5291 hw10

Yijin Wang

12/4/2020

Define 'AGE GROUP' as 'YOUNG" if 'age < 65", and OLD, otherwise.

1. Using a Cox proportional hazards model, estimate the hazard rate for old relative to young
2. Assess the validity of the proportional hazards assumption in (1)
3. Repeat 1, adjusting "Sex".

```
library(survival)
data(lung)
#1
lung["age"]<-ifelse(lung["age"]<65, "YOUNG", "OLD")
fitcox1<-coxph(Surv(time, status) ~ age, lung)
summary(fitcox1)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ age, data = lung)
##
##   n= 228, number of events= 165
##
##             coef exp(coef) se(coef)     z Pr(>|z|)
## ageYOUNG -0.2985    0.7419   0.1562 -1.91   0.0561 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## ageYOUNG    0.7419      1.348    0.5462     1.008
##
## Concordance= 0.538  (se = 0.022 )
## Likelihood ratio test= 3.62  on 1 df,   p=0.06
## Wald test            = 3.65  on 1 df,   p=0.06
## Score (logrank) test = 3.68  on 1 df,   p=0.06
```

2 1) The regression coefficients: The beta coefficient for age = -0.2985 indicates that younger people have lower risk of death (lower survival rates) than older people. 2) Hazard ratios: The exp(coef) indicates that being younger people reduces the hazard by a factor of 0.7419 or 26%. 3) Statistical significance: We have p-values for three alternative tests for overall significance of the model. The P-values are all above 0.05, which indicates that age group is not significant.

```
#3
fitcox2<-coxph(Surv(time, status) ~ sex, lung)
summary(fitcox2)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ sex, data = lung)
##
##   n= 228, number of events= 165
```

```
##
##        coef exp(coef) se(coef)      z Pr(>|z|)
## sex -0.5310    0.5880   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     0.588      1.701    0.4237     0.816
##
## 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
```

1) The regression coefficients: The beta coefficient for sex = -0.531 indicates that female have lower risk of death (lower survival rates) than male.
2) Hazard ratios: The exp(coef) indicates that being female reduces the hazard by a factor of 0.59 or 41%.
3) Statistical significance: We have p-values for three alternative tests for overall significance of the model. The P-values are all below 0.05, which indicates that sex is highly significant.