MTH6157 Survival Models

Questions – Week 4 – Covariates and Proportional Hazard Models

Solutions

Q1

(a) a covariate is any factor that might affect the rate of mortality or survival

(b) The biggest single weakness is that on the data collection date, workers who have recently suffered an accident will not be present causing the data to be biased. *You should be able to list other weaknesses related to the specific choice of covariates.*

(c) for the different covariates one possibility would be:

- months of experience – keep as numeric value in covariate matrix
- electrician qualification – 1=yes, 0=no
- hat colour – 0=white, 1=yellow, 2=other
- witnessed accidents – 0=0, 1=1, 2= more than 1

(d) advantages

- simpler likelihood equation
- might be easier to communicate the relationship

disadvantages

- lose some information by giving up the specific data
- the scoring system is subjective
- different modellers will come up with different scoring systems

Q2

(i) the baseline hazard applies to a life with all z=0 so a patient in Spain under age 55 treated immediately upon diagnosis

(ii) (a) $z_1 = 0 \quad z_2 = 2 \quad z_3 = 1$

$$\therefore \lambda_{SSb}(t) = \lambda_0(t)\, e^{0+2\times 0.4 - 1\times 0.3}$$

$$= \lambda_0(t)\, e^{0.5}$$

(b) $$S(t) = e^{-\int_0^t \lambda_i(t)\, ds}$$

$$= e^{-\int_0^t \lambda_0(s) e^{0.5} ds}$$

$$= \exp\left[-\int_0^t \lambda_0(s)\, ds\right]^{e^{0.5}}$$

(iii) for the Italian patient

$z_1 = 1 \quad z_2 = 1 \quad z_3 = 1$

$$\therefore \lambda_{Ibs}(t) = \lambda_0(t)\, e^{0.03+0.4-0.3}$$

$$= \lambda_0(t)\, e^{0.13}$$

3 months is 13 weeks

$$S_{Ibs}(13) = \exp\left[-\int_0^{13} \lambda_0(s)\, ds\right]^{e^{0.13}}$$

$$= 0.95$$

$$\therefore \exp\left[-\int_0^{13} \lambda_0(s)\, ds\right] = 0.95^{e^{-0.13}}$$

So for the 55b life in (ii)

$$S_{55b}(13) = \exp\left[-\int_0^{13} \lambda_0(s)\, ds\right]^{e^{0.5}}$$

from (ii)(b)

$$= \left(0.95\, e^{-0.13}\right)^{e^{0.5}}$$

$$= \underline{\underline{0.928}}$$

Q3

(i) Model 1 has 3 covariates: age, time, BMI

Model 2 has 6 covariates: age, time, BMI, hospital ID, blood pressure, smoker

let L_3 be the maximised likelihood statistic for model 1

and L'_6 is the maximised likelihood statistic for model 2

then the likelihood ratio statistic is $-2(\log L_3 - \log L'_6)$

(ii) we compare this with critical values of χ^2 on 6-3 = 3 degrees of freedom