MODELLI DI SOPRAVVIVENZA

- Modello di sopravvivenza continuo
- Modello di sopravvivenza discreto
- Rilevazione dei dati per la stima di un modello sopravvivenza
- Funzione di sopravvivenza empirica
- La stima di Kaplan-Meier per la funzione di sopravvivenza in un modello discreto
- La stima di Nelson-Aalen
- Il modello attuariale per la descrizione della sopravvivenza
- Tavole di mortalità
- Stima delle probabilità di decesso mediante lo stimatore di Kaplan-Meier
- Esposti al rischio
- Completamento del modello di sopravvivenza per età non intere
- Modello di sopravvivenza selezionato, tavole selezionate e tavole selezionate ridotte

MODELLO DI SOPRAVVIVENZA CONTINUO

Sia

T n.a. non negativo che esprime la durata aleatoria da un istante iniziale fino al verificarsi di un determinato evento

Esempi:

- durata di funzionamento di un'apparecchiatura dall'istante di accensione in cui l'apparecchiatura inizia a funzionare
- durata di vita di una cavia dall'istante di somministrazione di una determinata sostanza
- durata di vita dall'istante in cui ad un soggetto è diagnosticata una certa malattia
- durata di vita dalla nascita

Si indica con 0 l'istante iniziale

T ha determinazioni: $t > 0$

T ha supporto $[0,+\infty[$

Si definisce

Funzione di sopravvivenza $S(t) = P(T > t), \quad t \geq 0$

Si ha

- $S(0) = P(T > 0) = 1$
- $S(t)$ è una funzione non crescente: se $t_1 < t_2$ allora $S(t_1) \geq S(t_2)$
- $\lim_{t \to +\infty} S(t) = 0$, se si esclude l'esistenza di masse aderenti a $+\infty$

Si definisce

Funzione di ripartizione $F(t) = P(T \leq t)$

Poiché:

$$F(t) = P(T \leq t) = 1 - P(T > t) = 1 - S(t)$$

per assegnare la distribuzione di probabilità del n.a. T è equivalente assegnare la funzione di ripartizione oppure la funzione di sopravvivenza

Se la distribuzione del n.a. T è dotata di funzione di densità $f(t)$ continua si ha:

$$F(t) = \int_0^t f(u)du \qquad S(t) = 1 - F(t) = \int_t^{+\infty} f(u)du$$

inoltre

$$f(t) = \frac{d}{dt}F(t) = -\frac{d}{dt}S(t)$$

Quindi, per assegnare la distribuzione di probabilità del n.a. T è equivalente assegnare la funzione di ripartizione, oppure la funzione di sopravvivenza, oppure la funzione di densità

Si definisce

Funzione di rischio o hazard function

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t | T \ge t)}{\Delta t}$$

Si ha

$$\lambda(t) = \frac{f(t)}{S(t)} \quad \text{da cui} \quad f(t) = S(t)\lambda(t)$$

<u>Osservazione</u>: interpretazione del significato di $\lambda(t)$

Essendo la distribuzione del n.a. T è dotata di densità $f(t)$ continua si ha:

$$P(t \le T < t + \Delta t | T \ge t) = \frac{F(t+\Delta t) - F(t)}{S(t)} = \frac{f(t)\Delta t + o(\Delta t)}{S(t)} \qquad \text{con} \quad \lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$$

Si ha allora che a meno di infinitesimi di ordine superiore al primo

$$P(t \le T < t + \Delta t | T \ge t) \cong \frac{f(t)}{S(t)} \Delta t = \lambda(t) \Delta t$$

$\lambda(t)$ misura l'intensità istantanea dei decessi all'istante t per gli individui in vita fino all'istante t

Si ha inoltre

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \ln(S(t))$$

quindi, assegnata $S(t)$ si determina $\lambda(t)$

Essendo $f(t)$ continua, vale anche viceversa, infatti integrando si ottiene

$$\int_0^t \lambda(u)du = -\ln(S(t))$$

da cui si ottiene

$$S(t) = \exp\left(-\int_0^t \lambda(u)du\right)$$

In tali ipotesi è allora equivalente assegnare la distribuzione di probabilità del n.a. T assegnando la funzione di sopravvivenza oppure la funzione di rischio

Si definisce

Funzione di rischio integrato $\qquad \Lambda(t) = \int_0^t \lambda(u)du$

Si ha allora

$$S(t) = \exp(-\Lambda(t))$$

MODELLO DI SOPRAVVIVENZA DISCRETO

Sia

T n.a. non negativo che esprime la durata aleatoria da un istante iniziale fino al verificarsi di un determinato evento

Si indica con 0 l'istante iniziale

T ha determinazioni: $0 < t_1 < t_2 < \dots$

Esempi:

- osservazioni di durate con misurazioni arrotondate
- osservazioni di durate raggruppate in intervalli (si considera come durata il punto medio di ciascun intervallo)
- durate misurate con un numero intero di unità (per esempio, durate misurate in settimane)

Si definiscono le

probabilità $q_j = P(T = t_j), \qquad j = 1, 2, \ldots$

essendo $\sum_{j=1}^{+\infty} q_j = 1$

Si definisce

Funzione di sopravvivenza $S(t) = P(T > t) = \sum_{t_j > t} q_j \, , t \geq 0$

Si ha

- $S(t)$ è una funzione non crescente, continua a destra; infatti, se $t_k \leq t < t_{k+1}$ si ha

$$S(t) = P(T > t) = \sum_{j > k} q_j = S(t_k)$$

- per assegnare la distribuzione di probabilità del n.a. T è equivalete assegnare le probabilità oppure la funzione di sopravvivenza

Si definisce

Funzione di rischio o hazard function $\lambda(t_j) = P(T = t_j | T \geq t_j), \quad j = 1, 2, \ldots$

Si ha

$$\lambda(t_j) = \frac{q_j}{S(t_{j-1})}, \qquad j = 1, 2, \ldots$$

essendo

$$t_0 = 0 \qquad \text{e} \qquad S(t_0) = 1$$

Si ha allora che, assegnate le probabilità, è assegnata la funzione di sopravvivenza ed è pure assegnata la funzione di rischio

Poiché $q_j = S(t_{j-1}) - S(t_j)$ si ha

$$\lambda(t_j) = \frac{q_j}{S(t_{j-1})} = \frac{S(t_{j-1}) - S(t_j)}{S(t_{j-1})} = 1 - \frac{S(t_j)}{S(t_{j-1})}, \qquad j = 1, 2, \ldots$$

Proviamo che assegnata la funzione di rischio risulta assegnata anche la funzione di sopravvivenza.

Sia $t_k \leq t < t_{k+1}$ si ha

$$S(t) = S(t_k) = \frac{S(t_k)}{S(t_{k-1})} \cdot \frac{S(t_{k-1})}{S(t_{k-2})} \cdot \cdots \cdot \frac{S(t_1)}{S(t_0)}$$

Essendo $\dfrac{S(t_j)}{S(t_{j-1})} = 1 - \lambda(t_j), \; j = 1, 2, \ldots$ si ottiene

$$S(t) = S(t_k) = (1 - \lambda(t_k)) \cdot (1 - \lambda(t_{k-1})) \cdot \cdots \cdot (1 - \lambda(t_1)) = \prod_{t_j \leq t} (1 - \lambda(t_j))$$

È equivalente assegnare la distribuzione di probabilità del n.a. T assegnando la funzione di sopravvivenza oppure la funzione di rischio

<u>Osservazione</u>: interpretazione del significato di $\dfrac{S(t_j)}{S(t_{j-1})} = 1 - \lambda(t_j)$

Si ha

$$\frac{S(t_j)}{S(t_{j-1})} = \frac{P(T > t_j)}{P(T > t_{j-1})} = \frac{P(T > t_{j-1}, T > t_j)}{P(T > t_{j-1})} = P(T > t_j \mid T > t_{j-1})$$

Si può allora interpretare la formula

$$S(t) = S(t_k) = \prod_{t_j \le t} \left(1 - \lambda(t_j)\right) = \prod_{t_j \le t} P(T > t_j \mid T > t_{j-1})$$

come prodotto di probabilità condizionate di sopravvivenza.

RILEVAZIONE DEI DATI PER LA STIMA DI UN MODELLO DI SOPRAVVIVENZA

I dati per l’analisi della sopravvivenza si dicono **completi** se per ogni individuo osservato si osserva l’istante iniziale e l’evento; è quindi osservata la durata (time to event data).

Si possono avere due tipi di rilevazioni:

- le rilevazioni trasversali (cross-sectional studies)
- le rilevazioni longitudinali (longitudinal studies)

Rilevazioni trasversali

Si individua un <u>gruppo di studio</u>, cioè un gruppo di individui per i quali interessa studiare la sopravvivenza (per es. la popolazione di una certa zona o anche di un’intera nazione, gli assicurati di una compagnia di assicurazione, gli iscritti ad un fondo pensione, …).

Si fissa un <u>periodo di osservazione</u> durante il quale viene osservato il gruppo di studio.

All’inizio dell’osservazione ci saranno individui già presenti, ai quali se ne aggiungeranno altri durante il periodo di osservazione; alcuni individui possono uscire per causa diversa dal decesso durante l’osservazione; ci saranno individui ancora in vita al termine dell’osservazione.

Tipicamente, nelle rilevazioni trasversali, i dati sono **incompleti**:

se non è osservato l'istante iniziale, l'<u>osservazione</u> è detta <u>troncata a sinistra</u>

se non è osservato l'evento, l'<u>osservazione</u> è detta <u>censurata a destra</u>

Rilevazioni longitudinali

Sono tipicamente usate negli studi clinici, quando le durate, dall'istante iniziale al verificarsi dell'evento, non sono molto lunghe.

Si seleziona un <u>gruppo di studio</u>, cioè un gruppo di individui per i quali interessa studiare la sopravvivenza, e si osserva il gruppo fino a quando per tutti gli individui è osservato l'evento e sono quindi osservate le durate dall'istante iniziale al verificarsi dell'evento.

Negli studi clinici, si possono spesso avere **dati completi**, se per ogni individuo è osservato l'istante iniziale, che può coincidere con una stessa data di calendario, ma non necessariamente. Si parla di *cohort complete design*.

Se si termina l'osservazione prima di avere osservato tutti gli eventi, si hanno dati <u>incompleti</u> (censurati a destra). Infatti, se per alcuni individui non si osserva l'evento, si perdono alcune informazioni sulla sopravvivenza della collettività oggetto di studio.

FUNZIONE DI SOPRAVVIVENZA EMPIRICA

Rilevazione longitudinale con dati completi

<u>Osservazioni</u>: n individui osservati; $t_1 < t_2 < \ldots < t_n$ durate osservate

Si definisce **funzione di sopravvivenza empirica**

$$\hat{S}(t) = \begin{cases} 1 & t < t_1 \\ \dfrac{n-j}{n} & t_j \le t < t_{j+1},\ j = 1, 2, \ldots, n-1 \\ 0 & t \ge t_n \end{cases}$$

La funzione di sopravvivenza empirica fornisce una stima della funzione di sopravvivenza $S(t)$

Sia

N_t il n.a. di individui in vita all'istante t a partire da $N_0 = n$ individui osservati all'istante 0

Lo stimatore $\tilde{S}(t)$ del quale la funzione di sopravvivenza empirica è la stima per $S(t)$ è

$$\tilde{S}(t) = \frac{N_t}{n}$$

<u>Proprietà dello stimatore</u> $\tilde{S}(t)$

Nell'ipotesi che le durate aleatorie di vita degli $N_0 = n$ individui osservati all'istante 0 siano ugualmente distribuite ed indipendenti con funzione di sopravvivenza $S(t)$, fissato $t > 0$ il n.a. N_t ha distribuzione Binomiale($n, S(t)$)

Si ha allora

$$E(N_t) = n\,S(t) \qquad\qquad Var(N_t) = n\,S(t)\,(1 - S(t))$$

Risulta quindi che $\tilde{S}(t)$ è uno stimatore non distorto, infatti per ogni $t > 0$ si ha

$$E\left(\tilde{S}(t)\right) = E\left(\frac{N_t}{n}\right) = S(t)$$

Si ha inoltre

$$Var\left(\tilde{S}(t)\right) = Var\left(\frac{N_t}{n}\right) = \frac{S(t)\,(1 - S(t))}{n}$$

Quindi una stima di $Var\left(\tilde{S}(t)\right)$ è data da

$$\hat{Var}\left(\tilde{S}(t)\right) = \frac{\hat{S}(t)\,(1 - \hat{S}(t))}{n}$$

LA STIMA DI KAPLAN-MEIER PER LA FUNZIONE DI SOPRAVVIVENZA IN UN MODELLO DISCRETO

Rilevazione trasversale: per ogni individuo si osserva l'istante iniziale; alcuni individui danno luogo ad osservazioni censurate a destra. Quindi i dati sono incompleti.

<u>Osservazioni</u>: n individui osservati; m decessi osservati alle età $t_1 < t_2 < \ldots < t_k$

$d_1, d_2, \ldots, d_k$ numeri di decessi osservati alle età $t_1, t_2, \ldots, t_k$

$$d_1 + d_2 + \ldots + d_k = m$$

$n - m$ numero di osservazioni censurate

Sia t_0 l'età di ingresso in osservazione corrispondente all'istante iniziale

Si ripartisce l'intervallo $[t_0, +\infty[$ negli intervalli

$$[t_j, t_{j+1}[\qquad j = 0, 1, \ldots, k \qquad \text{con} \quad t_{k+1} = +\infty$$

Sia c_j, $j = 0, 1, \ldots, k$, il numero di osservazioni censurate nell'intervallo $[t_j, t_{j+1}[$

$$c_0 + c_1 + \ldots + c_k = n - m$$

Con riferimento all'intervallo $\left[t_j, t_{j+1}\right[$, $j = 0, 1, \ldots, k$, indichiamo con

$$t_{j1} \le t_{j2} \le \ldots \le t_{jc_j}$$

le età di uscita delle c_j osservazioni censurate

Sia

$$n_j, \;\; j = 1, \ldots, k,$$

il numero di individui in vita all'età t_j che danno luogo ai d_j decessi

Si ha

$$n_j = (d_j + d_{j+1} + \ldots + d_k) + (c_j + c_{j+1} + \ldots + c_k) \qquad j = 1, \ldots, k$$

<u>Obiettivo</u>: stimare la funzione di ripartizione $F(t)$
del n.a. T durata aleatoria di vita dall'età t_0

Disponendo delle osservazioni: decesso all'età t_j, $j = 1, \ldots, k$

osservazione censurata all'età

t_{ji} , $j = 0, \ldots, k$, $i = 1, \ldots, c_j$

Nell'ipotesi che i numeri aleatori osservati siano stocasticamente indipendenti ed ugualmente distribuiti con funzione di ripartizione $F(t)$ la verosimiglianza delle osservazioni è

$$L = \prod_{j=0}^{k} \left(\left[F(t_j) - F(t_j^-) \right]^{d_j} \cdot \prod_{i=1}^{c_j} \left[1 - F(t_{ji}) \right] \right)$$

essendo

$$F(t_j^-) = \lim_{t \to t_j^-} F(t) \qquad j = 1, \ldots, k$$

$$\left[F(t_0) - F(t_0^-) \right]^{d_0} = 1$$

<u>Obiettivo</u>: determinare la stima $\hat{F}(t)$ di massima verosimiglianza della funzione di ripartizione $F(t)$

<u>Osservazione</u>: $L \neq 0 \quad \Leftrightarrow \quad F(t_j) - F(t_j^-) > 0 \quad \forall\ j = 1, \ldots, k$, quindi

inoltre $1 - F(t_{ji})$ è massimo se $F(t_{ji}) = F(t_j)$

Quindi la verosimiglianza diventa

$$L = \prod_{j=1}^{k} \left[F(t_j) - F(t_j^-) \right]^{d_j} \cdot \prod_{j=0}^{k} \left[1 - F(t_j) \right]^{c_j}$$

Osservazioni:

- $F(t)$ è la funzione di ripartizione di un n.a. con distribuzione discreta e determinazioni $t_1, t_2, \ldots, t_k$
- per stimare $F(t)$ si tratta quindi di stimare le probabilità

$$q_j = F(t_j) - F(t_j^-) \qquad\qquad j = 1, \ldots, k$$

- ovvero di stimare la funzione di rischio

$$\lambda_j = \lambda(t_j) = \frac{q_j}{S(t_{j-1})} = \frac{F(t_j) - F(t_j^-)}{1 - F(t_{j-1})} \qquad\qquad j = 1, \ldots, k$$

essendo

$$S(t_j) = 1 - F(t_j) = \prod_{t_i \le t_j} (1 - \lambda(t_i)) = \begin{cases} 1 & \text{se } j = 0 \\ (1-\lambda_1)(1-\lambda_2)\cdots(1-\lambda_j) & \text{se } j = 1, \ldots, k \end{cases}$$

Si ottiene allora la seguente espressione per la verosimiglianza

$$L = \prod_{j=1}^{k} [\lambda(t_j)]^{d_j} \cdot \prod_{j=1}^{k} [1 - F(t_{j-1})]^{d_j} \cdot \prod_{j=0}^{k} [1 - F(t_j)]^{c_j} = \prod_{j=1}^{k} [\lambda(t_j)]^{d_j} \cdot \prod_{j=1}^{k} [1 - \lambda(t_j)]^{n_j - d_j}$$

Data la verosimiglianza

$$L = \prod_{j=1}^{k} \lambda_j^{d_j} \left[1-\lambda_j\right]^{n_j-d_j}$$

Osservazione

$\lambda_j^{d_j} \left[1-\lambda_j\right]^{n_j-d_j}$ è proporzionale alla verosimiglianza di un n.a. con distribuzione Binomiale(n_j, λ_j)

Sia D_j il n.a. dei decessi all'età t_j a partire dagli n_j individui in vita all'età t_j prima che si verifichino i d_j decessi

Si ipotizza per D_j la distribuzione Binomiale(n_j, λ_j)

La stima di massima verosimiglianza di λ_j è

$$\hat{\lambda}_j = \frac{d_j}{n_j}$$

La stima di Kaplan-Meier per la funzione di sopravvivenza è allora

$$\hat{S}(t) = \prod_{t_j \le t} \left(1 - \hat{\lambda}_j\right) = \prod_{t_j \le t} \left(1 - \frac{d_j}{n_j}\right)$$

<u>Osservazione</u>

Poiché

$$S(t) = \prod_{t_j \le t} \left(1 - \lambda(t_j)\right) = \prod_{t_j \le t} P(T > t_j \mid T > t_{j-1})$$

si può allora interpretare la formula

$$\hat{S}(t) = \prod_{t_j \le t} \left(1 - \hat{\lambda}_j\right) = \prod_{t_j \le t} \left(1 - \frac{d_j}{n_j}\right)$$

come prodotto delle stime delle probabilità condizionate di sopravvivenza.

La stima di massima verosimiglianza, di Kaplan-Meier, della funzione di ripartizione $F(t)$ è allora

$$\hat{F}(t) = 1 - \hat{S}(t) = 1 - \prod_{t_j \le t} \left(1 - \frac{d_j}{n_j}\right)$$

Esempio

j	t_j	n_j	d_j	c_j	$1-d_j/n_j$	$\hat{S}(t_j)$
0	0	8	0	1	1	1
1	1,5	7	1	0	0,857143	0,857143
2	2	6	2	1	0,666667	0,571429
3	3,5	3	1	2	0,666667	0,380952

Osservazioni

- La stima di Kaplan-Meier $\hat{F}(t)$ è definita per $t \le t_k$
- Se l'età massima osservata è un dato censurato, cioè se $c_k > 0$, la $\hat{F}(t)$ non è definita per $t > t_k$ ed è $\hat{F}(t_k) < 1$; una possibilità è definire $\hat{F}(t) = \hat{F}(t_k)$ per $t > t_k$ ma ciò comporta la presenza di masse aderenti a $+\infty$
- La stima di Kaplan-Meier $\hat{F}(t)$ può essere definita anche in presenza di osservazioni troncate a sinistra, cioè di osservazioni per le quali non si è osservato l'istante iniziale essendo entrate in osservazione ad un'età maggiore di t_0. Tuttavia, in tal caso $\hat{S}(t)$ è la stima della funzione di sopravvivenza condizionata alla età minima di ingresso osservata $t_{\min}$, essendo $t_0 \le t_{\min} < t_1$; quindi è una stima di $P(T > t | T > t_{\min}) = S(t)/S(t_{\min})$

Proprietà dello stimatore di Kaplan-Meier

$$\hat{S}(t) = \prod_{t_j \leq t} \left(1 - \frac{d_j}{n_j}\right) \qquad \text{è la stima di Kaplan-Meier di } S(t),\ t \geq 0$$

Sia

$$S(t) = \prod_{t_j \leq t} p_j \qquad \text{con} \qquad p_j = 1 - \lambda(t_j) = P(T > t_j \mid T > t_{j-1})$$

si ha

$$\hat{S}(t) = \prod_{t_j \leq t} \hat{p}_j \qquad \text{essendo} \qquad \hat{p}_j = 1 - \frac{d_j}{n_j}$$

dove $\hat{p}_j$ è una stima della probabilità $P(T > t_j \mid T > t_{j-1})$, $j = 1, \ldots, k$

Indichiamo con

$$\tilde{S}(t) = \prod_{t_j \leq t} \tilde{p}_j \ , \qquad t \geq 0$$

lo stimatore del quale $\hat{S}(t)$ è la stima.

Per valutare speranza matematica e varianza dello stimatore $\tilde{S}(t)$ occorre formulare delle ipotesi sui n.a. $\tilde{p}_j$, $j = 1, \ldots, k$

Osservazione

Sia

D_j il n.a. dei decessi da una popolazione di n_j individui in vita all'età t_j

nell'ipotesi che D_j abbia distribuzione Binomiale(n_j, q_j) si ha

$$E(D_j) = n_j\, q_j \qquad\qquad Var(D_j) = n_j\, q_j\, (1 - q_j)$$

si ha inoltre che il n.a.

$\dfrac{D_j}{n_j}$ ha distribuzione Binomiale scalata ed è

$$E\left(\frac{D_j}{n_j}\right) = q_j \qquad\qquad Var\left(\frac{D_j}{n_j}\right) = \frac{q_j\, (1 - q_j)}{n_j}$$

Sia $\mathcal{I} = \{n_0, n_1, , \ldots , n_k\}$ lo stato di informazione sugli individui presenti alle età t_j, $j = 1, \ldots, k$ subito prima che si osservino i decessi d_j, $j = 1, \ldots, k$

Si formulano le seguenti ipotesi condizionate a $\mathcal{I} = \{n_0, n_1, , \ldots , n_k\}$ sui n.a.

$$\tilde{p}_j = 1 - \frac{D_j}{n_j}, \qquad j = 1, \ldots, k$$

<u>Ipotesi</u>

Condizionatamente a $\mathcal{I} = \{n_0, n_1, , \ldots , n_k\}$, i n.a.

$$\tilde{p}_j = 1 - \frac{D_j}{n_j}, \qquad j = 1, \ldots, k$$ siano stocasticamente indipendenti

e siano

$$E\left(\tilde{p}_j \middle| \mathcal{I}\right) = p_j \qquad Var\left(\tilde{p}_j \middle| \mathcal{I}\right) = \frac{p_j\,(1 - p_j)}{n_j} \qquad j = 1, \ldots, k$$

Risulta allora che $\tilde{S}(t) = \prod_{t_j \le t} \tilde{p}_j$ è uno stimatore non distorto, infatti

$$E\left(\tilde{S}(t) \middle| \mathcal{I}\right) = E\left(\prod_{t_j \le t} \tilde{p}_j \middle| \mathcal{I}\right) = \prod_{t_j \le t} p_j = S(t) \qquad t \ge 0$$

La varianza dello stimatore $\tilde{S}(t)$ è

$$Var\left(\tilde{S}(t) \middle| \mathcal{I}\right) = [S(t)]^2 \left[\prod_{t_j \le t} \left(1 + \frac{1 - p_j}{p_j n_j}\right) - 1\right]$$

e può essere approssimata da

$$Var\left(\tilde{S}(t) \middle| \mathcal{I}\right) \cong [S(t)]^2 \sum_{t_j \le t} \frac{1 - p_j}{p_j n_j}$$

dalla quale si ottiene la **formula di Greenwood**, che fornisce una stima della varianza dello stimatore di Kaplan-Meier

$$V\hat{a}r\left(\tilde{S}(t) \middle| \mathcal{I}\right) = \left[\prod_{t_j \le t} \left(1 - \frac{d_j}{n_j}\right)\right]^2 \sum_{t_j \le t} \frac{d_j}{(n_j - d_j) n_j}$$

LA STIMA DI NELSON-AALEN PER LA FUNZIONE DI SOPRAVVIVENZA

Fornisce una formula approssimata per la stima della funzione di sopravvivenza di Kaplan-Meier.

Dalla stima di Kaplan-Meier della funzione di sopravvivenza $\hat{S}(t) = \prod_{t_j \le t} \left(1 - \frac{d_j}{n_j}\right) \quad t \ge 0$

si ottiene

$$\log\left(\hat{S}(t)\right) = \sum_{t_j \le t} \log\left(1 - \frac{d_j}{n_j}\right) \qquad t \ge 0$$

Considerando l'approssimazione lineare della funzione $f(x) = \log(1 - x)$ si ottengono le seguenti approssimazioni: $\log\left(1 - \frac{d_j}{n_j}\right) \cong -\frac{d_j}{n_j}$

Si ottiene allora la seguente stima della funzione di sopravvivenza, detta **stima di Nelson-Aalen**

$$\hat{S}(t) = \exp\left(-\sum_{t_j \le t} \frac{d_j}{n_j}\right) \qquad t \ge 0$$

IL MODELLO ATTUARIALE PER LA DESCRIZIONE DELLA SOPRAVVIVENZA

Si tratta di un modello di tipo continuo generalmente di tipo non parametrico, ovvero con funzione di sopravvivenza descritta mediante una tavola di sopravvivenza o tavola di mortalità.

Sia

T_0 n.a. non negativo che esprime la durata aleatoria di vita dalla nascita

Si definiscono

Funzione di sopravvivenza $S(t) = P(T_0 > t), \quad t \geq 0$

Funzione di ripartizione $F_0(t) = P(T_0 \leq t), \quad t \geq 0$

Nel caso di distribuzione dotata di funzione di densità $f_0(t)$ continua si ha:

$$F_0(t) = \int_0^t f_0(u)du \qquad t \geq 0$$

inoltre

$$f_0(t) = \frac{d}{dt} F_0(t) = -\frac{d}{dt} S(t)$$

Si definisce **intensità istantanea di mortalità** o **forza di mortalità**

$$\mu(t) = \lim_{\Delta t \to 0} \frac{P(T_0 \le t + \Delta t | T_0 > t)}{\Delta t}$$

Si ha

$$\mu(t) = \frac{f_0(t)}{S(t)} \qquad \text{da cui} \qquad f_0(t) = S(t)\mu(t)$$

<u>Osservazione</u>: a meno di infinitesimi di ordine superiore al primo si ha

$$P(T_0 \le t + \Delta t | T_0 > t) \cong \frac{f_0(t)}{S(t)} \Delta t = \mu(t)\Delta t$$

Si ha inoltre

$$\mu(t) = \frac{f_0(t)}{S(t)} = -\frac{d}{dt} \ln(S(t))$$

da cui si ottiene

$$S(t) = \exp\left(-\int_0^t \mu(u) du \right)$$

è allora equivalente assegnare il modello di sopravvivenza attraverso la funzione di sopravvivenza oppure l'intensità istantanea di mortalità

Si definisce **vita media alla nascita**

$$\bar{e}_0 = E(T_0) = \int_0^{+\infty} t \, dF_0(t)$$

Si ha

$$\bar{e}_0 = E(T_0) = \int_0^{+\infty} t \, dF_0(t) = \int_0^{+\infty} 1 - F_0(t) \, dt = \int_0^{+\infty} S(t) \, dt$$

Sia

T_x n.a. non negativo che esprime la durata aleatoria di vita per un individuo di età $x > 0$

Poiché

$$T_x = (T_0 - x) | T_0 > x$$

si può esprimere la distribuzione di probabilità del n.a. T_x attraverso la distribuzione di T_0

Sia $F_x(t) = P(T_x \leq t), \quad t \geq 0$

si ha

$$F_x(t) = 1 - \frac{S(x+t)}{S(x)}$$

Si definisce la funzione di sopravvivenza

$$_t p_x = P(T_x > t) = 1 - F_x(t) = \frac{S(x+t)}{S(x)}$$

Si definisce la funzione di ripartizione della distribuzione condizionata

$$_t q_x = 1 - {}_t p_x = F_x(t)$$

Si può esprimere inoltre la funzione di densità condizionata

$$f_x(t) = \frac{d}{dt} F_x(t) = -\frac{1}{S(x)} \frac{d}{dt} S(x+t) = \frac{f_0(x+t)}{S(x)}$$

Ricordando che l'intensità istantanea di mortalità relativa alla distribuzione di T_0 è

$$\mu(t) = \lim_{\Delta t \to 0} \frac{P(T_0 \leq t + \Delta t | T_0 > t)}{\Delta t} = \frac{f_0(t)}{S(t)}$$

per definire l'intensità istantanea di mortalità relativa alla distribuzione di T_x si considera

$$\lim_{\Delta t \to 0} \frac{P(T_x \leq t + \Delta t | T_x > t)}{\Delta t} = \frac{f_x(t)}{_t p_x} = \frac{f_0(x+t)}{S(x+t)} = \mu(x+t)$$

quindi l'intensità istantanea di mortalità della distribuzione condizionata coincide con l'intensità istantanea di mortalità della distribuzione di T_0

Poiché

$$\frac{f_x(t)}{{}_t p_x} = \frac{f_0(x+t)}{S(x+t)} = \mu(x+t)$$

si ha

$$f_x(t) = {}_t p_x \ \mu(x+t)$$

<u>Osservazione</u>: a meno di infinitesimi di ordine superiore al primo si ha

$$P\left(T_x \leq t + \Delta t \middle| T_x > t\right) \cong \mu(x+t)\Delta t$$

Si definisce **vita media residua** di una persona di età x

$$\bar{e}_x = E(T_x) = \int_0^{+\infty} t \ dF_x(t)$$

Si ha

$$\bar{e}_x = E(T_x) = \int_0^{+\infty} t \ dF_x(t) = \int_0^{+\infty} 1 - F_x(t) \ dt = \int_0^{+\infty} {}_t p_x \ dt = \int_0^{+\infty} \frac{S(x+t)}{S(x)} \ dt$$

Si definisce **tasso centrale di mortalità** relativo all'intervallo di età $(x, x+1)$

$$m_x = \frac{\int_0^1 \mu(x+t) S(x+t)\, dt}{\int_0^1 S(x+t)\, dt}$$

Più in generale si può definire il **coefficiente di mortalità** relativo all'intervallo di età $(x, x+n)$

$$_n m_x = \frac{\int_0^n \mu(x+t) S(x+t)\, dt}{\int_0^n S(x+t)\, dt}$$

TAVOLE DI MORTALITÀ

In ambito attuariale la funzione di sopravvivenza è generalmente descritta mediante una **tavola di mortalità** o **tavola di sopravvivenza**.

La tavola di sopravvivenza può essere interpretata come la tabulazione sugli interi di una funzione di sopravvivenza $S(x)$ definita per $x \geq 0$, quindi di tipo continuo.

x	$S(x)$
0	1
1	0,99172
2	0,99105
$\vdots$	$\vdots$
$\omega - 1$	0,00015
ω	0

ω è detta età estrema ed è tale che $S(\omega - 1) > 0$ ed $S(\omega) = 0$

Definiamo le grandezze che compaiono in una tavola di sopravvivenza.

x	l_x	d_x	q_x	L_x	m_x	T_x	$\overset{o}{e}_x$
0	100.000	828	0,008280	99.586,00	0,008314	7.881.389,00	78,81
1	99.172	67	0,000676	99.138,50	0,000676	7.781.803,00	78,47
2	99.105	42	0,000424	99.084,00	0,000424	7.682.664,50	77,52
3	99.063	32	0,000323	99.047,00	0,000323	7.583.580,50	76,55
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
109	30	15	0,500000	22,50	0,666667	30,00	1,00
110	15	15	1,000000	7,50	2,000000	7,50	0,50
111	0						

Si definiscono le seguenti grandezze

Def. $$l_x = l_0\, S(x) \qquad x = 0, 1, \dots, \omega$$

dove l_0 è detto radice della tavola ed è fissato opportunamente, per esempio $l_0 = 100.000$

l_x esprime il numero atteso di individui in vita all'età x, a partire da una collettività di l_0 neonati, nell'ipotesi che la sopravvivenza sia descritta dalla funzione di sopravvivenza $S(x)$

Def. $$d_x = l_x - l_{x+1} \qquad x = 0, 1, \dots, \omega$$

d_x esprime il numero atteso di decessi nell'intervallo di età $]x, x+1]$

Si ha

$$q_x = 1 - p_x = 1 - \frac{S(x+1)}{S(x)} = 1 - \frac{l_{x+1}}{l_x} = \frac{d_x}{l_x}$$

Per definire le seguenti grandezze

$$L_x \qquad m_x \qquad T_x \qquad \overset{o}{e}_x$$

occorre assumere che la distribuzione di probabilità, espressa attraverso la funzione di sopravvivenza $S(x)$, $x \geq 0$, sia dotata di funzione di densità continua

Si ha allora

$$f_0(x) = -\frac{d}{dx} S(x) = -\frac{d}{dx}\frac{l_x}{l_0} \qquad \mu(x) = \frac{f_0(x)}{S(x)} = -\frac{1}{l_x}\frac{d}{dx} l_x$$

$$f_x(t) = -\frac{d}{dt}{}_t p_x = -\frac{1}{l_x}\frac{d}{dt} l_{x+t} \qquad \mu(x+t) = \frac{f_x(t)}{{}_t p_x} = -\frac{1}{l_{x+t}}\frac{d}{dt} l_{x+t}$$

e da queste si può calcolare la vita media alla nascita $\bar{e}_0$ e la vita media residua $\bar{e}_x$ di un individuo di età x

Si ha

$$\bar{e}_0 = \int_0^{\omega} S(x)\, dx = \int_0^{\omega} \frac{l_x}{l_0}\, dx = \frac{\int_0^{\omega} l_x dx}{l_0}$$

Si definisce

Def. $$T_0 = \int_0^{\omega} l_x dx$$

Poiché

$$\bar{e}_0 = \frac{T_0}{l_0}$$

esprime la vita media alla nascita, allora

T_0 (NB: da non confondere con il n.a. durata aleatoria di vita alla nascita) rappresenta il numero atteso di anni vissuti da una collettività di l_0 neonati

Analogamente si ha

$$\bar{e}_x = \int_0^{\omega-x} {}_tp_x \, dt = \int_0^{\omega-x} \frac{l_{x+t}}{l_x} \, dt = \frac{\int_0^{\omega-x} l_{x+t} dt}{l_x} = \frac{\int_x^{\omega} l_y dy}{l_x}$$

Si definisce

Def. $$T_x = \int_x^{\omega} l_y dy$$

Poiché

$$\bar{e}_x = \frac{T_x}{l_x}$$

esprime la vita media residua di un individuo di età x, allora

T_x (NB: da non confondere con il n.a. durata aleatoria di vita per un individuo di età x) rappresenta il numero atteso di anni vissuti da una collettività di l_x individui di età x

Per il tasso centrale di mortalità relativo all'intervallo di età $(x, x+1)$ si ha

$$m_x = \frac{\int_0^1 \mu(x+t) S(x+t)\, dt}{\int_0^1 S(x+t)\, dt} = \frac{\int_0^1 f_0(x+t)\, dt}{\int_0^1 S(x+t)\, dt} = \frac{S(x) - S(x+1)}{\int_0^1 S(x+t)\, dt} = \frac{l_x - l_{x+1}}{\int_0^1 l_{x+t}\, dt} = \frac{d_x}{\int_0^1 l_{x+t}\, dt}$$

Si definisce

Def. $$L_x = \int_0^1 l_{x+t}\, dt$$

Poiché

$$L_x = T_x - T_{x+1}$$

L_x esprime il numero atteso di anni vissuti da una popolazione di l_x individui di età x, tra le età x ed $x+1$

Si ha allora

$$m_x = \frac{d_x}{L_x}$$

<u>Osservazione</u>: poiché nella tavola di sopravvivenza la funzione $S(x)$ è definita sugli interi $x = 0, 1, \dots, \omega$, per calcolare

$$L_x = \int_0^1 l_{x+t}\, dt$$

occorre approssimare l'integrale.

Se si approssima l'integrale mediante l'area del trapezio, ovvero si considera l'interpolante lineare per definire la funzione l_{x+t} per $t \in (0,1)$, si ottiene la seguente approssimazione

$$L_x \cong \frac{l_x + l_{x+1}}{2}$$

Essendo inoltre $T_x = \int_x^{\omega} l_y\, dy = L_x + L_{x+1} + \cdots + L_{\omega-1}$

si ottiene la seguente approssimazione

$$T_x \cong \sum_{h=0}^{\omega-x-1} \frac{l_{x+h} + l_{x+h+1}}{2}$$

e quindi l'approssimazione della vita media residua $\bar{e}_x$ mediante la **vita media completa**

$$\overset{o}{e}_x = \frac{1}{l_x} \sum_{h=0}^{\omega-x-1} \frac{l_{x+h} + l_{x+h+1}}{2}$$

<u>Osservazione</u>

poiché si ha

$$S(x) = \frac{S(x)}{S(x-1)} \cdot \frac{S(x-1)}{S(x-2)} \cdots \frac{S(1)}{S(0)} \qquad \text{per } x = 1, 2, \ldots, \omega - 1$$

ed essendo

$$\frac{S(x)}{S(x-1)} = \frac{P(T_0 > x)}{P(T_0 > x-1)} = \frac{P(T_0 > x, T_0 > x-1)}{P(T_0 > x-1)} = P(T_0 > x \mid T_0 > x-1) = p_{x-1}$$

per $x = 1, 2, \ldots, \omega - 1$

si ottiene

$$S(x) = p_{x-1} \cdot p_{x-2} \cdots p_0 \qquad \text{per } x = 1, 2, \ldots, \omega - 1$$

Quindi la stima della funzione di sopravvivenza, in un modello non parametrico, potrà essere ottenuta mediante la stima delle probabilità condizionate di sopravvivenza p_x ovvero mediante la stima delle probabilità condizionate di decesso q_x.

STIMA DELLE PROBABILITÀ DI DECESSO MEDIANTE LO STIMATORE DI KAPLAN-MEIER

Obiettivo: stimare le probabilità q_x , $x = 0,1,\ldots,\omega - 1$

Consideriamo la generica classe di età $]x, x+1]$ e supponiamo di avere osservato in tale classe di età

$d_1, d_2, \ldots, d_k$ decessi alle età $t_1, t_2, \ldots, t_k$ con $x < t_1 < t_2 < \ldots < t_k \leq x+1$

essendo

n_j, $j = 1, \ldots, k$, il numero di individui in vita all'età t_j che hanno dato luogo ai d_j decessi

Si noti che $t_0 = x$ rappresenta l'istante iniziale

La stima di Kaplan-Meier per la probabilità q_x è

$$\hat{q}_x = 1 - \hat{p}_x = 1 - \prod_{j=1}^{K} \left(1 - \frac{d_j}{n_j} \right)$$

ESPOSTI AL RISCHIO

Nella tavola di sopravvivenza sono riportate le seguenti grandezze

$q_x = \frac{d_x}{l_x}$ probabilità che un individuo in vita all'età x, deceda con età esatta in $]x, x+1]$

$m_x = \frac{d_x}{L_x}$ tasso centrale di mortalità, è un'intensità

Per esprimere q_x dato m_x e viceversa occorre stabilire un legame tra l_x e L_x

Sussistono le seguenti relazioni

$$L_x = \int_0^1 l_{x+t}\, dt = \int_x^{x+1} l_y\, dy = T_x - T_{x+1} = l_{x+1} + l_x \int_0^1 t\, f_x(t)\, dt$$

L_x esprime il numero atteso di anni vissuti nell'intervallo di età $]x, x+1]$ dagli l_x individui in vita all'età x; è anche detto **esposizione** o **numero di esposti al rischio** nell'intervallo di età $]x, x+1]$

Per giustificare formalmente tale interpretazione di L_x consideriamo i seguenti n.a.

$T^{(i)}$ durata aleatoria di vita dell'individuo i, $i=1, 2, \dots, l_x$, nell'intervallo di età $\left]x, x+1\right]$

<u>Nota</u>: $T^{(i)}$ ha determinazioni $\left]0, 1\right]$

Indicato con $T_x^{(i)}$ la durata aleatoria di vita dell'i-esimo individuo in vita all'età x si ha

$$T^{(i)} = \min\left(T_x^{(i)}, 1\right)$$

Nell'ipotesi che le durate aleatorie di vita degli l_x individui siano ugualmente distribuite con funzione di ripartizione $F_x(t)$ si ha

$$P(T^{(i)} \le t) = \begin{cases} 0 & t \le 0 \\ F_x(t) & 0 < t < 1 \\ 1 & t \ge 1 \end{cases}$$

ed è inoltre $P(T^{(i)} = 1) = 1 - F_x(1) = p_x$

Il numero atteso L_x di anni vissuti nell'intervallo di età $\left]x, x+1\right]$ dagli l_x individui in vita all'età x è allora

$$E\left(\sum_{i=1}^{l_x} T^{(i)}\right) = \sum_{i=1}^{l_x} E\left(T^{(i)}\right) = \sum_{i=1}^{l_x}\left[\int_0^1 t\, dF_x(t) + 1\, p_x\right] = l_x \int_0^1 t\ f_x(t)\, dt + l_x \frac{l_{x+1}}{l_x} = l_x \int_0^1 t\ f_x(t)\, dt + l_{x+1}$$

Posto

$$\bar{t}_x = \frac{l_x \int_0^1 t\, f_x(t)\, dt}{d_x}$$

la durata attesa di vita di ciascun individuo che decede in $]x, x+1]$

dalla relazione

$$L_x = l_{x+1} + l_x \int_0^1 t\, f_x(t)\, dt$$

si ottiene

$$L_x = l_x - d_x(1 - \bar{t}_x)$$

Si possono esprimere quindi le seguenti relazioni

$$m_x = \frac{d_x}{L_x} = \frac{q_x}{1-(1-\bar{t}_x)q_x} \qquad q_x = \frac{d_x}{l_x} = \frac{m_x}{1+(1-\bar{t}_x)m_x}$$

Se $\bar{t}_x = 1/2$ (ciò si ha in ipotesi di distribuzione uniforme dei decessi) si ha

$$m_x = \frac{d_x}{L_x} = \frac{q_x}{1-1/2\, q_x} \qquad q_x = \frac{d_x}{l_x} = \frac{m_x}{1+1/2\, m_x}$$

COMPLETAMENTO DEL MODELLO DI SOPRAVVIVENZA PER ETÀ NON INTERE

Nella tavola di mortalità sono riportati l_x $x = 0, 1, \dots, \omega$

In molti casi si ha la necessità di definire l_x per qualsiasi età $x \geq 0$

<u>Obiettivo</u>: definire l_{x+t} , $x = 0, 1, \dots, \omega$ e $0 \leq t \leq 1$ tale che

l_{x+t} continua in $[0, 1]$ e derivabile in $(0, 1)$

Interpolazione lineare o ipotesi di distribuzione uniforme dei decessi

$$l_{x+t} = a + b\, t \quad \Rightarrow \quad l_{x+t} = l_x - t\, d_x$$

$$\Rightarrow \quad l_{x+t} = (1 - t)\, l_x + t\, l_{x+1}$$

In tale ipotesi si ha

$$_{s-r}p_{x+r} = \frac{1 - s\, q_x}{1 - r\, q_x} \qquad _s q_x = s\, q_x \qquad \mu(x+t) = \frac{q_x}{1 - t\, q_x} \qquad f_x(t) = q_x$$

$$L_x = l_{x+1} + \frac{1}{2} d_x = \frac{l_{x+1} + l_x}{2}$$

Interpolazione esponenziale o ipotesi di intensità di mortalità costante

$$l_{x+t} = a \cdot b^t \qquad \Rightarrow \quad l_{x+t} = l_x \, (p_x)^t$$

$$\Rightarrow \quad l_{x+t} = (l_x)^{1-t} \cdot (l_{x+1})^t$$

In tale ipotesi si ha

$$\mu(x+t) = -\log p_x$$

$${}_{s-r}p_{x+r} = e^{-\mu_x (s-r)} \qquad \text{con} \qquad \mu_x = -\log p_x$$

$$f_x(t) = e^{-\mu_x t} \mu_x \qquad L_x = \frac{d_x}{\mu_x} \qquad m_x = \mu_x$$

Interpolazione iperbolica

$$l_{x+t} = \frac{1}{a + b\,t} \qquad \Rightarrow \quad \frac{1}{l_{x+t}} = \frac{1}{l_x} + \left(\frac{1}{l_{x+1}} - \frac{1}{l_x} \right) t$$

$$\Rightarrow \quad \frac{1}{l_{x+t}} = \frac{1}{l_{x+1}}\, t + \frac{1}{l_x}(1 - t)$$

In tale ipotesi si ha

$$_{1-r}q_{x+r} = (1 - r)\, q_x \qquad _t p_x = \frac{1}{1 + \left(\frac{l_x}{l_{x+1}} - 1 \right) t}$$

$$\mu(x+t) = \frac{q_x}{p_x + t\, q_x}$$

nota: decrescente con t

MODELLO DI SOPRAVVIVENZA SELEZIONATO

Si utilizza un modello selezionato per descrivere la durata aleatoria di vita da un'assegnata età in cui l'individuo è selezionato; per esempio, nel caso delle assicurazioni, l'età di selezione è l'età (intera) di ingresso in assicurazione.

Funzione di sopravvivenza $S(t;x), \quad t \geq 0, \qquad x = a, a+1, \ldots$

Nel caso di distribuzione dotata di funzione di densità continua, per assegnare la distribuzione di probabilità si può assegnare l'intensità istantanea di mortalità

$$\mu_{[x]}(t), \quad t \geq 0, \qquad x = a, a+1, \ldots$$

Si ha $\mu_{[x]}(t) = -\frac{d}{dt} \ln(S(t;x))$

da cui si ottiene

$$ {}_t p_{[x]} = S(t;x) = \exp\left(-\int_0^t \mu_{[x]}(u) du \right)$$

Si definisce **aspettativa di vita** per una persona entrata in assicurazione all'età x

$$\overline{e}_{[x]} = \int_0^{+\infty} S(t;x)\, dt$$

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

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

dove $$l_{[x]+t} = l_{[x]} \cdot S(t;x) \qquad x = a, a+1, \ldots \qquad t = 0,1,\ldots$$

$l_{[x]}$ è la radice della tavola di sopravvivenza degli assicurati entrati in assicurazione all'età x

$l_{[x]+t}$ rappresenta il numero atteso di assicurati in vita all'età $x+t$ a partire da una collettività di $l_{[x]}$ assicurati entrati in assicurazione all'età x; t è l'antidurata

Si ha $$_t p_{[x]} = \frac{l_{[x]+t}}{l_{[x]}} \qquad\qquad _t q_{[x]} = 1 - {}_t p_{[x]}$$

Per effetto della selezione sulla mortalità si ha

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \dots$$

Poiché l'effetto della selezione sulla mortalità tende ad esaurirsi dopo un certo numero di anni, da un certa antidurata t' in poi la sopravvivenza viene fatta dipendere soltanto dall'età raggiunta

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

Si costruiscono allora le **tavole di mortalità 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

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

RIFERIMENTI BIBLIOGRAFICI

Klein. J.P., Moeschberger M. L. , Survival Analysis – Techniques for Censored and Truncated Data, Springer-Verlag, 1997 (Cap. 2, 3, 4)

D. London, Survival models and their estimation, Actex publications, 1997 (Cap. 1, 2, 3)

E. Pitacco, Matematica e tecnica attuariale delle assicurazioni sulla durata di vita, Lint, 2002 (Cap. 2, App. A)