

# Analisi di tempi di comparsa di eventi ricorrenti in bio-medicina

Programmo-  
teca

**Maria Grazia Valsecchi**, Istituto di Statistica Medica e Biometria e Centro Zambon, Università di Milano

In questo lavoro si descrive un programma che consente di eseguire l'analisi di dati emergenti da situazioni cliniche in cui la malattia studiata presenta manifestazioni patologiche ricorrenti.

Riferendosi a metodologie tipiche dell'analisi della sopravvivenza, esso permette di descrivere l'esperienza di vita dei pazienti tenendo conto non solo del numero di recidive che ognuno di essi presenta durante il follow-up, ma anche del loro tempo di comparsa. Esso permette di valutare se e come varia per un soggetto la possibilità di manifestare un evento patologico in relazione al numero di eventi precedenti.

## 1. Introduzione

Nel campo dell'analisi statistica di dati clinici si incontrano a volte situazioni in cui la malattia studiata presenta manifestazioni patologiche ricorrenti quali le emorragie nella cirrosi epatica, le ricorrenze dolorose nella pancreatite cronica ecc.. In tali casi al clinico può essere utile sapere se e come varia per il paziente la probabilità di avere una recidiva in relazione al numero di recidive precedenti.

Le tecniche statistiche che sembrano più appropriate sono quelle che, ricorrendo alle metodologie su cui si basa quella branca della statistica medica che si occupa dell'analisi della sopravvivenza, permettono di descrivere l'esperienza di vita dei pazienti tenendo conto non solo del numero di recidive che ognuno di essi presenta durante lo studio, ma anche del tempo in cui esse si manifestano. La variabile di risposta considerata è, dunque, il tempo in cui compare per la prima volta e per tutte le volte successive l'evento ricorrente che si intende studiare.

In questo lavoro si descrive un programma che consente di eseguire l'analisi di dati emergenti da situazioni cliniche del tipo sopra descritto. Prima di entrare nella descrizione dettagliata della procedura adottata si vuole porre in evidenza la sua caratteristica peculiare che permette, senza ricorrere all'uso di variabili tempo-dipendenti (Cox, 1972), di tenere conto del tempo di comparsa delle recidive.

Il paziente che entra nello studio al momento della prima diagnosi appartiene all'insieme di soggetti che non hanno avuto alcuna recidiva. Nel momento in cui, durante il periodo di osservazione, egli presenti una prima recidiva, diviene un membro del gruppo di soggetti caratterizzati dall'aver manifestato una sola volta tale evento patologico; qualora in un altro momento di osservazione presenti una seconda recidiva, di-

viene, a quel punto, un membro del gruppo di pazienti con due recidive e così via. Si permette quindi ad un paziente che durante il periodo di osservazione presenta  $n$  recidive di passare attraverso  $n + 1$  gruppi in modo tale che gli  $n$  passaggi avvengano esattamente al tempo di comparsa dell'evento ricorrente.

Se un soggetto, ad un certo momento durante lo studio, muore o è perso di vista, il tempo in cui ciò accade è considerato troncato (censored), poiché da quel momento non è più possibile osservare sul paziente il manifestarsi dell'evento patologico studiato.

## 2. Descrizione del metodo

Il metodo adottato fa riferimento a quello introdotto da Chan *et al.* (1982) per lo studio di eventi patologici ricorrenti in malattie congenite. Esso, con opportune modificazioni, è qui applicato anche nell'analisi di risposte di pazienti in un follow-up concernente anche malattie non congenite.

Nel caso in cui la malattia studiata sia congenita, si prefissano, sull'asse dei tempi, intervalli di ampiezza costante opportunamente scelta che rappresentano le classi di età.

Un paziente affetto da tale malattia entra nello studio a partire dall'intervallo di tempo in cui cade la sua età alla diagnosi; in seguito, se il paziente presenta l'evento patologico studiato, si considera come variabile di risposta la sua età al verificarsi dell'evento e si effettua quindi il suo passaggio da un gruppo al successivo nell'intervallo corrispondente alla sua classe di età. In questo modo si tiene conto del carattere congenito della malattia poiché l'origine dell'asse dei tempi rappresenta l'anno di nascita ed il rischio di presentare l'evento patologico è « corretto » per la variabile: età del soggetto (Chang *et al.*, 1982). Questo tipo di utilizzo dell'asse dei tempi è applicabile anche a malattie non congenite qualora per queste si riferisca l'origine dell'asse dei tempi all'inizio del follow-up e si prefissino intervalli di lunghezza costante opportuna.

Ogni soggetto cui è diagnosticata la malattia (non congenita) entra nello studio al tempo zero e, qualora presenti un evento patologico, si considera come variabile di risposta il tempo intercorrente tra la diagnosi e la comparsa di tale evento. Questo valore individua l'intervallo di tempo in cui il paziente esce dal gruppo cui appartiene ed entra in quello successivo.

In tutte le notazioni che compaiono nel prosieguo l'indice denota l'intervallo di tempo, e ad esso è sempre associato un indice  $n$  che denota che il gruppo considerato è quello dei soggetti che hanno manifestato  $n$  recidive. Infatti un soggetto appartiene, appena entrato nello studio, al

gruppo di coloro che contano  $n = 0$  recidive; ogni qualvolta si verifica un evento nell'intervallo  $i$ -mo, egli passa al gruppo con  $n = 1, 2, \dots, n$  recidive.

Per ciascuno dei  $n + 1$  gruppi così definiti si costruisce una tabella i cui ingredienti sono:

$(i, i + 1], i = 0, 1, 2, \dots, I$	sono gli intervalli di tempo prefissati. Essi possono essere scelti della durata di maggior interesse clinico;
$n = 0, 1, \dots, N$	numero degli eventi ricorrenti;
$l_{i,n}$	numero di soggetti presenti all'inizio dell'intervallo $(i, i + 1]$ nel gruppo con $n$ recidive;
$b_{in}$	numero di persone che entrano durante l'intervallo $(i, i + 1]$ nel gruppo con $n$ recidive;
$b^*_{in}$	numero di persone che entrano nel gruppo con $n$ recidive nell'intervallo $(i, i + 1]$ e presentano la $n + 1$ -ma nello stesso intervallo;
$w_{in}$	numero di persone che escono dal gruppo con $n$ recidive nell'intervallo $(i, i + 1]$ perché sono morte o perse di vista (a causa della fine della sperimentazione o altro);
$d_{in}$	numero di persone con $n$ recidive che presentano la $n + 1$ -ma nell'intervallo $(i, i + 1]$ ( $d_{in} = b_{in+1} + b^*_{in+1}$ );
$\vartheta_i^{b_{in}}$	la proporzione di giorni trascorsi dall'inizio dell'intervallo $(i, i + 1]$ all'entrata del soggetto $j$ -mo, contato nei $b_{in}$ , nel gruppo con $n$ recidive;
$\vartheta_i^{b^*_{in}}$	la proporzione di giorni trascorsi dall'inizio dell'intervallo $(i, i + 1]$ all'entrata del soggetto $j$ -mo, contato nei $b^*_{in}$ , nel gruppo con $n$ recidive;
$\vartheta_i^{w_{in}}$	la proporzione di giorni trascorsi dall'inizio dell'intervallo $(i, i + 1]$ fino all'uscita del soggetto $j$ -mo, contato nei $w_{in}$ , dal gruppo con $n$ recidive;
$\vartheta_i^{d_{in}}$	la proporzione di giorni trascorsi dall'inizio dell'intervallo $(i, i + 1]$ al giorno in cui l'individuo $j$ -mo, appartenente al gruppo con $n$ recidive, presenta la $n + 1$ -ma.

Si nota che se  $L$  è il numero dei soggetti entrati in uno studio relativo a malattie non congenite, allora è  $l_{in} = L$  per  $i = 0$ ,  $n = 0$ , e:

$$l_{i+1,n} = l_{in} + b_{in} - w_{in} - d_{in} \quad (1)$$

Nel caso in cui lo studio sia relativo a malattie congenite e quindi gli intervalli di tempo rappresentino classi di età è

$$l_{in} = 0 \quad \text{per } i = 0 \quad \text{ed } n = 0$$

mentre per passare da  $l_i$  a  $l_{i+1}$  vale ancora la (1).

Sulla base di queste quantità si possono calcolare per il gruppo di soggetti con  $n$  recidive:

— il numero effettivo di persone-unità di tempo che sviluppano la  $n + 1$ -ma recidiva nell'intervallo  $(i, i + 1]$

$$D_{in} = \sum_{j=1}^{d_{in}} \vartheta_j^{d_{in}} - \sum_{j=1}^{b_{in}^*} \vartheta_j^{b_{in}^*}$$

— il numero effettivo di persone-unità di tempo che sono esposte al rischio di sviluppare la  $n + 1$ -ma recidiva nell'intervallo  $(i, i + 1]$

$$E_{in} = l_{in} + \sum_{j=1}^{b_{in}} (1 - \vartheta_j^{b_{in}}) - \sum_{j=1}^{w_{in}} (1 - \vartheta_j^{w_{in}}) - \sum_{j=1}^{d_{in}} (1 - \vartheta_j^{d_{in}}).$$

La probabilità condizionale di sviluppare nell'intervallo  $(i, i + 1]$  la  $n + 1$ -ma recidiva, dato che un soggetto ne ha già avute  $n$ , è

$$q_{in} = \frac{D_{in}}{E_{in}}$$

e la relativa funzione di probabilità cumulativa per il gruppo con  $n$  recidive è

$$Q_{in} = \prod_{k=0}^{i-1} q_{kn} \quad \text{per } i > 0$$

con

$$Q_{0n} = 1.$$

La probabilità condizionale di non sviluppare la  $n + 1$ -ma recidiva nell'intervallo  $(i, i + 1]$  è

$$P_{in} = 1 - q_{in}$$

e la relativa funzione di probabilità cumulativa è

$$P_{in} = \prod_{k=0}^{i-1} p_{kn} \quad \text{per } i > 0$$

con

$$P_{0n} = 1.$$

Il rischio di sviluppare l'evento  $n + 1$ -mo nel punto medio  $m_i$  dell'intervallo  $(i, i + 1]$  è

$$\lambda(t_{m_i, n}) = \frac{2 q_{in}}{b_i (1 + p_{in})}$$

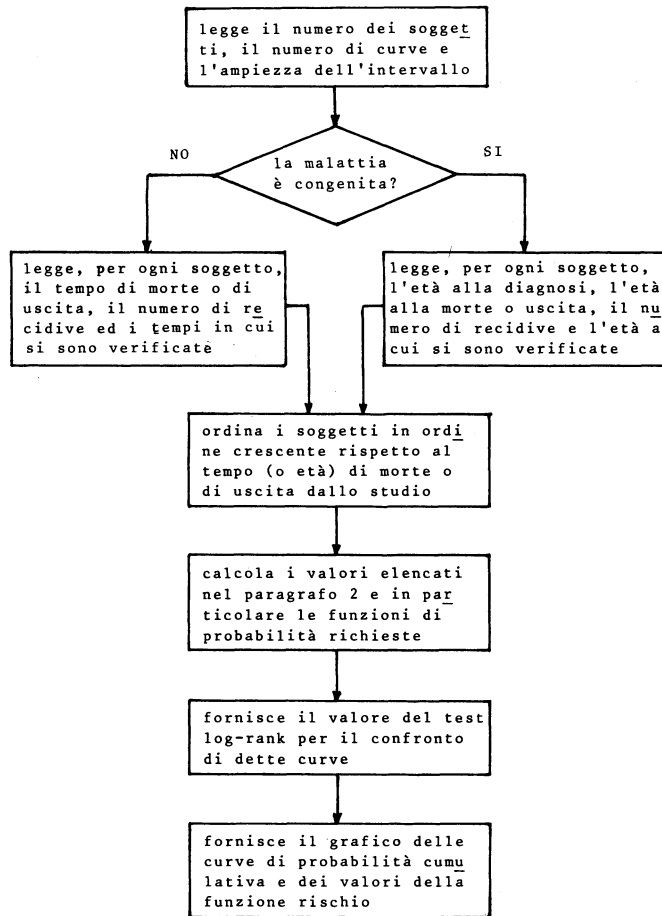


Fig. 1 - Flow-chart del programma.

dove  $b_i$  è l'ampiezza dell'intervallo.

La mediana della curva relativa al gruppo con  $n$ -recidive, se esiste, è

$$Q_{50} = (t_i - t_0) + \frac{p_{in} - 1/2}{p_{in} - p_{i+1n}}$$

dove

$$p_{in} \geq \frac{1}{2} \text{ e } p_{i+1n} < \frac{1}{2} .$$

Il confronto tra le curve di probabilità cumulativa  $P_{in}$ , con  $n$  che

varia da 0 fino al numero fissato dall'utente (che è comunque inferiore o tutt'al più uguale a  $N - 1$ ) è eseguito per mezzo del log-rank test (Mantel, 1966).

Si fa notare che la funzione  $P_{iN}$  è una funzione di valore costante uguale a 1 poiché non si osserva nessun altro evento sui soggetti che, durante lo studio, hanno avuto il numero massimo  $N$  di eventi ricorrenti. Per la stessa ragione la funzione  $Q_{iN}$  ha valore costante uguale a 0.

### 3. Descrizione del programma

#### 3.1. Descrizione generale

Un flow-chart che descrive in modo sintetico la struttura del programma è riportato in Fig. 1.

#### 3.2. Input

Il programma richiede in entrata le seguenti quantità:

- 1 - numero dei soggetti;
- 2 - numero di curve che l'utente desidera considerare; esso può essere minore o tutt'al più uguale al numero massimo di eventi ricorrenti;
- 3 - lunghezza dell'intervallo di tempo fissato;
- 4 - un indice di valore 0 se tali intervalli sono classi d'età (per malattie congenite), di valore 1 se sono intervalli di tempo che hanno origine all'inizio del follow-up;
- 5 - se l'indice di cui al punto 4 è zero, per ogni soggetto: età alla diagnosi, età alla morte o uscita dallo studio, numero di recidive ed età a cui esse si sono verificate;  
se l'indice di cui al punto 4 è uguale a 1, per ogni soggetto: il tempo di morte o di uscita dallo studio, il numero di recidive ed i loro tempi di occorrenza.

#### 3.3. Output

Il programma fornisce, per ogni gruppo di pazienti, a partire da quello con 0 recidive:

- 1 - una tabella in cui, per ogni intervallo di tempo, si dà il numero dei pazienti « a rischio », il numero dei pazienti che presentano l'evento ricorrente (e quindi passano ad un altro gruppo), di quelli che sono morti o persi di vista, di quelli che provengono da un altro gruppo;
- 2 - a discrezione dell'utente le quantità del punto 1 espresse in persone-unità di tempo calcolate come indicato nel paragrafo 2;
- 3 - la funzione cumulativa  $P_{in}$  e la sua mediana (se esistente);
- 4 - i valori della funzione rischio nei punti medi degli intervalli.

Tab. I - Per ognuno dei 41 pazienti: numero di eventi emorragici (N. Emor.), tempo di sopravvivenza o di uscita dallo studio (T. Sopr.), tempo di comparsa del primo episodio emorragico (I° Emor.), tempo di comparsa dell'eventuale secondo episodio emorragico (II° Emor.).

N. Emor.	T. Sopr.	I°Emor.	II°Emor.
2	55	30	45
1	167	27	0
1	198	180	0
1	280	210	0
1	297	272	0
2	315	135	240
1	315	150	0
1	315	270	0
1	417	372	0
1	440	420	0
1	497	455	0
1	551	540	0
1	577	510	0
1	585	401	0
2	585	390	450
2	605	180	330
2	605	191	450
1	675	360	0
2	700	660	676
1	700	664	0
1	700	660	0
1	715	570	0
1	781	750	0
1	799	786	0
1	852	840	0
1	859	450	0
2	1058	720	780
1	1066	1050	0
1	1118	793	0
2	1118	780	1050
2	1118	785	990
2	1118	780	960
1	1119	270	0
2	1119	281	990
2	1119	275	840
2	1119	270	720
1	1138	1110	0
2	1147	1110	1123
2	1215	600	840
2	1228	990	1200
1	1324	1290	0

- Fornisce inoltre, a discrezione dell'utente, il grafico delle stime:
- delle funzioni cumulative  $P_{in}$ ,  $n = 0, \dots, N-1$  (probabilità di non sviluppare la  $n+1$ -ma recidiva);
  - delle funzioni cumulative  $Q_{in}$ ,  $n = 0, \dots, N-1$  (probabilità di sviluppare la  $n+1$ -ma recidiva);
  - dei valori  $(t_{m,n})$  assunti dalla funzione rischio nel punto centrale dell'intervallo.

È quindi riportato il risultato del log-rank test per il confronto delle curve preceduto dalla rispettiva tabella con gli eventi osservati e attesi per gruppo.

Tab. II - Gruppo di pazienti che non hanno avuto emorragie.

Curva numero J= 1 paz. con J-1 eventi a priori

N. Int.	N. Rischio	N. Eventi	N. Persi	N. Entrati	PROB.	PROB. Cum	RISCHIO
1	41	2	0	0	.9535	.9535	.0061
2	39	0	0	0	1.0000	.9535	.0000
3	39	0	0	0	1.0000	.9535	.0000
4	39	0	0	0	1.0000	.9535	.0000
5	39	2	0	0	.9610	.9164	.0051
6	37	2	0	0	.9459	.8669	.0070
7	35	2	0	0	.9602	.8324	.0052
8	33	0	0	0	1.0000	.8324	.0000
9	33	3	0	0	.9091	.7567	.0116
10	30	3	0	0	.9783	.7403	.0029
11	27	0	0	0	1.0000	.7403	.0000
12	27	1	0	0	.9630	.7128	.0048
13	26	2	0	0	.9449	.6736	.0071
14	24	2	0	0	.9415	.6342	.0076
15	22	1	0	0	.9545	.6053	.0059
16	21	1	0	0	.9917	.6003	.0011
17	20	1	0	0	.5500	.5703	.0065
18	19	1	0	0	.9474	.5403	.0068
19	18	1	0	0	.9444	.5103	.0077
20	17	1	0	0	.9412	.4803	.0076
21	16	0	0	0	1.0000	.4803	.0000
22	16	2	0	0	.8750	.4702	.0156
23	14	1	0	0	.9898	.4160	.0013
24	13	1	0	0	.9231	.3840	.0099
25	12	1	0	0	.9167	.3520	.0106
26	11	2	0	0	.8182	.2880	.0220
27	9	3	0	0	.8824	.2541	.0140
28	6	1	0	0	.8333	.2117	.0204
29	5	0	0	0	1.0000	.2117	.0000
30	5	0	0	0	1.0000	.2117	.0000
31	5	0	0	0	1.0000	.2117	.0000
32	5	0	0	0	1.0000	.2117	.0000
33	5	1	0	0	.8060	.1694	.0240
34	4	0	0	0	1.0000	.1694	.0000
35	4	1	0	0	.7590	.1270	.0292
36	3	0	0	0	1.0000	.1270	.0000
37	3	2	0	0	.3333	.0423	.0593
38	1	0	0	0	1.0000	.0423	.0000
39	1	0	0	0	1.0000	.0423	.0000
40	1	0	0	0	1.0000	.0423	.0000
41	1	0	0	0	1.0000	.0423	.0000
42	1	0	0	0	1.0000	.0423	.0000
43	1	1	0	0	.0000	.0000	.0667
44	0	0	0	0	1.0000	.0000	.0000
45	0	0	0	0	1.0000	.0000	.0000

Mediana Curva J-MA 580. 2779

Per ogni mese sono riportati il numero dei pazienti a rischio (N. Rischio), il numero di quelli che presentano il primo episodio emorragico (N. Eventi), il numero di quelli che sono morti o persi di vista (N. Persi), e di quelli che provengono dal gruppo « precedente » (N. Entrati); in questo caso sempre chiaramente nullo, la probabilità di non sviluppare la prima emorragia (Prob.), la rispettiva funzione di probabilità cumulativa (Prob. Cum) ed il valore della funzione rischio nei punti centrali (Rischio).

#### 4. Esempio

L'esempio verte su un sottoinsieme di 41 soggetti estratto da una casistica di 125 pazienti affetti da cirrosi epatica e che hanno presentato emorragie da varici esofagee, raccolta presso l'Istituto di Medicina Generale e Pneumologia di Palermo.

La variabile di risposta studiata è il tempo di occorrenza del primo o dei primi due episodi di emorragia esofagea.

Su ognuno dei pazienti si è rilevato il tempo di sopravvivenza dalla



Tab. III - Gruppo di pazienti che hanno avuto 1 emorragia.

Curva numero J= 2 paz. con J-1 eventi a priori							
N. Int.	N. Rischio	N. Eventi	N. Persi	N. Entrati	PROB.	PROB. Cum	RISCHIO
1	0	0	0	2	1.0000	1.0000	.0000
2	2	1	0	0	.6667	.6667	.0370
3	1	0	0	0	1.0000	.6667	.0000
4	1	0	0	0	1.0000	.6667	.0000
5	1	0	0	2	1.0000	.6667	.0000
6	3	0	1	2	1.0000	.6667	.0000
7	4	0	1	2	1.0000	.6667	.0000
8	5	1	0	0	.8000	.5333	.0240
9	4	0	0	3	1.0000	.5333	.0000
10	7	0	2	3	1.0000	.5333	.0000
11	8	1	2	0	.8571	.4571	.0177
12	5	0	0	1	1.0000	.4571	.0000
13	6	0	0	2	1.0000	.4571	.0000
14	8	0	1	2	1.0000	.4571	.0000
15	4	2	1	1	.7692	.3516	.0272
16	7	0	0	1	1.0000	.3516	.0000
17	8	0	1	1	1.0000	.3516	.0000
18	8	0	0	1	1.0000	.3516	.0000
19	9	0	1	1	1.0000	.3516	.0000
20	9	0	2	1	1.0000	.3516	.0000
21	8	0	0	0	1.0000	.3516	.0000
22	8	0	0	2	1.0000	.3516	.0000
23	10	1	1	1	.9461	.3327	.0070
24	9	1	3	1	.8667	.2883	.0166
25	6	0	0	1	1.0000	.2883	.0000
26	7	1	0	2	.8571	.2472	.0177
27	8	0	2	3	1.0000	.2472	.0000
28	9	2	0	1	.7778	.1922	.0263
29	8	0	2	0	1.0000	.1922	.0000
30	6	0	0	0	1.0000	.1922	.0000
31	6	0	0	0	1.0000	.1922	.0000
32	6	1	0	0	.8333	.1602	.0204
33	5	2	0	1	.6000	.0961	.0427
34	4	0	0	0	1.0000	.0961	.0000
35	4	1	0	1	.7500	.0721	.0292
36	4	0	1	0	1.0000	.0721	.0000
37	3	0	0	2	1.0000	.0721	.0000
38	4	1	0	0	.8523	.0614	.0182
39	1	0	0	0	1.0000	.0614	.0000
40	1	1	0	0	.0000	.0000	.0667
41	0	0	0	0	1.0000	.0000	.0000
42	0	0	0	0	1.0000	.0000	.0000
43	0	0	0	1	1.0000	.0000	.0000
44	1	0	0	0	1.0000	.0000	.0000
45	1	0	1	0	1.0000	.0000	.0000

Mediana Curva J-MA 313.1350

Per ogni mese sono riportati il numero dei pazienti a rischio (N. Rischio), il numero di quelli che presentano la seconda emorragia (N. Eventi), il numero di quelli che sono morti o persi di vista (N. Persi), e di quelli che provengono dal gruppo « precedente » (N. Entrati), la probabilità di non sviluppare la seconda emorragia (Prob.), la rispettiva funzione di probabilità cumulativa (Prob. Cum.), ed il valore della funzione rischio nei punti centrali (Rischio).

diagnosi o, per i persi di vista, il tempo in cui sono stati in osservazione durante il follow-up ed il tempo in cui si sono manifestati i primi due episodi emorragici.

In Tab. I sono riportati il numero delle emorragie ed, in giorni dalla diagnosi, i tempi osservati.

Nell'esempio si sono fissati intervalli di tempo dell'ampiezza di 30 giorni e si sono richieste le tabelle e le funzioni di probabilità inerenti il gruppo di pazienti con nessuno ed uno episodio emorragico rispettiva-

Tab. IV - Valore del log-rank test per il confronto delle curve di probabilità di non comparsa dell'evento emorragico.

---

CONFRONTO DI CURVE DI SOPRAVVIVENZA

Eventi osservati per gruppo	31. 23	14. 47
Eventi attesi per gruppo	31. 32	14. 38
Osservati/attesi per gruppo	1. 00	1. 01

MATRICE DI DISPERSIONE

7. 7579564  
 —7. 7579563      7. 7579563

INVERSA DELLA MATRICE OTTENUTA TOGLIENDO L'ULTIMA RIGA E L'ULTIMA COLONNA DELLA MATRICE DI DISPERSIONE  
 .12889992

LOG-RANK TEST       $X^2 =$       .0011      G.D.L. = 1

---

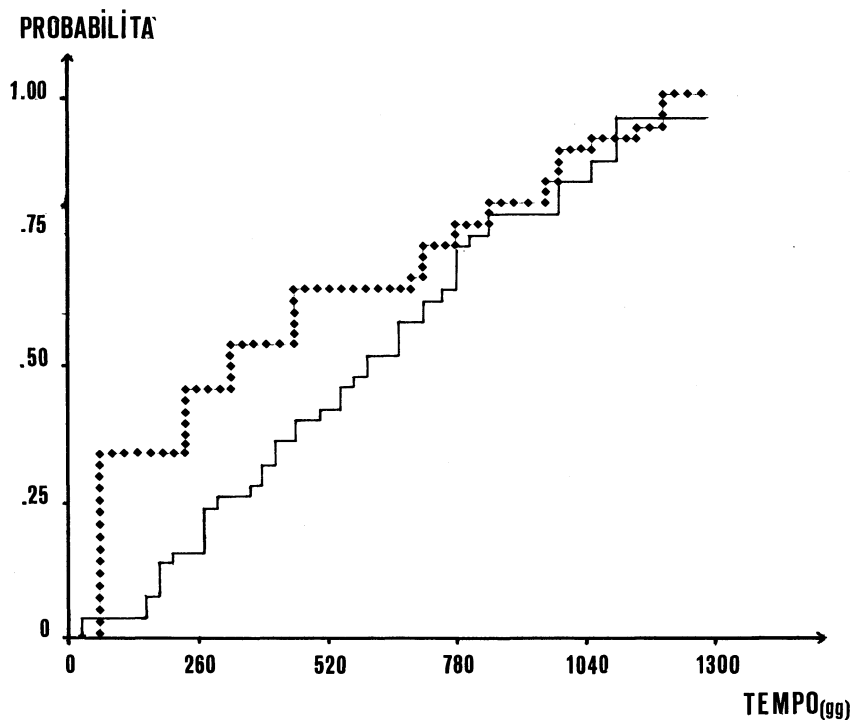


Fig. 2 - Probabilità di comparsa dell'evento emorragico in pazienti con nessuna emorragia (curva —) ed in pazienti con una emorragia (curva ●).

mente. Le tabelle ottenute sono riportate nelle Tab. II e III ed il grafico delle funzioni di probabilità è disegnato nella Fig. 2.

Il valore del log-rank test, riportato in Tab. IV, mostra che le funzioni di probabilità di non occorrenza dell'evento emorragico non sono in questo caso significativamente differenti.

## 5. Caratteristiche software

Il programma scritto in linguaggio FORTRAN ASCII è stato implementato su un calcolatore UNIVAC 1100/80. Il numero degli intervalli di tempo è limitato ad un massimo di 60 ed il numero di possibili passaggi dei soggetti da un gruppo ad un successivo e quindi di possibili curve è al massimo 5. Entrambi questi limiti possono essere facilmente modificati.

Il numero dei soggetti non ha limitazioni.

Con i suddetti limiti il programma occupa 2536 I BANK - 12671 D BANK.

## Riferimenti bibliografici

- Chan L.S., Powars D., Lee J. and Weiss J., A modified lifetable method to study congenital genetic disorders: an application in sickle cell anemia. *J. Chron. Dis.*, **35**, 1982, pp. 401-409.
- Cox D.R., Regression models and life tables (with discussion). *J. Royal Statist. Soc. (B)*, **34**, 1972, pp. 187-219.
- Mantel N., Evaluation of survival data and two new rank order statistics arising in the consideration. *Cancer Chemother. Rep.*, **50**, 1966, pp. 163-170.

## Summary

This paper describes a program that can be helpful in studying clinical situations which present recurrent complications or events, allowing the user to determine the conditional probabilities of developing such complications.

The main feature of the adopted procedure is that it takes account not only of the number of recurrent events but also of their time of occurrence.