Longitudinal data analysis : examen

Course materials are allowed

Exercice 1

In a study of time to first opening of an account on a social media, 5032 children were asked "when did open your first account on a social media?"

- Some do not have an account.
- Some answers were "I have an account but I cannot remember when I opened it".
- Some remembered when they first opened their account.

Which observations are left-censored, which are right-censored ?

Exercice 2

Let T the time of interest has a probability density function given by

$$x\lambda^2 \exp(-\lambda x) \text{ for all } x > 0 \text{ where } \lambda > 0$$

ans C a censoring time with exponential distribution $\mathcal{E}(\theta)$ ($\theta > 0$) and assume that T and C are independent. Define $\left(T^C = T \wedge C, \delta = \mathbb{1}_{T \leq C}\right)$

1. Compute $\mathbb{P}(\delta = 1)$.
2. Compute $\mathbb{E}(T^C)$.
3. Check that $\mathbb{E}(T) \geq \mathbb{E}(T^C)$.
4. Let $(T_1^C, \delta_1), \ldots, (T_n^C, \delta_n)$ be an i.i.d. sample in this model. Compute the maximum likelihood estimator of θ.

Exercice 3

In the Stanford Heart Transplant data (Kalbfleisch and Prentice 2011), survival of patients on the waiting list for the Stanford heart transplant program were recorded along with the variables

- fustat: dead or alive
- surgery: prior bypass surgery
- age: age (in years)

- futime: follow-up time
- wait.time: time before transplant
- transplant: transplant indicator
- accept.yr: acceptance into program .

1. Which variables are time-dependent ?
2. Explain which transformation has to be applied to the data on patient 4. The initial data are

```
##   fustat surgery      age futime wait.time transplant accept.yr
## 1      1       0 30.84463     49        NA          0      1967
## 2      1       0 51.83573      5        NA          0      1968
## 3      1       0 54.29706     15         0          1      1968
## 4      1       0 40.26283     38        35          1      1968
## 5      1       0 20.78576     17        NA          0      1968
## 6      1       0 54.59548      2        NA          0      1968
```

3. Which of these two Cox models is correct ? Would you conclude that the transplant has an influence on the risk of death ?

   - Model 1

```
## coxph(formula = Surv(futime, fustat) ~ surgery + transplant +
##     age, data = jasa)
##
##                coef exp(coef) se(coef)     z       p
## age          0.0589    1.0607   0.0150  3.91 9.1e-05
## surgery     -0.4190    0.6577   0.3712 -1.13    0.26
## transplant -1.7171    0.1796   0.2785 -6.16 7.1e-10
```

   - Model 2

```
## coxph(formula = Surv(start, stop, event) ~ age + surgery +
##      transplant, data = jasa1)
##
##                coef exp(coef) se(coef)     z     p
## age          0.0306    1.0310   0.0139  2.20 0.028
## surgery     -0.7733    0.4615   0.3597 -2.15 0.032
## transplant  0.0141    1.0142   0.3082  0.05 0.964
```

References

[KP11] John D Kalbfleisch and Ross L Prentice. *The statistical analysis of failure time data*. Vol. 360. John Wiley & Sons, 2011.