Applied Regression II Final - Part Two

Nick Williams

Model summaries for both model one and model two are found in the summary section.

Model One

Model one uses a Cox model to model the time until a child experiences depression as a function of parental history of depression:
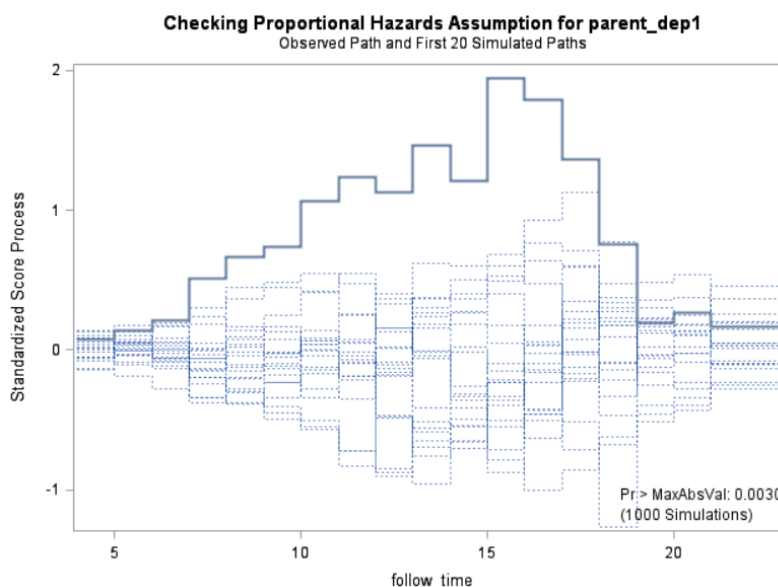
$$h(t, x) = h_0(t)h(\beta_1 x_{1i})$$

$$\text{where, } x_1 = \begin{cases} 0 \text{ if no history of parental depression} \\ 1 \text{ if history of parental depression} \end{cases}$$

SAS code for the model is as follows:

```
proc phreg data = depression;
    class parent_dep (ref = '0') / param = ref;
    model follow_time * child_dep(0) = parent_dep / ties = efron risklimits;
    assess ph / resample;
run;
```

I tested the proportional hazards assumption for parental depression and found that it was violated (p = 0.003). Because the proportional hazards assumption is violated, parental depression is time-dependent and this needs to be accounted for.

Checking Proportional Hazards Assumption for parent_dep1
Observed Path and First 20 Simulated Paths

Model Two

A check of the proportional hazards assumption in model one showed that the effect of parental history on the time until a child experiences depression is not constant over time. As such, in model 2 I have introduced a time-dependent covariate ("early_onset") that indicates if the age that depression started or the age at which censoring occurred was pre-pubertal (< 13 years old) or not (13 years or older):

$$h(t, x) = h_0(t)h(\beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i}x_{2i})$$

$$\text{where, } x_1 = \begin{cases} 0 \text{ if no history of parental depression} \\ 1 \text{ if history of parental depression} \end{cases}$$

$$x_2 = \begin{cases} 0 \text{ if follow time is } \geq 13 \text{ years} \\ 1 \text{ if follow time is } < 13 \text{ years} \end{cases}$$

SAS code for model two and plot is as follows:

```
data depression;
    set depression;
    if follow_time >= 13 then early_onset = 0;
        else early_onset = 1;
run;

proc phreg data = depression;
    class parent_dep (ref = '0') / param = ref;
    class early_onset (ref = '0') / param = ref;
    model follow_time * child_dep(0) = parent_dep early_onset parent_dep*early_onset / ties = efron;
    hazardratio parent_dep / diff = ref;
run;

data plot;
    input id parent_dep early_onset;
    datalines;
    1 1 1
    2 0 1
    3 1 0
    4 0 0
    ;
run;

proc phreg data = depression plots(overlay) = survival;
    model follow_time * child_dep(0) = parent_dep early_onset parent_dep*early_onset / ties = efron;
    hazardratio parent_dep / diff = ref;
    baseline covariates = plot / rowid = id;
    title "Survival curves of different child categories";
run;
```

Summary

Using likelihood ratio tests, at the 5% significance level both model one and model two were statistically significant (Model one: $\chi_1^2 = 7.77, p = 0.005$, Model two: $\chi_3^2 = 98.89, p < 0.001$).

Table 1 provides hazard ratios and corresponding 95% confidence intervals for model one and model two. Using model one, children with parental history of depression have 2.01 times the risk of experiencing depression themselves compared to children without parental history of depression (95% CI: 1.209, 3.332). However, a check of the proportional hazards assumption for model one revealed that the effect of parental depression was not constant over time (p = 0.002).

Model two includes a time dependent covariate that indicates if depression onset was pre-pubertal or not (depression onset occurs before 13) and an interaction term between this variable and parental depression ($\chi_1^2 = 3.92, p = 0.048$). Assessing risk of depression using this interaction term revealed that children with a parental history of depression have 1.32 (95% CI: 0.73, 2.38) times the risk of experiencing non-pre-pubertal depression compared to children without a parental history of depression. However, children with parental

history of depression have 5.16 (95% CI: 1.53, 17.42) times the risk of experiencing pre-pubertal depression than children without a parental history of depression.

Based on these analysis, children of depressed parents do experience a higher risk of depression themselves before the age of 13 compared to children without depressed parents. However, the risk of depression after the age of 13 is the same among both groups of children. Survival curves for model two can be found in Figure 1.

Table 1: Hazard ratios comparing risk of depression according to parental depression history

| Covariate	HR (95% CI)	p-Value
Model 1: Parental depression status only		
Parental Depression	2.01 (1.21, 3.33)	0.007
No Parental Depression	1.00 (reference)	
Model 2: Pre-pubertal onset interaction		
No parental depression, pre-pubertal onset	1.00 (reference)	
Parental depression, pre-pubertal onset	5.16 (1.53, 17.42)	
No parental depression, not pre-pubertal onset	1.00 (reference)	
Parental depression, not pre-pubertal onset	1.32 (0.73, 2.38)	
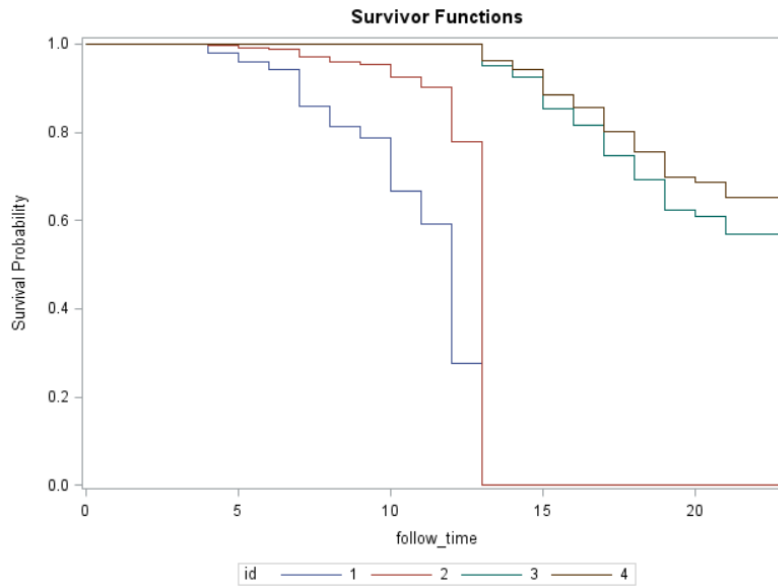
Figure 1: id 1: history of parental depression, pre-pubertal onset of depression; id 2: no history of parental depression, pre-pubertal depression; id 3: history of parental depression, no pre-pubertal onset; id 4: no history of parental depression, no pre-pubertal onset