	48	98285
5	99668	16	99582	27	99402	38	99075	49	98156
6	99660	17	99570	28	99383	39	99028	50	98015
7	99652	18	99556	29	99363	40	98976	51	97854
8	99646	19	99541	30	99341	41	98918	52	97672
9	99639	20	99526	31	99317	42	98855	53	97479
10	99632	21	99509	32	99291	43	98784	54	97271

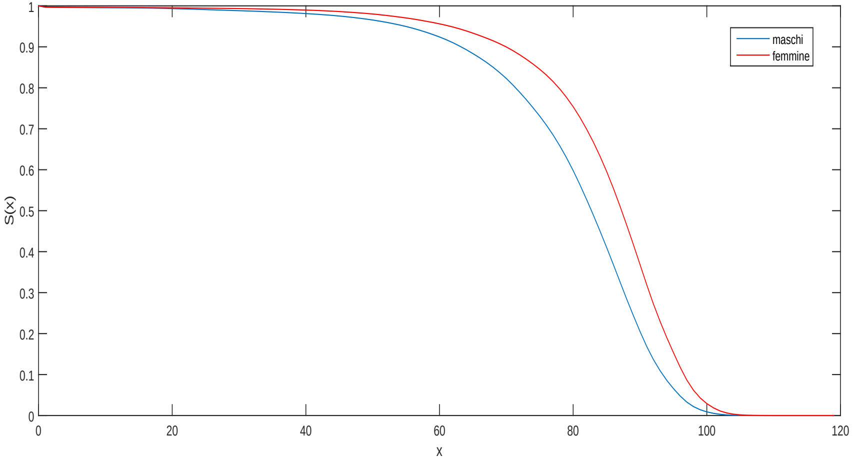
Tavola ISTAT della mortalità italiana, anno 2013, femmine

Le tavole ISTAT 2013

Età	ℓ_x	**Età**	ℓ_x	**Età**	ℓ_x	**Età**	ℓ_x	**Età**	ℓ_x
56	96811	67	92143	78	79714	89	41785	99	4313
57	96541	68	91490	79	77670	90	36890	100	2910
58	96249	69	90762	80	75390	91	31964	101	1871
59	95944	70	89970	81	72809	92	27270	102	1110
60	95613	71	89055	82	69946	93	23000	103	622
61	95248	72	88057	83	66816	94	19063	104	328
62	94845	73	86967	84	63370	95	15382	105	162
63	94392	74	85791	85	59609	96	11839	106	74
64	93893	75	84518	86	55530	97	8693	107	32
65	93342	76	83108	87	51146	98	6198	108	12

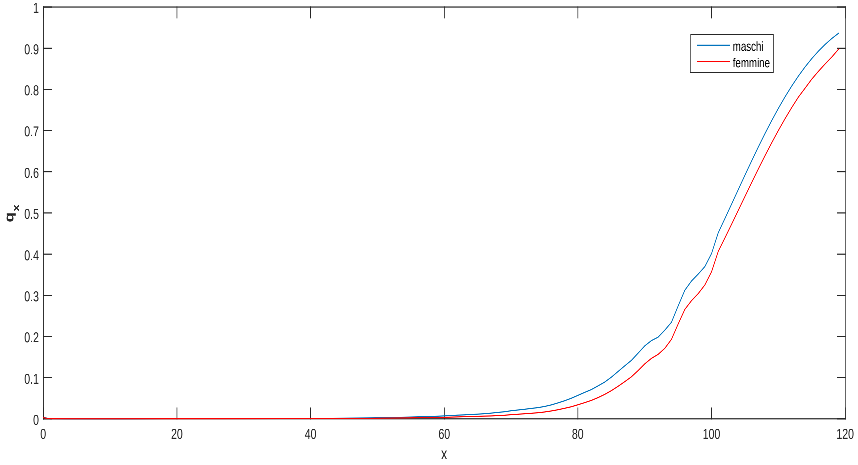
Tavola ISTAT della mortalità italiana, anno 2013, femmine

Confronto tra le funzioni di sopravvivenza



Confronto tra le funzioni di sopravvivenza dei maschi e delle femmine nelle tavole ISTAT 2013

Il tasso di mortalità



Confronto tra i tassi di mortalità dei maschi e delle femmine nelle tavole ISTAT 2013

Alcuni esempi numerici

Utilizzando le tavole ISTAT 2013 calcoliamo alcune probabilità:

$$\mathbf{{}_{20}p_{20}} = \frac{\ell_{40}}{\ell_{20}} = \begin{cases} \frac{98120}{99332} = 98,78\% & \text{maschi} \\ \frac{98976}{99526} = 99,45\% & \text{femmine} \end{cases}$$

$$\mathbf{{}_{20}q_{20}} = \frac{\ell_{20} - \ell_{40}}{\ell_{20}} = \begin{cases} \frac{99332-98120}{99332} = 1,12\% & \text{maschi} \\ \frac{99526-98976}{99526} = 0,55\% & \text{femmine} \end{cases}$$

$$\mathbf{{}_{25}p_{45}} = \frac{\ell_{70}}{\ell_{45}} = \begin{cases} \frac{82269}{97521} = 84,36\% & \text{maschi} \\ \frac{89970}{98613} = 91,24\% & \text{femmine} \end{cases}$$

$$\mathbf{{}_{20|10}q_{20}} = \frac{\ell_{40} - \ell_{50}}{\ell_{20}} = \begin{cases} \frac{98120-96545}{99332} = 1,59\% & \text{maschi} \\ \frac{98976-98015}{99526} = 0,97\% & \text{femmine} \end{cases}$$

Tavole di mortalità selezionate

Oltre all'età, la mortalità ovviamente dipende da molti fattori (genere, fumatori/non fumatori,...). Empiricamente, si osserva che la mortalità di chi sottoscrive un contratto di assicurazione è tipicamente, nei primi anni dopo la sottoscrizione, più bassa della media. Per tenere conto di questo fenomeno gli attuari utilizzano le cosiddette tavole di mortalità selezionate, in cui l'età dell'assicurato x viene decomposta in età della sottoscrizione e tempo trascorso dalla sottoscrizione. Ad esempio, la notazione standard

$$\mathbf{q_{[40]+5}}$$

indica la probabilità di morte per un 45enne che abbia sottoscritto la polizza 5 anni fa, cioè all'età di 40 anni. L'effetto della selezione tende a ridursi e a scomparire dopo qualche anno, quindi si assume solitamente che

$$\mathbf{q_{[x]}} < \mathbf{q_{[x-1]+1}} < \mathbf{q_{[x-2]+2}} < \cdots < \mathbf{q_x}.$$

Esercizi

Esercizio

Usando le tavole di mortalità ISTAT 2013 maschi fornite nelle slides precedenti, calcolare le seguenti probabilità e descriverne il significato:

a) $\mathbf{p_{30}}$

b) $\mathbf{q_{30}}$

c) $\mathbf{{}_{10}p_{30}}$

d) $\mathbf{{}_{10}q_{30}}$

e) $\mathbf{{}_{i|1}q_{30}}$ per $i = 0, \ldots, 9$.

Verificare che la somma delle probabilità calcolate al punto e) coincide con la probabilità calcolata al punto d).

Esercizio

Sapendo che $\mathbf{p_x} = 0.99$, $\mathbf{p_{x+1}} = 0.985$, $\mathbf{{}_3p_{x+1}} = 0.95$, $\mathbf{q_{x+3}} = 0.02$, calcolare:

a) $\mathbf{p_{x+3}}$
b) $\mathbf{{}_2p_x}$
c) $\mathbf{{}_2p_{x+1}}$
d) $\mathbf{{}_3p_x}$
e) $\mathbf{{}_{1|2}q_x}$.

Esercizi

Esercizio

Considerare la funzione

$$S_0(x) = \frac{1}{10}\sqrt{100 - x}, \quad x = 0, \ldots, 100.$$

a) Disegnare il grafico di S_0 e verificare che può rappresentare una funzione di sopravvivenza
b) Quanto vale l'età massima ω in questo caso?
c) Calcolare la probabilità che un nuovo nato muoia tra le età di 19 e 36 anni

Esercizi

Esercizio

Considerare la funzione

$$S_0(x) = \frac{18000 - 110x - x^2}{18000}.$$

a) Disegnare il grafico di S_0 e verificare che può rappresentare una funzione di sopravvivenza
b) Quanto vale l'età massima ω in questo caso?
c) Calcolare ${}_{20}\mathbf{p}_0$
d) Calcolare $S_{20}(t)$
e) Calcolare la probabilità che un individuo di venti anni muoia tra le età di trenta e quaranta anni.