The Kaplan-Meier Survival Test

Survival analysis is concerned with studying the occurrence of an event such as death or change in a subject or object at various times following the beginning of the study. Survival curves show the percentage of subjects surviving at various times as the study progresses. In many cases, it is desired to compare survival of an experimental treatment with a control treatment. This method is heavily used in medical research but is not restricted to that field. For example, one might compare the rate of college failure among students in an experimental versus a control group.

To obtain a survival curve you need only two items of information in your data file for each subject: the survival time and a code for whether or not an event occurred or the subject has been lost from the study moved, disappeared, etc. (censored.) If an event such as death occurred, it is coded as a 1. If censored it is coded as a 2.

```
CASES FOR FILE C:\OpenStat\KaplanMeier1.LAZ
```

0	Time	Event_Censored
1	1	2
2	3	2
3	5	2
4	6	1
5	6	1
6	6	1
7	6	1
8	6	1
9	6	1
10	8	1
11	8	1
12	9	2
13	10	1
14	10	1
15	10	2
16	12	1
17	12	1
18	12	1
19	12	1
20	12	1
21	12	1
22	12	2
23	12	2
24	13	2
25	15	2
26	15	2
27	16	2
28	16	2
29	18	2
30	18	2
31	20	1
32	20	2
33	22	2
34	24	1
35	24	1
36	24	2
37	27	2
38	28	2

39	28	2
40	28	2
41	30	1
42	30	2
43	32	1
44	33	2
45	34	2
46	36	2
47	36	2
48	42	1
49	44	2

We are really recording data for the "Time" variable that is sequential through the data file. We are concerned with the percent of survivors at any given time period as we progress through the observation times of the study. We record the "drop-outs" or censored subjects at each time period also. A unit cannot be censored and be one of the deaths - these are mutually exclusive.

Next we show a data file that contains both experimental and control subjects:

```
CASES FOR FILE C:\OpenStat\KaplanMeier2.TEX
```

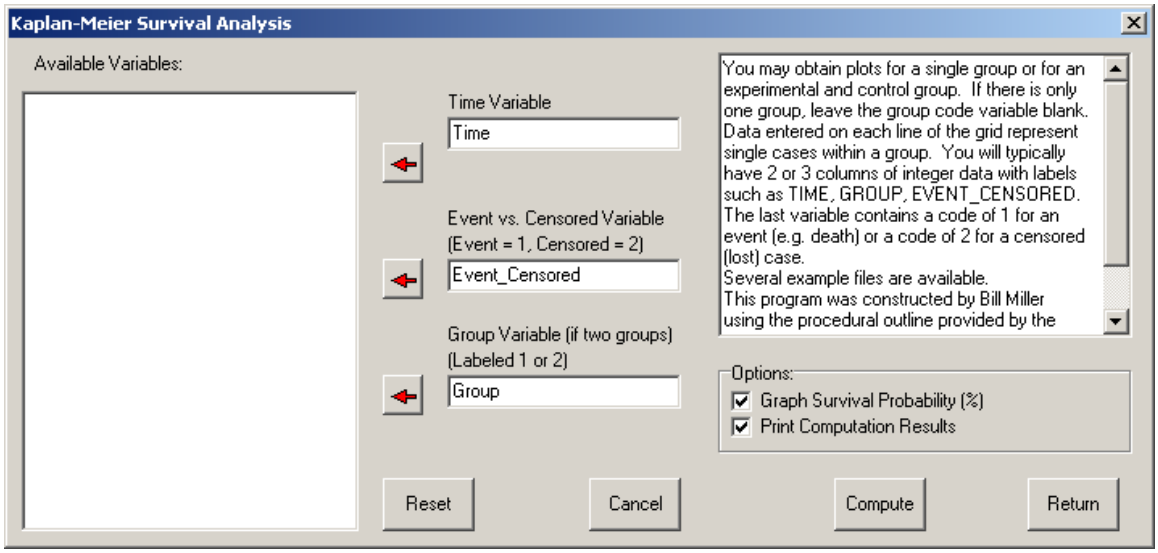
0	Time	Group	Event_Censored
1	1	1	2
2	3	2	2
3	5	1	2
4	6	1	1
5	6	1	1
6	6	2	1
7	6	2	1
8	6	2	1
9	6	2	1
10	8	2	1
11	8	2	1
12	9	1	2
13	10	1	1
14	10	1	1
15	10	1	2
16	12	1	1
17	12	1	1
18	12	1	1
19	12	1	1
20	12	2	1
21	12	2	1
22	12	1	2
23	12	2	2
24	13	1	2
25	15	1	2
26	15	2	2
27	16	1	2
28	16	2	2
29	18	2	2
30	18	2	2
31	20	2	1
32	20	1	2
33	22	2	2
34	24	1	1

35	24	2	1
36	24	1	2
37	27	1	2
38	28	2	2
39	28	2	2
40	28	2	2
41	30	2	1
42	30	2	2
43	32	1	1
44	33	2	2
45	34	1	2
46	36	1	2
47	36	1	2
48	42	2	1
49	44	1	2

In this data we code the groups as 1 or 2. Censored cases are always coded 2 and Events are coded 1. This data is, in fact, the same data as shown in the previous data file. Note that in time period 6 there were 6 deaths (cases 4-9.) Again, notice that the time periods are in ascending order.

Shown below is the specification dialog for this second data file. This is followed by the output obtained when you click the compute button.

Figure 1. The Kaplan-Meier Dialog



Comparison of Two Groups Methd

TIME	GROUP	CENSORED	TOTAL AT RISK	EVENTS	AT RISK IN GROUP 1	EXPECTED NO. EVENTS IN 1	AT RISK IN GROUP 2	EXPECTED NO. EVENTS IN 2
0	0	0	49	0	25	0.0000	24	0.0000
1	1	1	49	0	25	0.0000	24	0.0000
3	2	1	48	0	24	0.0000	24	0.0000
5	1	1	47	0	24	0.0000	23	0.0000
6	1	0	46	6	23	3.0000	23	3.0000
6	1	0	40	0	21	0.0000	19	0.0000
6	2	0	40	0	21	0.0000	19	0.0000
6	2	0	40	0	21	0.0000	19	0.0000
6	2	0	40	0	21	0.0000	19	0.0000
6	2	0	40	0	21	0.0000	19	0.0000
8	2	0	40	2	21	1.0500	19	0.9500
8	2	0	38	0	21	0.0000	17	0.0000

9	1	1	38	0	21	0.0000	17	0.0000
10	1	0	37	2	20	1.0811	17	0.9189
10	1	0	35	0	18	0.0000	17	0.0000
10	1	1	35	0	18	0.0000	17	0.0000
12	1	0	34	6	17	3.0000	17	3.0000
12	1	0	28	0	13	0.0000	15	0.0000
12	1	0	28	0	13	0.0000	15	0.0000
12	1	0	28	0	13	0.0000	15	0.0000
12	2	0	28	0	13	0.0000	15	0.0000
12	2	0	28	0	13	0.0000	15	0.0000
12	1	1	28	0	13	0.0000	15	0.0000
12	2	1	27	0	12	0.0000	15	0.0000
13	1	1	26	0	12	0.0000	14	0.0000
15	1	1	25	0	11	0.0000	14	0.0000
15	2	1	24	0	10	0.0000	14	0.0000
16	1	1	23	0	10	0.0000	13	0.0000
16	2	1	22	0	9	0.0000	13	0.0000
18	2	1	21	0	9	0.0000	12	0.0000
18	2	1	20	0	9	0.0000	11	0.0000
20	2	0	19	1	9	0.4737	10	0.5263
20	1	1	18	0	9	0.0000	9	0.0000
22	2	1	17	0	8	0.0000	9	0.0000
24	1	0	16	2	8	1.0000	8	1.0000
24	2	0	14	0	7	0.0000	7	0.0000
24	1	1	14	0	7	0.0000	7	0.0000
27	1	1	13	0	6	0.0000	7	0.0000
28	2	1	12	0	5	0.0000	7	0.0000
28	2	1	11	0	5	0.0000	6	0.0000
28	2	1	10	0	5	0.0000	5	0.0000
30	2	0	9	1	5	0.5556	4	0.4444
30	2	1	8	0	5	0.0000	3	0.0000
32	1	0	7	1	5	0.7143	2	0.2857
33	2	1	6	0	4	0.0000	2	0.0000
34	1	1	5	0	4	0.0000	1	0.0000
36	1	1	4	0	3	0.0000	1	0.0000
36	1	1	3	0	2	0.0000	1	0.0000
42	2	0	2	1	1	0.5000	1	0.5000
44	1	1	0	0	1	0.0000	0	0.0000

TIME	DEATHS	GROUP	AT RISK	PROPORTION SURVIVING	CUMULATIVE PROP.SURVIVING
1	0	1	25	0.0000	1.0000
3	0	2	24	0.0000	1.0000
5	0	1	24	0.0000	1.0000
6	6	1	23	0.9130	0.9130
6	0	1	21	0.0000	0.9130
6	0	2	19	0.0000	0.8261
6	0	2	19	0.0000	0.8261
6	0	2	19	0.0000	0.8261
6	0	2	19	0.0000	0.8261
8	2	2	19	0.8947	0.7391
8	0	2	17	0.0000	0.7391
9	0	1	21	0.0000	0.9130
10	2	1	20	0.9000	0.8217
10	0	1	18	0.0000	0.8217
10	0	1	18	0.0000	0.8217
12	6	1	17	0.7647	0.6284
12	0	1	13	0.0000	0.6284
12	0	1	13	0.0000	0.6284
12	0	1	13	0.0000	0.6284
12	0	2	15	0.0000	0.6522
12	0	2	15	0.0000	0.6522
12	0	1	13	0.0000	0.6284
12	0	2	15	0.0000	0.6522
13	0	1	12	0.0000	0.6284
15	0	1	11	0.0000	0.6284
15	0	2	14	0.0000	0.6522
16	0	1	10	0.0000	0.6284
16	0	2	13	0.0000	0.6522
18	0	2	12	0.0000	0.6522

18	0	2	11	0.0000	0.6522
20	1	2	10	0.9000	0.5870
20	0	1	9	0.0000	0.6284
22	0	2	9	0.0000	0.5870
24	2	1	8	0.8750	0.5498
24	0	2	7	0.0000	0.5136
24	0	1	7	0.0000	0.5498
27	0	1	6	0.0000	0.5498
28	0	2	7	0.0000	0.5136
28	0	2	6	0.0000	0.5136
28	0	2	5	0.0000	0.5136
30	1	2	4	0.7500	0.3852
30	0	2	3	0.0000	0.3852
32	1	1	5	0.8000	0.4399
33	0	2	2	0.0000	0.3852
34	0	1	4	0.0000	0.4399
36	0	1	3	0.0000	0.4399
36	0	1	2	0.0000	0.4399
42	1	2	1	0.0000	0.0000
44	0	1	1	0.0000	0.4399

Total Expected Events for Experimental Group = 11.375
Observed Events for Experimental Group = 10.000
Total Expected Events for Control Group = 10.625
Observed Events for Control Group = 12.000
Chisquare = 0.344 with probability = 0.442
Risk = 0.778, Log Risk = -0.250, Std.Err. Log Risk = 0.427
95 Percent Confidence interval for Log Risk = (-1.087,0.586)
95 Percent Confidence interval for Risk = (0.337,1.796)

EXPERIMENTAL GROUP CUMULATIVE PROBABILITY

CASE	TIME	DEATHS	CENSORED	CUM.PROB.
1	1	0	1	1.000
3	5	0	1	1.000
4	6	6	0	0.913
5	6	0	0	0.913
12	9	0	1	0.913
13	10	2	0	0.822
14	10	0	0	0.822
15	10	0	1	0.822
16	12	6	0	0.628
17	12	0	0	0.628
18	12	0	0	0.628
19	12	0	0	0.628
22	12	0	1	0.628
24	13	0	1	0.628
25	15	0	1	0.628
27	16	0	1	0.628
32	20	0	1	0.628
34	24	2	0	0.550
36	24	0	1	0.550
37	27	0	1	0.550
43	32	1	0	0.440
45	34	0	1	0.440
46	36	0	1	0.440
47	36	0	1	0.440
49	44	0	1	0.440

CONTROL GROUP CUMULATIVE PROBABILITY

CASE	TIME	DEATHS	CENSORED	CUM.PROB.
2	3	0	1	1.000

6	6	0	0	0.826
7	6	0	0	0.826
8	6	0	0	0.826
9	6	0	0	0.826
10	8	2	0	0.739
11	8	0	0	0.739
20	12	0	0	0.652
21	12	0	0	0.652
23	12	0	1	0.652
26	15	0	1	0.652
28	16	0	1	0.652
29	18	0	1	0.652
30	18	0	1	0.652
31	20	1	0	0.587
33	22	0	1	0.587
35	24	0	0	0.514
38	28	0	1	0.514
39	28	0	1	0.514
40	28	0	1	0.514
41	30	1	0	0.385
42	30	0	1	0.385
44	33	0	1	0.385
48	42	1	0	0.000

The chi-square coefficient as well as the graph indicates no difference was found between the experimental and control group beyond what is reasonably expected through random selection from the same population.