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;
```

Analysis of Maximum Likelihood Estimates										
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits		Label
mi_ord	1	1	0.42665	0.13910	9.4081	0.0022	1.532	1.167	2.012	mi_ord 1

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

1. $H_0 : \beta_1 = 0$
   $H_A : \beta_1 \neq 0$
2. Using a Wald Chi-squared test:

$$Z^2 = 9.4081 \sim \chi_1^2 \text{ under the null}$$

$$p = P(\chi_1^2 \geq 9.4081) = 0.0022$$

3. $0.0022 < 0.05 \rightarrow$ 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 rate of death among patients with recurrent MI is 1.53 times the rate 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;
```

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	2454.641	2445.498
AIC	2454.641	2447.498
SBC	2454.641	2450.869

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	2454.641	2305.777
AIC	2454.641	2319.777
SBC	2454.641	2343.372

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

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

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

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

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

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

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

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 = ref;
    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;
```

Analysis of Maximum Likelihood Estimates

Parameter			DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits		Label
mi_ord	1		1	1.09033	0.47976	5.1649	0.0230	.	.	.	mi_ord 1
age_c	2		1	0.20705	0.49322	0.1762	0.6746	.	.	.	age_c 2
age_c	3		1	1.90747	0.36561	27.2193	<.0001	.	.	.	age_c 3
age_c	4		1	2.36623	0.35174	45.2554	<.0001	.	.	.	age_c 4
bmi			1	-0.04925	0.01529	10.3695	0.0013	0.952	0.924	0.981	
year	2		1	0.20725	0.17375	1.4228	0.2329	1.230	0.875	1.729	year 2
year	3		1	0.64636	0.19217	11.3132	0.0008	1.909	1.310	2.782	year 3
mi_ord*age_c	1	2	1	0.40491	0.67369	0.3612	0.5478	.	.	.	mi_ord 1 * age_c 2
mi_ord*age_c	1	3	1	-0.95485	0.54160	3.1083	0.0779	.	.	.	mi_ord 1 * age_c 3
mi_ord*age_c	1	4	1	-1.27263	0.51765	6.0440	0.0140	.	.	.	mi_ord 1 * age_c 4

Estimated Covariance Matrix

Parameter		mi_ord1	age_c2	age_c3	age_c4	bmi	year2	year3	mi_ord1age_c2	mi_ord1age_c3	mi_ord1age_c4
mi_ord1	mi_ord 1	0.2301724303	0.1011386374	0.1047166024	0.1047723611	0.0000367613	-.0037433195	-.0006356784	-.2270303048	-.2299477413	-.2298536805
age_c2	age_c 2	0.1011386374	0.2432679183	0.1012418318	0.1015865825	0.0001523104	-.0015633804	-.0009063901	-.2434335201	-.1010566605	-.1010465950
age_c3	age_c 3	0.1047166024	0.1012418318	0.1336718772	0.1069845905	0.0005629523	-.0007073904	0.0014341397	-.1006223294	-.1325350771	-.1045216180
age_c4	age_c 4	0.1047723611	0.1015865825	0.1069845905	0.1237215584	0.0010291190	-.0002558398	0.0003758684	-.1000088118	-.1048210668	-.1190616461
bmi		0.0000367613	0.0001523104	0.0005629523	0.0010291190	0.0002339034	-.0000118128	0.0000024728	0.0002546695	-.0000849050	-.0000155606
year2	year 2	-.0037433195	-.0015633804	-.0007073904	-.0002558398	-.0000118128	0.0301889609	0.0175055751	0.0055558723	0.0027929079	0.0016911615
year3	year 3	-.0006356784	-.0009063901	0.0014341397	0.0003758684	0.0000024728	0.0175055751	0.0369282524	0.0019771253	0.0000257457	0.0010877189
mi_ord1age_c2	mi_ord 1 * age_c 2	-.2270303048	-.2434335201	-.1006223294	-.1000088118	0.0002546695	0.0055558723	0.0019771253	0.4538638978	0.2266262215	0.2266398801
mi_ord1age_c3	mi_ord 1 * age_c 3	-.2299477413	-.1010566605	-.1325350771	-.1048210668	-.0000849050	0.0027929079	0.0000257457	0.2266262215	0.2933279012	0.2296499751
mi_ord1age_c4	mi_ord 1 * age_c 4	-.2298536805	-.1010465950	-.1045216180	-.1190616461	-.0000155606	0.0016911615	0.0010877189	0.2266398801	0.2296499751	0.2679641417

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

87	1	1	4	30	3	363	0.54179	0.073016	0.41602	0.70558

The probability, from SAS, that an individual who is older than 80, with a BMI equal to 30, and was in the 2001 cohort study survives more than 365 days (the closest unique event time point is 363 days) is 0.54179.