## Stima con il metodo della massima verosimiglianza

Con riferimento alla classe di età $]x, x+1]$ per scrivere la verosimiglianza delle osservazioni

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

definiamo, per ogni $i = 1, \dots, n_x$, i n.a.

$T^{(i)}$ durata aleatoria di vita dell’individuo $i$ nell’intervallo di età $]x, x+1]$

Nota: $T^{(i)}$ ha determinazioni $]r_i, s_i]$

Indicato con $T^{(i)}_{x+r_i}$ la durata aleatoria di vita dell’$i$-esimo individuo in vita all’età $x+r_i$ si ha

$$T^{(i)} = \min\left(T^{(i)}_{x+r_i} + r_i, s_i\right)$$

Nell'ipotesi che per ogni individuo $i$ la durata aleatoria di vita sia descritta dallo stesso modello di sopravvivenza, dotato di funzione di densità, si ha

$$P(T^{(i)} \leq t) = \begin{cases} 0 & t \leq r_i \\ F_{x+r_i}(t - r_i) & r_i < t < s_i \\ 1 & t \geq s_i \end{cases}$$

ed è inoltre $P(T^{(i)} = s_i) = 1 - F_{x+r_i}(s_i - r_i) = 1 - {}_{s_i - r_i}q_{x+r_i} = {}_{s_i - r_i}p_{x+r_i}$

Quindi se l'individuo $i$ decede con età esatta $x + t_i$ la verosimiglianza di tale osservazione è

$$f_{x+r_i}(t_i - r_i) = {}_{t_i - r_i}p_{x+r_i}\ \mu(x + t_i)$$

Se invece l'individuo $i$ raggiunge in vita l'età di uscita pianificata $x + s_i$, la verosimiglianza di tale osservazione è

$${}_{s_i - r_i}p_{x+r_i}$$

Si definiscono

$S = \{ i \mid \text{l'individuo } i \text{ è in vita all'età } x + s_i \}$ *survival*

$D = \{ i \mid \text{l'individuo } i \text{ esce per morte all'età } x + t_i \}$ *death*

In ipotesi di indipendenza stocastica dei n.a. $T^{(i)}$ la verosimiglianza delle osservazioni è

$$L = \prod_{i \in S} {}_{s_i - r_i}p_{x+r_i} \cdot \prod_{i \in D} f_{x+r_i}(t_i - r_i) = \prod_{i \in S} {}_{s_i - r_i}p_{x+r_i} \cdot \prod_{i \in D} {}_{t_i - r_i}p_{x+r_i}\, \mu(x + t_i)$$

A) Nell’ipotesi di <u>interpolazione esponenziale</u> $\mu(x+t) = \mu_x$ , $0 < t \le 1$

$${}_{s_i - r_i}p_{x+r_i} = e^{-\mu_x (s_i - r_i)}$$

la verosimiglianza è

$$L = \prod_{i \in S} e^{-\mu_x (s_i - r_i)} \cdot \prod_{i \in D} \left( e^{-\mu_x (t_i - r_i)}\, \mu_x \right) = \exp\left[ -\mu_x \left( \sum_{i \in S} (s_i - r_i) + \sum_{i \in D} (t_i - r_i) \right) \right] \cdot (\mu_x)^{d_x}$$

dove $d_x = \# D$

La log-verosimiglianza è allora

$$l = \log L = \left[ -\mu_x \left( \sum_{i \in S} (s_i - r_i) + \sum_{i \in D} (t_i - r_i) \right) \right] + d_x \log(\mu_x)$$

Risolvendo l'equazione di verosimiglianza si trova

$$\hat{\mu}_x = \frac{d_x}{\sum_{i \in S} (s_i - r_i) + \sum_{i \in D} (t_i - r_i)}$$

dove $\sum_{i \in S} (s_i - r_i) + \sum_{i \in D} (t_i - r_i)$ è detta **esposizione esatta**

Se poniamo

$$E_x^C = \sum_{i \in S} (s_i - r_i) + \sum_{i \in D} (t_i - r_i)$$

si ha

$$L = e^{-\mu_x E_x^C} \cdot (\mu_x)^{d_x}$$

Osservazione

Sia $D_x$ il n.a. dei decessi nella classe di età $]x, x+1]$, in ipotesi di distribuzione di Poisson di parametro $\mu_x E_x^C$ si ha

$$P(D_x = d_x) = \frac{\left(\mu_x E_x^C\right)^{d_x}}{d_x!} e^{-\mu_x E_x^C} = \frac{\left(E_x^C\right)^{d_x}}{d_x!} e^{-\mu_x E_x^C} \mu_x{}^{d_x} \propto L$$

Pertanto ai fini della stima di massima verosimiglianza dell'intensità istantanea di mortalità sono equivalenti le ipotesi esponenziale e di distribuzione di Poisson per il n.a. dei decessi $D_x$ con $E(D_x) = \mu_x E_x^C$

B)Nell'ipotesi di interpolazione lineare

$$_{s_i - r_i}p_{x+r_i} = \frac{1 - s_i\ q_x}{1 - r_i\ q_x} \qquad\qquad \mu(x + t_i) = \frac{q_x}{1 - t_i\ q_x}$$

indicato con $d_x = \# D$, la verosimiglianza è

$$L = \prod_{i \in S} \frac{1 - s_i\ q_x}{1 - r_i\ q_x} \cdot \prod_{i \in D} \left( \frac{1 - t_i\ q_x}{1 - r_i\ q_x} \cdot \frac{q_x}{1 - t_i\ q_x} \right) = \prod_{i \in S \cup D} \left(1 - r_i\ q_x\right)^{-1} \prod_{i \in S} \left(1 - s_i\ q_x\right) \quad q_x{}^{d_x}$$

Dalla log-verosimiglianza

$$l = \log L = -\sum_{i \in S \cup D} \log(1 - r_i\, q_x) + \sum_{i \in S} \log(1 - s_i\, q_x) + d_x \log(q_x)$$

si ottiene l'equazione di log-verosimiglianza che può essere risolta per via numerica

$$\sum_{i \in S \cup D} \frac{r_i}{1 - r_i\, q_x} - \sum_{i \in S} \frac{s_i}{1 - s_i\, q_x} + \frac{d_x}{q_x} = 0$$

<u>Osservazione</u>

Soluzioni in forma chiusa possono essere ottenute per dati particolari (p. es. se $r_i = 0$ e $s_i = 1$ per ogni $i$) e nel caso di particolari dati raggruppati (per es. se nel caso $s_i = 1$ per ogni $i$, si considera una comune età media di ingresso $r$ con $0 < r < 1$ per ogni $i$, oppure nel caso $r_i = 0$ per ogni $i$, se si considera una comune età media di uscita pianificata $s$ con $0 < s < 1$ per ogni $i$)

Vedi London, cap. 7