MODELLI MULTISTATO

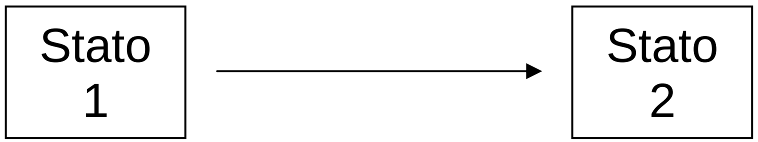
- Introduzione ai modelli multistato
- Esempio di modello multistato per descrivere la progressione di una malattia
- I modelli multistato
- Un modello multistato per l'assicurazione malattia

INTRODUZIONE AI MODELLI MULTISTATO

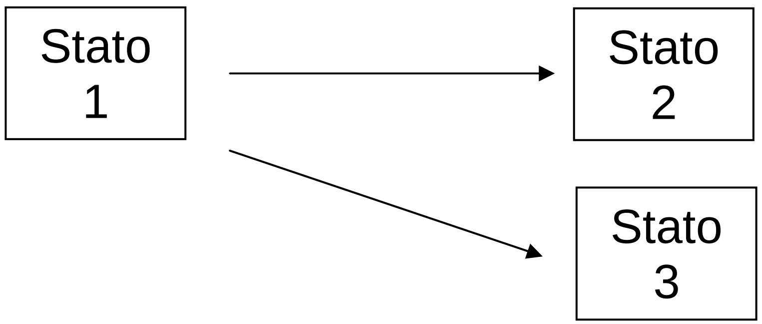
Il modello per la durata aleatoria fino al verificarsi di un determinato evento può essere visto come un particolare modello multistato con due stati soltanto:

Stato 1 l'individuo è in vita
Stato 2 l'individuo è deceduto

In tale modello è possibile soltanto il passaggio dallo stato 1 allo stato 2



Un modello con due cause di uscita può invece essere descritto mediante un modello multistato con tre stati:

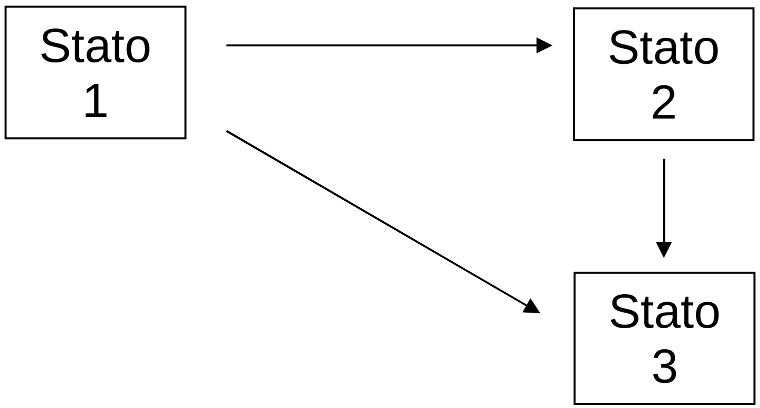


Stato 3 l'individuo è uscito per altra causa

In tali modelli le probabilità di passaggio tra gli stati possono essere descritte mediante le intensità (di mortalità e di uscita per altra causa).

In situazioni più generali si parla di intensità di passaggio da uno stato all’altro.

Un altro esempio di modello a tre stati è l’assicurazione di invalidità permanente.



Stato 1	l’individuo è attivo
Stato 2	l’individuo è invalido
Stato 3	l’individuo è deceduto

La probabilità di passaggio dallo stato 1 allo stato 2 dipende dall’intensità di invalidità; la probabilità di passaggio dallo stato 2 allo stato 3 dipende dall’intensità di mortalità per gli invalidi e la probabilità di passaggio dallo stato 1 allo stato 3 dipende dall’intensità di mortalità per gli attivi.

ESEMPIO DI MODELLO MULTISTATO
PER DESCRIVERE LA PROGRESSIONE DI UNA MALATTIA

Il modello di Panjer per descrivere la progressione dell’AIDS è un modello a 6 stati

Stato 1a l’individuo è non infetto
Stato 1b l’individuo è sieropositivo ma non è ammalato

I successivi tre stati descrivono i progressivi stadi della malattia:

Stato 2a *Stato 2b* *Stato 3*

L’ultimo stato riguarda il decesso:

Stato 4 l’individuo è deceduto

Stato 1a → Stato 1b → Stato 2a → Stato 2b → Stato 3 → Stato 4

Poiché sono possibili solo i passaggi da uno stato a quello successivo, le durate di permanenza nei singoli stati sono descritte da modelli ad una sola causa di uscita e le intensità di uscita sono dette <u>intensità di progressione</u>

Sia

T_j durata aleatoria di permanenza nello stato j, $\quad j = 1a, 1b, 2a, 2b, 3$

Ipotesi:

T_j , $j = 1a, 1b, 2a, 2b, 3$ stocasticamente indipendenti

intensità di progressione costanti: μ_j , $\quad j = 1a, 1b, 2a, 2b, 3$

In tali ipotesi, la durata di permanenza in uno stato non dipende dalla durata di permanenza negli altri stati.

$$P(T_j > t) = e^{-\mu_j t} \qquad\qquad f_{T_j}(t) = e^{-\mu_j t}\mu_j$$

Sia

$p_{j,\, j+1}(t)$ probabilità che un individuo presente nello stato j, sia nello stato $j+1$ dopo t anni

$$p_{j,\, j+1}(t) = \int_0^t e^{-\mu_j r}\mu_j\; e^{-\mu_{j+1}(t-r)}dr$$

Per la stima del modello, se si dispone delle informazioni sulle durate esatte in cui si sono avuti i passaggi tra i vari stati, le stime di massima verosimiglianza delle intensità di progressione sono date dai rapporti tra le numerosità dei passaggi, da uno stato all'altro, e le esposizioni totali esatte.

Spesso però informazioni di questo tipo non sono disponibili; per esempio, nello studio di Panjer i dati disponibili erano:

- durata di tempo in cui l'individuo è stato osservato, raggruppate in 4 gruppi di durate ($i = 1, 2, 3, 4$):

  3-6 mesi; 6-12 mesi; 12-24 mesi; 24-36 mesi
- stato in cui l'individuo si trovava all'inizio dell'osservazione
- informazione se durante l'osservazione l'individuo è rimasto nello stesso stato oppure se ne è uscito

I parametri $\mu_j, \; j = 1a, 1b, 2a, 2b, 3,$ possono essere stimati separatamente massimizzando le verosimiglianze

$$L(\mu_j) \qquad j = 1a, 1b, 2a, 2b, 3$$

Con riferimento allo stato j, $j = 1a, 1b, 2a, 2b, 3$, siano

- n_i il numero di individui osservati appartenenti all'i-esimo gruppo di durate, $i = 1, 2, 3, 4$

- d_i il numero di individui relativi all'i-esimo gruppo di durate, $i = 1, 2, 3, 4$, usciti dallo stato j durante l'osservazione

- r_i la durata media di osservazione per gli individui relativi all'i-esimo gruppo di durate, $i = 1, 2, 3, 4$, e non usciti dallo stato j durante l'osservazione (per semplicità si assume $r_1 = 4{,}5$ mesi, $r_2 = 9$ mesi, $r_3 = 18$ mesi, $r_4 = 30$ mesi)

- p_i la probabilità che un individuo appartenente all'i-esimo gruppo di durate, $i = 1, 2, 3, 4$, rimanga nello stato j durante l'osservazione

La verosimiglianza in funzione del parametro μ_j è allora

$$L(\mu_j) = \prod_{i=1}^{4} (1 - p_i)^{d_i} \, p_i^{\,n_i - d_i}$$

essendo $p_i = e^{-\mu_j r_i}$

Dalla logverosimiglianza

$$l(\mu_j) = \sum_{i=1}^{4} d_i \log(1 - p_i) + (n_i - d_i) \log(p_i)$$

si ottiene l'equazione di verosimiglianza che si può risolvere per via numerica.

$$\sum_{i=1}^{4} -\frac{\partial p_i}{\partial \mu_j} \left(\frac{d_i}{1 - p_i} - \frac{n_i - d_i}{p_i} \right) = 0$$

I MODELLI MULTISTATO

Sia

$\mathscr{S} = \{1, 2, \ldots, N\}$ l'insieme degli stati

$\mathscr{T}$ l'insieme delle transizioni

$\mathscr{T} \subseteq \{(i, j) | i \neq j; i, j \in \mathscr{S}\}$

Se lo stato 1 indica lo stato iniziale all'istante 0, si ipotizza che tutti gli stati $j \in \mathscr{S}$ siano raggiungibili dallo stato 1 mediante transizioni dirette o indirette.

Si definisce **modello multistato** la coppia $(\mathscr{S}, \mathscr{T})$

Con riferimento ad un rischio (per esempio un individuo) sia

$S(t)$ lo stato occupato dal rischio al tempo $t \geq 0$

Si ipotizza $S(0) = 1$

$\{S(t); t \geq 0\}$ è un processo stocastico a parametro continuo con valori in $\mathscr{S}$

si indica con $s(z)$ lo stato occupato dal processo nell'istante $z > 0$ in una realizzazione $\{s(t)\}$ del processo stocastico $\{S(t); t \geq 0\}$

Si dirà che il processo $\{S(t); t \geq 0\}$ è una **catena markoviana a parametro continuo** se

per ogni $0 \leq \tau < u$ e per ogni i, j, $s(z)$ con $0 \leq z < \tau$, $i, j, s(z) \in \mathscr{S}$ tali che

$$P\big((S(z) = s(z)) \wedge (S(\tau) = i) \wedge (S(u) = j)\big) > 0$$

si ha

$$P\big(S(u) = j \big| (S(z) = s(z)) \wedge (S(\tau) = i)\big) = P\big(S(u) = j \big| S(\tau) = i\big)$$

Le probabilità condizionate

$$P_{ij}(t,u) = P\big(S(u) = j \big| S(t) = i\big) \qquad 0 \leq t \leq u$$

sono dette <u>probabilità di transizione</u>; essendo $P_{ij}(t,t) = 0$ per $i \neq j$ e $P_{ij}(t,t) = 1$ per $i = j$

Le probabilità di transizione soddisfano le seguenti proprietà:

$$0 \leq P_{ij}(t,u) \leq 1 \qquad \text{per ogni } i, j;\ 0 \leq t \leq u$$

$$\sum_{j \in \mathscr{S}} P_{ij}(t,u) = 1 \qquad \text{per ogni } i;\ 0 \leq t \leq u$$

Si definiscono inoltre le <u>probabilità di permanenza</u>

$$P_{\underline{ii}}(t,u) = P\big(S(z) = i \text{ per ogni } z \in [t,\, u] \,\big|\, S(t) = i\big)$$

Le probabilità di transizione soddisfano le <u>equazioni di Chapman-Kolmogorov</u>

$$P_{ij}(t,u) = \sum_{k \in \mathscr{S}} P_{ik}(t,w) P_{kj}(w,u) \qquad\qquad t \leq w \leq u$$

Si definiscono le <u>intensità di transizione</u>

$$\mu_{ij}(t) = \lim_{u \to t} \frac{P_{ij}(t,u)}{u - t}$$

Si prova che

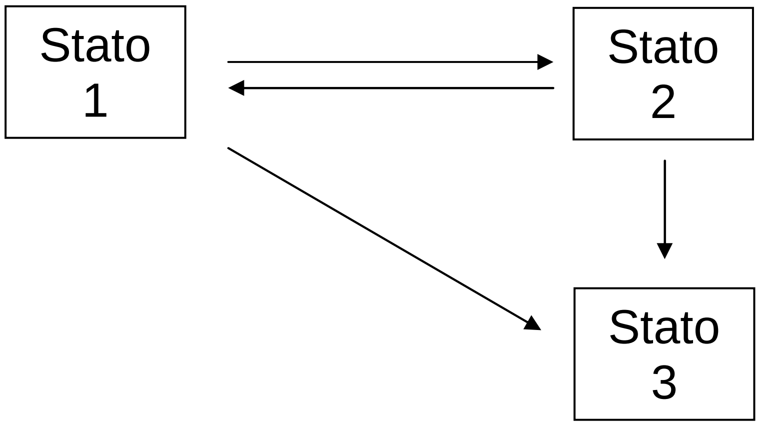
$$P_{\underline{ii}}(z,t) = \exp\left[-\int_z^t \mu_i(u)\, du \right]$$

Si prova inoltre che sotto condizioni generali per le intensità di transizione si possono determinare, a partire da esse, le probabilità di transizione.

<u>Transition intensity approach</u>: assegnate le intensità di transizione, da queste si determinano le probabilità di transizione.

UN MODELLO MULTISTATO PER L'ASSICURAZIONE MALATTIA

Un esempio importante è il modello per l'assicurazione malattia (Permanent Health Insurance – PHI)



Stato 1 l'individuo è attivo

Stato 2 l'individuo è invalido

Stato 3 l'individuo è deceduto

Supponiamo che le intensità di transizione siano costanti, indipendenti dall'età raggiunta:

$$\mu_{12}(t) = \mu_{12} \qquad \mu_{13}(t) = \mu_{13} \qquad \mu_{21}(t) = \mu_{21} \qquad \mu_{23}(t) = \mu_{23}$$

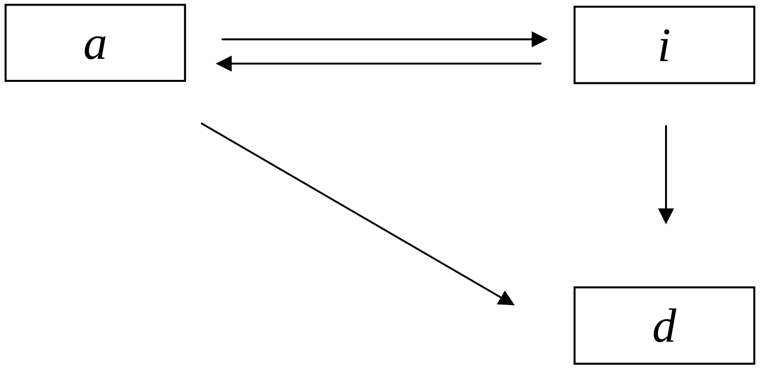
da esse si possono ottenere le probabilità di permanenza nello stato di attivo o di invalido e le probabilità di decesso, rispettivamente dallo stato di attivo e dallo stato di invalido.

Il modello può essere utilizzato per stimare le intensità di transizione dipendenti soltanto dall'età e costante a tratti su ogni intervallo di età $(x, x+1)$

Siano

$$\mu_x^{ai} \qquad \mu_x^{ia} \qquad \mu_x^{ad} \qquad \mu_x^{id}$$

le intensità di transizione relative alla classe di età $(x, x+1)$



Supponiamo che si disponga di osservazioni su individui nella classe di età e che per ogni individuo siano rilevati

le durate di tempo tra due transizioni successive
i numeri di transizioni di ciascun tipo

Si prova che la verosimiglianza delle osservazioni è

$$L\left(\mu_x^{ai},\mu_x^{ia},\mu_x^{ad},\mu_x^{id}\right)=\exp\left[-\left(\mu_x^{ai}+\mu_x^{ad}\right)c\right]\cdot\exp\left[-\left(\mu_x^{ia}+\mu_x^{id}\right)w\right]\cdot(\mu_x^{ad})^{d}\cdot(\mu_x^{id})^{u}\cdot(\mu_x^{ai})^{s}\cdot(\mu_x^{ia})^{r}$$

dove

- c è il tempo totale osservato nello stato di attivo
- w è il tempo totale osservato nello stato di invalido
- d è il numero totale di transizioni dallo stato di attivo allo stato di deceduto
- u è il numero totale di transizioni dallo stato di invalido allo stato di deceduto
- s è il numero totale di transizioni dallo stato di attivo allo stato di invalido
- r è il numero totale di transizioni dallo stato di invalido allo stato di attivo

Si ottengono le seguenti stime di massima verosimiglianza:

$$\hat{\mu}_x^{ad}=\frac{d}{c}\qquad\hat{\mu}_x^{id}=\frac{u}{w}\qquad\hat{\mu}_x^{ai}=\frac{s}{c}\qquad\hat{\mu}_x^{ia}=\frac{r}{w}$$

Anche per tale modello si pone il problema della perequazione delle stime ottenute. In letteratura sono stati proposti a tale scopo modelli GLM.

RIFERIMENTI BIBLIOGRAFICI

S. Haberman, E. Pitacco, Actuarial models for disability insurance, Chapman & Hall, 1999 (Par. 1.1, 1.3, 1.4, 4.1, 4.2)

D. London, Survival models and their estimation, Actex publications, 1997 (Par. 10.1, 10.2)