

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.

Answer Question 1

```
1 # Installing eha package
2 if (!requireNamespace("eha", quietly = TRUE)) {
3   install.packages("eha")
4 }
5
6 # Load the eha package
7 library(eha)
8
9 # Load extra survival package, required for survival analysis functions
10 library(survival)
11
12 # Load the dataset
13 data(child)
```

```

14
15 # Inspect the structure of the dataset to make variables names visible
16 str(child)
17
18 # Fit the Cox Proportional Hazards model
19 cox_model <- coxph(Surv(exit, event) ~ m.age + sex, data = child)
20
21 # Print the summary of the Cox model
22 summary(cox_model)

```

Results of Cox Proportional Hazards Model

The Cox Proportional Hazards model was fitted using mother's age and infant's gender as covariates. The dataset consisted of 26,574 observations with 5,616 events. Below are the detailed results of the model:

Variable	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 **

Table 1: Coefficients and hazard ratios from the Cox model.

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Variable	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

Table 2: Hazard ratios and 95% confidence intervals for Cox model covariates.

Model Fit Statistics:

- Concordance: 0.519 (SE = 0.004)
- Likelihood ratio test: 22.52 on 2 df, p=1e-05
- Wald test: 22.52 on 2 df, p=1e-05
- Score (logrank) test: 22.53 on 2 df, p=1e-05