Problem Set 4

Applied Stats II

Due: April 12, 2024

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 23:59 on Friday April 12, 2024. No late assignments will be accepted.

Question 1

We're interested in modeling the historical causes of child mortality. We have data from 26855 children born in Skellefteå, Sweden from 1850 to 1884. Using the "child" 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.

```
1 # a) Using the Surv() function to build a survival object
2 child_surv <- with(child, Surv(enter, exit, event))
3
4
5 # b) Run a Cox Proportional Hazard regression
6 cox <- coxph(child_surv ~ m.age + sex, data = child)
7
8 summary(cox)
```

- Concordance = 0.519 (SE = 0.004)

- Likelihood ratio test = 22.52 (df = 2, p = 0.00001)

- Wald test = 22.52 (df = 2, p = 0.00001)

Table 1: Cox Proportional Hazard Model Results

Variable	Coefficient	Exp(coef)	Standard Error
m.age	0.007617	1.007646	0.002128
sexfemale	-0.082215	0.921074	0.026743

	z	Pr(>|z|)	95% CI Lower	95% CI Upper
m.age	3.580	0.000344	1.003	1.0119
sexfemale	-3.074	0.002110	0.874	0.9706

- Score (logrank) test = 22.53 (df = 2, p = 0.00001)

There is a -0.082215 decrease in the expected log of the hazard for female babies compared to male, holding mother's age constant. There is a 0.007617 increase in the expected log of the hazard for one increase in mother's age, holding sex constant.