5291

Yi Chen(yc3356)

April 18, 2018

Homework 7

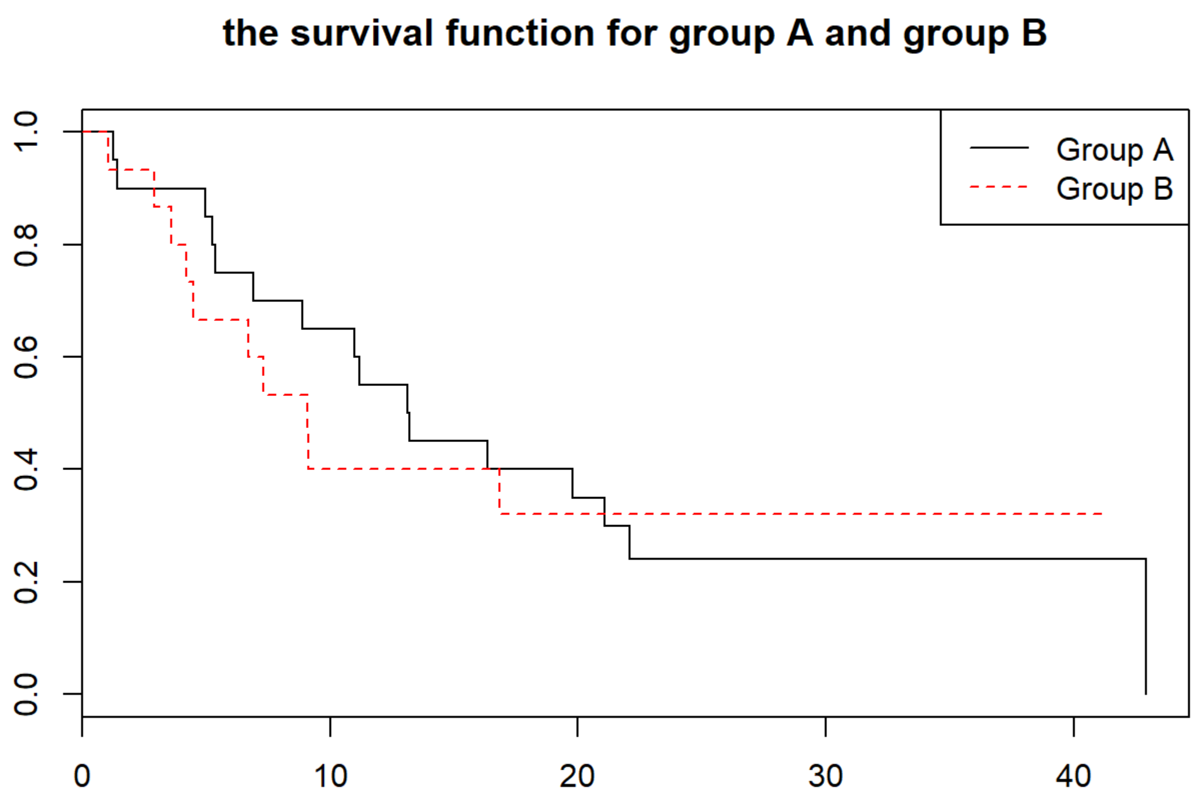
```
# read the data
times <- c(1.25,1.41, 4.98, 5.25, 5.38, 6.92, 8.89, 10.98, 11.18, 13.11, 13.21, 16.33, 19.77, 21.08,21.84,
 22.07, 31.38, 32.61, 37.18, 42.92, 1.05, 2.92, 3.61, 4.20, 4.49, 6.72, 7.31, 9.08, 9.11, 14.49, 16.85, 1
8.82, 26.59, 30.26, 41.34)
status <- c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0
, 0, 0, 0)
group <- c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
 2, 2, 2)
```

1. Obtain and plot the survival fucntions S_A and S_B of group A and B.

```
library(survival)
```

```
## Warning: package 'survival' was built under R version 3.4.4
```

```
fit <- survfit(Surv(times,status)~group)
plot(fit,lty = 1:2, col = 1:2,main = "the survival function for group A and group B")
legend("topright",legend=c("Group A", "Group B"), col = 1:2, lty=1:2, cex=1)
```




```
print("the survival function for group A")
```

```
## [1] "the survival function for group A"
```

```
fit$surv[1:20]
```

```
##  [1] 0.95 0.90 0.85 0.80 0.75 0.70 0.65 0.60 0.55 0.50 0.45 0.40 0.35 0.30
## [15] 0.30 0.24 0.24 0.24 0.24 0.00
```

```
print("the survival function for group B")
```

```
## [1] "the survival function for group B"
```

```
fit$surv[21:length(fit$surv)]
```

```
##  [1] 0.9333333 0.8666667 0.8000000 0.7333333 0.6666667 0.6000000 0.5333333
##  [8] 0.4666667 0.4000000 0.4000000 0.3200000 0.3200000 0.3200000 0.3200000
## [15] 0.3200000
```

2. Estimate $S_A(10)$ and $S_B(10)$ using a 95% confidence interval.

$$S\hat{(t)} \pm Z_{\alpha/2}\sigma\hat{(t)}$$

```
summary(fit)
```

```
## Call: survfit(formula = Surv(times, status) ~ group)
##
##                 group=1
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   1.25     20       1     0.95  0.0487        0.859        1.000
##   1.41     19       1     0.90  0.0671        0.778        1.000
##   4.98     18       1     0.85  0.0798        0.707        1.000
##   5.25     17       1     0.80  0.0894        0.643        0.996
##   5.38     16       1     0.75  0.0968        0.582        0.966
##   6.92     15       1     0.70  0.1025        0.525        0.933
##   8.89     14       1     0.65  0.1067        0.471        0.897
##  10.98     13       1     0.60  0.1095        0.420        0.858
##  11.18     12       1     0.55  0.1112        0.370        0.818
##  13.11     11       1     0.50  0.1118        0.323        0.775
##  13.21     10       1     0.45  0.1112        0.277        0.731
##  16.33      9       1     0.40  0.1095        0.234        0.684
##  19.77      8       1     0.35  0.1067        0.193        0.636
##  21.08      7       1     0.30  0.1025        0.154        0.586
##  22.07      5       1     0.24  0.0980        0.108        0.534
##  42.92      1       1     0.00     NaN           NA           NA
##
##                 group=2
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   1.05     15       1    0.933  0.0644        0.815        1.000
##   2.92     14       1    0.867  0.0878        0.711        1.000
##   3.61     13       1    0.800  0.1033        0.621        1.000
##   4.20     12       1    0.733  0.1142        0.540        0.995
##   4.49     11       1    0.667  0.1217        0.466        0.953
##   6.72     10       1    0.600  0.1265        0.397        0.907
##   7.31      9       1    0.533  0.1288        0.332        0.856
##   9.08      8       1    0.467  0.1288        0.272        0.802
##   9.11      7       1    0.400  0.1265        0.215        0.743
##  16.85      5       1    0.320  0.1239        0.150        0.684
```

As we can see from the summary of the fit result.

$$S_B(10) \in [0.215, 0.743]$$

$$S_A(10) \in [0.471, 0.897]$$

3. Test H0: $S_A = S_B$

```
survdiff(Surv(times,status)~group, rho=0)
```

```
## Call:
## survdiff(formula = Surv(times, status) ~ group, rho = 0)
##
##          N Observed Expected (O-E)^2/E (O-E)^2/V
## group=1 20       16    16.66    0.0261    0.0749
## group=2 15       10     9.34    0.0466    0.0749
##
##  Chisq= 0.1  on 1 degrees of freedom, p= 0.784
```

As we can see from the test: the p = 0.784 which means that we fail to reject the null hypothesis that the two survival functions are equal.

4 Cox proportional hazard model

```
fit2 <- coxph(formula = Surv(times, status) ~ factor(group))
summary(fit2)
```

```
## Call:
## coxph(formula = Surv(times, status) ~ factor(group))
##
##   n= 35, number of events= 26
##
##                  coef exp(coef) se(coef)     z Pr(>|z|)
## factor(group)2 0.1120    1.1185   0.4096 0.273    0.784
##
##                exp(coef) exp(-coef) lower .95 upper .95
## factor(group)2     1.119      0.894    0.5012     2.496
##
## Concordance= 0.539  (se = 0.054 )
## Rsquare= 0.002   (max possible= 0.987 )
## Likelihood ratio test= 0.07  on 1 df,   p=0.7853
## Wald test            = 0.07  on 1 df,   p=0.7845
## Score (logrank) test = 0.07  on 1 df,   p=0.7844
```

Interpretation: We have 95% confidence to conclude that: if we fixed other protential variables, approximately, at any instant in time, the group 1 would be between 0.5012 and 2.496 times likely to die thant than group 2.

5. Test H0: $\beta = 0$

As we can see from the result of the test. The p-value is 0.784 which means that we fail to reject the H0 which means that the $\beta = 0$. Besides, we can see that the confidence interval for the question d include the point 1 which also means that the value of β could be equal to 0.