FREQUENZE DI DECESSO PER TAVOLE SELEZIONATE

Un modello di sopravvivenza selezionato è definito mediante una famiglia di funzioni di sopravvivenza

$$S(t;x) \qquad t \geq 0 \qquad x = a, a+1, \dots$$

dove

x è l'età (intera) di ingresso in assicurazione

t è l'antidurata dell'assicurazione

Una **tavola di mortalità selezionata** è definita da un insieme di sequenze del tipo

$$\begin{array}{llll} l_{[a]} & l_{[a]+1} & l_{[a]+2} & \dots \\ l_{[a+1]} & l_{[a+1]+1} & l_{[a+1]+2} & \dots \\ \dots & & & \\ l_{[x]} & l_{[x]+1} & l_{[x]+2} & \dots \\ \dots & & & \end{array}$$

dove

$$l_{[x]+t} = l_{[x]} \cdot S(t;x)$$

$$x = a, a+1, \dots \qquad t = 0, 1, \dots$$

Per stimare un modello di sopravvivenza selezionato si determinano le frequenze di decesso per le diverse età di ingresso in assicurazione e antidurate.

Poiché l'età (intera) di ingresso in assicurazione è l'età arrotondata all'emissione della polizza e l'antidurata è il numero di anni in cui l'individuo è presente in assicurazione, si prende come riferimento l'anno di polizza.

Con riferimento all'età arrotondata x di ingresso in assicurazione ed all'intervallo di antidurate $]t, t+1]$,il vettore delle durate $(r_i, s_i, t_i, k_i,)$ dell'individuo i che contribuisce alla osservazione per l'intervallo di antidurate $]t, t+1]$ è così definito

r_i con $0 \le r_i < 1$ tale che $t + r_i$ è l'antidurata esatta all'ingresso in osservazione nell'intervallo di antidurate $]t, t+1]$

s_i con $0 < s_i \le 1$ tale che $t + s_i$ è l'antidurata esatta di uscita dall'osservazione dell'intervallo di antidurate $]t, t+1]$

t_i con $0 < t_i \le 1$ tale che $t + t_i$ è l'antidurata esatta di uscita per morte

k_i con $0 < k_i \le 1$ tale che $t + k_i$ è l'antidurata esatta di uscita per altra causa

Si definisce

$\theta_{[x]+t}$ il numero di decessi osservati per gli assicurati entrati in assicurazione all'età arrotondata x e con antidurata esatta in $]t, t+1]$

$E_{[x]+t}$ il numero iniziale di esposti al rischio nell'anno di polizza $]t, t+1]$, per gli assicurati entrati in assicurazione all'età arrotondata x

$$E_{[x]+t} = \sum_{i \in S \cup D \cup W} (1 - r_i) - \sum_{i \in S} (1 - s_i) - \sum_{i \in W} (1 - k_i)$$

Si definisce la **frequenza di decesso**

$$\overset{o}{q}_{[x]+t} = \frac{\theta_{[x]+t}}{E_{[x]+t}}$$

che fornisce una stima di $q_{[x+f]+t}$ con $-\frac{1}{2} \leq f < \frac{1}{2}$

Nel caso di distribuzione uniforme dei compleanni nell'anno di polizza si può assumere

$f = 0$ e quindi $\overset{o}{q}_{[x]+t}$ fornisce una stima di $q_{[x]+t}$

Indicato con

$E^{C}_{[x]+t}$ il numero centrale di esposti al rischio nell'anno di polizza $]t, t+1]$, per gli assicurati entrati in assicurazione all'età arrotondata x

$$E^{C}_{[x]+t} = \sum_{i \in S \cup D \cup W} (1 - r_i) - \sum_{i \in S} (1 - s_i) - \sum_{i \in W} (1 - k_i) - \sum_{i \in D} (1 - t_i)$$

si definisce la **frequenza centrale di decesso**

$$\overset{o}{m}_{[x]+t} = \frac{\theta_{[x]+t}}{E^{C}_{[x]+t}}$$

che fornisce una stima di $m_{[x+f]+t}$

Poiché l'effetto della selezione si esaurisce entro un certo numero t' di anni

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \ldots < q_{[x-t']+t'} = q_{[x-t'-1]+t'+1} = q_{[x-t'-2]+t'+2} = \ldots = q_{[a]+x-a}$$

si definiscono le tavole selezionate ridotte

$$\begin{array}{cccccc}
l_{[a]} & l_{[a]+1} & l_{[a]+2} & \cdots & l_{[a]+t'-1} & l_{(a+t')} \\
l_{[a+1]} & l_{[a+1]+1} & l_{[a+1]+2} & \cdots & l_{[a+1]+t'-1} & l_{(a+t'+1)} \\
\vdots & & & & & \vdots \\
l_{[x]} & l_{[x]+1} & l_{[x]+2} & \cdots & l_{[x]+t'-1} & l_{(x+t')} \\
\vdots & & & & & \vdots
\end{array}$$

dove $l_{[a]}$ è la radice della tavola e $l_{[x]}$ è tale che

$$l_{[x]} \cdot S(t'; x) = l_{(x+t')} \qquad x = a+1, a+2, \ldots$$

Indichiamo con

$$q_{(x)} = q_{[x-t']+t'} = q_{[x-t'-1]+t'+1} = \ldots = q_{[a]+x-a}$$ con $x = a+t$ e $t = t', t'+1, \ldots$

Per stimare

$$q_{(x)} = q_{[x-t']+t'} = q_{[x-t'-1]+t'+1} = \ldots = q_{[a]+x-a} \qquad \text{con } x = a+t \quad \text{e} \quad t = t', t'+1, \ldots$$

si continua a prendere come riferimento l'anno di polizza e si considerano le frequenze di decesso

$$\overset{o}{q}_{(x)} = \frac{\theta_{(x)}}{E_{(x)}}$$

essendo

$$\theta_{(x)} = \theta_{[x-t']+t'} + \theta_{[x-t'-1]+t'+1} + \ldots + \theta_{[a]+x-a}$$

$$E_{(x)} = E_{[x-t']+t'} + E_{[x-t'-1]+t'+1} + \ldots + E_{[a]+x-a}$$

oppure le frequenze centrali di decesso

$$\overset{o}{m}_{(x)} = \frac{\theta_{(x)}}{E^{C}_{(x)}}$$

essendo

$$E^{C}_{(x)} = E^{C}_{[x-t']+t'} + E^{C}_{[x-t'-1]+t'+1} + \ldots + E^{C}_{[a]+x-a}$$

STIMA DI MODELLI DI SOPRAVVIVENZA NON PARAMETRICI
USCITE SOLTANTO PER MORTE

Con riferimento alla classe di età $]x, x+1]$ supponiamo di avere osservato n_x individui e di disporre di dati individuali esatti riassunti, per ogni individuo i che contribuisce alla osservazione per tale classe di età, dal vettore delle durate

$$(r_i, s_i, t_i) \qquad i = 1, \ldots, n_x$$

essendo

$x + r_i$ l'età di ingresso in osservazione nella classe di età $]x, x+1]$ con $0 \le r_i < 1$

$x + s_i$ l'età di uscita pianificata dalla osservazione per la classe di età $]x, x+1]$ con $0 < s_i \le 1$

$x + t_i$ l'età di uscita per morte se $\theta_i = x + t_i$, altrimenti $t_i = 0$

<u>Nota</u>: se il riferimento è l'anno di vita, $x + r_i$, $x + s_i$ e $x + t_i$ sono età esatte, se il riferimento è l'anno di polizza (o l'anno di calendario), x è l'età arrotondata all'anniversario di polizza ed r_i, s_i e t_i sono durate riferire all'anno di polizza (o all'anno di calendario).

Stima con il metodo dei momenti

Si definiscono i n.a.

$$D_i = \begin{cases} 1 & \text{se l'individuo } i \text{ decede nella classe di età } x \\ 0 & \text{altrimenti} \end{cases} \qquad i = 1, \ldots, n_x$$

Si definisce il n.a. D_x dei decessi nella classe di età $]x, x+1]$

$$D_x = \sum_{i=1}^{n_x} D_i$$

Nell'ipotesi che il modello di descrizione della sopravvivenza sia lo stesso per ogni i si ha

$$E(D_x) = \sum_{i=1}^{n_x} E(D_i) = \sum_{i=1}^{n_x} {}_{s_i - r_i} q_{x+r_i}$$

Sia d_x il numero dei decessi osservati nella classe di età $]x, x+1]$;
si può scrivere quindi l'equazione dei momenti

$$E(D_x) = d_x \qquad \Leftrightarrow \qquad \sum_{i=1}^{n_x} {}_{s_i - r_i} q_{x+r_i} = d_x$$

A) Nell'ipotesi $r_i = 0$ e $s_i = 1$ per ogni i si ha

$$\sum_{i=1}^{n_x} {}_{s_i - r_i}q_{x+r_i} = d_x \quad \Leftrightarrow \quad q_x \, n_x = d_x$$

$$\Rightarrow \quad \hat{q}_x = \frac{d_x}{n_x}$$

La stima ottenuta coincide con la stima di massima verosimiglianza di q_x in ipotesi di distribuzione Binomiale(n_x, q_x) per il n.a. D_x

B) Nell'ipotesi di interpolazione lineare con $r_i = 0$ per ogni i si ha

$$\sum_{i=1}^{n_x} {}_{s_i - r_i}q_{x+r_i} = d_x \quad \Leftrightarrow \quad \sum_{i=1}^{n_x} {}_{s_i}q_x = d_x \quad \Leftrightarrow \quad \sum_{i=1}^{n_x} s_i \, q_x = d_x$$

$$\Rightarrow \quad \hat{q}_x = \frac{d_x}{\sum_{i=1}^{n_x} s_i}$$

essendo in tale ipotesi ${}_s q_x = s \, q_x$

C) Nell'ipotesi di interpolazione esponenziale

$$ {}_{s-r}q_{x+r} = 1 - \left(1 - q_x\right)^{s-r} $$

che richiede di risolvere l'equazione dei momenti numericamente

D) Nell'ipotesi di interpolazione iperbolica con $s_i = 1$ per ogni i si ha

$$ \sum_{i=1}^{n_x} {}_{s_i - r_i}q_{x+r_i} = d_x \quad \Leftrightarrow \quad \sum_{i=1}^{n_x} {}_{1-r_i}q_{x+r_i} = d_x \quad \Leftrightarrow \quad \sum_{i=1}^{n_x} (1 - r_i)\, q_x = d_x $$

$$ \Rightarrow \quad \hat{q}_x = \frac{d_x}{\sum_{i=1}^{n_x} (1 - r_i)} $$

essendo in tale ipotesi ${}_{1-r}q_{x+r} = (1 - r)\, q_x$

Per risolvere l'equazione dei momenti in forma chiusa

$$ \sum_{i=1}^{n_x} {}_{s_i - r_i}q_{x+r_i} = d_x $$

si formula, in generale, la seguente ipotesi: ${}_{s-r}q_{x+r} = (s - r)\, q_x$

Si ha allora

$$\sum_{i=1}^{n_x} {}_{s_i - r_i} q_{x+r_i} = d_x \quad \Leftrightarrow \quad \sum_{i=1}^{n_x} (s_i - r_i)\, q_x = d_x$$

$$\Rightarrow \quad \hat{q}_x = \frac{d_x}{\sum_{i=1}^{n_x} (s_i - r_i)}$$

dove $\sum_{i=1}^{n_x} (s_i - r_i)$ è detta **esposizione totale pianificata** nella classe di età $]x, x+1]$

Sia

$$\tilde{q}_x = \frac{D_x}{\sum_{i=1}^{n_x} (s_i - r_i)}$$

lo stimatore del quale la stima $\hat{q}_x$ è il valore osservato

Sotto l'ipotesi

$${}_{s-r} q_{x+r} = (s - r)\, q_x$$

$\tilde{q}_x$ è non distorto, infatti $E(\tilde{q}_x) = q_x$

Inoltre, in ipotesi di indipendenza stocastica dei n.a. D_i, $i=1,\dots,n_x$, si ha

$$Var(\tilde{q}_x)=\frac{\sum_{i=1}^{n_x} {}_{s_i-r_i}q_{x+r_i}\left(1-{}_{s_i-r_i}q_{x+r_i}\right)}{\left[\sum_{i=1}^{n_x}(s_i-r_i)\right]^2}=\frac{\sum_{i=1}^{n_x}(s_i-r_i)q_x\left(1-(s_i-r_i)q_x\right)}{\left[\sum_{i=1}^{n_x}(s_i-r_i)\right]^2}=\frac{q_x\sum_{i=1}^{n_x}(s_i-r_i)-q_x^2\sum_{i=1}^{n_x}(s_i-r_i)^2}{\left[\sum_{i=1}^{n_x}(s_i-r_i)\right]^2}$$

Sostituendo al posto di q_x il valore stimato $\hat{q}_x$ si ottiene una stima di $Var(\tilde{q}_x)$.

Se si considera invece l’ipotesi cosiddetta “binomiale”, cioè

$$E(\tilde{q}_x)=q_x \qquad Var(\tilde{q}_x)=\frac{q_x(1-q_x)}{\sum_{i=1}^{n_x}(s_i-r_i)}$$

si ottiene la seguente stima della $Var(\tilde{q}_x)$

$$\hat{Var}(\tilde{q}_x)=\frac{\hat{q}_x(1-\hat{q}_x)}{\sum_{i=1}^{n_x}(s_i-r_i)}$$