Il modello probabilistico per la durata della vita umana

Dr. Salvatore Scognamiglio

Università degli Studi di Napoli "Parthenope"

Lezioni di tecniche attuariali per le assicurazioni

La Durata della vita umana come variabile casuale T_0

Si consideri fissato un individuo e si indichi con T_0 la *durata della sua vita* espressa in anni, cioè l'età dell'individuo alla data di decesso, è una variabile aleatoria che assume valori reali e positivi.

Anche se è ragionevole assumere che la variabile aleatoria T_0 sia superiormente limitata, cioè che l'individuo prima o poi muoia, a volte vengono utilizzati dei modelli probabilistici in cui non si pone limite alla durata della vita umana. Per questo motivo indicheremo con $\omega = sup T_0$, lasciando aperta la possibilità che sia $\omega < +\infty$ o che sia $\omega = +\infty$.

In entrambi i casi si assumerà che un individuo non può raggiungere l'età ω. Quindi in entrambi i casi, si ha che $T_0 < \omega$ e quindi $T_0 \in (0, \omega)$.

La Funzione di Sopravvivenza

Si indichi, per $x \in [0, \omega)$, con $F_0(x) = \mathbb{P}_t(T_0 \leq x)$ la *funzione di ripartizione* della variabile aleatoria T_0, che esprime la probabilità che l'individuo raggiunga al più l'età x, cioè che muoia entro il suo x-esimo anno di vita.

La *funzione di sopravvivenza* è la funzione di ripartizione complementare:

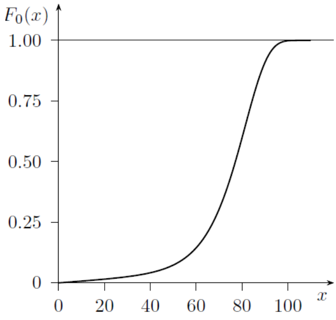
$$S(x) = \mathbb{P}_t(T_0 > x) = 1 - F_0(x)$$

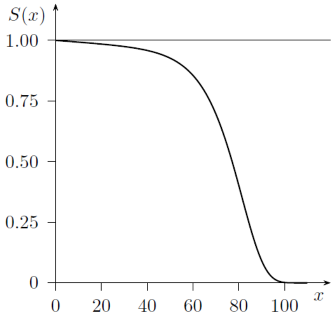
ed esprime quindi la probabilità che l'individuo non muoia nei primi x anni di vita, cioè che sia in vita all'età x.

La funzione $F_0(x)$ è monotona non decrescente mentre la funzione $S(x)$ è monotona non crescente e risulta

$$F_0(0) = 0, \qquad \lim_{x \to \omega^-} F_0(x) = 1,$$

$$S(0) = 1, \qquad \lim_{x \to \omega^-} S(x) = 0.$$





La Variabile Casuale T_x

Si fissi un istante di valutazione t ("oggi") e si consideri un individuo che, al tempo t, sia in vita ed abbia età $x \geq 0$ anni. Si indichi con:

$$T_x = T_0 - x$$

la durata residua della sua vita. É una variabile aleatoria che assume le sue determinazioni nell'intervallo $(0, \omega_x)$, dove:

$$\omega_x = \begin{cases} \omega - x & se \quad \omega < +\infty \\ +\infty & se \quad \omega = +\infty \end{cases}$$

è il limite superiore alla durata residua della vita dell'individuo di età x alla data di valutazione.

Osservazione: Assumeremo che la variabilie aleatoria T_x, sia definita solo se l'individuo sia in vita ed abbia età x alla data di valutazione, cioè quando $T_0 > x$, lasciandola non definita nel caso, peraltro poco interessante, che $T_0 \leq x$.

La Funzione di Ripartizione F_x

Alla data di valutazione la funzione di ripartizione della variabile aleatoria T_x definita per $k \in [0, \omega_x)$, e

$$F_x(k) = \mathbb{P}_t(T_x \leq k) = \mathbb{P}_t(T_0 \leq x + k | T_0 \geq x).$$

Essa misura la probabilità che l'individuo raggiunga al più l'età $x + k$, condizionata all'informazione disponibile alla data di valutazione, cioè all'evento che abbia raggiunto l'età x. Per tenere conto dell'informazione $T_0 > x$ si è usato la probabilità condizionata.

Alcune Proprietà

Usando la definizione di probabilità condizionata si può esprimere la funzione di ripartizione F_x tramite la funzione di ripartizione F_0 o tramite la funzione di sopravvivenza S:

$$\begin{aligned} F_x(k) &= \frac{\mathbb{P}_t(x < T_0 \leq x+k)}{\mathbb{P}_t(T_0 > x)} \\ &= \frac{F_0(x+k) - F_0(x)}{1 - F_0(x)} \\ &= \frac{S(x) - S(x+k)}{S(x)} \\ &= 1 - \frac{S(x+k)}{S(x)}. \end{aligned}$$

Relazioni che mostrano che la distribuzione di probabilità della durata residua della vita di un individuo di età corrente x è completamente nota quando si conosce una delle seguenti:

- la funzione di ripartizione $F_x(k)$ per $k \in [0, \omega_x)$;
- la funzione di ripartizione $F_0(k)$ per $k \in [x, \omega)$;
- la funzione di sopravvivenza $S(k)$ per $k \in [x, \omega)$.

Notazione Attuariale Standard (1)

Nella matematica attuariale sono in uso da molti anni e in tutto il mondo alcune notazioni standard per indicare la probabilità di sopravvivenza e di morte.

Sia $k \in [0, \omega_x)$; le probabilità di sopravvivenza dopo k anni e di morte entro k anni dell'individuo di età x alla data di valutazione si indicano con:

$$ {}_k p_x = \mathbb{P}_t(T_x > k) = 1 - F_x(k) = \frac{S(x+k)}{S(x)}, $$

$$ {}_k q_x = \mathbb{P}_t(T_x \leq k) = F_x(k) = 1 - {}_k p_x = 1 - \frac{S(x+k)}{S(x)}. $$

Notazione Attuariale Standard (2)

Si usa la convenzione di omettere l'indice sinistro nel caso di $k = 1$ e si scrive semplicemente:

$$q_x = {}_1q_x \qquad p_x = {}_1p_x\,.$$

Dalle proprietà della funzione di ripartizione, agli estremi del dominio della variabile k risulta:

$$\begin{aligned} {}_0p_x &= 1, \qquad & \lim_{k \to \omega_x} {}_kp_x &= 0, \\ {}_0q_x &= 0, \qquad & \lim_{k \to \omega_x} {}_kq_x &= 1. \end{aligned}$$

Il comportamento di ${}_kp_x$ e ${}_kq_x$ all'estremo ω_x esprime il fatto che il nostro individuo morirà al più tardi fra ω_x anni.

La probabilità di Morte Differita

La probabilità di morte ${}_hq_x$ è la probabilità che l'individuo muoia in un intervallo di tempo di durata h anni che parte dalla data di valutazione.

Nella trattazione delle polizze vita occorrerà considerare anche la probabilità che l'individuo muoia in un intervallo di tempo di durata h anni, ma con inizio dopo k anni dalla data di valutazione.

Si chiama *probabilità di morte differita* e si indica con:

$$ {}_{k|h}q_x = \mathbb{P}_t(k < T_x \leq k+h) = \mathbb{P}_t(x+k < T_0 \leq x+k+h |\, T_0 > x). $$

La sbarretta verticale nell'indice sinistro separa l'entità del differimento k dalla durata h del periodo di riferimento per il calcolo della probabilità.

Alcune Proprietà

Usando le proprietà della funzione di ripartizione F_x risulta che:

$$
\begin{aligned}
{}_{k|h}q_x &= F_x(k+h) - F_x(k) \\
&= 1 - \frac{S(x+k+h)}{S(x)} - \left(1 - \frac{S(x+k)}{S(x)}\right) \\
&= \frac{S(x+k) - S(x+k+h)}{S(x)}.
\end{aligned}
$$

Si noti che, per definizione:

$$ {}_{0|h}q_x = {}_hq_x $$

e valgono inoltre le seguenti identità notevoli:

$$
\begin{aligned}
{}_{k|h}q_x &= {}_kp_x - {}_{k+h}p_x\,, \\
{}_{k|h}q_x &= {}_{k+h}q_x - {}_kq_x\,.
\end{aligned}
$$

La prima può essere interpretata osservando che l'individuo dovesse morire tra $t+k$ e $t+k+h$, allora dovrà essere stato vivo in $t+k$ e morto (non vivo) in $t+k+h$, la seconda discende dalla prima passando alle probabilità complementari.

La speranza matematica di vita alla nascita e_0

La **speranza matematica di vita** (o **vita media**) alla nascita, indicata con e_0, rappresenta il numero di anni, in media, che un individuo alla nascita si aspetta di vivere.

Questa grandezza, che non è altro che un valore atteso, può essere calcolata:

$$e_0 = \sum_{h=0}^{\omega-1} h \cdot {}_{h|1}q_0 \, ,$$

sviluppando i termini:

$$e_0 = 0 \cdot {}_{|1}q_0 + 1 \cdot {}_{1|1}q_0 + 2 \cdot {}_{2|1}q_0 + \cdots + (\omega - 1) \cdot {}_{\omega-1|1}q_0 \, .$$

La speranza matematica di vita residua e_x

Più in generale, la speranza matematica di vita residua (o vita media residua), per una testa di età x, indicata con e_x è data da:

$$e_x = \sum_{h=0}^{\omega-x-1} h \cdot {}_{h|1}q_x .$$

Si osservi che l'aspettativa di vita residua diminuisce all'aumenta dell'età x.

Inoltre, per una testa di età x, tra la speranza matematica di vita alla nascita e quella all'eta x sussiste la seguente relazione:

$$e_0 < x + e_x.$$

Alcune semplificazioni del modello probabilistico della Vita Umana (1)

Il modello probabilistico è stato sviluppato fino ad ora in riferimento ad:

- una fissata data di valutazione t,
- un fissato individuo, di età x alla data di valutazione.

In teoria occorrerebbe quindi avere a disposizione una diversa funzione di sopravvivenza per ogni distinto individuo che si intende considerare, ma questo è inconcepibile nella pratica: una compagnia di assicurazioni ha tipicamente da alcune decine di migliaia a svariante centinaia di migliaia di assicurati.

Alcune semplificazioni del modello probabilistico della Vita Umana (2)

Un'estensione universalmente accettata è quella di considerare collettivi omogenei (per esempio per sesso) di coetanei ad assegnargli un unica funzione di sopravivvenza, tipicamente stimata con tecniche statistiche.

Possono tuttavia esserci difficoltà anche in questo approccio, come accade in Italia dove non sono disponibili statistiche su base generazionale.
Per questo motivo si ricorre spesso a una ulteriore approssimazione e si una stessa funzione di sopravvivenza "media" per individui (o collettivi) di età diversa alla data di valutazione.

Alcune semplificazioni del modello probabilistico della Vita Umana (3)

Per illustrare le conseguenze di quest'ipotesi, si considerino due individui di età x e $y = x + k$ alla data di valutazione t con $k > 0$. Se si indicano con S^x e S^y le funzioni di sopravvivenza dei due individui, la probabilità del primo di essere in vita in $t + k + h$, condizionata alla sua sopravvivenza dopo i primi k anni è:

$$ {}_h p_{x+k} = \frac{S^x(x+k+h)}{S^x(x+k)} $$

mentre la probabilità del secondo di essere in vita $t + h$ è:

$$ {}_h p_y = \frac{S^y(y+h)}{S^y(y)} = \frac{S^y(x+k+h)}{S^y(x+k)} $$

le due espressioni differiscono unicamente per la diversa funzione di sopravvivenza, che è quella specifica di ciascuno dei due individui.

Se invece si usa un'unica funzione di sopravvivenza per entrambi, allora $S^x = S^y$ e le due probabilità coincidono.
In particolare ciò comporta che la distribuzione di probabilità della durata residua della vita dell'individuo con età maggiore coincide con la distribuzione di probabilità della durata di vita dell'individuo più giovane, condizionata al fatto che egli risulti in vita all'età $x + k$.

La distribuzione di probabilità $\mathbb{P}_s$ (1)

Le distribuzioni di probabilità che abbiamo fino a qui considerato sono tutte legate ad una data di valutazione t.

Sorge spesso la necessità di considerare la distribuzione di probabilità della durata della vita di un individuo ad una data futura s.

Naturalmente, alla data t, la distribuzione di probabilità $\mathbb{P}_s$ è aleatoria, perchè dipende da eventi che possono accadere in $[t, s]$ e che possono modificare le opinioni probabilistiche sulla durata della vita dell'individuo in particolare o di una collettività in generale. Se però si assume che queste non mutino, e quest'ipotesi adotta si costantemente per le basi tecniche del I ordine, allora la distribuzione di probabilità della durata della vita al tempo s dipende solo dall'età dell'individuo a quella data ed è determinata dalla stessa funzione di sopravvivenza.

La distribuzione di probabilità $\mathbb{P}_s$ (2)

In particolare la probabilità ${}_hp_{x+k}$, che nell'ipotesi di un'unica funzione di sopravvivenza per tutte le generazioni, assumeva i due significati coincidenti di probabilità alla data t per l'individuo di età $x+k$ e di probabilità condizionata alla data t per l'individuo di età x, assuma ora il terzo significato di probabilità alla data s per l'individuo di età x alla data t.

Tavole di Mortalità (1)

Spesso la funzione di sopravvivenza viene assegnata per valori interi di x; viene cioè tabulata per $x = 0, 1, \ldots, \omega$ e questa tabulazione prende il nome di **tavola di mortalità o sopravvivenza**.

Le tavole di mortalità si costruiscono a partire da una collettività I_α di persone coetanee di età $\alpha \in [0, \omega]$. Si supponga che tale collettività sia chiusa a nuovi ingressi (si tratta quindi di una generazione) e che l'unica causa di uscita sia data dal decesso.

Al crescere di t, la probabilità di sopravvivenza si riduce, tendendo verso zero. Naturalmente vale che $S(0) = 1$ e $S(\omega) = 0$.

Tavole di Mortalità (2)

Esistono diversi tipi di tavole di mortalità:

- *tavole di "generazione"* (rilevamento longitudinale): data una generazione nata in $t = 0$ si rilevano i superstiti della collettività ad ogni epoca fino alla totale estinzione della popolazione considerata. Tale rilevazione richiede un periodo di osservazione pari a ω anni (solitamente 100-110).
- *tavole di "contemporanei"* per una data collettività, ad esempio residenti in un paese, in un periodo di osservazione di durata fissata, ad esempio 10 anni, è rilevata la mortalità delle persone appartenenti alla collettività ed aventi tutte le varie età $(0 < x < \omega - 1)$.

Tavole di Mortalità (3)

Nella sezione “demo.istat” vengono raccolti e messi a disposizione i dati ufficiali piu recenti sulla popolazione residente nei comuni Italiani derivanti dalle indagini effetuate presso gli Uffici di Anagrafe.

L’ISTAT colleziona dati e mette a disposizione tavolte di mortalità di “contemporanei” dal 1974 al 2018 sia per la popolazione totale che effettuando la distinzione fra maschi e femmine.

Esempio: Tavole di mortalità (Italia - Maschi - 2016)

x	l_x	d_x	$q_x \cdot 10^4$	e_x
0	100000	323	3.23192	80.562
1	99677	17	0.17512	79.822
2	99659	14	0.14490	78.836
3	99645	12	0.11881	77.847
4	99633	10	0.09838	76.856
5	99623	9	0.08618	75.864
6	99615	8	0.07736	74.870
7	99607	8	0.07979	73.876
...	...	...	...	...
...	...	...	...	...
100	1204	465	386.45284	1.881
101	739	301	407.42077	1.751
102	438	193	440.22362	1.611
103	245	116	473.68132	1.485
104	129	65	507.52015	1.372
105	64	34	541.44804	1.270
106	29	17	575.17224	1.178
107	12	8	608.40239	1.097
108	5	3	640.86841	1.024
109	2	1	672.31966	0.959
110	1	0	702.53907	0.902
...	...	...	...	...
...	...	...	...	...

Tavole di Mortalità (4)

Le grandezze che entrano in gioco in una tavola di mortalità sono:

- Il numero di sopravviventi, l_x sono coloro che, provenienti dalla popolazione iniziale fittizia di 100.000 nati, sopravvivono ai vari compleanni. I sopravviventi sono legati alle probabilità di morte dalla seguente relazione:
$$l_{x+1} = l_x(1 - q_x)$$
- Il numero di decessi d_x, sono coloro che muoiono tra il compleanno x e il compleanno $x + 1$ e pertanto:
$$d_x = l_x q_x$$
- la probabilità di morte q_x calcolata come:
$$q_x = \frac{d_x}{l_x}$$
- speranza di vita: rappresenta il numero medio di anni che restano da vivere ai sopravviventi all'età x.

Tavole di Mortalità (5)

Le immagini della funzione di sopravvivenza $S(x)$ possono essere ottenuti per valori $x = 0, 1, \ldots, \omega$ a partire dalla tavola di mortalità:

$$S(x) = \frac{l_x}{l_0},$$

Dai cui è facile desumere i valori di $F_0(\cdot)$ e $F_x(\cdot)$ per $x = 0, 1, \ldots, \omega$.

Esempio Numerico (1)

Si consideri la SIM2016 pubblicata dall'ISTAT.

Determinare le probabilità che una testa maschile di età 25:

1. viva almeno un anno,
2. sia in vita all'età 60,
3. non superi i 60 anni,
4. viva solo un anno,
5. muoia all'età 60,
6. muoia fra l'età di 60 e 70.

Esempio Numerico 1: Risultati (1)

1. la probabilità che viva almeno un anno si può calcolare utilizzando la formula (8):

$$
{}_1p_{25} = \mathbb{P}_t(T_{25} > 1) = \mathbb{P}_t(T_0 > 26 | T_0 > 25) = \frac{S(25+1)}{S(25)} =
$$

$$
= \frac{S(26)}{S(25)} = 0.999546
$$

2. Lo stesso vale per la probabilità che la testa sia in vita all'età 60:

$$
{}_{35}p_{25} = \mathbb{P}_t(T_{25} > 35) = \mathbb{P}_t(T_0 > 60 | T_0 > 25) =
$$

$$
= \frac{S(25+35)}{S(25)} = \frac{S(60)}{S(25)} = 0.9368282
$$

3. la probabilità di morire non dopo i 60 anni è:

$$
{}_{35}q_{25} = \mathbb{P}_t(T_{25} \leq 35) = \mathbb{P}_t(T_0 \leq 60 | T_0 > 25) = 1 - {}_{35}p_{25} = 0.063172
$$

Esempio Numerico 1: Risultati (2)

4. la probabilità che viva solo un anno si può ottenere calcolando la probabilità di morte differita con $x = 25, k = 1, h = 1$ applicando la formula in (1):

$$ {}_{1|1}q_{25} = \mathbb{P}_t(1 < T_{25} \leq 2) = \mathbb{P}_t(26 < T_0 \leq 27 | T_0 > 25) = $$

$$ = \frac{S(26) - S(27)}{S(25)} = 0.0004537 $$

5. mentre la probabilità che muoia all'età 60:

$$ {}_{35|1}q_{25} = \mathbb{P}_t(35 < T_{25} \leq 36) = \mathbb{P}_t(60 < T_0 \leq 61 | T_0 > 25) = $$

$$ = \frac{S(60) - S(61)}{S(25)} = 0.0062617 $$

6. in analogo, la probabilità che muoia fra l'età di 60 e 70:

$$ {}_{35|10}q_{25} = \mathbb{P}_t(35 < T_{25} \leq 45) = \mathbb{P}_t(60 < T_0 \leq 70 | T_0 > 25) = $$

$$ = \frac{S(60) - S(70)}{S(25)} = 0.095702 $$

Esempio Numerico (2)

Si consideri la SIM2016 pubblicata dall'ISTAT.

Determinare le probabilità che una testa femminile di età 40:

1. viva almeno due anni,
2. sia in vita all'età 70,
3. non superi l'età 70,
4. viva solo due anni,
5. muoia all'età 70,
6. muoia fra l'età di 70 e 75.

Esempio Numerico 2: Risultati (1)

1. la probabilità che viva almeno due anni è:

$$ {}_{2}p_{40} = \mathbb{P}_t(T_{40} > 2) = \mathbb{P}(T_0 > 42 | T_0 > 40) = \frac{S(40+2)}{S(40)} = $$

$$ = \frac{S(42)}{S(40)} = 0.998778 $$

2. Lo stesso vale per la probabilità che la testa sia in vita all'età 70:

$$ {}_{30}p_{40} = \mathbb{P}_t(T_{40} > 30) = \mathbb{P}(T_0 > 70 | T_0 > 40) = \frac{S(40+30)}{S(40)} = $$

$$ = \frac{S(70)}{S(40)} = 0.913278 $$

3. la probabilità di non superare i 70 anni è:

$$ {}_{30}q_{40} = \mathbb{P}_t(T_{40} \leq 30) = \mathbb{P}(T_0 \leq 70 | T_0 > 40) = 1 - {}_{30}p_{40} = 0.086722 $$

Esempio Numerico 2: Risultati (2)

4. la probabilità che viva solo due anni si può ottenere calcolando la probabilità di morte differita con $x = 40$, $k = 2$, $h = 1$ applicando la formula in (1):

$$ {}_{2|1}q_{40} = \mathbb{P}_t(2 < T_{40} \leq 3) = \mathbb{P}(42 < T_0 \leq 43 | T_0 > 40) = $$

$$ \frac{S(42) - S(43)}{S(40)} = 0.000677 $$

5. mentre la probabilità che muoia all'età 70:

$$ {}_{30|1}q_{40} = \mathbb{P}_t(30 < T_{40} \leq 31) = \mathbb{P}(70 < T_0 \leq 71 | T_0 > 40) = $$

$$ \frac{S(70) - S(71)}{S(40)} = 0.008353 $$

6. in analogo, la probabilità che muoia fra l'età di 60 e 70:

$$ {}_{30|5}q_{40} = \mathbb{P}_t(30 < T_{40} \leq 35) = \mathbb{P}(70 < T_0 \leq 75 | T_0 > 40) = $$

$$ \frac{S(70) - S(75)}{S(40)} = 0.053502 $$

Esempio Numerico (3)

Completare con i valori mancanti la seguente tavola di mortalità:

x	l_x	d_x	q_x
55		976	
56	87165		0.01228
57			
58	84940	1235	
59			
60	82345		0.01756

Esempio Numerico (3): Risultati

Completare con i valori mancanti la seguente tavola di mortalità:

x	l_x	d_x	q_x
55	**88141**	976	**0.01107**
56	87165	**1070**	0.01228
57	**86095**	**1155**	**0.01341**
58	84940	1235	**0.01454**
59	**83705**	**1360**	**0.01525**
60	82345	**1446**	0.01756