

Session 4: Lecture notes exercise solutions

Exercise 4.1

The first event occurs at time 1 for an individual with $x = 1$. There are 10 individuals in the risk set at time 1; 5 of these have $x = 1$ and 5 have $x = 0$. The contribution to the partial likelihood from time 1 is:

$$\frac{e^{\beta}}{5 + 5e^{\beta}}.$$

The second event time is time 3. The person who has the event has $x = 1$ and the risk set comprises 9 individuals; 4 with $x = 1$ and 5 with $x = 0$. The contribution to the partial likelihood at time 3 is therefore:

$$\frac{e^{\beta}}{5 + 4e^{\beta}}.$$

The third event time is time 5. The person who has the event has $x = 0$ and the risk set comprises 7 individuals; 4 with $x = 0$ and 3 with $x = 1$. The contribution to the partial likelihood at time 5 is:

$$\frac{1}{4 + 3e^{\beta}}.$$

Similarly, the contributions to the partial likelihood from the 4th and 5th events times are respectively:

$$\frac{e^{\beta}}{3 + 2e^{\beta}}, \quad \frac{1}{3 + e^{\beta}}.$$

The full partial likelihood is the product of the contributions from each event time:

$$L_P = \frac{e^{\beta}}{(5 + 5e^{\beta})} \frac{e^{\beta}}{(5 + 4e^{\beta})} \frac{1}{(4 + 3e^{\beta})} \frac{e^{\beta}}{(3 + 2e^{\beta})} \frac{1}{(3 + e^{\beta})}.$$

Exercise 4.2

The hazard ratio (HR) estimate is 2.66. The hazard for death in the IM-positive group is estimated to be 2.66 times that in the IM-negative group. The 95% CI for the HR is (1.13, 6.25) and the p-value is 0.024. There is evidence at the $< 5\%$ level of an association between IM status and the hazard for death.

Exercise 4.3

- After conditioning on smoking status and family history, a 10 grams per day increase in alcohol intake was associated with a 14% increase in the hazard rate for breast cancer and this was highly statistically significant ($p=0.005$).
- After conditioning on alcohol intake and smoking status, a family history of breast cancer was associated with an almost 80% increase in the hazard rate of breast cancer (estimated hazard ratio 1.766), and this was highly statistically significant ($p < 0.001$).

- After conditioning on alcohol intake and family history, smoking status was not associated with breast cancer risk. Although the hazard ratio in former smokers compared to never smoker was 0.875, suggesting a reduced risk in former smokers, the result was not statistically significant ($p=0.197$). The hazard ratio in current smokers relative to never smokers was close to 1 with a wide confidence interval which includes 1.

Exercise 4.4

The survivor curves in the two treatment groups (Figure 4.8) are similar up to about 20 days, after which they diverge. In the right hand plot we are looking for the two curves to be approximately parallel over time if the proportional hazards assumption was appropriate. This is clearly not the case. There is clear evidence here that the relative effects of the two treatments on survival changes over time.

Figure 4.9 compares the estimated survivor curves obtained after fitting the Cox proportional hazards model with the Kaplan-Meier estimates of the survivor curves in the two treatment groups. The survivor curves from the Cox model do not match closely with the Kaplan-Meier estimates, providing further evidence that the proportional hazards assumption is not appropriate.