Hw7

Zhuodiao Kuang

2023-11-10

Problem 1

Control the age at transplant for those subjects who received the transplant in the heart transplant data set.

The age should be the age at baseline before transplant. The age at the time of transplant should be calculated using baseline_age+wait_time. **Clarification:** The age variable in the heart transplant dataset is the age when the subject enrolled to the study. The risk of death is likely to be associated with age. The treatment benefit of heart transplant, if any, may also depend upon the age at transplant.

Problem 1 in Homework 7 asks to understand the transplant's effect when controlling the age at transplant in the time-varying covariate model. In the Cox model, age should be the age at enrollment of the study before transplant. **If subjects received transplants, the age should be changed to the age at the time of transplant.**)

Solution 1

Import and clean the data.

Here I did some data manipulation to get the wait time for each patient. I noticed that the age was centered by 48, therefore I recovered it to it original scale. For age at transplant, for those who received transplantation, I use the original scale of the baseline age plus the wait time divided by 365.25, that is, I assume the wait time was in days; For those who didn't received transplantation, the age at transplant is the same to the baseline age.

```r
heart_dat = readxl::read_excel("Datasets.xlsx", sheet = "Heartdata")
heart_dat$WAITING_TIME_FOR_TRANSPLANT =
  ifelse(is.na(
    heart_dat$WAITING_TIME_FOR_TRANSPLANT), 0,
    heart_dat$WAITING_TIME_FOR_TRANSPLANT)
heart_dat$AGE_T = heart_dat$AGE + ((heart_dat$WAITING_TIME_FOR_TRANSPLANT)/365.25)

heart_dat$TRANSPLANT_STATUS = as.factor(heart_dat$TRANSPLANT_STATUS)
heart_dat$SURVIVAL_STATUS = as.numeric(heart_dat$SURVIVAL_STATUS)
heart_dat$SURVIVAL_TIME = as.numeric(heart_dat$SURVIVAL_TIME)
heart_dat$PATIENT_ID = as.integer(heart_dat$PATIENT_ID)
```

Two covariates, the indicator of transplantation and the time-varying variable `AGE_T` are included when building the cox model:

```r
heart.fit = coxph(Surv(SURVIVAL_TIME, SURVIVAL_STATUS==1) ~
                    TRANSPLANT_STATUS + AGE_T, data = heart_dat)
heart.sum = summary(heart.fit)
rownames(heart.sum$coefficients) = c("Transplant Status = 1", "Age(t)")
heart.sum$coefficients |>
  kable("latex",
        digits = 4,
        escape = F,
        booktabs = T,
        caption = "Regresion Coefficients Estimates of the Time-Varying Cox Model") |>
  kable_styling(position = "center", latex_options = "hold_position")
```

Table 1: Regresion Coefficients Estimates of the Time-Varying Cox Model

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
Transplant Status = 1	-1.7947	0.1662	0.2717	-6.6048	0e+00
Age(t)	0.0599	1.0617	0.0153	3.9223	1e-04

From the regression coefficients result above, we can see that when controlling the age at transplant, the transplant's effect shows significant, and it significantly lower the risk of event, the risk reduction is about 83.4%.

Problem 2

Check proportional hazard assumption between the two sex groups in the PBC dataset. Provide the results of the checking.

a.Using log{-logS(t,Z)}

b.Plot the observed and fitted

c.Including a couple of continuous variables and check the interaction with time

d.Plot the Schoenfeld residual of the fitted model with two continuous covariates in problem 2.c.