BIOS 622 Homework 7

Instructor: Sijin Wen

April 6, 2025

1. The following table gives a small data set of survival times and a covariate z:

patient ID	survival time in years	δ	z
1	7	1	4
2	8	1	3
3	9	0	5
4	10	1	6

where $\delta = 0$ means a right censored observation. Assuming a Cox proportional hazards model:

$$\lambda(t|z) = \lambda_0(t)e^{\beta z}$$

where $\lambda_0(t)$ is baseline hazard function (sometimes, we write it as $h_0(t)$ below in Hint, for example):

(a) Write down the partial likelihood of β.

(b) In R, plot the log partial likelihood of β in [-8, 3], and convince yourself that this function is concave.

(c) In R, find $\hat{\beta}$ that maximize this log partial likelihood function, calculate the second derivative of the log partial likelihood function at $\hat{\beta}$.

(d) (Optional if you have SAS) Use Phreg in SAS (sample code from 'Cox model rationale' in Week 10 folder) to fit the above proportional hazards model to the data. How do your results compare to those from the SAS output?

(e) Redo (d) using coxph in R. What is the estimated hazard ratio (i.e., $exp(\hat{\beta})$) and its 95%CI?

(f) Redo (e) using R code assuming a Weibull model (it is also a proportional hazards model). What is the estimated hazard ratio (i.e., $exp(\hat{\beta})$) and its 95%CI?

(g) What is your interpretation about the hazard ratio $exp(\hat{\beta})$? Is the covariate z a continuous variable or a categorical variable?

Hint: here is the example we will discuss in the lecture (on next page):

patient ID	survival time	δ	z
1	2	1	2
2	2	0	2
3	5	1	1
4	7	1	3

The partial likelihood is

$$\begin{aligned} PL(\beta) &= \frac{h_0(2)e^{2\beta}}{h_0(2)e^{2\beta} + h_0(2)e^{2\beta} + h_0(2)e^{\beta} + h_0(2)e^{3\beta}} \times \frac{h_0(5)e^{\beta}}{h_0(5)e^{\beta} + h_0(5)e^{3\beta}} \times \frac{h_0(7)e^{3\beta}}{h_0(7)e^{3\beta}} \\ &= \frac{e^{2\beta}}{e^{2\beta} + e^{2\beta} + e^{\beta} + e^{3\beta}} \times \frac{e^{\beta}}{e^{\beta} + e^{3\beta}} \times \frac{e^{3\beta}}{e^{3\beta}} \end{aligned}$$

The key question is 'what is β to maximize the above function'? Notice that the survival time is not shown in the PL in the final calculation since $h_0(t)e^{\beta z}$ is in the PL but $h_0(t)$ has been canceled out in each components, which implies that we don't need to specify the survival distribution in the Cox model.