

## Exercise 23: Comparing survival curves

The dataset `gehan` consist of data from a trial of 42 leukaemia patients<sup>14</sup>. `Cens=1` corresponds to no censoring. Some were treated with the drug 6-mercaptopurine and the rest are controls.

1. For both groups, calculate the Kaplan Meier estimate and its standard error  
Also estimate the median survival time for each of the groups.
2. How can we obtain a 95% “confidence curves” for the estimated curves? Try to make a plot of both the survival curves and the confidence curves for both groups in the same figure. Would you say there is any difference in the two curves?
3. Use the log-rank test to test if there is a significant difference between the two groups.
4. Fit a Cox regression model with treatment as a binary covariate and interpret the result.

Control		Drug	
Time	Cens	Time	Cens
1	1	6	1
1	1	6	1
2	1	6	1
2	1	6	0
3	1	7	1
4	1	9	0
4	1	10	1
5	1	10	0
5	1	11	0
8	1	13	1
8	1	16	1
8	1	17	0
8	1	19	0
11	1	20	0
11	1	22	1
12	1	23	1
12	1	25	0
15	1	32	0
17	1	32	0
22	1	34	0
23	1	35	0

---

<sup>14</sup>From Venables, W. N and Ripley, B. D. (1994). Modern Applied Statistics with S-Plus. Springer-Verlag.

