La modellizzazione della mortalità

Fabio Bellini

Università di Milano-Bicocca

fabio.bellini@unimib.it

9 novembre 2020

La forza di mortalità

Torniamo al caso di un modello probabilistico in cui la durata della vita alla nascita è espressa dalla variabile casuale T_0 con funzione di sopravvivenza $S_0(x)$ e funzione di densità

$$f_0(x) = \frac{d}{dx}F_0(x) = \frac{d}{dx}(1 - S_0(x)) = -\frac{d}{dx}S_0(x).$$

La *forza di mortalità* μ_x è definita come il tasso di mortalità q_x per unità di tempo, per intervalli di tempo molto piccoli. Più precisamente, definiamo

$$\begin{aligned}\mu_x &= \lim_{\Delta x \to 0} \frac{\mathbb{P}(x \leq T_x \leq x + \Delta x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{F_x(x + \Delta x) - F_x(x)}{\Delta x} = \\ &= \lim_{\Delta x \to 0} \frac{F_0(x + \Delta x) - F_0(x)}{S_0(x)\Delta x} = \frac{f_0(x)}{S_0(x)}.\end{aligned}$$

La quantità μ_x prende anche il nome di *hazard rate*.

La distribuzione esponenziale

Calcoliamo la forza di mortalità nel caso della distribuzione esponenziale. Se $T_0 \sim Exp(\lambda)$, allora

$$f_0(x) = \lambda \exp(-\lambda x), \text{ con } x \geq 0 \text{ e } \lambda > 0,$$

e si ha semplicemente

$$S_0(x) = 1 - F_0(X) = \exp(-\lambda x).$$

Applicando la definizione di forza di mortalità, si ha che

$$\mu_x = \frac{f_0(x)}{S_0(x)} = \frac{\lambda \exp(-\lambda x)}{\exp(-\lambda x)} = \lambda,$$

cioè *la distribuzione esponenziale ha una forza di mortalità costante*. Si tratta della cosiddetta *proprietà di assenza di memoria* della distribuzione esponenziale; in altri termini, in questo modello non è presente *invecchiamento*.

Esempio

Riprendiamo l'esempio

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

In questo caso la funzione di densità è data da

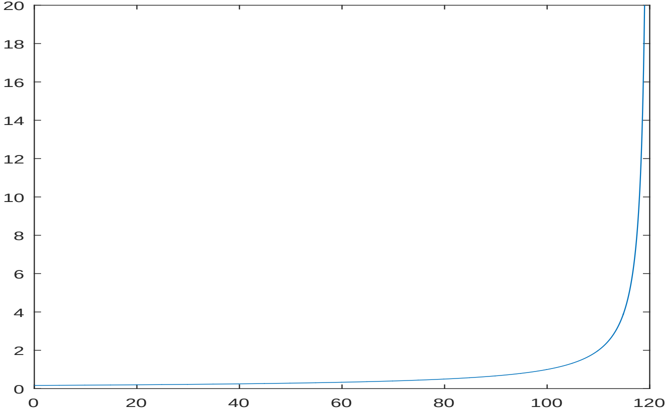
$$f_0(x) = -\frac{d}{dx} S_0(x) = \frac{1}{720} \left(1 - \frac{x}{120}\right)^{1/6-1}, \quad 0 \leq x \leq 120,$$

da cui applicando la definizione della forza di mortalità otteniamo

$$\mu_x = \frac{\frac{1}{720}\left(1 - \frac{x}{120}\right)^{1/6-1}}{\left(1 - \frac{x}{120}\right)^{1/6}} = \frac{1}{6(120 - x)}.$$

In questo caso la forza di mortalità è crescente e presenta addirittura un asintoto verticale per $x = \omega = 120$ (vedere il grafico nella slide successiva).

Esempio



Forza di mortalità μ_x con asintoto verticale per $\omega = 120$.

Dalla forza di mortalità alla funzione di sopravvivenza

Conoscendo la forza di mortalità μ_x è possibile ricavare la corrispondente funzione di sopravvivenza $S_0(x)$. Si ha innanzitutto che la forza di mortalità può anche essere scritta come

$$\mu_x = -\frac{d}{dx} \log S_0(x),$$

in quanto calcolando la derivata a secondo membro con il teorema di derivazione delle funzioni composte otteniamo

$$-\frac{d}{dx} \log S_0(x) = -\frac{1}{S_0(x)} \frac{d}{dx} S_0(x) = \frac{f_0(x)}{S_0(x)},$$

che è proprio la definizione della forza di mortalità. Integrando tra 0 e t entrambi i membri della prima equazione, otteniamo

$$\int_0^t \mu_x \, dx = -\int_0^t \frac{d}{dx} \log S_0(x) = -[\log S_0(x)]_0^t = -\log S_0(t) + \log S_0(0) =$$
$$= -\log S_0(t),$$

Dalla forza di mortalità alla funzione di sopravvivenza

poiché $S_0(0) = 1$. Cambiando i segni e facendo l'esponenziale di entrambi i membri otteniamo infine la formula

$$S_0(t) = \exp\left(-\int_0^t \mu_x \, dx\right).$$

Notate la analogia con la matematica finanziaria tradizionale:

$$\begin{aligned} \text{funzione di sopravvivenza } S_0 &\longleftrightarrow \text{ fattore di attualizzazione } v \\ \text{forza di mortalità } \mu_x &\longleftrightarrow \text{ forza di interesse } \delta \\ \text{distribuzione esponenziale } &\longleftrightarrow \text{ regime sconto composto} \\ \text{proprietà di assenza di memoria } &\longleftrightarrow \text{ scindibilità} \end{aligned}$$

Nel caso $\mu_x = \lambda =$ costante, possiamo infatti verificare che

$$S_0(t) = \exp\left(-\int_0^t \mu_x \, dx\right) = \exp(-\lambda t).$$

Dalla forza di mortalità alla funzione di sopravvivenza

Dalla forza di mortalità μ_x possiamo anche ricavare anche una formula per la funzione di sopravvivenza condizionata $S_x(t)$:

$$S_x(t) = \frac{S_0(x+t)}{S_0(x)} = \frac{\exp\left(-\int_0^{x+t} \mu_s \, ds\right)}{\exp\left(-\int_0^{x} \mu_s \, ds\right)} = \exp\left(-\int_x^{x+t} \mu_s \, ds\right)$$

Spesso le leggi di mortalità sono espresse direttamente in funzione di μ_x, piuttosto che di $S_0(t)$, in quanto la forza di mortalità presenta un più chiaro significato modellistico. La funzione di sopravvivenza alla nascita $s_0(t)$ e la funzione di sopravvivenza condizionata $S_x(t)$ possono essere poi calcolate con le formule appena ricavate. Nelle slides successive vediamo i due esempi più semplici: la legge di Gompertz e la legge di Makeham.

La legge di Gompertz

La *legge di Gompertz* è data da

$$\mu_x = Bc^x,$$

dove le costanti B e c soddisfano $0 < B < 1$ e $c > 1$. Questa legge cattura nel modo più semplice possibile il fenomeno dell'invecchiamento, con una forza di mortalità che cresce esponenzialmente in funzione della età x. Ricaviamo la espressione della funzione di sopravvivenza alla nascita $S_0(t)$ utilizzando la relazione

$$S_0(t) = \exp\left(-\int_0^t \mu_s \, ds\right).$$

La legge di Gompertz

Si ottiene

$$S_0(t) = \exp\left(-\int_0^t Bc^s\,ds\right) = \exp\left(-B\int_0^t c^s\,ds\right) = \exp\left(-B\left[\frac{c^s}{\log c}\right]_0^t\right) =$$
$$= \exp\left(-B\left[\frac{c^t-1}{\log c}\right]\right) = \exp\left(\frac{B}{\log c}\left[1-c^t\right]\right).$$

Verificate per esercizio che $S_0(0) = 1$, $\lim_{t\to+\infty} S_0(t) = 0$, e che $S_0(t)$ è una funzione decrescente. In modo analogo partendo dalla relazione

$$S_x(t) = \exp\left(-\int_x^{x+t} \mu_s\,ds\right),$$

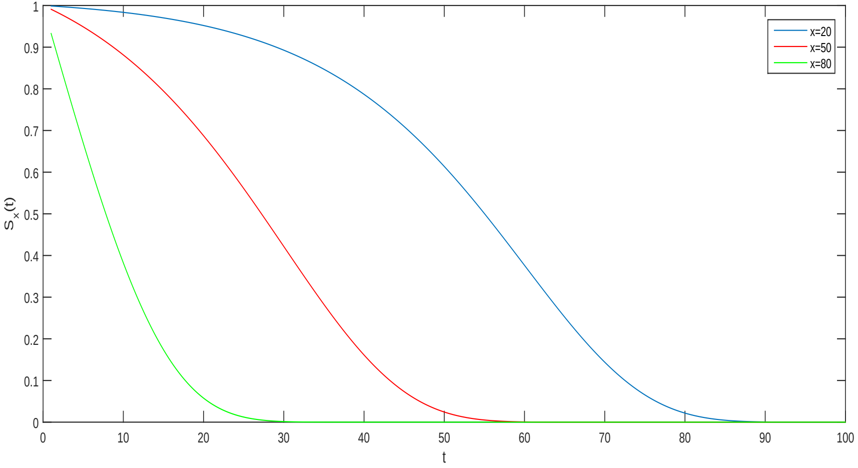
è possibile ricavare nel modello di Gompertz la funzione di sopravvivenza condizionata

La legge di Gompertz

$$
\begin{aligned}
S_x(t) &= \exp\left(-\int_x^{x+t} Bc^s\, ds\right) = \exp\left(-B\int_x^{x+t} c^s\, ds\right) = \\
&= \exp\left(-B\left[\frac{c^s}{\log c}\right]_x^{x+t}\right) = \exp\left(-B\left[\frac{c^{x+t}-c^x}{\log c}\right]\right) = \\
&= \exp\left(\frac{Bc^x}{\log c}\left[1-c^t\right]\right).
\end{aligned}
$$

Anche qui è facile verificare che $S_x(0) = 1$, $\lim_{t\to+\infty} S_x(t) = 0$, e che $S_x(t)$ è una funzione decrescente. Possiamo inoltre osservare che all'aumentare di x, per t fissato, la funzione di sopravvivenza condizionata $S_x(t)$ è decrescente, in quanto il termine $[1 - c^t]$ è negativo; questo modello presenta dunque invecchiamento. Nella slide successiva riportiamo i grafici di $S_x(t)$ nel caso $B = 0.0003$ e $c = 1.07$, per $x = 20, 50, 80$.

La legge di Gompertz



La legge di Makeham

La *legge di Makeham* è una immediata generalizzazione delle legge di Gompertz, con un parametro aggiuntivo $A > 0$:

$$\mu_x = A + Bc^x$$

Con lo stesso procedimento seguito nel caso della legge di Gompertz ricaviamo

$$S_0(t) = \exp\left(-At + \frac{B}{\log c}\left[1 - c^t\right]\right)$$
$$S_x(t) = \exp\left(-At + \frac{Bc^x}{\log c}\left[1 - c^t\right]\right).$$

Aspettativa di vita

Ricordiamo che la variabile casuale T_x rappresenta la vita residua di un individuo di età x. Nella terminologia attuariale tradizionale la sua media prende il nome di *aspettativa di vita completa* e si indica con la notazione attuariale $\mathring{e}_x$:

$$\mathring{e}_x = \mathbb{E}[T_x].$$

In molte applicazioni è utile considerare anche la parte intera della variabile casuale T_x, detta "vita incompleta", indicata con

$$K_x = \lfloor T_x \rfloor,$$

dove la funzione $\lfloor x \rfloor$ indica appunto la parte intera di x. La media di K_x prende il nome di *aspettativa di vita incompleta* e si indica con la notazione attuariale e_x:

$$e_x = \mathbb{E}[K_x].$$

Aspettativa di vita completa

Per calcolare la aspettativa di vita completa utilizziamo una formula per il calcolo della media di qualsiasi variabile casuale X positiva con funzione di sopravvivenza $S(x)$ e funzione di densità $f(x)$. Si ha

$$\mathbb{E}[X] = \int_0^{+\infty} x\, f(x)\, dx,$$

e poiché $-S(x)$ è una primitiva di $f(x)$, integrando per parti otteniamo

$$\int_0^{+\infty} xf(x)\, dx = [-xS(x)]_0^{+\infty} - \int_0^{+\infty} -S(x)\, dx = \int_0^{+\infty} S(x)\, dx,$$

sotto la ipotesi

$$\lim_{x \to +\infty} xS_0(x) = 0^+,$$

che è sempre verificata se la media di X è finita.

Aspettativa di vita completa

Concludiamo che per una generica variabile casuale positiva

$$\mathbb{E}[X] = \int_0^{+\infty} S(x)\, dx,$$

cioè la media coincide con l'integrale tra 0 e $+\infty$ della funzione di sopravvivenza. Ne segue che la aspettativa di vita completa è data da

$$\mathring{e}_x = \int_0^{+\infty} S_x(t)\, dt = \int_0^{+\infty} {}_{\mathbf{t}}\mathbf{p}_{\mathbf{x}}\, dt.$$

Anche il momento secondo di una variabile casuale positiva può essere calcolato integrando per parti:

$$\mathbb{E}[X^2] = \int_0^{+\infty} x^2 f(x)\, dx = \left[-x^2 S(x)\right]_0^{+\infty} - \int_0^{+\infty} -S(x) 2x\, dx =$$

$$= 2\int_0^{+\infty} xS(x)\, dx$$

Varianza della vita completa

sotto la ipotesi

$$\lim_{x \to +\infty} x^2 S(x) = 0^+,$$

che è sempre verificata se $\mathbb{E}[X^2] < +\infty$. Utilizzando questa formula possiamo calcolare il momento secondo e la varianza di T_x:

$$\mathbb{E}[T_x^2] = 2 \int_0^{+\infty} tS_x(t)\, dt = 2 \int_0^{+\infty} t \cdot {}_{\mathbf{t}}\mathbf{p}_{\mathbf{x}}\, dt.$$

$$Var[T_x] = 2 \int_0^{+\infty} tS_x(t)\, dt - (\mathring{e}_x)^2 = 2 \int_0^{+\infty} t \cdot {}_{\mathbf{t}}\mathbf{p}_{\mathbf{x}}\, dt - (\mathring{e}_x)^2.$$

Esempio

Calcoliamo l'aspettativa di vita completa nell'esempio iniziale

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

La funzione di sopravvivenza condizionata è data da

$$S_x(t) = \frac{S(x+t)}{S(x)} = \frac{\left(1 - \frac{x+t}{120}\right)^{1/6}}{\left(1 - \frac{x}{120}\right)^{1/6}} = \left(\frac{120 - (x+t)}{120 - x}\right)^{1/6} =$$

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

Notate che il numero massimo di anni che può vivere l'individuo di età x è pari a $120 - x$, in accordo con $\omega = 120$.

Esempio

Ne segue che

$$\mathring{e}_x = \int_0^{+\infty} S_x(t)\, dt = \int_0^{120-x} \left(1 - \frac{t}{120-x}\right)^{1/6} dt,$$

e l'integrale può essere calcolato con la sostituzione

$$y = 1 - \frac{t}{120-x}, \quad dy = -\frac{dt}{120-x}, \quad dt = -(120-x)dy$$

da cui otteniamo

$$\mathring{e}_x = \int_0^{120-x} \left(1 - \frac{t}{120-x}\right)^{1/6} dt = -\int_1^0 y^{1/6}(120-x)\, dy =$$
$$= (120-x)\int_0^1 y^{1/6}\, dy = \frac{6}{7}(120-x).$$

Esempio

Il momento secondo di T_x è dato da

$$\mathbb{E}[T_x^2] = 2\int_0^{+\infty} tS_x(t)\,dt = 2\int_0^{120-x} t\left(1 - \frac{t}{120-x}\right)^{1/6} dt,$$

e utilizzando come prima la sostituzione

$$y = 1 - \frac{t}{120-x}, \quad t = (120-x)(1-y)$$
$$dy = -\frac{dt}{120-x}, \quad dt = -(120-x)dy,$$

Esempio

otteniamo

$$\begin{aligned}\mathbb{E}[T_x^2] &= 2\int_0^{120-x} t\left(1-\frac{t}{120-x}\right)^{1/6} dt = \\ &= -2(120-x)^2\int_1^0 (1-y)y^{1/6}\,dy = \\ &= 2(120-x)^2\int_0^1 (1-y)y^{1/6}\,dy = 2(120-x)^2\left(\frac{6}{7}-\frac{6}{13}\right),\end{aligned}$$

da cui si ha

$$Var[T_x] = (120-x)^2\left(2(6/7-6/13)-(6/7)^2\right) \simeq 0.05651\cdot(120-x)^2.$$

Nella tabella della prossima slide riportiamo come esempio alcuni valori.

Esempio

x	$\overset{\circ}{e}_x$	$Var[T_x]$	$Std[T_x]$	$x + \overset{\circ}{e}_x$
20	85.71	56.51	7.52	105.71
30	77.14	45.77	6.77	107.14
40	68.57	36.17	6.01	108.57
50	60.00	27.69	5.26	110.00
60	51.43	20.34	4.51	111.43
70	42.86	14.13	3.76	112.86
80	34.29	9.04	3.01	114.29
90	25.71	5.09	2.26	115.71

Possiamo verificare che:

- $\overset{\circ}{e}_x$ è una funzione decrescente di x
- $Std[T_x]$ è una funzione decrescente di x in quanto ci si avvicna a ω
- $x + \overset{\circ}{e}_x$ è una funzione crescente di x.

Aspettativa di vita incompleta

Veniamo ora al calcolo della aspettativa di vita incompleta. K_x è una variabile casuale discreta, che può assumere i valori $0, 1, 2, \ldots$. Si ha

$$e_x = \mathbb{E}[K_x] = \sum_{k=0}^{+\infty} k\mathbb{P}(K_x = k),$$

e inoltre

$$\mathbb{P}(K_x = k) = \mathbb{P}(k < T_x \leq k+1) = {}_{\mathbf{k}|\mathbf{1}}\mathbf{q}_{\mathbf{x}},$$

la probabilità di morte differita tra le età $x + k$ e $x + k + 1$. Ne segue che

$$\begin{aligned} e_x &= \sum_{k=0}^{+\infty} k\mathbb{P}(K_x = k) = \sum_{k=1}^{+\infty} k \cdot {}_{\mathbf{k}|\mathbf{1}}\mathbf{q}_{\mathbf{x}} = \sum_{k=1}^{+\infty} k \cdot ({}_{\mathbf{k}}\mathbf{p}_{\mathbf{x}} - {}_{\mathbf{k+1}}\mathbf{p}_{\mathbf{x}}) = \\ &= ({}_{\mathbf{1}}\mathbf{p}_{\mathbf{x}} - {}_{\mathbf{2}}\mathbf{p}_{\mathbf{x}}) + 2 \cdot ({}_{\mathbf{2}}\mathbf{p}_{\mathbf{x}} - {}_{\mathbf{3}}\mathbf{p}_{\mathbf{x}}) + 3 \cdot ({}_{\mathbf{3}}\mathbf{p}_{\mathbf{x}} - {}_{\mathbf{4}}\mathbf{p}_{\mathbf{x}}) + 4 \cdot ({}_{\mathbf{4}}\mathbf{p}_{\mathbf{x}} - {}_{\mathbf{5}}\mathbf{p}_{\mathbf{x}}) + \ldots \\ &= {}_{\mathbf{1}}\mathbf{p}_{\mathbf{x}} + {}_{\mathbf{2}}\mathbf{p}_{\mathbf{x}} + {}_{\mathbf{3}}\mathbf{p}_{\mathbf{x}} + \cdots = \sum_{k=1}^{+\infty} {}_{\mathbf{k}}\mathbf{p}_{\mathbf{x}}, \end{aligned}$$

Aspettativa di vita incompleta

in quanto alcuni termini della serie si cancellano tra di loro, in analogia a quanto accade con le serie telescopiche. Osserviamo anche che la formula

$$e_x = \sum_{k=1}^{+\infty} {}_{\mathbf{k}}\mathbf{p}_{\mathbf{x}}$$

è semplicemente l'analogo per una distribuzione discreta della formula

$$\overset{\circ}{e}_x = \int_0^{+\infty} {}_{\mathbf{t}}\mathbf{p}_{\mathbf{x}}\, dt$$

vista per le distribuzione continue, in cui l'integrale è sostituito da una serie. Osserviamo inoltre che spezzando l'intervallo di integrazione possiamo scrivere

$$\overset{\circ}{e}_x = \sum_{j=0}^{+\infty} \int_j^{j+1} {}_{\mathbf{t}}\mathbf{p}_{\mathbf{x}}\, dt,$$

Relazione tra aspettativa di vita completa e incompleta

e approssimando l'integrale con la media della funzione sugli estremi dell'intervallo di integrazione come segue

$$\int_j^{j+1} {}_t\mathbf{p}_\mathbf{x}\, dt \approx \frac{{}_\mathbf{j}\mathbf{p}_\mathbf{x} + {}_\mathbf{j+1}\mathbf{p}_\mathbf{x}}{2}$$

si ottiene

$$\mathring{e}_x \approx \sum_{j=0}^{+\infty} \frac{{}_\mathbf{j}\mathbf{p}_\mathbf{x} + {}_\mathbf{j+1}\mathbf{p}_\mathbf{x}}{2} = \frac{1}{2} + \sum_{j=1}^{+\infty} {}_\mathbf{j}\mathbf{p}_\mathbf{x} = \frac{1}{2} + e_x,$$

cioè la aspettativa di vita completa è maggiore di 6 mesi rispetto alla aspettativa di vita incompleta, una approssimazione spesso usata in pratica.