Answer Key: Problem Set 4

Applied Stats II

Jeffrey Ziegler

Instructions

- *Please show your work! You may lose points by simply writing in the answer. If the problem requires you to execute commands in* `R`*, please include the code you used to get your answers. Please also include the* `.R` *file that contains your code. If you are not sure if work needs to be shown for a particular problem, please ask.*
- *Your homework should be submitted electronically on GitHub in* `.pdf` *form.*
- *This problem set is due before class on Monday April 4, 2022. No late assignments will be accepted.*
- *Total available points for this homework is 80.*

Question 1

We're interested in modeling the historical causes of infant mortality. We have data from 5641 first-born in seven Swedish parishes 1820-1895. Using the "infants" dataset in the `eha` *library, fit a Cox Proportional Hazard model using mother's age and infant's gender as covariates. Present and interpret the output.*

First, let's estimate our Cox proportional hazard model:

```
# load data on infant mortality by mother's background
data("infants")

# estimate duration Cox PH model that includes both predictors (child, mother)
infantMorality <- coxph(Surv(enter, exit, event) ~ age + sex, data = infants)
```

Looking at Table 1, we can interpret the estimated coefficient for gender as the logged hazard ratio of boys with respect to girls for infants with mothers of the same age. If we exponentiate the estimated coefficient of gender (exp(coef)= 0.616), we see that boys are about 40% more likely to survive than girls with mothers of the same age (HR $<$ 1 = reduction in the hazard). Looking at the age coefficient, we can see that having an older mother (holding infant gender

Table 1: Cox Proportional Hazard model results

age	−0.04
	(0.05)
sexboy	−0.49
	(0.44)
AIC	171.25
# events	21
N	105

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

constant) is also a "good prognostic factor" that reduces the risk of death, though this is not a statistically non-zero effect.

To visually investigate the relationship between gender and mothers' age further, we can see that in Figure 1, children generally survive longer when they have older mothers (which may be symptomatic of access to healthcare and economic class). Moreover, we can see that boys generally survive longer than girls (most likely due to selective gender preferences of children in the 19th century).

Figure 1: Predicted hazard ratio by gender and mother age.