hw4

Nathan Hawkins

2/9/2022

The following problems will use the Ovarian Cancer data available on Learning Suite.

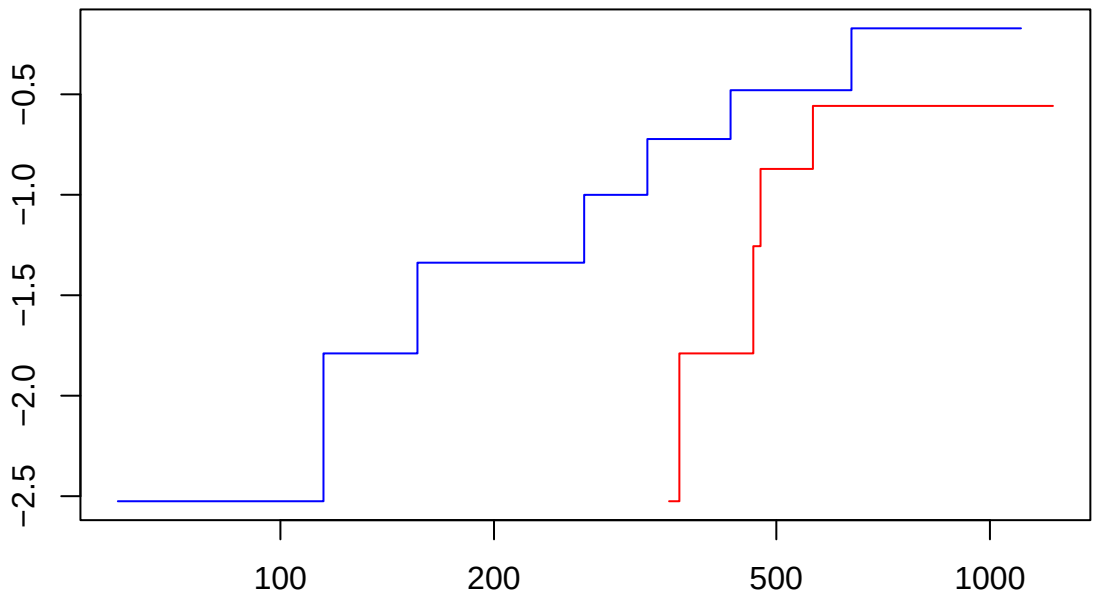
1. Create a complimentary log-log plot of the data comparing the 2 treatment groups. Does it appear that the proportional hazards assumption holds? What will the consequences be if we fit a Cox proportional hazards model to this data?

The lines don't look exactly proportional. We may have trouble with the estimates as our estimates are linear but the data don't appear to be. Though this may be due to a small sample size.

```
library(asaur)
library(survival)
ov = read.csv("ovarian.csv")
head(ov)
```

```
##   Patient time status treatment age
## 1       1  156      1         1  66
## 2       2 1040      0         1  38
## 3       3   59      1         1  72
## 4       4  421      0         2  53
## 5       5  329      1         1  43
## 6       6  769      0         2  59
```

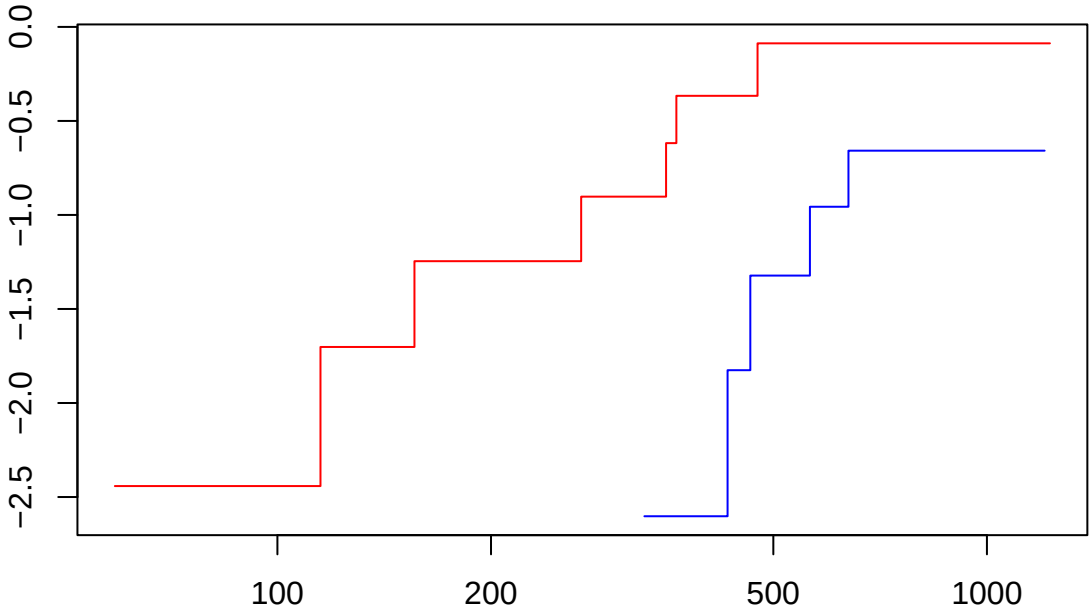
```
survfit(Surv(time, status)~treatment,
        data=ov) |>
  plot(fun='cloglog', col=c('blue','red'))
```



2. Create the same plot, but comparing patients up to age 56 to those over 56. Does proportional hazards hold in this case?

When using age over 56 as a variable, the hazards look more proportional than in the previous case.

```
ov$age_cat = ifelse(ov$age > 56, 1, 0)
survfit(Surv(time, status)~  age_cat,
        data=ov) |>
  plot(fun='cloglog', col=c('blue','red'))
```



3. Fit a Cox proportional hazards model to the Ovarian data using treatment and age as predictor variables. What are the hazard ratios of each predictor? Interpret the HRs in context including the confidence intervals.

Being in the treatment group implies that the subject's hazard is 0.45 (or 45%) of the hazard of the non treatment. With 95% confidence the hazard is between (0.13, 1.56). A person 1 year older in age implies a hazard 1.15 times greater than the younger person. With 95% confidence this hazard is between 1.06 and 1.27. Treatment is not significant in this model, age is significant at the 0.05 level.

```
fit1 <- coxph(Surv(time, status) ~ treatment + age,
              data=ov)

fit1
```

```
## Call:
## coxph(formula = Surv(time, status) ~ treatment + age, data = ov)
##
##               coef exp(coef) se(coef)      z       p
## treatment -0.79593   0.45116  0.63294 -1.258 0.20857
## age        0.14657   1.15786  0.04585  3.196 0.00139
##
## Likelihood ratio test=15.82  on 2 df, p=0.0003666
## n= 26, number of events= 12
```

```
exp(confint(fit1))
```

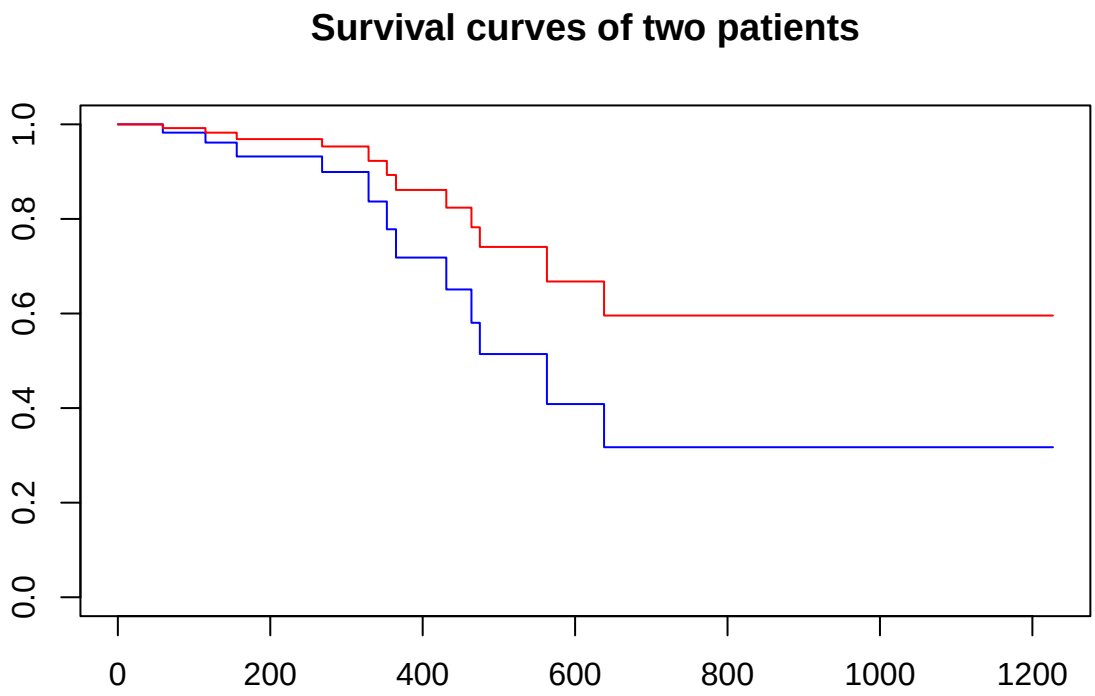
```
##              2.5 %   97.5 %
## treatment 0.130488 1.559881
## age       1.058336 1.266733
```

```
summary(fit1)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ treatment + age, data = ov)
##
##   n= 26, number of events= 12
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## treatment -0.79593   0.45116  0.63294 -1.258  0.20857
## age        0.14657   1.15786  0.04585  3.196  0.00139 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           exp(coef) exp(-coef) lower .95 upper .95
## treatment    0.4512     2.2165    0.1305     1.560
## age          1.1579     0.8637    1.0583     1.267
##
## Concordance= 0.794  (se = 0.079 )
## Likelihood ratio test= 15.82  on 2 df,   p=4e-04
## Wald test            = 13.55  on 2 df,   p=0.001
## Score (logrank) test = 18.61  on 2 df,   p=9e-05
```

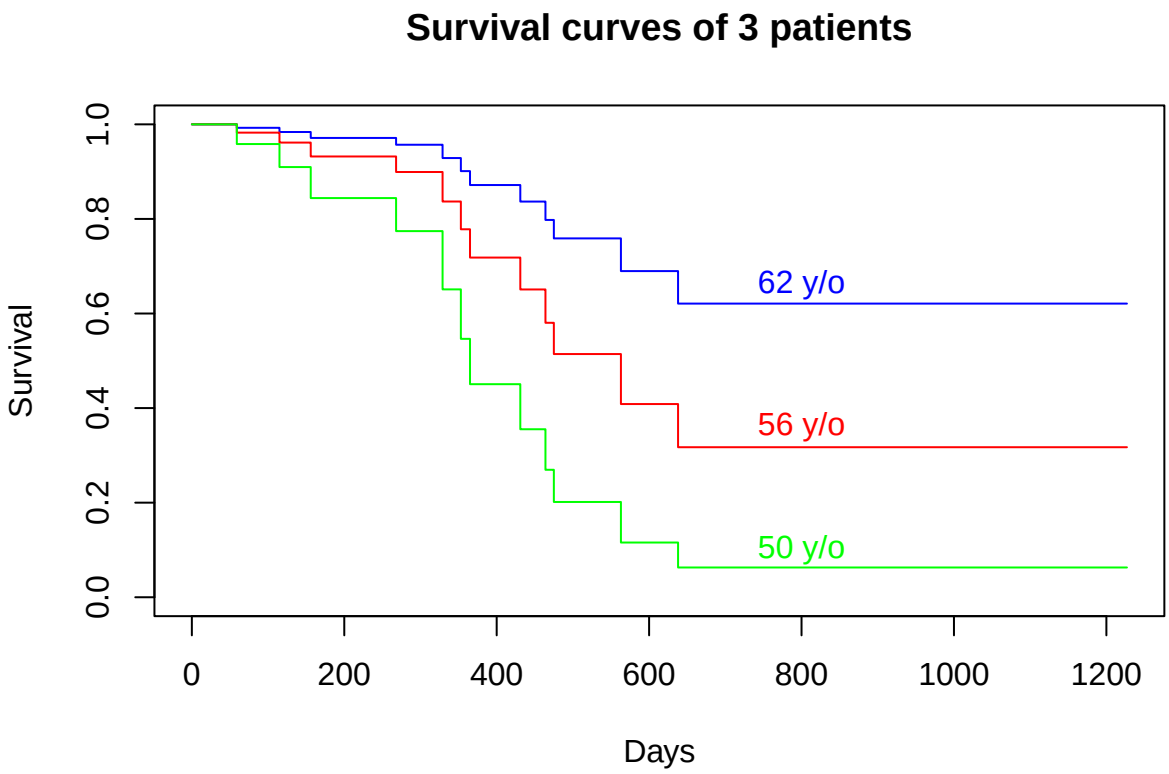
4. Create a plot of the predicted survival curves (based on the ph model) for 2 56 year old patients, one in the treatment group, one in the control group.

```
tmpdat <- data.frame(treatment=c(1, 2), age = c(56, 56))

plot(survfit(fit1, tmpdat), col=c('blue','red'), main = "Survival curves of two patients")
```



5. Create a plot of the predicted survival curves (based on the ph model) for 3 women in the control group, one each of the ages 50, 56, and 62. Label the curves, or include a legend to identify which line is which. Give a 1-2 sentence description of the comparison between the ages.

```
tmpdat <- data.frame(treatment=c(1, 1, 1), age = c(50, 56, 62))

plot(survfit(fit1, tmpdat), col=c('blue','red', 'green'),
     main = "Survival curves of 3 patients",
     xlab = "Days",
     ylab = "Survival")
text(800, 0.66, "62 y/o", col = "blue")
text(800, 0.36, "56 y/o", col = "red")
text(800, 0.1, "50 y/o", col = "green")
```


Survival curves of 3 patients
Survival
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000
1200
Days
62 y/o
56 y/o
50 y/o