GR5291 Homework 10

Hanao Li hl3202

11/22/2019

Problem 1

```
if(!require("pacman")) install.packages("pacman")
p_load(survminer, survival, dplyr)
```

```
data = lung
data$sex_group = ifelse(data$sex==1, "Male", "Female")
data$sex_group <- as.factor(data$sex_group)
surv = Surv(time = data$time, event = data$status)
cox = coxph(surv ~ sex_group, data = data)
summary(cox)
```

```
## Call:
## coxph(formula = surv ~ sex_group, data = data)
##
##   n= 228, number of events= 165
##
##                 coef exp(coef) se(coef)     z Pr(>|z|)
## sex_groupMale 0.5310    1.7007   0.1672 3.176  0.00149 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## sex_groupMale     1.701      0.588     1.226      2.36
##
## Concordance= 0.579  (se = 0.021 )
## Likelihood ratio test= 10.63  on 1 df,   p=0.001
## Wald test            = 10.09  on 1 df,   p=0.001
## Score (logrank) test = 10.33  on 1 df,   p=0.001
```

From the results, we could see that the coefficient is 0.5310 which means that the hazard ratio for female relatie to the male is exp(0.5310) that is 1.701

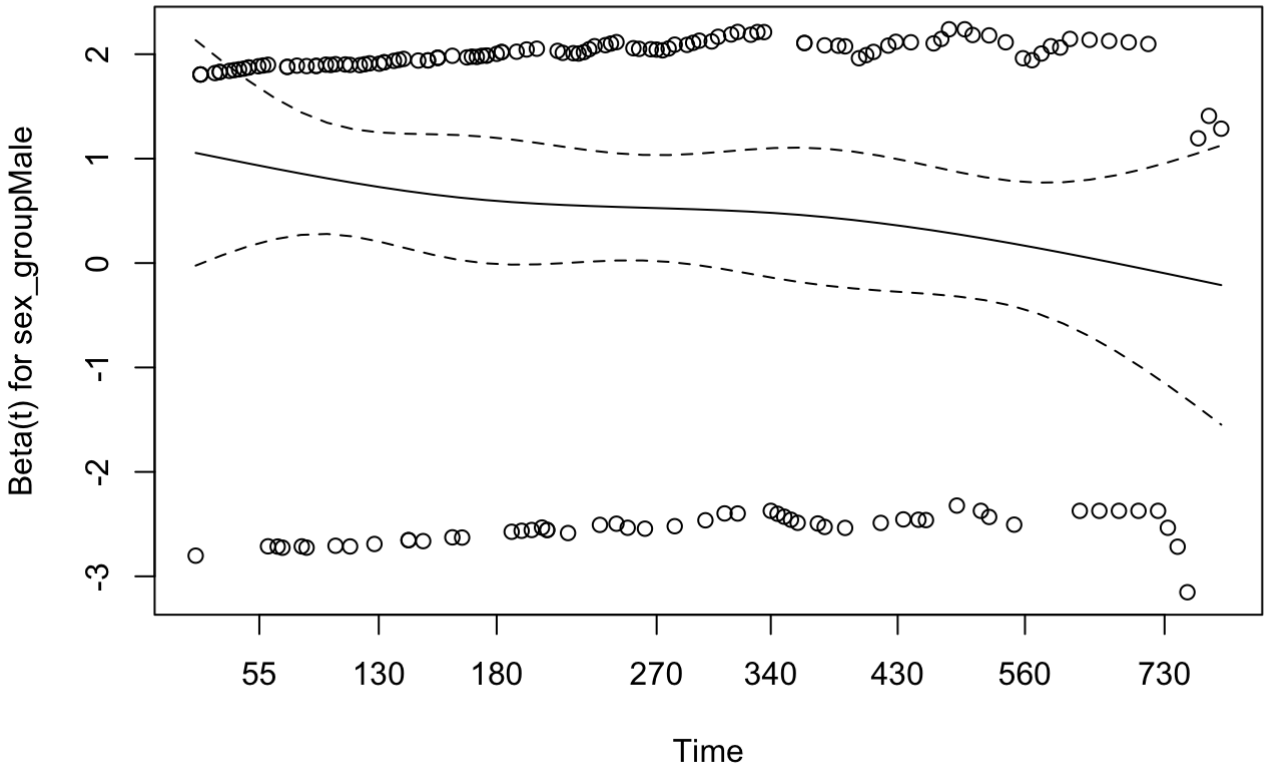
Problem 2

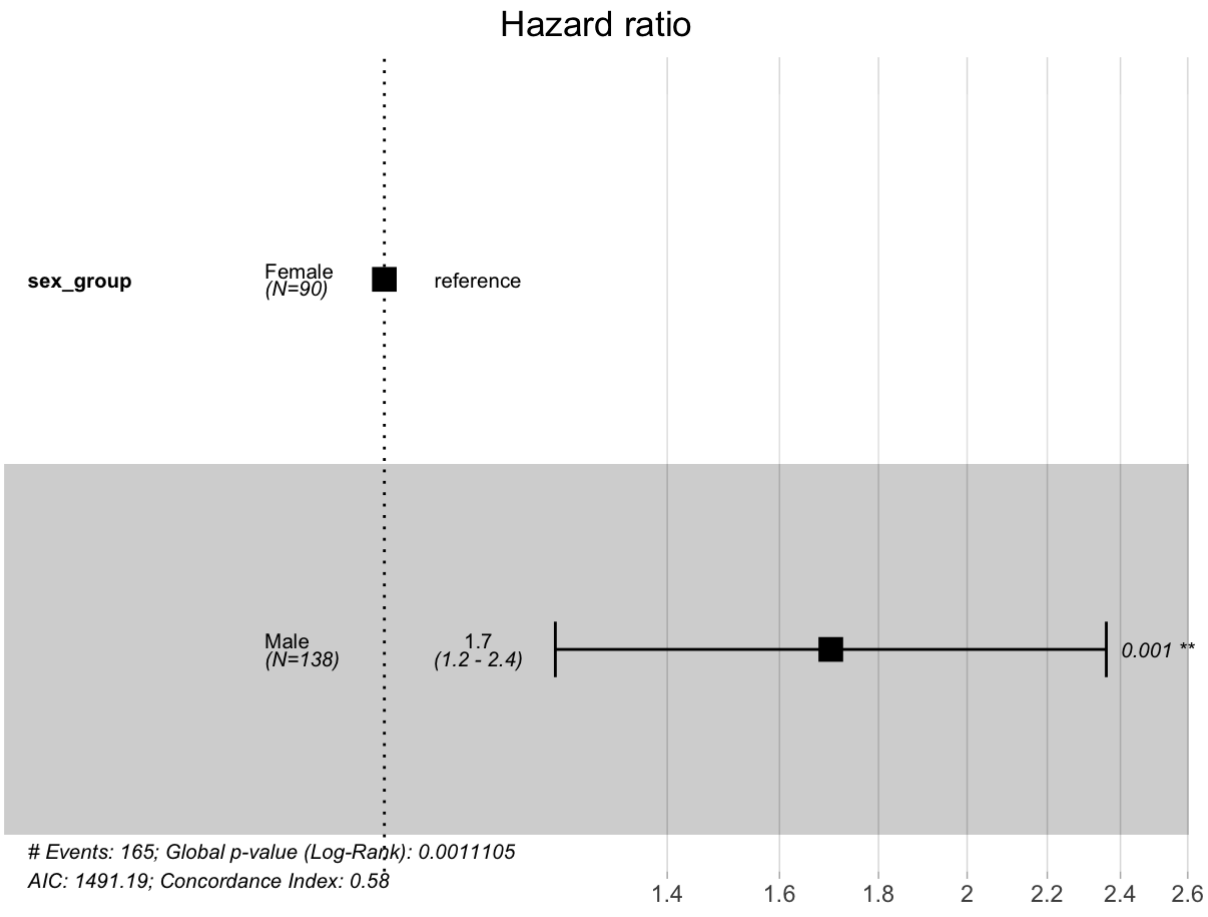
```
p1 = cox.zph(cox)
plot(p1)
p1
```

```
##                  rho chisq      p
## sex_groupMale -0.131  2.77 0.0962
```

```
ggforest(cox, data = data)
```

```
## Warning: Removed 1 rows containing missing values (geom_errorbar).
```





From the results, we could see that the p value we get form the cox zph test is larger than 0.05 so we do not reject the

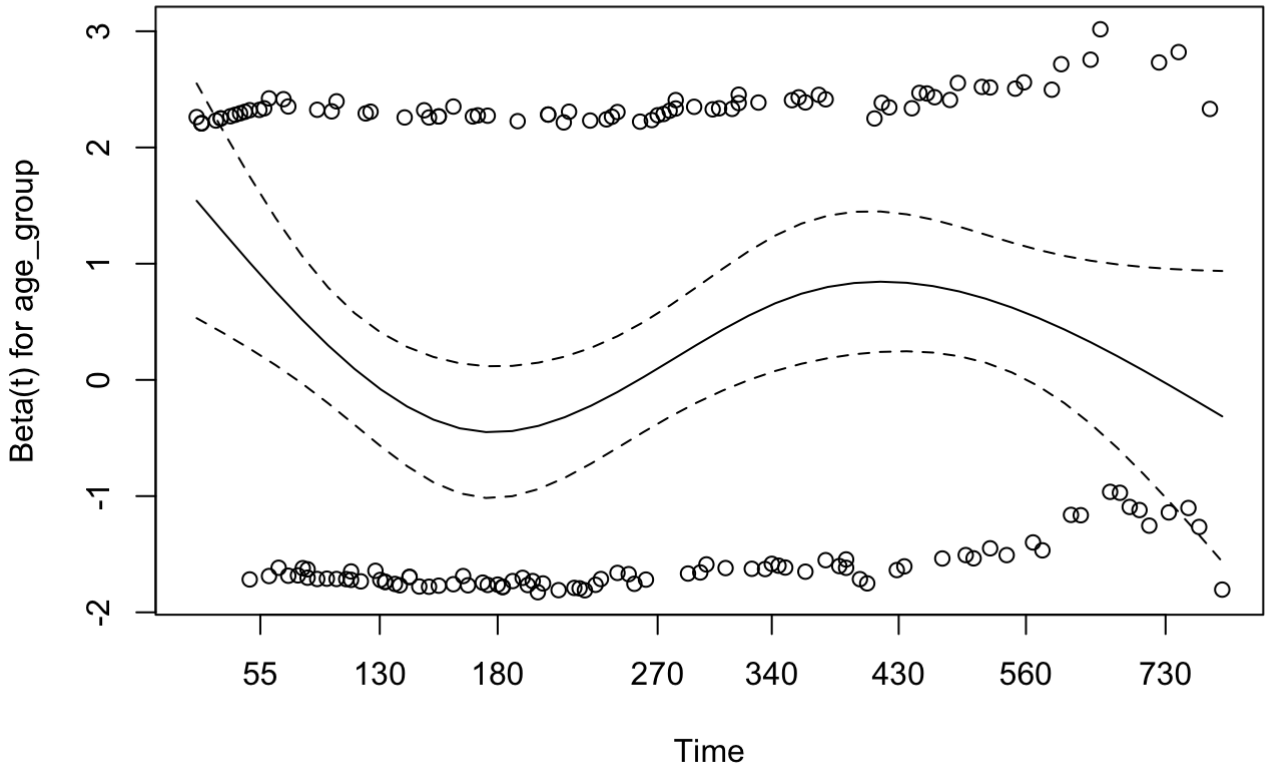
null hypothesis so we can assume the proportional hazards.

Problem 3

```
data$age_group = 1
data$age_group[which(data$age < 65)] = 0
cox = coxph(surv ~ age_group +sex_group, data = data)
summary(cox)
```

```
## Call:
## coxph(formula = surv ~ age_group + sex_group, data = data)
##
##   n= 228, number of events= 165
##
##                 coef exp(coef) se(coef)     z Pr(>|z|)
## age_group     0.2892    1.3354   0.1564 1.849  0.06449 .
## sex_groupMale 0.5257    1.6916   0.1673 3.143  0.00167 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## age_group         1.335     0.7489    0.9828     1.815
## sex_groupMale     1.692     0.5911    1.2188     2.348
##
## Concordance= 0.596  (se = 0.024 )
## Likelihood ratio test= 14.03  on 2 df,   p=9e-04
## Wald test            = 13.49  on 2 df,   p=0.001
## Score (logrank) test = 13.76  on 2 df,   p=0.001
```

```
p2 = cox.zph(cox)
plot(p2)
```

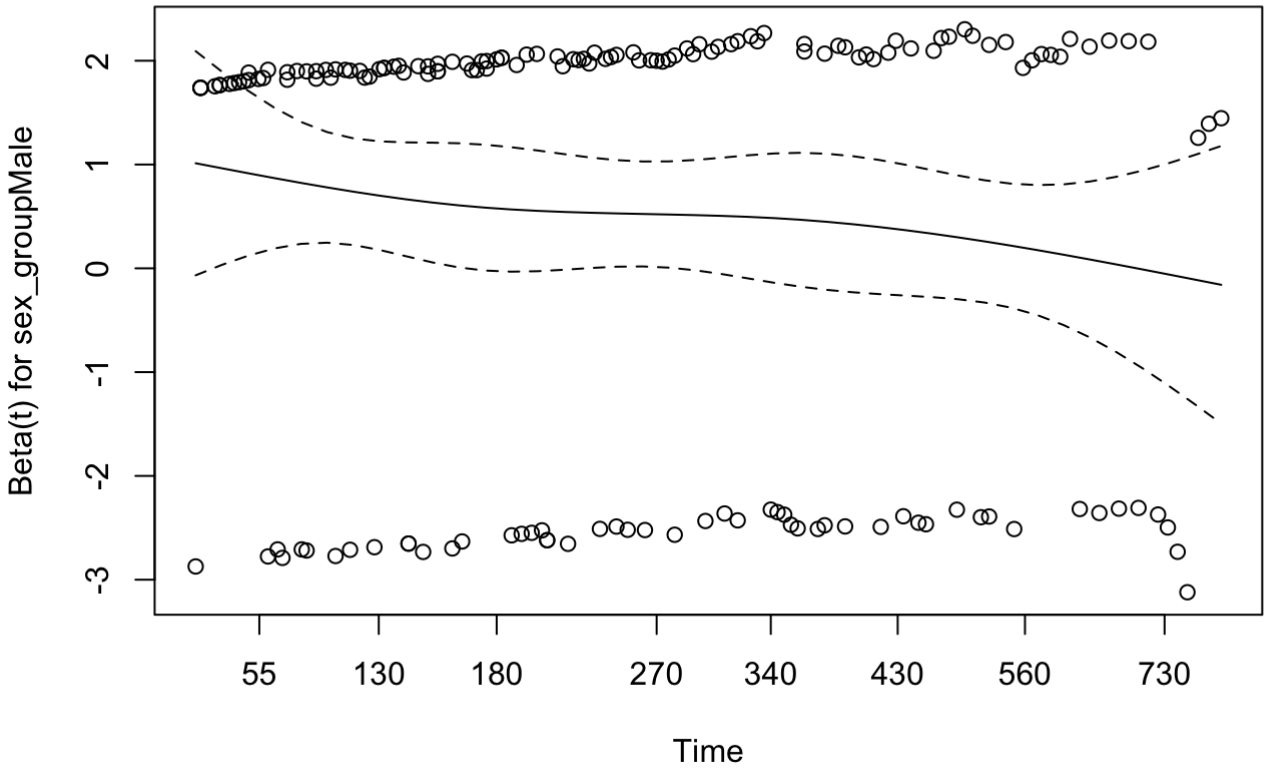


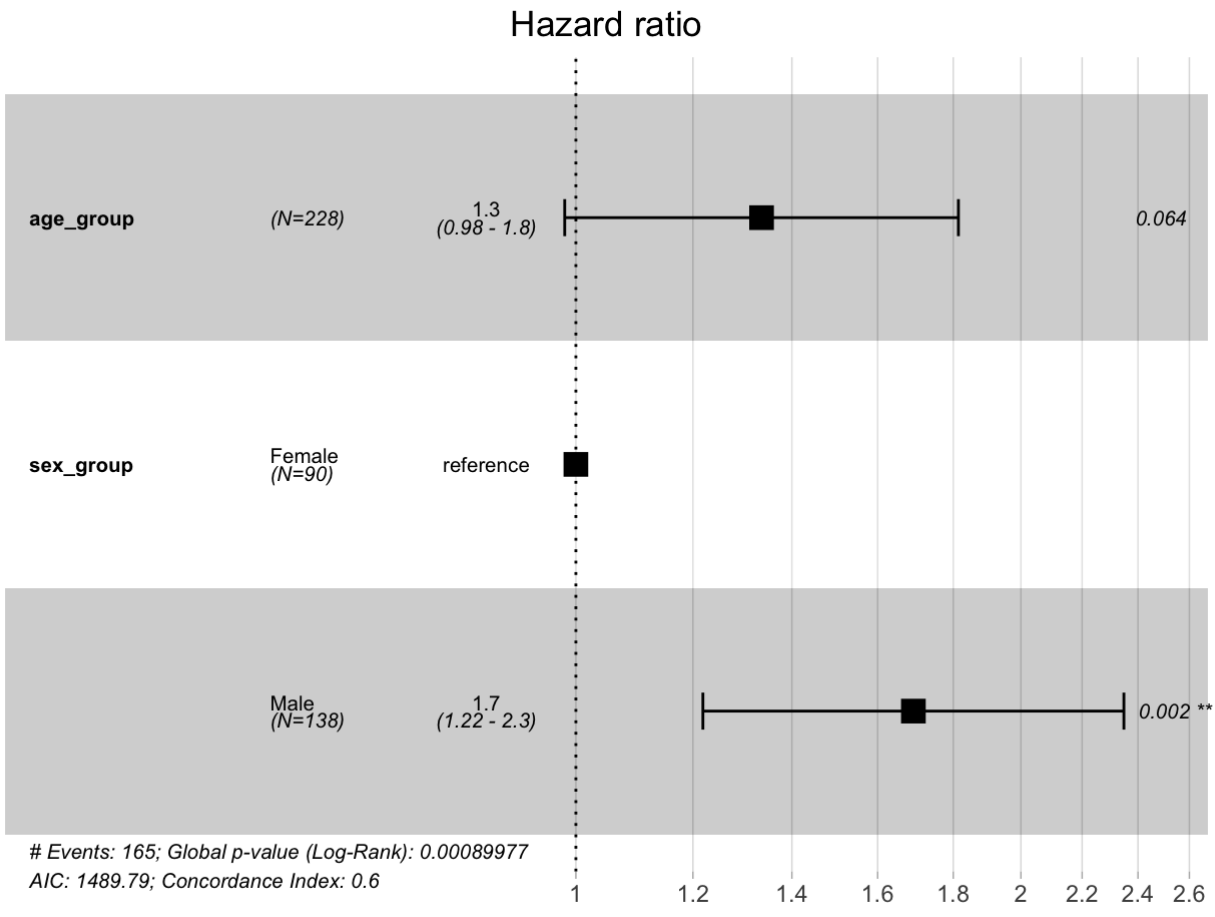

```
p2
```

```
##                   rho   chisq     p
## age_group      0.0033 0.00177 0.966
## sex_groupMale -0.1192 2.28315 0.131
## GLOBAL             NA 2.28343 0.319
```

```
ggforest(cox, data = data)
```

```
## Warning: Removed 1 rows containing missing values (geom_errorbar).
```





From the results above, we could see that sex_groupMale is significant. It has a coeficient of 0..5257 and exp(0.5257)

that is 1.6932 which means that the result is significant relationship between patient's sex and decreased risk of death. The hazard ratio for female relative to the male is 1.6932 adjusting for age. From the cox zph test, we could see that none of the variables are significant so we could not reject the null and we can conclude we can assume the proportional hazards.