Problem Set 4

Applied Stats II

Due: April 16, 2023

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 Sunday April 16, 2023. No late assignments will be accepted.

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.

```
install.packages("eha")
install.packages("survival")
library("survival")
library(eha)
library("stargazer")
data("infants")
infants

fit Cox Proportional Hazard model
mod <- coxph(Surv(enter, exit, event) ~ age + sex, data = infants)
summary(mod)

mod2 <- coxph(Surv(enter, exit, event) ~ sex, data = infants)
```

```
sub_infants <- with(infants,
data.frame(
sex = c("girl", "boy")))

plot survival proportion by sex
plot(survfit(mod2, newdata = sub_infants),
conf.int = F,
ylim = c(0.2, 1),
col = c("red", "blue"),
xlab = "Time (Days)",
ylab = "Survival Proportion",
main = "Survival Plot by Sex")
legend("bottomleft",
legend=c("Girl", "Boy"),
lty = 1,
col = c("red", "blue"),
text.col = c("red", "blue"))

exponentiate estimates
exp(-0.04)
exp(-0.485)
```

Although the results are not statistically different from zero, interpreting the coefficients for boys indicates a decrease of 0.485 in the log likelihood of infant mortality compared to girls, holding the age of the mother constant. Taking the exponent, the hazard ratio for boys is 0.61 that of girls, so 61 boys die in infancy for every hundred girls. Age, again not significant, but a one unit increase in the age of the mother is associated with a decrease of 0.04 in the log likelihood of infant mortality, holding the sex of the child constant.