Problem Set 4

Idi Amin Da Silva - Student Number: 23372225, 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.

```
table(child$sex)
#  male     female
# 13676     12898
###
# Fit Cox Proportional Hazard model
##
#The 'child' dataset from the 'eha' package is a dataset of 26,574 children born in
# Skellefte?, Sweden, 1850-1884. Children are followed for fifteen years or until death or
# outmigration.
####################
# I am interested in  Run a Cox Proportional Hazard regression on the data, using an additive model with
##'m.age' and 'sex' as explanatory variables. Also I will Run a test to assess the quality of the
```

```r
## model. How can I interpret the coefficients?  and Plot the model.
##

# Fit Cox Proportional Hazard model
cox_model <- coxph(Surv(enter, exit, event) ~ m.age + sex, data = child)

summary(cox_model) ## Print the summary of the model
## Answer : Table Output:
```

Table 1: Cox Proportional Hazard Model Summary

	coef	exp(coef)	se(coef)	z	Pr(>—z—)
m.age	0.007617	1.007646	0.002128	3.580	0.000344
sexfemale	-0.082215	0.921074	0.026743	-3.074	0.002110

	exp(coef)	exp(-coef)	lower .95	upper .95
m.age	1.0076	0.9924	1.003	1.0119
sexfemale	0.9211	1.0857	0.874	0.9706

	Value
Concordance	0.519
Likelihood ratio test	22.52
Wald test	22.52
Score (logrank) test	22.53

With every additional increase in mother's age i.e., increases of one unit of age for mother is associated with a increase of 0.007617 in the expected log of the hazard. There is a 0.082215 decrease in the expected log of the hazard for female babies.

in other words I can interpret this outcome as the following: Interpretation: For each unit increase in mother's age, the hazard (risk) of the event increases by a factor of approximately 1.0076, holding all other variables constant. This effect is statistically significant at the 0.05 level (indicated by ***).

Interpretation: Being female (compared to male) decreases the hazard (risk) of the event by a factor of approximately 0.9211, holding all other variables constant. This effect is statistically significant at the 0.05 level (indicated by **)

stargazer($\cos_m odel$)
The code above stargazer($\cos_m odel$)$Willproducethefollowingtableoutcome$:

Table 2:

	Dependent variable:
	enter
m.age	0.008^{***}
	(0.002)
sexfemale	-0.082^{***}
	(0.027)
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

m.age: The coefficient is 0.008 with a standard error of 0.002. This suggests that for each unit increase in mother's age, the entry time increases by 0.008 units. The coefficient is statistically significant at the 1sexfemale: The coefficient is -0.082 with a standard error of 0.027. This suggests that being female (compared to male) decreases the entry time by 0.082 units. The coefficient is statistically significant at the 1