

**METODI NON PARAMETRICI PER
LA STIMA E IL CONFRONTO DELLA
SOPRAVVIVENZA TRA GRUPPI**

ESEMPIO 1:

I dati sono riportati nel libro “E. Marubini, M.G. Valsecchi. Analysing survival data from clinical trials and observational studies. pag. 41” e si riferiscono ad una sperimentazione clinica multicentrica sul mantenimento della remissione in bambini con leucemia linfoblastica acuta (Freireich, 1963).

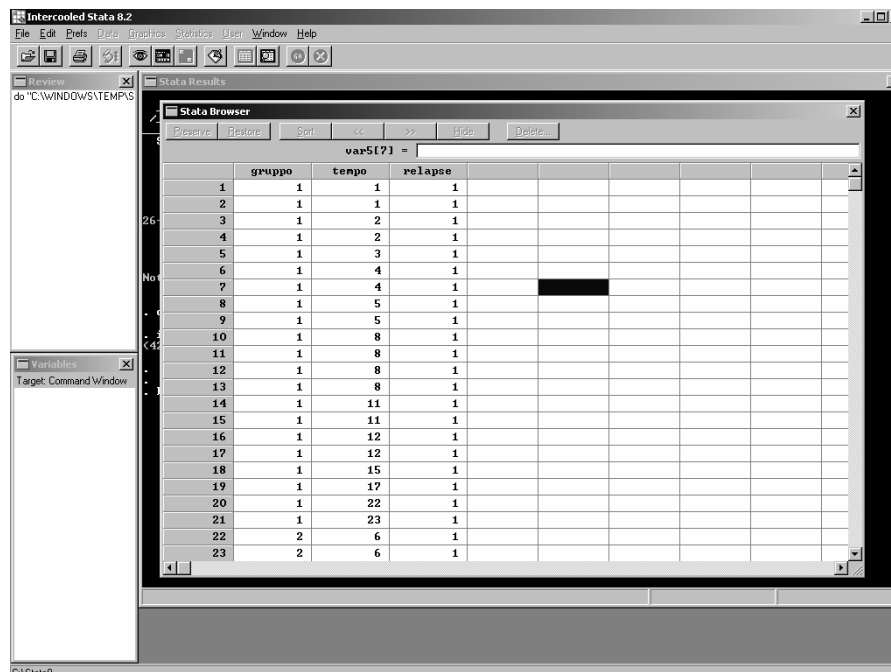
I dati sono forniti su floppy disk (nome file: **freireich.dat**).

Il contenuto del file è il seguente:

variabili: gruppo di terapia (variabile ‘**gruppo**’: **1 = placebo** **2 = 6-MP**)
 tempo di sopravvivenza in settimane (variabile ‘**tempo**’)
 status (variabile ‘**relapse**’: **0 = censored** **1 = relapse**)
numero casi: **42**

Caricare in memoria un file di dati nel formato ASCII (.dat):

infile gruppo tempo relapse using a:\freireich.dat



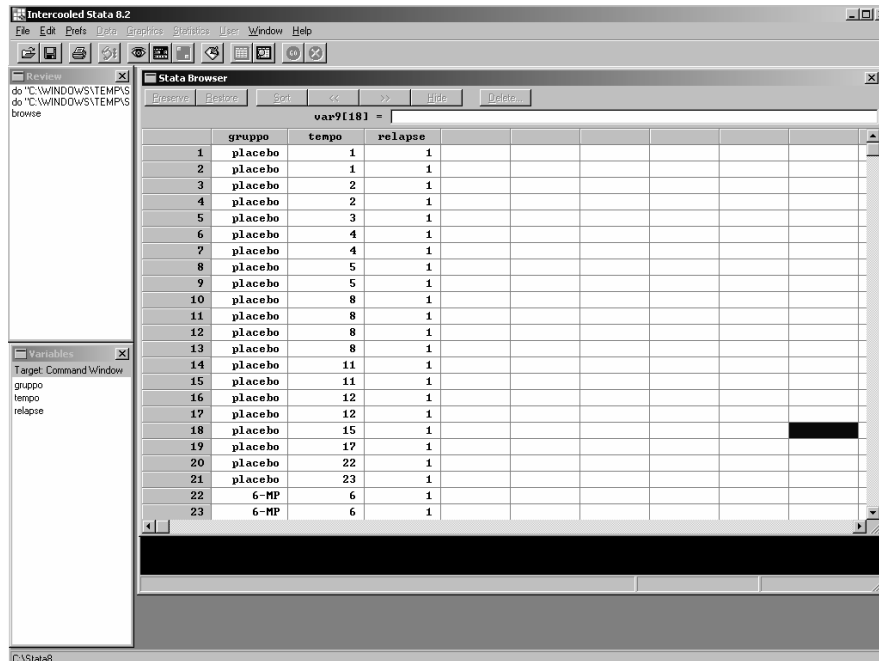
The screenshot shows the Intercooled Stata 8.2 interface. The main window displays a data browser for the file 'a:\freireich.dat'. The table has four columns: 'gruppo', 'tempo', and 'relapse'. The data is as follows:

	gruppo	tempo	relapse
1	1	1	1
2	1	1	1
3	1	2	1
4	1	2	1
5	1	3	1
6	1	4	1
7	1	4	1
8	1	5	1
9	1	5	1
10	1	8	1
11	1	8	1
12	1	8	1
13	1	8	1
14	1	11	1
15	1	11	1
16	1	12	1
17	1	12	1
18	1	15	1
19	1	17	1
20	1	22	1
21	1	23	1
22	2	6	1
23	2	6	1

Per migliorare la presentazione dei risultati è consigliabile esplicitare il significato dei valori della variabile gruppo di terapia tramite delle label.

Assegnare label ai valori della variabile gruppo di terapia:

```
label define tr_code 1 "placebo" 2 "6-MP"  
label values gruppo tr_code
```



The screenshot shows the Stata 8.2 interface. The main window displays a data browser for 'var9[18]'. The table contains the following data:

	gruppo	tempo	relapse
1	placebo	1	1
2	placebo	1	1
3	placebo	2	1
4	placebo	2	1
5	placebo	3	1
6	placebo	4	1
7	placebo	4	1
8	placebo	5	1
9	placebo	5	1
10	placebo	8	1
11	placebo	8	1
12	placebo	8	1
13	placebo	8	1
14	placebo	11	1
15	placebo	11	1
16	placebo	12	1
17	placebo	12	1
18	placebo	15	1
19	placebo	17	1
20	placebo	22	1
21	placebo	23	1
22	6-MP	6	1
23	6-MP	6	1

E' necessario 'informare' il software che i dati in memoria sono dati di sopravvivenza, specificando quali sono le variabili chiave (**tempo di sopravvivenza** e **variabile indicatrice dell'evento**) ed il codice dell'evento.

Definire i dati di sopravvivenza in STATA:

```
stset tempo, failure(relapse==1)
```

```
failure event: relapse == 1  
obs. time interval: (0, tempo]  
exit on or before: failure
```

```
42 total obs.  
0 exclusions
```

```
42 obs. remaining, representing  
30 failures in single record/single failure data  
541 total analysis time at risk, at risk from t = 0  
earliest observed entry t = 0  
last observed exit t = 35
```

Quindi si procede con la **stima della funzione di sopravvivenza** separatamente nei due gruppi di trattamento (6-MP e placebo).

Stimare la funzione di sopravvivenza tramite la procedura di Kaplan-Meier:

sts list, by(gruppo)

Intercooled Stata 8.2

File Edit Prefs Data Graphics Statistics User Window Help

Review

```
do "C:\WINDOWS\TEMP\S
do "C:\WINDOWS\TEMP\S
do "C:\WINDOWS\TEMP\S
do "C:\WINDOWS\TEMP\S
```

Stata Results

```
. sts list, by(gruppo)

      failure _d: relapse == 1
      analysis time _t: tempo
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
placebo							
1	21	2	0	0.9048	0.0641	0.6700	0.9753
2	19	2	0	0.8095	0.0857	0.5689	0.9239
3	17	1	0	0.7619	0.0929	0.5194	0.8933
4	16	2	0	0.6667	0.1029	0.4254	0.8250
5	14	2	0	0.5714	0.1080	0.3380	0.7492
8	12	4	0	0.3810	0.1060	0.1831	0.5778
11	8	2	0	0.2857	0.0986	0.1166	0.4818
12	6	2	0	0.1905	0.0857	0.0595	0.3774
15	4	1	0	0.1429	0.0764	0.0357	0.3212
17	3	1	0	0.0952	0.0641	0.0163	0.2612
22	2	1	0	0.0476	0.0465	0.0033	0.1970
23	1	1	0	0.0000	-	-	-
6-MP							
6	21	3	1	0.8571	0.0764	0.6197	0.9516
7	17	1	0	0.8067	0.0869	0.5631	0.9228
9	16	0	1	0.8067	0.0869	0.5631	0.9228
10	15	1	1	0.7529	0.0963	0.5032	0.8894
11	13	0	1	0.7529	0.0963	0.5032	0.8894
13	12	1	0	0.6902	0.1068	0.4316	0.8491
16	11	1	0	0.6275	0.1141	0.3675	0.8049
17	10	0	1	0.6275	0.1141	0.3675	0.8049
19	9	0	1	0.6275	0.1141	0.3675	0.8049
20	8	0	1	0.6275	0.1141	0.3675	0.8049
22	7	1	0	0.5378	0.1282	0.2678	0.7468
23	6	1	0	0.4482	0.1346	0.1881	0.6801
25	5	0	1	0.4482	0.1346	0.1881	0.6801
32	4	0	2	0.4482	0.1346	0.1881	0.6801
34	2	0	1	0.4482	0.1346	0.1881	0.6801
35	1	0	1	0.4482	0.1346	0.1881	0.6801

end of do-file

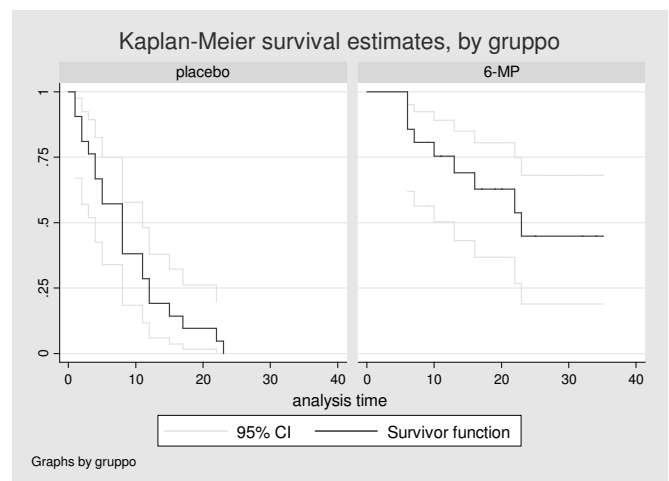
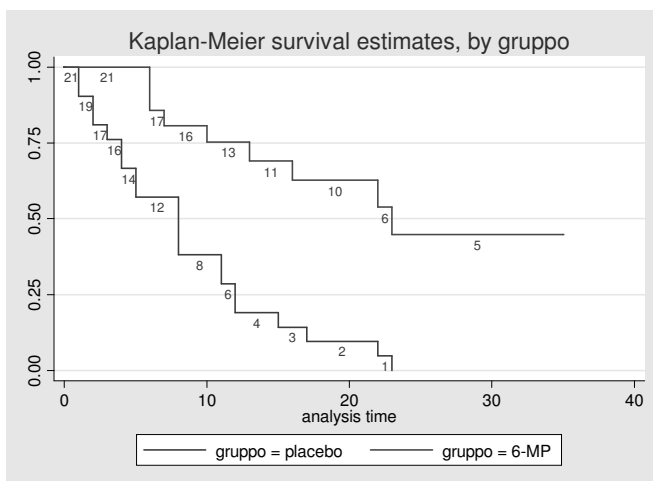
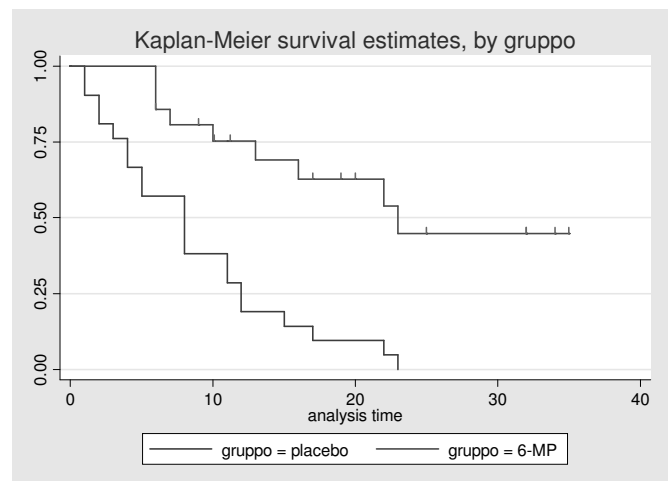
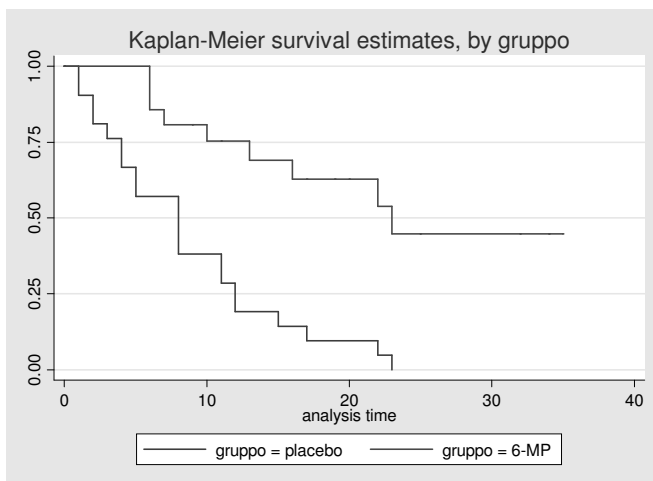
Stata Command

C:\Stata8

E' possibile **representare graficamente la stima della funzione di sopravvivenza** separatamente nei due gruppi di trattamento (6-MP e placebo), eventualmente riportando sulle curve i tempi di sopravvivenza troncati, il numero di soggetti a rischio oppure l'intervallo di confidenza (formula di Greenwood) per ciascuna stima.

Tracciare il grafico della funzione di sopravvivenza stimata tramite la procedura di Kaplan-Meier:

```
sts graph, by(gruppo)
sts graph, by(gruppo) censored(multiple)
sts graph, by(gruppo) atrisk
sts graph, by(gruppo) gwood
```



Non è conveniente “leggere” gli intervalli di confidenza direttamente dal grafico. E’ invece possibile **calcolare direttamente l’intervallo di confidenza** per la probabilità di sopravvivenza ad un determinato tempo ed eventualmente modificare il livello di confidenza.

Calcolare l’intervallo di confidenza per la probabilità di sopravvivenza:

sts list, by(gruppo) at(4 8 12)

```
failure _d: relapse == 1
analysis time _t: tempo
```

Time	Beg. Total	Fail	Survivor Function	Std. Error	[95% Conf. Int.]	

placebo						
4	16	7	0.6667	0.1029	0.4254	0.8250
8	12	6	0.3810	0.1060	0.1831	0.5778
12	6	4	0.1905	0.0857	0.0595	0.3774
6-MP						
4	0	0	1.0000	.	.	.
8	17	4	0.8067	0.0869	0.5631	0.9228
12	13	1	0.7529	0.0963	0.5032	0.8894

Note: Survivor function is calculated over full data and evaluated at indicated times; it is not calculated from aggregates shown at left.

sts list, by(gruppo) at(4 8 12) level(90)

```
failure _d: relapse == 1
analysis time _t: tempo
```

Time	Beg. Total	Fail	Survivor Function	Std. Error	[90% Conf. Int.]	

placebo						
4	16	7	0.6667	0.1029	0.4685	0.8051
8	12	6	0.3810	0.1060	0.2121	0.5484
12	6	4	0.1905	0.0857	0.0750	0.3460
6-MP						
4	0	0	1.0000	.	.	.
8	17	4	0.8067	0.0869	0.6125	0.9102
12	13	1	0.7529	0.0963	0.5511	0.8736

Note: Survivor function is calculated over full data and evaluated at indicated times; it is not calculated from aggregates shown at left.

Il software permette di **salvare alcune statistiche** (stima della probabilità di sopravvivenza, errore standard, limiti di confidenza, ...) per effettuare altri calcoli o per produrre grafici supplementari.

Salvare statistiche prodotte tramite la procedura di Kaplan-Meier:

```
sts generate surv=s, by(gruppo)
sts generate lowb=lb(s), by(gruppo)
sts generate uppb=ub(s), by(gruppo)
```

E' possibile calcolare alcune **statistiche di sintesi della distribuzione del tempo di sopravvivenza**, quali il tempo di sopravvivenza mediano e medio.

Calcolare il tempo di sopravvivenza mediano:

```
stci, by(gruppo) median
```

```
failure _d: relapse == 1
analysis time _t: tempo
```

gruppo	no. of subjects	50%	Std. Err.	[95% Conf. Interval]	
placebo	21	8	1.20344	4	11
6-MP	21	23	1.187919	13	.
total	42	12	1.716577	8	17

Calcolare il tempo di sopravvivenza medio:

stci, by(gruppo) rmean

```
failure _d: relapse == 1
analysis time _t: tempo
```

gruppo	no. of subjects	restricted mean	Std. Err.	[95% Conf. Interval]
placebo	21	8.666667	1.37739	5.96703 11.3663
6-MP	21	23.2874 (*)	2.827468	17.7457 28.8291
total	42	15.33933 (*)	1.860218	11.6934 18.9853

(*) largest observed analysis time is censored, mean is underestimated.

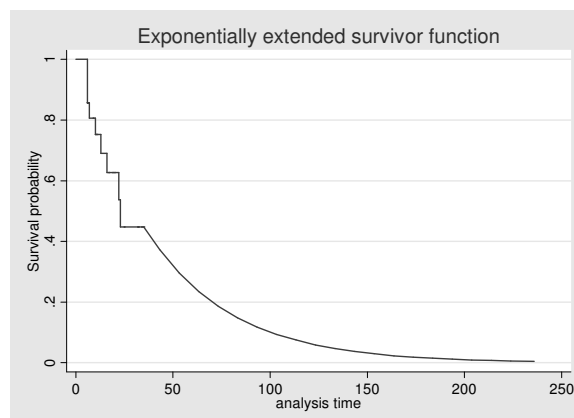
stci, by(gruppo) emean

```
failure _d: relapse == 1
analysis time _t: tempo
```

gruppo	no. of subjects	extended mean
placebo	21	8.666667 (*)
6-MP	21	42.83265
total	42	19.32587

(*) no extension needed

stci if gruppo==2, emean graph



Infine, è possibile scegliere tra diversi **test per il confronto della sopravvivenza tra gruppi** (log-rank, Breslow, Tarone-Ware, ...).

Confrontare la sopravvivenza tra gruppi diversi (I):

sts test gruppo

```
failure _d: relapse == 1
analysis time _t: tempo
```

Log-rank test for equality of survivor functions

gruppo	Events observed	Events expected
placebo	21	10.75
6-MP	9	19.25
Total	30	30.00

```
chi2(1) = 16.79
Pr>chi2 = 0.0000
```

sts test gruppo, wilcoxon

```
failure _d: relapse == 1
analysis time _t: tempo
```

Wilcoxon (Breslow) test for equality of survivor functions

gruppo	Events observed	Events expected	Sum of ranks
placebo	21	10.75	271
6-MP	9	19.25	-271
Total	30	30.00	0

```
chi2(1) = 13.46
Pr>chi2 = 0.0002
```

Confrontare la sopravvivenza tra gruppi diversi (II):

sts test gruppo, tware

```
failure _d: relapse == 1  
analysis time _t: tempo
```

Tarone-Ware test for equality of survivor functions

gruppo	Events observed	Events expected	Sum of ranks
placebo	21	10.75	51.162748
6-MP	9	19.25	-51.162748
Total	30	30.00	0

chi2(1) = 15.12
Pr>chi2 = 0.0001