P8110 Homework 5

Nick Williams

October 9, 2018

```
data hw5;
    infile "C:\Users\niwi8\OneDrive\Documents\fall_2018\regression\homework\p8110_hw5\HW5data.csv"
        delimiter= ',' missover dsd;
    input id len_follow final_stat mi_ord bmi year age_c;
run;
```

Problem 1

```
proc phreg data = hw5;
    class mi_ord (ref = "0") / param = ref;
    model len_follow * final_stat(0) = mi_ord / risklimits covb ties = efron;
    title "Cox model, length of follow-up as a function of MI order";
run;
```

$$h(t, x, \beta) = h_0(t)\exp(\beta_1(x_1)) \quad \text{where, } x_1 = \begin{cases} 0 \text{ if first MI} \\ 1 \text{ if recurrent MI} \end{cases}$$

1. $H_0 : \beta_1 = 0$
   $H_A : \beta_1 \neq 0$
2. Using a likelihood ratio test:
   $$G = 9.1431 \sim \chi^2_1 \text{ under the null}$$
   $$p = P(\chi^2_1 \geq 9.1431) = 0.0025$$
3. $0.0025 < 0.05 \rightarrow$ reject the null hypothesis
4. There is sufficient evidence to conclude that, at the 5% significance level, the rate of death is different between patients with first time MI and patients with recurrent MI. The risk of death among patients with recurrent MI is 1.53 times the risk of death among patients experiencing MI for the first time.

Problem 2

```
proc phreg data = hw5;
    class mi_ord (ref = "0")
          age_c (ref = "1")
          year (ref = "1") / param = ref;
    model len_follow * final_stat(0) = mi_ord age_c bmi year
          / risklimits covb ties = efron;
    title "Cox model, length of follow-up as a function of MI order,
           age category, bmi, and cohort year";
run;
```

$$h(t, x_i, \beta_i) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7) \text{ where,}$$

$$x_1 = \begin{cases} 1 \text{ if first MI} \\ 0 \text{ if recurrent MI} \end{cases} \quad x_2 = \begin{cases} 1 \text{ if } 60 \leq \text{age} < 70 \\ 0 \text{ otherwise} \end{cases} \quad x_3 = \begin{cases} 1 \text{ if } 70 \leq \text{age} < 80 \\ 0 \text{ otherwise} \end{cases}$$

$$x_4 = \begin{cases} 1 \text{ if age} \geq 80 \\ 0 \text{ otherwise} \end{cases} \quad x_5 = \text{BMI} \quad x_6 = \begin{cases} 1 \text{ if cohort 2} \\ 0 \text{ otherwise} \end{cases}$$

$$x_7 = \begin{cases} 1 \text{ if cohort 3} \\ 0 \text{ otherwise} \end{cases}$$

1. $H_0 : \beta_2 = \beta_3 = ... = \beta_7$
   $H_A :$ not H_0
2. Using likelihood ratio test:

$$\begin{aligned} G &= -2\log[l_p(\hat{\beta}) - l_q(\hat{\beta})] \\ &= 2445.498 - 2305.777 \\ &= 139.721 \sim \chi^2_6 \text{ under the null} \end{aligned}$$

$$p = P(139.721 \geq \chi^2_6) < 0.001$$

3. $0.001 < 0.05 \rightarrow$ reject the null
4. At the 5% signficance level, there is sufficient evidence to conclude that model 2, the adjusted model, is a better predictor of survival time than model 1, the crude model.

Problem 3

```
proc phreg data = hw5;
    class mi_ord (ref = "0")
          age_c (ref = "1")
          year (ref = "1") / param = ref;
    model len_follow * final_stat(0) = mi_ord age_c bmi year mi_ord * age_c
          / risklimits covb ties = efron;
    hazardratio mi_ord / at (age_c = "2") diff = all;
    baseline out = model_3 survival = surv lower = lcl upper = ucl;
    title "Cox model, length of follow-up as a function of MI order,
           age category, bmi, cohort year, interaction between
           MI order and age category";
run;
```

$h(t, x_i, \beta_i) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_1 x_2 + \beta_9 x_1 x_3 + \beta_{10} x_1 x_4)$ where,

$$x_1 = \begin{cases} 1 \text{ if first MI} \\ 0 \text{ if recurrent MI} \end{cases} \quad x_2 = \begin{cases} 1 \text{ if } 60 \leq \text{age} < 70 \\ 0 \text{ otherwise} \end{cases} \quad x_3 = \begin{cases} 1 \text{ if } 70 \leq \text{age} < 80 \\ 0 \text{ otherwise} \end{cases}$$

$$x_4 = \begin{cases} 1 \text{ if age} \geq 80 \\ 0 \text{ otherwise} \end{cases} \quad x_5 = \text{BMI} \quad x_6 = \begin{cases} 1 \text{ if cohort 2} \\ 0 \text{ otherwise} \end{cases}$$

$$x_7 = \begin{cases} 1 \text{ if cohort 3} \\ 0 \text{ otherwise} \end{cases}$$

Hazard ratio

$$\begin{aligned} HR\{\text{recurrent MI vs. first MI, age} = 60-70\} &= \frac{h(t, x_1 = 1, x_2 = 1, x_3 = x_4 = ... = x_7 = 0)}{h(t, x_1 = 0, x_2 = 1, x_3 = x_4 = ... = x_7 = 0)} \\ &= \frac{h_0(t)\exp(\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_8)}{h_0(t)\exp(\hat{\beta}_2)} \\ &= \exp(\hat{\beta}_1 + \hat{\beta}_8) \\ &= \exp(1.09033 + 0.4091) \\ &= 4.4791 \end{aligned}$$

95% CI

$$\begin{aligned} \hat{\text{Var}}(\hat{\beta}_1 + \hat{\beta}_8) &= \hat{\text{Var}}(\hat{\beta}_1) + \hat{\text{Var}}(\hat{\beta}_8) + 2\hat{\text{Cov}}(\hat{\beta}_1 + \hat{\beta}_8) \\ &= 0.2302 + 0.4539 + 2(-0.227) \\ &= 0.2301 \\ \hat{\text{SE}}(\hat{\beta}_1 + \hat{\beta}_8) &= \sqrt{0.2301} = 0.4797 \end{aligned}$$

$$\begin{aligned} 95\%\text{CI} &= \exp[(\hat{\beta}_1 + \hat{\beta}_8) \pm 1.96 \times \hat{\text{SE}}(\hat{\beta}_1 + \hat{\beta}_8)] \\ &= \exp(1.49943 \pm 0.9402) \\ &= \exp(0.5592, 2.439) \\ &= (1.749, 11.469) \end{aligned}$$

The death rate for patients between ages 60 and 70 with recurrent MI is 4.48 times the death rate compared to patients between the ages of 60 and 70 experiencing their first MI throughout the study period. Furthermore, we are 95% confident that this hazard ratio is between 1.749 and 11.469.

Problem 4

```
data surv_pred;
    input id mi_ord age_c bmi year;
    cards;
    1 1 4 30 3
    ;
run;

proc phreg data = hw5 plots(cl) = survival;
    class mi_ord (ref = "0")
          age_c (ref = "1")
          year (ref = "1") / param = ref;
    model len_follow * final_stat(0) = mi_ord age_c bmi year mi_ord * age_c
          / risklimits covb ties = efron;
    baseline covariates = surv_pred out = pred survival = _all_ / rowid = id;
run;

proc print data = pred;
run;
```