Survival Analysis HW 6

Mario Ibanez

March 26, 2016

Question 9.1

In Exercise 8.1, a proportional hazards model was fit to data from a study of the effects of ploidy on survival for patients with cancer of the tongue. A single binary covariate was used. Using an appropriate time-dependent covariate, test the hypothesis that the hazard rates for the two groups are proportional.

Answer

Below is a small view of the data for this question:

```
head(tongue)
```

```
##   type time delta      Type
## 1    1    1     1 Aneuploid
## 2    1    3     1 Aneuploid
## 3    1    3     1 Aneuploid
## 4    1    4     1 Aneuploid
## 5    1   10     1 Aneuploid
## 6    1   13     1 Aneuploid
```

There are a few ways to test whether the hazard rates for the two groups are proportional, namely, adding time depending covariates to the model and testing their significance. This can be done manually as was done in examples in the book. A new covariate can be defined as:

$$Z_2(t) = Z_1 * g(t)$$

where $g(t)$ is a function of time such as $ln(time)$. Below is the output of the model with this extra covariate:

```
tongue$Z2_time <- tongue$type * log(tongue$time)
cox_fit <- coxph(Surv(time = time, event = delta) ~ Type + Z2_time,
                 data = tongue, ties = "breslow")
```

```
summary(cox_fit)
```

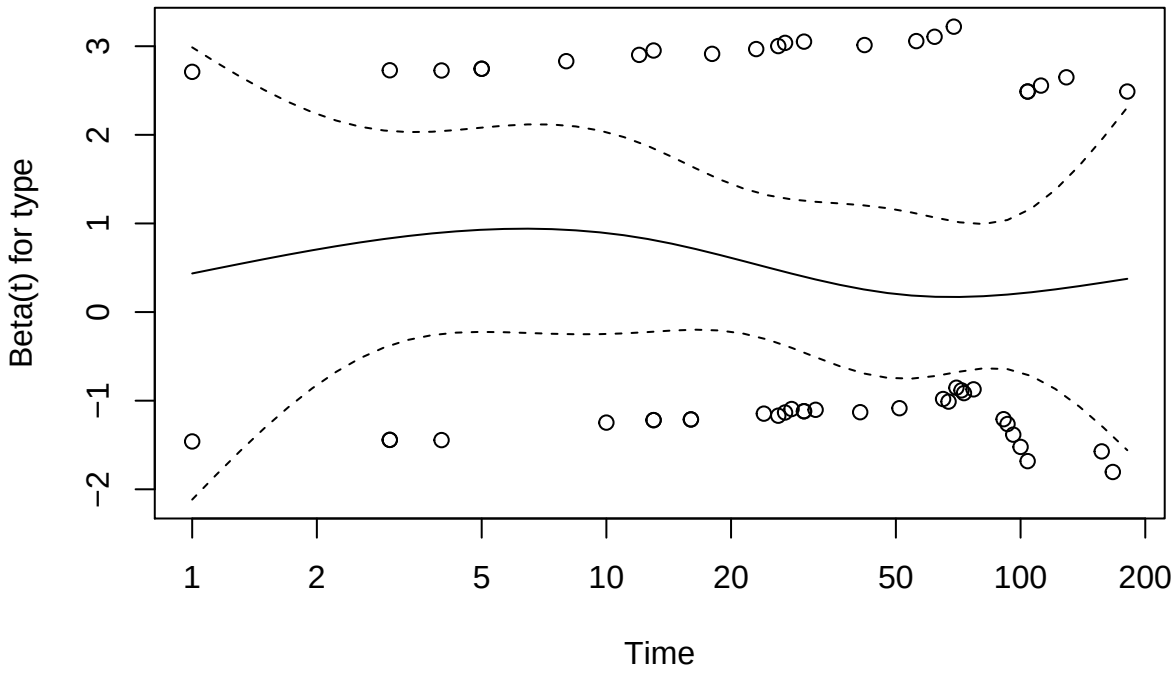
```
## Call:
## coxph(formula = Surv(time = time, event = delta) ~ Type + Z2_time,
##     data = tongue, ties = "breslow")
##
##   n= 80, number of events= 53
##
##                   coef exp(coef)  se(coef)      z Pr(>|z|)
## TypeDiploid    7.1332 1252.9045    0.9640  7.399 1.37e-13 ***
## Z2_time       -1.8679    0.1544    0.2614 -7.147 8.87e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##             exp(coef) exp(-coef) lower .95 upper .95
## TypeDiploid 1252.9045  0.0007981 189.37326 8289.2890
## Z2_time        0.1544  6.4746938   0.09254    0.2578
##
## Concordance= 0.956  (se = 0.045 )
## Rsquare= 0.796   (max possible= 0.993 )
## Likelihood ratio test= 127  on 2 df,   p=0
## Wald test            = 55.8  on 2 df,   p=7.643e-13
## Score (logrank) test = 159.5  on 2 df,   p=0
```

The coefficient for the parameter $type * ln(time)$ has a very small p-value, which is evidence that the hazard is not proportional through time.

Another method is to use the built in function in *R* called *cox.zph*. This function automatically performs the method used above to test whether the assumptions hold:

```
cox_fit <- coxph(Surv(time = time, event = delta) ~ type,
                 data = tongue, ties = "breslow")
cox_zph <- cox.zph(cox_fit, transform = log, global = TRUE)
print(cox_zph); plot(cox_zph)
```

```
##         rho chisq     p
## type -0.103 0.544 0.461
```



Using this test, it appears that the assumption that the hazard rate is proportional through time is reasonable. In the plot above, a line that is approximately horizontal is what would be expected if the assumption of proportional hazard was met.

Question 9.2

In Exercise 8.2, a proportional hazards model was fit to data from a study of the survival of rats implanted with F98 glioma cells in their brains. Three groups of rats were considered: control rats, rats given radiation only, and rats given radiation plus BPA. Using an appropriate set of time-dependent covariates, test that the hazard rates of the three groups are proportional.

Answer

Below is a small view of the data for this question:

```
head(bnct)
```

```
##   trt time death Treatment
## 1   1   20     1 Untreated
## 2   1   21     1 Untreated
## 3   1   23     1 Untreated
## 4   1   24     1 Untreated
## 5   1   24     1 Untreated
## 6   1   26     1 Untreated
```

As in the previous question, a new covariate can be added to the model and tested for significance:

$$Z_2(t) = Z_1 * g(t)$$

Below is the output of fitting a Cox model with this extra time dependent covariate:

```
bnct$Z2_time <- log(bnct$time) * bnct$trt
cox_fit <- coxph(Surv(time = time, event = death) ~ Treatment + Z2_time,
                 data = bnct)
summary(cox_fit)
```

```
## Call:
## coxph(formula = Surv(time = time, event = death) ~ Treatment +
##     Z2_time, data = bnct)
##
##   n= 30, number of events= 27
##
##                              coef  exp(coef)   se(coef)      z Pr(>|z|)
## TreatmentRadiated+BPA  9.761e+01  2.468e+42  1.975e+01  4.942 7.74e-07 ***
## TreatmentUntreated    -9.331e+01  3.005e-41  1.900e+01 -4.911 9.04e-07 ***
## Z2_time               -2.808e+01  6.406e-13  5.681e+00 -4.942 7.72e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                       exp(coef) exp(-coef) lower .95 upper .95
## TreatmentRadiated+BPA 2.468e+42  4.052e-43 3.797e+25 1.604e+59
## TreatmentUntreated    3.005e-41  3.328e+40 2.027e-57 4.455e-25
## Z2_time               6.406e-13  1.561e+12 9.353e-18 4.388e-08
##
## Concordance= 0.993  (se = 0.069 )
## Rsquare= 0.965   (max possible= 0.992 )
## Likelihood ratio test= 101  on 3 df,   p=0
## Wald test            = 25.27  on 3 df,   p=1.357e-05
## Score (logrank) test = 49.89  on 3 df,   p=8.418e-11
```

The p-value for the time dependent coefficient is very low which is evidence that the proportional hazards assumption is not met.

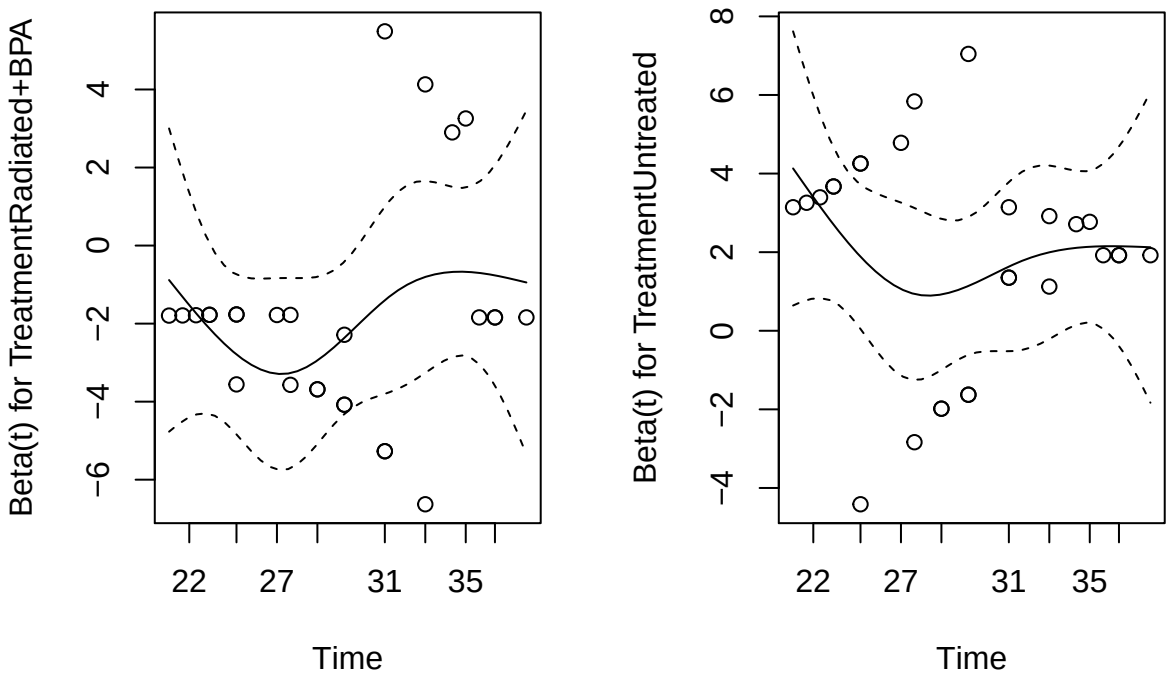
The built in R method can also be used to check this assumption:

```
cox_fit <- coxph(Surv(time = time, event = death) ~ Treatment, data = bnct)
cox.zph(cox_fit)
```

```
##                           rho chisq     p
## TreatmentRadiated+BPA  0.1968 0.746 0.388
## TreatmentUntreated    -0.0871 0.171 0.679
## GLOBAL                    NA 1.087 0.581
```

These p-values are consistent with the assumption that the proportional hazards assumption holds. The plot below is also used to check the assumption:

```
par(mfrow=c(1, 2))
plot(cox.zph(cox_fit))
```



There are two plots because the covariate *Treatment* has three levels. Horizontal lines are consistent with the assumption of proportional hazards. The lines in these two plots show some departure from being horizontal, which may be evidence that the assumption is not valid.