STAT 2261 Final

Orly Olbum

4/15/2021

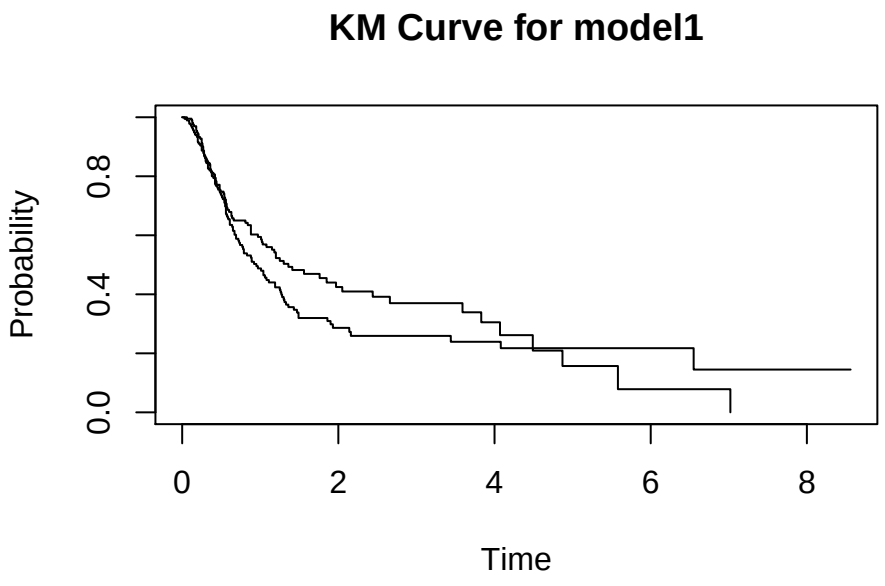
Problem 1

Examine thrombotic-event-free survival curves for those who received high dose of Heparin and for those who were randomized to the low dose group and perform a two-sample test to compare the efficacy of using a high dose versus a low dose in preventing major thrombotic events or death.

```
model1 = survdiff(Surv(time, event) ~ as.factor(trt), data = covid)
model1
```

```
## Call:
## survdiff(formula = Surv(time, event) ~ as.factor(trt), data = covid)
##
##                     N Observed Expected (O-E)^2/E (O-E)^2/V
## as.factor(trt)=0 211      113    103.7     0.831      1.73
## as.factor(trt)=1 189       89     98.3     0.877      1.73
##
##  Chisq= 1.7  on 1 degrees of freedom, p= 0.2
```

```
plot(survfit(Surv(time, event) ~ as.factor(trt), data = covid), ylab = "Probability",
     xlab = "Time", main = "KM Curve for model1")
```



The log rank tests a null hypothesis that there is no difference in survival curves between the groups specified (here, high dose vs. low dose) against the alternative hypothesis that there is evidence of some difference in our sample. The test statistic provides a p-value of 0.2, which is above our threshold for alpha = 0.05, so we fail to reject the null hypothesis and conclude that we don't have evidence to suggest there is a difference in using high vs. low dose in preventing major thrombotic events or death.

Problem 2

Fit a Cox proportional hazard model with treatment, two dummy variables for D-dimer with the low level as the reference, severity and age as predictors. After adjusting for other risk factors, does the high dose of Heparin seem to help prevent major thrombotic events or death?

```
model2 = coxph(Surv(time, event) ~ as.factor(trt) + as.factor(Ddimer) +
                 as.factor(severity) + age, data = covid)
model2
```

```
## Call:
## coxph(formula = Surv(time, event) ~ as.factor(trt) + as.factor(Ddimer) +
##     as.factor(severity) + age, data = covid)
##
##                          coef exp(coef) se(coef)      z        p
## as.factor(trt)1      -0.06720   0.93501  0.14481 -0.464   0.6426
## as.factor(Ddimer)1    0.10595   1.11177  0.18949  0.559   0.5761
## as.factor(Ddimer)2    0.51852   1.67954  0.21514  2.410   0.0159
## as.factor(severity)1  0.84578   2.32978  0.14708  5.751 8.89e-09
## age                   0.34318   1.40943  0.07842  4.376 1.21e-05
##
## Likelihood ratio test=55.57  on 5 df, p=9.965e-11
## n= 400, number of events= 202
```

```
model2.2 = coxph(Surv(time, event) ~ as.factor(Ddimer) + as.factor(severity) +
                   age, data = covid)
model2.2
```

```
## Call:
## coxph(formula = Surv(time, event) ~ as.factor(Ddimer) + as.factor(severity) +
##     age, data = covid)
##
##                         coef exp(coef) se(coef)     z        p
## as.factor(Ddimer)1   0.1096    1.1159   0.1893 0.579    0.562
## as.factor(Ddimer)2   0.5226    1.6865   0.2149 2.432    0.015
## as.factor(severity)1 0.8576    2.3576   0.1450 5.915 3.32e-09
## age                  0.3459    1.4132   0.0782 4.423 9.74e-06
##
## Likelihood ratio test=55.35  on 4 df, p=2.739e-11
## n= 400, number of events= 202
```

After fitting all variables in the Cox model, the high does of Heparin does not seem to help prevent major thrombotic events/death, in the full model it is not a significant predictor. We can also use the likelihood ratio test to see if adding or keeping out treatment (Heparin dose) in the model has any effect. We get a LR test stat of 55.57 - 55.35 = 0.22 on 1 df, which follows chi-sq distribution under the null hypothesis, and does not produce a statistically significant p-value, which says that keeping it in the model does not change probability of thrombotic event.

Problem 3

There is some concern that those who need the ICU level care may benefit differently from Heparin than those who do not need the ICU level care. Test whether this is true by adding a severity and treatment interaction to your model.

```
model3 = coxph(Surv(time, event) ~ as.factor(trt) + as.factor(Ddimer) +
                 as.factor(severity) + age + as.factor(trt)*as.factor(severity),
               data = covid)
model3
```

```
## Call:
## coxph(formula = Surv(time, event) ~ as.factor(trt) + as.factor(Ddimer) +
##     as.factor(severity) + age + as.factor(trt) * as.factor(severity),
##     data = covid)
##
##                                        coef exp(coef) se(coef)      z
## as.factor(trt)1                     -0.64185   0.52632  0.20941 -3.065
## as.factor(Ddimer)1                   0.11750   1.12468  0.18983  0.619
## as.factor(Ddimer)2                   0.64554   1.90701  0.21811  2.960
## as.factor(severity)1                 0.35541   1.42676  0.19212  1.850
## age                                  0.36661   1.44283  0.08015  4.574
## as.factor(trt)1:as.factor(severity)1 1.14199   3.13299  0.29328  3.894
##                                         p
## as.factor(trt)1                      0.00218
## as.factor(Ddimer)1                   0.53594
## as.factor(Ddimer)2                   0.00308
## as.factor(severity)1                 0.06433
## age                                 4.79e-06
## as.factor(trt)1:as.factor(severity)1 9.86e-05
##
## Likelihood ratio test=70.9  on 6 df, p=2.666e-13
## n= 400, number of events= 202
```

The interaction term for trt*severity is a significant predictor in the model. We can also justify this with a LR test, with a test statistic of 70.9 - 55.57 = 15.33 on 1 df, which gives a p-value of < 0.05. It appears that those who need ICU level care do benefit differently from Heparin than those who do not need ICU level care.

Problem 4

Can the variable Ddimer be included in the Cox model as a continuous variable? Justify your answer.

If we include a factor variable as continuous, the model will interpret the increase in unit (i.e., from 0 to 1, from 1 to 2) as increase in value, rather than a change of category for the variable. Because of this, we treat it as two dummy variables (or, for coding, use "as.factor" when including in the model). Additionally, we cannot have Ddimer levels between the numbers allotted. 0, 1, and 2 are designated to refer to categorical levels, and we cannot have a level 0.4 or 1.7 (an example of this is below - if we fit Ddimer as a continuous variable - however, it does not have an accurate interpretation). Thus, we must treat Ddimer and other variables like it as factor (categorical) variables.

```
model4 = coxph(Surv(time, event) ~ as.factor(trt) + Ddimer + as.factor(severity) +
                 age + as.factor(trt)*as.factor(severity), data = covid)
model4
```

```
## Call:
## coxph(formula = Surv(time, event) ~ as.factor(trt) + Ddimer +
##     as.factor(severity) + age + as.factor(trt) * as.factor(severity),
##     data = covid)
##
##                                         coef exp(coef) se(coef)      z
## as.factor(trt)1                      -0.61823   0.53890  0.20866 -2.963
## Ddimer                                0.34202   1.40779  0.11425  2.994
## as.factor(severity)1                  0.37376   1.45319  0.19183  1.948
## age                                   0.35689   1.42887  0.07964  4.481
## as.factor(trt)1:as.factor(severity)1  1.10221   3.01081  0.29143  3.782
##                                           p
## as.factor(trt)1                      0.003048
## Ddimer                               0.002756
## as.factor(severity)1                 0.051366
## age                                  7.42e-06
## as.factor(trt)1:as.factor(severity)1 0.000156
##
## Likelihood ratio test=68.88  on 5 df, p=1.752e-13
## n= 400, number of events= 202
```

Problem 5

Based on your final model discuss how each risk factor in your model is associated with the composite outcome.

Using model3 which includes the interaction term between treatment and severity, the risk factors are associated with the outcome in the following ways:

- The relative risk for major thrombotic event/death for patients taking high doses of Heparin compared to low doses is exp(-0.642) = 0.526, and is a statistically significant predictor in the model.
- The relative risk for major thrombotic event/death for patients of medium Ddimer level compared to low Ddimer level is 1.125, and is not a statistically significant predictor in the model.
- The relative risk for major thrombotic event/death for patients of high Ddimer level compared to low Ddimer level is 1.907, and is a statistically significant predictor in the model.
- The relative risk for major thrombotic event/death for patients who need ICU level of care compared to patients who do not need ICU level of care is 1.427 and is not a statistically significant predictor in the model.
- The relative risk for major thrombotic event/death for patients as they get older (standardized) is 1.443 for every unit increase in age and is a statistically significant predictor in the model.
- The interaction between treatment (dose of Heparin) and severity (need of ICU care vs. not) is a statistically significant predictor in the model, and indicates that the relative risk for thrombotic event/death for patients who both are treated with a high dose of Heparin and who need ICU level of care is 3.133 compared to patients who either have gotten a low dose of Heparin, require non-ICU level of care, or both.