

## CONFRONTO DELLA TAVOLA DI SOPRAVVIVENZA STIMATA CON UNA TAVOLA STANDARD

Siano

$$\hat{q}_x, \quad x = a, a + 1, \dots, \omega - 1$$

le stime delle probabilità di morte  $q_x$ ,  $x = a, a + 1, \dots, \omega - 1$ , di un modello di sopravvivenza non parametrico.

Ci si pone il problema se il fenomeno della mortalità osservata nella collettività possa essere descritto da una tavola di mortalità proveniente da altre esperienze statistiche. Tale tavola viene allora detta **tavola “standard”** e la indichiamo con

$$q'_x, \quad x = a, a + 1, \dots, \omega - 1$$

Per verificare se la tavola standard accosta bene le osservazioni si sottopone a verifica d'ipotesi la seguente ipotesi nulla:

$$H_0 : \quad q_x = q'_x \quad x = a, a + 1, \dots, \omega - 1$$

## Confronto della tavola di sopravvivenza stimata con una tavola standard

Con riferimento alla classe di età  $]x, x + 1]$  siano

$D_x$	n.a. di decessi
$d_x$	numero di decessi osservati
$E_x$	il numero di esposti al rischio

Per costruire la funzione test formuliamo le seguenti ipotesi

$$E(D_x) = E_x q_x \quad \text{Var}(D_x) = E_x q_x (1 - q_x)$$

Sotto l'ipotesi nulla i n.a.

$$Z_x = \frac{D_x - E_x q'_x}{\sqrt{E_x q'_x (1 - q'_x)}} \quad x = a, a + 1, \dots, \omega - 1$$

hanno distribuzione approssimata  $N(0, 1)$ . In ipotesi di indipendenza stocastica dei n.a.  $Z_x$   $x = a, a + 1, \dots, \omega - 1$ , sia ha che il n.a.

$$\sum_{x=a}^{\omega-1} (Z_x)^2$$

ha distribuzione approssimata chi-quadrato con  $n$  gradi di libertà, essendo  $n$  il numero di classi di età.

## Confronto della tavola di sopravvivenza stimata con una tavola standard

Fissato un livello di significatività  $\alpha$  si determina  $\chi_{n,1-\alpha}^2$  tale che  $P(\chi_n^2 > \chi_{n,1-\alpha}^2) = \alpha$  si rifiuta l'ipotesi nulla se

$$\sum_{x=a}^{\omega-1} (z_x)^2 > \chi_{n,1-\alpha}^2$$

essendo  $\sum_{x=a}^{\omega-1} (z_x)^2$  la determinazione osservata del n.a.  $\sum_{x=a}^{\omega-1} (Z_x)^2$

Tale test potrebbe non rilevare la mancanza di buon accostamento della tavola standard ai dati osservati, e quindi non fare rifiutare l'ipotesi nulla, nelle seguenti situazioni:

- esistenza di scostamenti eccessivamente elevati per alcune età, controbilanciati da scostamenti molto ridotti per altre età;
- numero eccessivo di scostamenti tutti dello stesso segno (conseguenza di una mortalità rilevata “uniformemente” maggiore o minore di quella attesa in base alla tavola standard);
- gruppi eccessivamente numerosi di età consecutive con scostamenti tutti dello stesso segno.

Confronto della tavola di sopravvivenza stimata con una tavola standard

Per evidenziare tali problematiche si utilizzano altri test, quale per esempio il test delle deviazioni cumulate.

Sotto l'ipotesi nulla

$$H_0: \quad q_x = q'_x \quad x = a, a + 1, \dots, \omega - 1$$

si ha

$$E(D_x - E_x q'_x) = 0 \quad \text{per ogni } x = a, a + 1, \dots, \omega - 1$$

Consideriamo le deviazioni cumulate nell'intervallo di età da  $x_1$  a  $x_2$

$$\sum_{x=x_1}^{x_2} (D_x - E_x q'_x)$$

In ipotesi di indipendenza stocastica dei n.a.  $D_x$  si ha

$$E \left[ \sum_{x=x_1}^{x_2} (D_x - E_x q'_x) \right] = 0 \quad \text{Var} \left[ \sum_{x=x_1}^{x_2} (D_x - E_x q'_x) \right] = \sum_{x=x_1}^{x_2} \text{Var}(D_x) = \sum_{x=x_1}^{x_2} E_x q'_x (1 - q'_x)$$

Confronto della tavola di sopravvivenza stimata con una tavola standard

La distribuzione della deviazione cumulata standardizzata

$$\frac{\sum_{x=x_1}^{x_2} (D_x - E_x q'_x)}{\sqrt{\sum_{x=x_1}^{x_2} E_x q'_x (1 - q'_x)}}$$

può essere approssimata mediante una  $N(0, 1)$  e quindi, fissato un livello di significatività  $\alpha$  si determina il quantile  $z_{1-\alpha/2}$  della distribuzione normale standard e si rifiuta l'ipotesi nulla se

$$\left| \frac{\sum_{x=x_1}^{x_2} (d_x - E_x q'_x)}{\sqrt{\sum_{x=x_1}^{x_2} E_x q'_x (1 - q'_x)}} \right| > z_{1-\alpha/2}$$

Tale analisi va ripetuta su diversi intervalli di età che evidenziano criticità nell'accostamento della tavola standard ai dati osservati.

## PEREQUAZIONE MEDIANTE TAVOLE STANDARD

Se è stata rifiutata l'ipotesi nulla sulla bontà di accostamento della tavola standard ai dati, si può decidere di “adattare” la tavola standard ai dati osservati.

Si ipotizza quindi un legame funzionale tra le probabilità di morte della collettività in esame

$$q_x, \quad x = a, a + 1, \dots, \omega - 1,$$

e le probabilità di morte riportate nella tavola standard

$$q'_x, \quad x = a, a + 1, \dots, \omega - 1$$

Si assume che tale funzione

$$q_x = f(x, q'_x)$$

dipenda da alcuni parametri che devono essere stimati, per esempio con il metodo dei minimi quadrati pesati.

## Perequazione mediante tavole standard

Siano

$$\hat{q}_x, \quad x = a, a + 1, \dots, \omega - 1$$

le stime delle probabilità di morte  $q_x$ ,  $x = a, a + 1, \dots, \omega - 1$ , di un modello di sopravvivenza non parametrico.

Dall'analisi grafica dei rapporti  $\frac{\hat{q}_x}{q'_x}$ ,  $x = a, a + 1, \dots, \omega - 1$

si individua un possibile legame funzionale  $q_x = f(x, q'_x)$ .

Per esempio, nel caso di andamento approssimativamente lineare si può ipotizzare

$$q_x = q'_x(a + b x)$$

oppure

$$q_x = a q'_x + b$$

Più in generale si può ipotizzare anche un legame con due tavole standard  $\{q'_x\}$  e  $\{q''_x\}$

$$q_x = a' q'_x + a'' q''_x$$

## Perequazione mediante tavole standard

In alternativa, Lidstone ha proposto di considerare per l'analisi grafica la seguente trasformazione

$$\log\left(\frac{p'_x}{\hat{p}_x}\right), \quad x = a, a + 1, \dots, \omega - 1$$

con

$$p'_x = 1 - q'_x \quad \hat{p}_x = 1 - \hat{q}_x$$

in quanto i  $\frac{p'_x}{\hat{p}_x}$  presentano un andamento più regolare rispetto ai  $\frac{\hat{q}_x}{q'_x}$

Se evidenziano un andamento approssimativamente costante si può ipotizzare

$$\log\left(\frac{p'_x}{\hat{p}_x}\right) = c \quad \Rightarrow \quad q_x = 1 - \frac{p'_x}{e^c}$$



## Perequazione mediante tavole standard

Dopo avere individuato la funzione  $f(x; q'_x, a, b, \dots)$  che esprime il legame tra  $q_x$  e  $q'_x$ , per la stima dei parametri della funzione si può utilizzare, per esempio, il metodo dei minimi quadrati.

$$\min_{a, b, \dots} F(a, b, \dots) \quad \text{con} \quad F(a, b, \dots) = \sum_{x=a}^{\omega-1} w_x [\hat{q}_x - f(x; q'_x, a, b, \dots)]^2$$

essendo

$w_x = 1$  nel caso di minimi quadrati non pesati

$w_x = \frac{n'_x}{\hat{q}_x}$  nel caso di minimi quadrati pesati

Si noti che nell'ipotesi  $Var(D_x) = n'_x q_x (1 - q_x)$  si ha

$$Var\left(\frac{D_x}{n'_x}\right) = \frac{q_x (1 - q_x)}{n'_x} \cong \frac{q_x}{n'_x}$$

quindi il peso  $w_x = \frac{n'_x}{\hat{q}_x}$  è approssimativamente pari al reciproco della varianza dello stimatore di  $q_x$

## Perequazione mediante tavole standard

Se la funzione è lineare, le stime dei minimi quadrati dei parametri si ottengono agevolmente risolvendo un sistema lineare.

Sia  $u_x$  tale che il legame funzionale tra  $q_x$  e  $q'_x$  sia espresso mediante la funzione lineare

$$f(x; a, b) = a + b x$$

Per esempio  $u_x = \frac{q_x}{q'_x}$  nel caso in cui sia  $q_x = q'_x(a + b x)$

Si ha allora

$$\min_{a, b} F(a, b) \quad \text{con } F(a, b) = \sum_{x=a}^{\omega-1} w_x [\hat{u}_x - (a + b x)]^2$$

essendo  $\hat{u}_x = \frac{\hat{q}_x}{q'_x}$

Si ottiene:

$$\begin{cases} \frac{\partial F}{\partial a} = 0 \\ \frac{\partial F}{\partial b} = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{x=a}^{\omega-1} w_x [\hat{u}_x - (a + b x)] = 0 \\ \sum_{x=a}^{\omega-1} w_x [\hat{u}_x - (a + b x)] x = 0 \end{cases}$$