Problem Set 4

Applied Stats II

Due: April 4, 2022

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 class on Monday April 4, 2022. No late assignments will be accepted.
- Total available points for this homework is 80.

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.

Installing packages
install.packages("eha")
install.packages("stargazer")
library(stargazer)
library(eha)
library(survival)
library(ggplot2)
library(tidyverse)
data("infants")

summary(infants)

fitting a cox model with infants gender and mothers age as covariates
add-mor ¡- coxph(Surv(enter, exit, event) sex + age, data = infants)
summary(add-mor)

There is a 0.49 decrease in the expected log
of the hazard of male infants compared to females,
holding mother's age constant.
There is a 0.04 decrease in the expected log
of the hazard for infants of difference
in mother's age, holding the infants sex constant.

Gender hazard ratio
exp(-0.48518)
0.6155864
The hazard ratio of male infants
is 0.62 that of female infants,
ie. male infants are less likely to die
(62 male infants die for every 100 female infants;
male deaths are 38 percent lower)
where mother's age is equal across infants.

At a given instant in time,
male infants are 0.616 times as likely to die
as females within a 95 percent confidence interval
(0.2587, 1.465), adjusting for mother's age.

From likelihood ratio test,
we can see that the model
is not statistically significant
since pvalue = 0.4.
pvalue needs to be less than 0.0.5 to be significant

Neither sex nor age are significant to the model, since pvalues ¿ 0.05.
drop1(add-mor, test = "Chisq")

Summary of finding above, no new information. Tidies the output.
stargazer(add-mor, type = "text")

Adding an interaction
cox.int ¡- coxph(Surv(enter, exit, event) sex * age, data = infants)
summary(cox.int)

Read for 3 'coef' values in output
There is a 1.89 increase in the expected
log of the hazard of male infants compared to females,
holding mother's age constant.
There is a -0.001519 decrease in the expected log
of the hazard for infants
of difference in mother's age,
holding the infants sex constant.
there is a -0.09128 decrease in the expected log
of the hazard of male infants
when both the gender and the mothers
age are taken into account
Model not sig, since p¿0.05.

drop1(cox.int, test = "Chisq")
Interaction term not sig since p greater than 0.05.

stargazer(cox.int, type = "text")

plot of the cumulative hazard function
plot-CoxPH ¡- coxreg(Surv(enter, exit, event) sex + age, data = infants)
plot(plot-CoxPH)

Plotting model
cox-fit ¡- survfit(add-mor)
autoplot(cox-fit)

newdat ¡- with(infants, data.frame(sex = c("boy", "girl"), age ="26"))
newdat

All of the code above allows us to produce figure 1 plot

plot(survfit(add-mor, newdata = newdat), xscale = 12,
conf.int = T,
ylim = c(0.6, 1),
col = c("red", "blue"),
xlab = "Time",
ylab = "Survival proportion",
main = "")
legend("bottomleft",
legend=c("boy", "girl"),
lty = 1,
col = c("red", "blue")
, text.col = c("red", "blue"))

running this code you get the following plots:
if you include the following code instead
of running the original packages it allows

autoplot to work and you get the autoplot
figure 2 that is shown below

rm(list=ls())
detach all libraries
detachAllPackages ¡- function()
basic.packages ¡- c("package:stats"
"package:graphics", "package:grDevices",
"package:utils", "package:datasets",
"package:methods", "package:base")
package.list ¡-
search()[ifelse(unlist(gregexpr("package:",
search()))==1, TRUE, FALSE)]
package.list ¡- setdiff(package.list, basic.packages)
if (length(package.list)¿0) for (package in package.list) detach(package, character.only=TRUE)

detachAllPackages()

load libraries

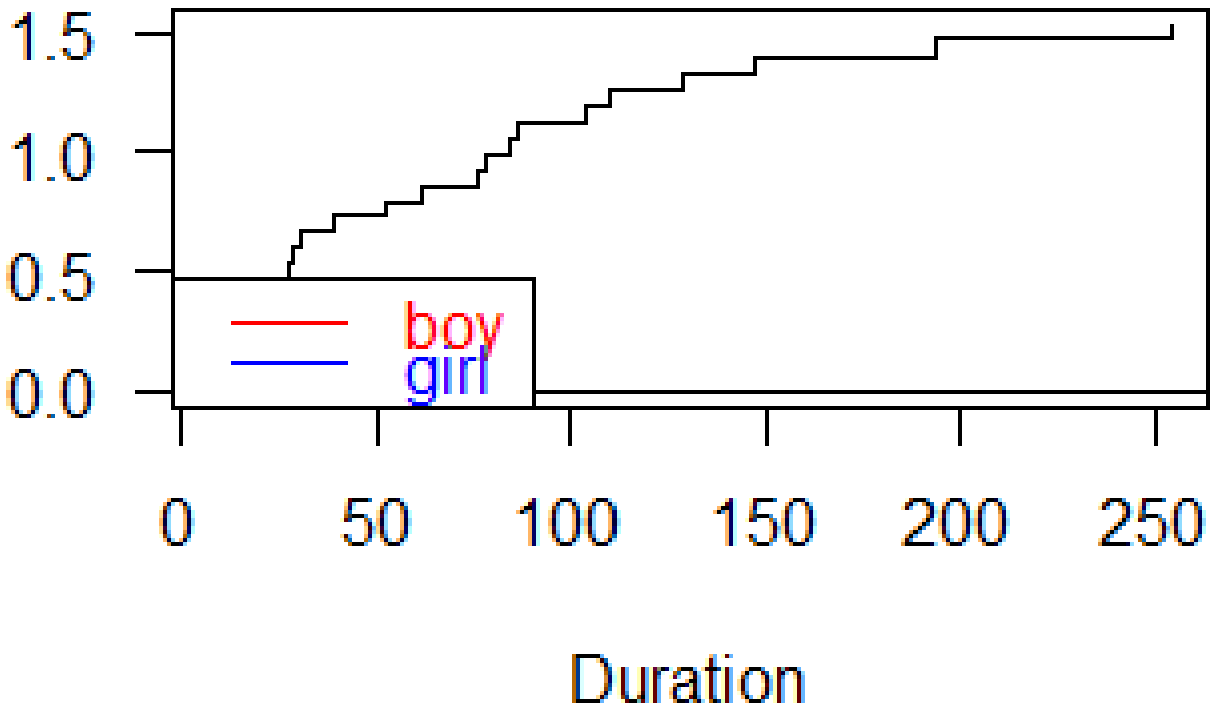

Figure 1: Caption

```
pkgTest ¡- function(pkg)
new.pkg ¡- pkg[!(pkg (dollar)in(dollar) installed.packages()[, "Package"])]
if (length(new.pkg))
install.packages(new.pkg, dependencies = TRUE)
sapply(pkg, require, character.only = TRUE)


here is where you load any necessary packages
ex: stringr
lapply(c("stringr"), pkgTest)

lapply(c("survival", "eha", "tidyverse", "ggfortify", "stargazer"), pkgTest)
```

1 Introduction

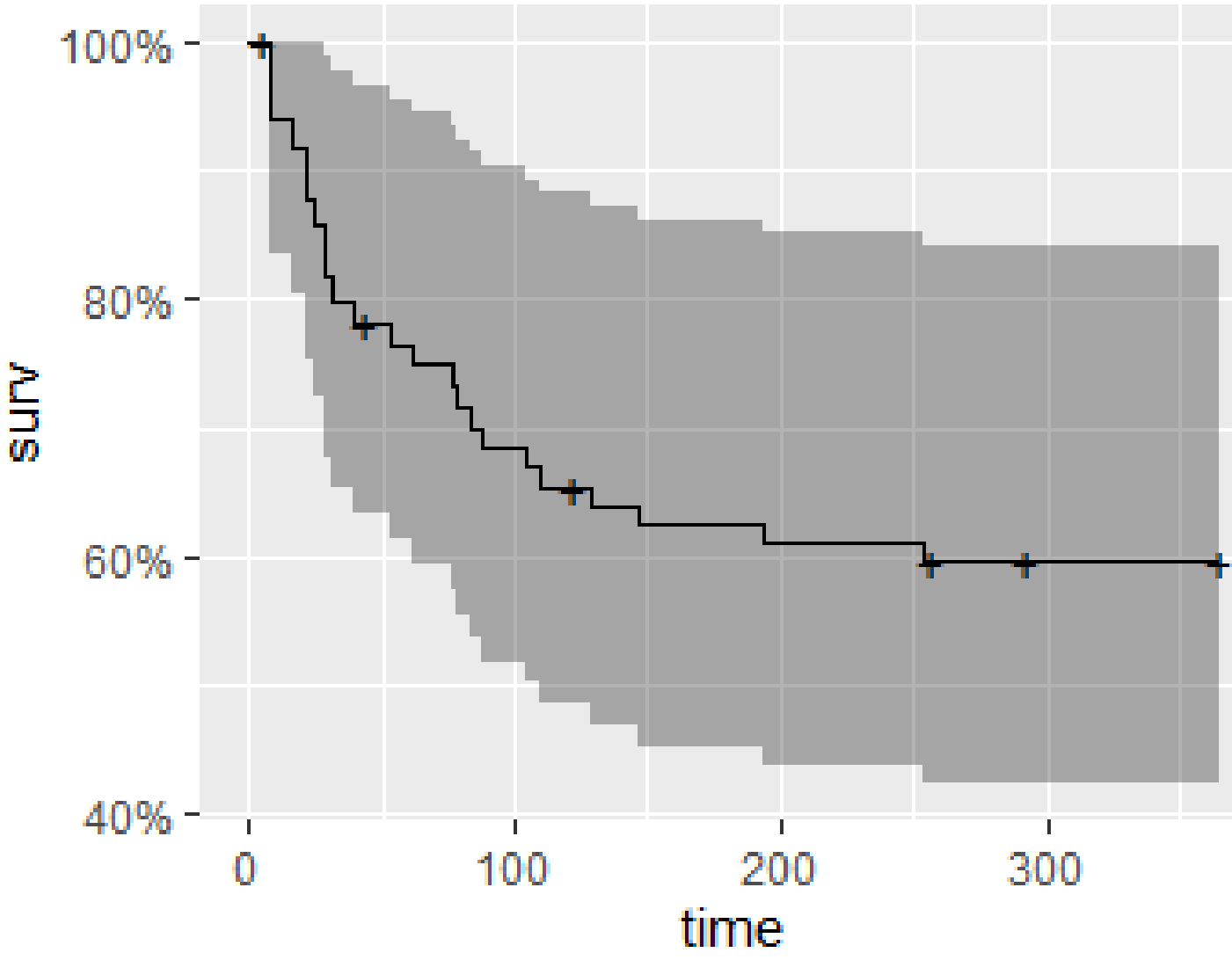

Figure 2: Caption