Vietnam Institute for Advanced Study in Mathematics

Survival analysis

Practical work 3: Logrank tests

(Lecturers: Agnès LAGNOUX & Jean-François DUPUY)

Exercise 1: Logrank test

We want to realize the logrank test on the data of Freireich (presented in the course). Remind that Freireich, in 1963, realized a therapeutic trial in order to compare the remission durations (in weeks) of patients that suffer from leukemia. The patients are divided into two subgroups: some of them received a medicine (6-MP) and the others a placebo. The results are presented in the following tabular:

6-MP	6	6	6	6^+	7	9^+	10	10^+	11^+	13	16
	17^+	19^+	20^+	22	23	25^+	32^+	32^+	34^+	35^+	
Placebo	1	1	2	2	3	4	4	5	5	8	8
	8	8	11	11	12	12	15	17	22	23	

The patients with a + sign correspond to lost subjects at the considered time of observation: they are censored, "excluded-alive" of the study and one only knows that their remission duration is greater than the observed delay.
One may use the following tabular to lead the different tests

	6-MP		Placebo			Weights w_i		
						Logrank	Gehan	Peto-Prentice
$T_{(i)}$	m_{Ai}	n_{Ai}	m_{Bi}	n_{Bi}	e_{Bi}	1	n_i	S_i^*
1								
2								
⋮								
23								

and present the results in

Test	Test stat.	p	RR_1
Logrank LR^2			
Approx. Logrank LRA^2			
Gehan			
Peto-Prentice			

Exercise 2: Logrank test

The following tabular presents the survival times after mastectomy of 45 women that suffers from a breast cancer. The patients are divided into two groups according to the presence or not of metastases. A + indicates a censored data. You can find the data in `breast.txt`.

No metastases	23	47	69	70+	71+	100+	101+	148	181	198+	208+	212+	224+
	5	8	10	13	18	24	26	26	31	35	40	41	48
Metastases	50	59	61	68	71	76+	105+	107+	109+	113	116+	118	143
	154+	162+	188+	212+	217+	225+							

1. Determine using `R` the Kaplan-Meier estimations of the survival functions in each group. Plot these estimations on the same figure adding the confidence intervals.

2. Compare the survival of the two groups using the the classical logrank test (function `survdiff`).

3. Use now the weighted logrank test with $w_i = \hat{S}_{KM}(t_i)$ (obtained for $\rho = 1$) to realize another comparison. Conclusion?

Exercise 3: Stratified logrank test

We consider a clinical trial conducted by Peto (1979) on comparison of the survival functions of two groups. We have an extra information: the kidney function that is known to influence the survival:

Participation time	Group	Kidney function	Participation time	Group	Kidney function
8	A	A	220	A	N
8	A	N	365^+	A	N
13	B	A	632	B	N
18	B	A	700	B	N
23	B	A	852^+	A	N
52	A	A	1296	B	N
63	A	A	1296^+	A	N
63	A	A	1328^+	A	N
70	B	N	1460^+	A	N
76	B	N	1976^+	A	N
180	B	N	1990^+	B	N
195	B	N	2240^+	B	N
210	B	N			

The letter A (respectively N) means an abnormal (resp. N) kidney function. The censored data are indicated by a +.

1. Check by the logrank test that the kidney function influences the survival. You can also plot the survival Kaplan-Meier function according to the kidney function.

2. Compare using a logrank test the survival functions of the two groups. Validate your results using the argument `subset` of the function `survdiff`. Conclusion?

3. Compare using a logrank test the survival functions of the two groups, separately for the patients with a normal kidney function and the patients with an abnormal one. Validate your results using the argument `subset` of the function `survdiff`. Conclusion?

4. Compare using a logrank test the survival functions of the two groups, stratifying on the kidney function. Validate your results using the option `+strata` of the function `survdiff`. Conclusion?

One may use the following tabular to lead the results analytically:

Death times t_i	Treatment A N m_{A_i}	A N n_{A_i}	A AN m_{A_i}	A AN n_{A_i}	B N m_{B_i}	B N n_{B_i}	B AN m_{B_i}	B AN n_{B_i}	Kidney function N e_{B_i}	N v_{B_i}	AN e_{B_i}	AN v_{B_i}	Total e_{B_i}	Total v_{B_i}
8														
13														
18														
23														
52														
63														
70														
76														
180														
195														
210														
220														
632														
700														
1296														
Total														

Exercise 4: Comparison of three subgroups

The data analyzed in this example concern three small (fictive) samples corresponding to the three different treatment doses (Thomas 1977). The survival and censoring are represented in the following tabular.

Group	N_j	Dose x_j	Survival and censoring									
A	9	0	73^+	74^+	75^+	76	76	76^+	99	166	246^+	
B	10	1.5	43^+	44^+	45^+	67	68^+	136	136	150	150	150
C	10	2	41^+	41^+	47	47^+	47^+	58	58	58	100^+	117

The censored data are indicated by a $+$.

One may use the following tabular to compute the heterogeneity and trend statistics.

t_i	Group A m_{A_i}	A n_{A_i}	B m_{B_i}	B n_{B_i}	C m_{C_i}	C n_{C_i}	e_{B_i}	v_{B_i}	e_{C_i}	v_{C_i}	c_{B_i,C_i}
47											
58											
67											
76											
99											
117											
136											
150											
166											
Total											

The expectations $e.$, variances $v.$ and covariances $c_{.,.}$ of the death numbers under H_0 are indicated only for groups B and C, since these quantities are not necessary to the computation of the statistics. The reader can compute E_A and check that

$$E_A + E_B + E_C = O_A + O_B + O_C = 15.$$