HW6

Code Appendix

2023-11-13

```
data(pbc, package="survival")
data = pbc |>
  mutate(time = time/30,
         status = ifelse(status != 0, 1, 0),
         edema = ifelse(edema == 0.5, 1, edema),
         sexn = case_when(sex == "m" ~ 1,sex == "f" ~ 0)) |>
  drop_na()
s = coxph(Surv(time, status) ~ sexn + edema + bili + albumin + copper + ast + protime + stage, data)
base = data.frame(sexn = 0,
                  edema = 0,
                  bili = 0,
                  albumin = 0,
                  copper = 0,
                  ast = 0,
                  protime = 0,
                  stage = 1)
survival = basehaz(s, newdata = base) |>
  mutate(survival = exp(-hazard)) |>
  filter(survival >= 0.8976 & survival <= 0.9005) |>
  knitr::kable()
```

```
b = data.frame(sexn = 1,
               edema = 1,
               bili = median(data$bili),
               albumin = median(data$albumin),
               copper = median(data$copper),
               ast = median(data$ast),
               protime = median(data$protime),
               stage = 4)
survival_2 = basehaz(s, newdata = b) |>
  mutate(survival = exp(-hazard)) |>
  filter(survival >= 0.895 & survival <= 0.9005) |>
  knitr::kable()
```

0.1

1. Potential confounding factors for the observed dose-response in efficacy:

- First, identify potential confounding factors by: imbalance baseline characteristics between two doses with a 1-side significant level of 0.05.
- Then, identify prognostic factors by: test each factor individually by controlling the dose and keep significant factors as the confounding factors.
- Lastly, control all the selected confounding factors.

2. If there were any dose by baseline interactions. Use Cox model to compare the survival function of different (independent) groups:

- Use a set of K covariates to represent the K+1 groups, select a reference group where k=0;
- Fit Cox model, and use likelihood ratio test for the results.

0.2

0.2.1

```
survival
```

hazard	time	survival
0.1048102	76.26667	0.9004954
0.1048102	76.46667	0.9004954
0.1080196	76.56667	0.8976100
0.1080196	76.70000	0.8976100
0.1080196	77.26667	0.8976100
0.1080196	77.66667	0.8976100
0.1080196	77.73333	0.8976100

The survival time that 90% subjects survived for baseline survival function is 76.6 month.

0.2.2

```
survival_2
```

hazard	time	survival
0.1050784	17.76667	0.9002540
0.1108590	18.30000	0.8950649

The survival time for male subjects who had edema and stage=4 and taking median for all other covariates in the model is 18.3 month.

0.3

After backward selection, the variables I chose are:
sex, edema, bili, albumin, copper, ast, protime, stage.

- Here's the code for SAS:

```
libname mylib "/home/u61890905/EPG1V2";
proc import datafile="/home/u61890905/pbc.xlsx"
out=pbc
dbms=xlsx replace;
run;
proc phreg data=pbc;
class trt sex;
model time*status(0)=trt sex
ascites hepato spiders edema
bili chol albumin copper
alk.phos ast trig platelet
protime stage
/selection=backward
slentry=0.25
slstay=0.15
details;
run;
```