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.

CODE:

```
data(child)
cox_model <- coxph(Surv(enter, exit, event) ~ sex + m.age, data = child)
summary(child_mortality_cox)
```

INTERPRETATION:

Holding mothers age constant a change from child sex of female to male is associated with an estimated decrease in the log hazard ratio by 0.082215. This value is highley statistically significant.

exp(-0.082215) = 0.9210739
0.9210739 - 1 = -0.07892609
Girls are about 7.8 percent more likley to survivie than boys.

Holding child gender constant a one year increase in mothers age is associated with an increasein the log hazard ratio of 0.007617. This value is highley statistically significant. In plane english older mothers tend to have a increased chance of child mortality.

exp(0.007617) = 1.007646
exp(0.007617) - 1 = 0.007646083
For every one-year increase in a mother's age, there is an estimated 0.007 percent increase in the risk of child mortality, when the sex of the child is held constant. However, I highley doubt that this relationship is liner. I would suspect a change from age 20 to 21 would not be the same as age 35-36. Reguardless, as age increases we see that the child mortality rate tends to increase.
Thanks for a great two semsters!

Table 1:

	Dependent variable:
	enter
sexfemale	-0.082^{***}
	(0.027)
m.age	0.008^{***}
	(0.002)
Observations	26,574
R^2	0.001
Max. Possible R^2	0.986
Log Likelihood	−56,503.480
Wald Test	22.520^{***} (df = 2)
LR Test	22.518^{***} (df = 2)
Score (Logrank) Test	22.530^{***} (df = 2)
Note:	*p<0.1; **p<0.05; ***p<0.01