

Lin_ST625_HW5

Frances Lin

2/5/2022

1

Suppose we have

Group A: 14, 25, 30, 34+, 46, 56+

Group B: 25+, 34, 34, 36, 43+, 46,

1a

Log-rank test is calculated by

1. List out the uncensored survival time of both groups (e.g. $t = 14, 25, 30, 34, 36, 46$)
2. At each t , construct a table with # of death (d_{i1} , d_{i2}) and # of alive ($Y_{i1} - d_{i1}$, $Y_{i2} - d_{i2}$) and their row and column total for each $j = 1, 2$ group
3. Compute $e_{i1} = d_i * (y_{i1}/y_i)$ and $v_{i1} = d_i * (y_{i1}/y_i) * ((y_i - y_{i1})/(y_i)) * ((y_i - d_i)/(y_i - 1))$
4. Each row of **1b** contains d_{i1} , e_{i1} and v_{i1} of each table

At each t_i ,

# death	# alive	total
di1	yi1 - di1	yi1
di2	yi2 - di2	yi2
di	yi - di	yi

$t = 14$

d	y_d	y
1	5	6
0	6	6
1	11	12

$t = 25$

d	y_d	y
1	4	5
0	6	6
1	10	11

$t = 30$

d	y_d	y
1	3	4
0	5	5
1	8	9

$t = 34$

d	y_d	y
0	3	3
2	3	5
2	6	8

$t = 36$

d	y_d	y
0	2	2
1	2	3
1	4	5

Finally, $t = 46$

d	y_d	y
1	1	2
1	0	1
2	1	3

1b

t	d_i_1	e_i_1	v_i_1
14	1	0.5	0.25
25	1	0.4545	0.2479
30	1	0.4444	0.2469
34	0	0.75	0.4018
36	0	0.4	0.24
46	1	1.333	0.2222
sum	4	3.882	1.609

1c

$$\text{Logrank (test statistics)} = \frac{(d-e)^2}{v} = (4 - 3.882)^2 / 1.609 =$$

[1] 0.008607251

1d

Recall that if H_0 is true, then it follows χ_1^2 , so $\text{p-val} = P(\chi_1^2 \geq \text{Logrank}) =$

[1] 0.9260821

Since $\text{p-val} > 0.05$ ($\alpha = 0.05$), we conclude that there is no significant evidence that the survival of the two groups are different.

2

Drug A: 16, 16+, 18+, 19+, 20, 28, 32+

Drug B: 10, 14, 15, 18, 20+, 21

2a

$t = 10$

d	y_d	y
0	7	7
1	5	6
1	12	13

$t = 14$

d	y_d	y
0	7	7
1	4	5
1	11	12

$t = 15$

d	y_d	y
0	7	7
1	3	4
1	10	11

$t = 16$

d	y_d	y
1	6	7
0	3	3
1	9	10

$t = 18$

d	y_d	y
0	5	5
1	2	3
1	7	8

$t = 20$

d	y_d	y
1	2	3
0	2	2
1	4	5

$t = 21$

d	y_d	y
0	2	2
1	0	1
1	2	3

$t = 28$

d	y_d	y
1	1	2
0	0	0
1	1	2

2b

t	d_i_1	e_i_1	v_i_1
10	0	0.5385	0.2485
14	0	0.5833	0.2431
15	0	0.6364	0.2314
16	1	0.7	0.21
18	0	0.625	0.2344
20	1	0.6	0.24
21	0	0.6667	0.2222
28	1	1	0
sum	3	5.35	1.63

2c

$$\text{Logrank (test statistics)} = \frac{(d-e)^2}{v} = (3 - 5.35)^2 / 1.63 =$$

[1] 3.388409

2d

$$\text{p-val} = P(\chi_1^2 \geq \text{Logrank}) =$$

[1] 0.06565627

For $\alpha = 0.05$ at least, since $\text{p-val} > 0.05$, we conclude that there is no significant evidence that the survival of the two groups are different. At $\alpha = 0.1$, there is some evidence.

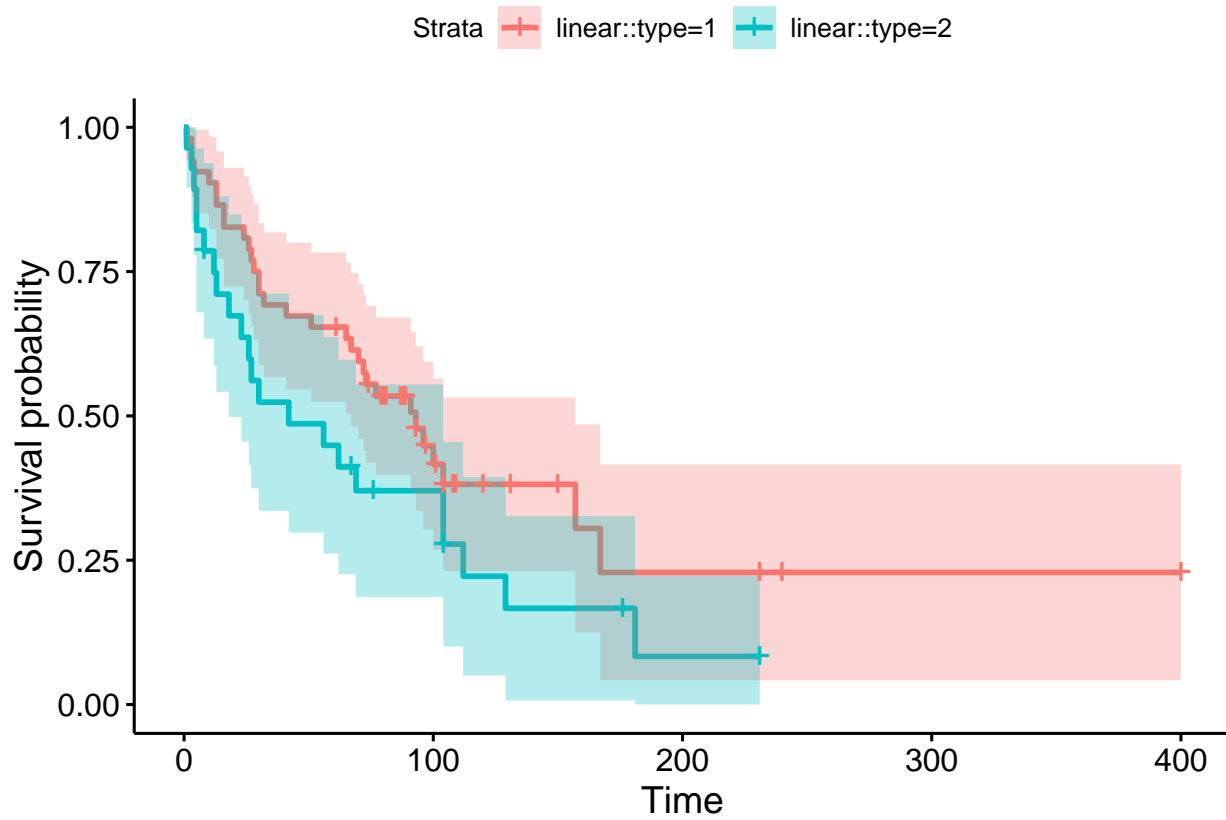
It is not consistent with my finding in Homework 4. Perhaps I made some errors somewhere.

3a

type	time	delta
1	1	1
1	3	1
1	3	1
1	4	1
1	10	1
1	13	1

3b

Kaplan-Meier plot with point-wise CIs for Aneuploid Tumor (type = 1) and Diploid Tumor (type = 2) is given:



3c

H_0 : There is no difference in survival functions between two groups (Aneuploid Tumor vs. Diploid Tumor)
(or $H_0 : S_1 = S_2$)

H_a : There is some difference in survival functions between two groups (Aneuploid Tumor vs. Diploid Tumor)

Log-rank test results:

```
## Call:
## survdiff(formula = Surv(time, delta) ~ type, data = Tongue)
```

```
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## type=1 52      31      36.6      0.843      2.79
## type=2 28      22      16.4      1.873      2.79
##
## Chisq= 2.8  on 1 degrees of freedom, p= 0.09

## Call:
## survdiff(formula = Surv(time, delta) ~ type, data = Tongue, rho = 1)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## type=1 52      20.2      24.4      0.731      3.3
## type=2 28      15.1      10.9      1.643      3.3
##
## Chisq= 3.3  on 1 degrees of freedom, p= 0.07
```

$Chisq = 2.8$ and $p = 0.09$ for the log-rank test and $Chisq = 3.3$ and $p = 0.07$ for the Peto-Peto test. At the $\alpha = 0.05$, we fail to reject the H_0 and conclude that there is not enough evidence to say that the survival functions of two groups differ. At the $\alpha = 0.1$, we would say there is some evidence.

3d

It is not quite consistent with the plots given in 3b. However, as Lab5-R-1.pdf points out that KM estimator has larger variability at later times. In addition, the 95% CI band of `type = 2` does overlap with the survival curve of `type = 1`.

4

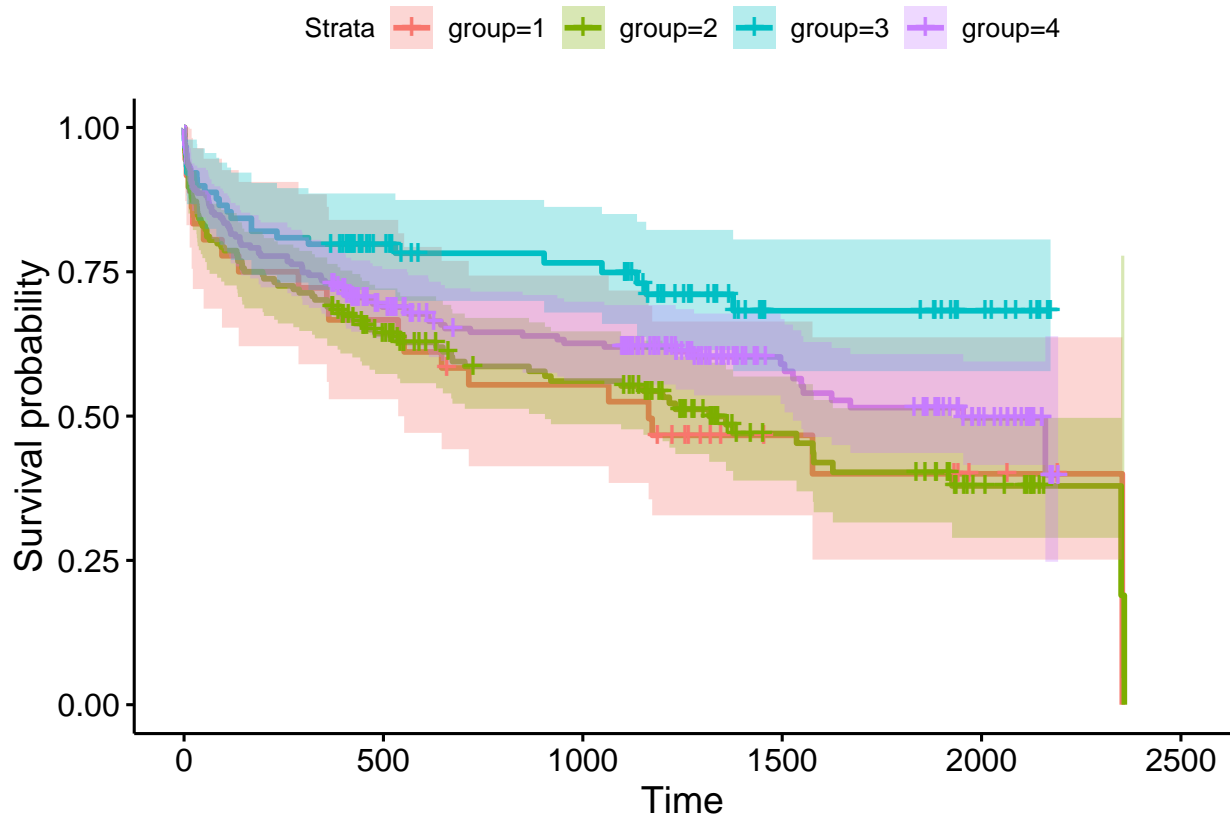
4a

Admitdate	Fdate	time
1997-01-13	2002-12-31	2178 days
1997-01-19	2002-12-31	2172 days
1997-01-01	2002-12-31	2190 days

4b

Don't think this is right. Perhaps need to create a group variable.

name	gender	cvd	n
M0	0	0	89
M1	0	1	211
F0	1	0	36
F1	1	1	164



4c

H_0 : There is no difference in survival functions among the four groups (F0, F1, M0, M2) (or $H_0 : S_1 = S_2 = S_3 = S_4$)

H_a : There is at least one difference in survival functions among the four groups (F0, F1, M0, M2)

Let **group=1** = F0 (Female with no history of Cardiovascular), **group=2** = F1, **group=3** = M0, etc

Log-rank test results:

```
## Call:
## survdiff(formula = Surv(time, fstat) ~ group, data = df)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## group=1  36      21     16.4      1.315      1.453
## group=2 164      83     67.9      3.352      4.987
## group=3  89      24     39.5      6.069      7.479
## group=4 211      87     91.2      0.198      0.348
##
##  Chisq= 11.1  on 3 degrees of freedom, p= 0.01
```

Since $Chisq = 11.1$ and $p = 0.01$, at the $\alpha = 0.05$, we reject the H_0 and conclude that there is evidence to say that at least one of the survival functions of the four groups differs.

```
##           chiSq df pChisq
## 1           11.1470 3      2
## n           7.8125 3      6
## sqrtN       9.3779 3      5
## S1          9.5028 3      3
## S2          9.4896 3      4
## FH_p=1_q=1 12.4503 3      1
## $tft
##           Q          Var      Z pNorm
## 1          2.8613e+01 2.1873e+02 1.9347      1
## n          9.1960e+03 3.0725e+07 1.6590      6
## sqrtN      5.0696e+02 7.7575e+04 1.8202      5
## S1          2.1213e+01 1.3545e+02 1.8227      4
## S2          2.1165e+01 1.3461e+02 1.8242      3
## FH_p=1_q=1 4.7986e+00 6.3355e+00 1.9064      2
##
## $scores
## [1] 1 2 3 4
```

PS. I know we are supposed to use the `comp` function of the `survMisc` package. However, there is perhaps a bug in there. There is an opened issue on Github: [here](#). All $pChisq > 1$ for Lab5-R-1.Rmd too. All $pChisq > 1$ here too.

4d

Let **group=1** = F0 (Female with no history of Cardiovascular), **group=2** = F1, **group=3** = M0, etc

```
##
## Pairwise comparisons using Log-Rank test
##
## data:  df and group
##
##      1      2      3
## 2 1.000 -      -
## 3 0.077 0.012 -
## 4 1.000 0.547 0.292
##
## P value adjustment method: bonferroni
```

Among them, **group=2** (F1: Female with history of Cardiovascular) and **group=3** (M0: Male with no history of Cardiovascular) are significantly different ($p = 0.012$). All the other pairs are not significantly different.