Cox model with time-varying covariates

Jean Peyen

Department of Statistics – University of Leeds

July 2023

Cox proportional hazards model

Hazard rates
Each subject $i \in [\![1, n]\!]$ is characterised by a set of p covariates $\mathbf{x}_i = (x_{i1}, x_{i2}, \cdots, x_{ip})$ and a terminal time t_i.

The hazard rate function takes the following form:

$$h_i(t) = \underbrace{\exp(\beta^t \mathbf{x}_i)}_{\text{relative hazard}} \cdot \underbrace{h_0(t)}_{\text{baseline hazard}} .$$

Partial log-likelihood

$$\log L(\beta) = \sum_{i=1}^{n} \delta_i \left(\beta^t \mathbf{x}_i - \log \sum_{\ell \in R(t_i)} \exp(\beta^t \mathbf{x}_\ell) \right)$$

Example (part 1)

Dataset

stanford_heart_transplants from the Python package lifelines

event	$\cdots$	1	0	0	$\cdots$
transplant	$\cdots$	0	1	0	$\cdots$
stop	$\cdots$	21	39	31	$\cdots$

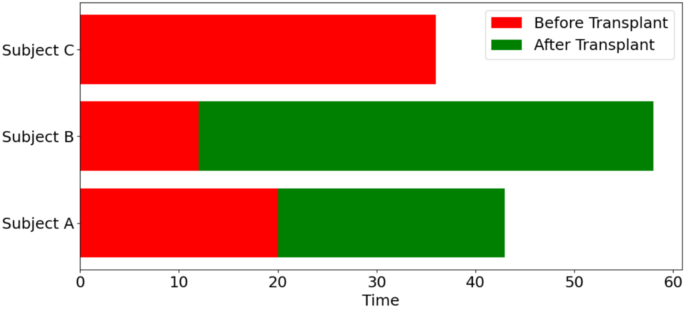
Fit summary

```
---
             coef  exp(coef)   se(coef)   coef lower 95%   coef upper 95%
covariate
transplant  -1.32       0.27       0.24            -1.80            -0.85

                 p
covariate
transplant  <0.005
---
```

Immortal bias

Patients must survive for a certain period until they receive a transplant, thus introducing an artificial survival advantage for the transplant group.



Cox model with time-varying covariates

Hazard rates

Each subject $i \in [\![1, n]\!]$ is characterised by a set of p time-varying covariates $\mathbf{x}_i(t) = (x_{i1}(t), x_{i2}(t), \cdots, x_{ip})$.

$$h_i(t) = \underbrace{\exp(\beta^t \mathbf{x}_i(t))}_{\text{relative hazard}} \cdot \underbrace{h_0(t)}_{\text{baseline hazard}} .$$

The proportionality assumption is no longer valid when the covariates are time-varying.

Two types of time-varying covariates:

- internal : evolution is affected by the survival of the subject
- external : do not require the survival of the subject for their existence

Partial log-likelihood

$$\log L(\beta) = \sum_{i=1}^{n} \delta_i \left(\beta^t \mathbf{x}_i(t_i) - \log \sum_{\ell \in R(t_i)} \exp(\beta^t \mathbf{x}_\ell(t_i)) \right)$$

Example (part 2)

Dataset

id	⋯	98	98	99	100	100	⋯
event	⋯	0	0	1	0	0	⋯
transplant	⋯	0	1	0	0	1	⋯
start	⋯	0	96	0	0	38	⋯
stop	⋯	96	109	21	38	39	⋯

Fit summary

```
---
             coef  exp(coef)   se(coef)   coef lower 95%   coef upper 95%
covariate
transplant   0.13       1.14       0.30            -0.46             0.72

                p
covariate
transplant   0.67
---
```

Sources

- David Collett *Modelling Survival Data in Medical Research*, First edition 1994 by Chapman & Hall
- Cameron Davidson-Pilon, *Lifelines: survival analysis in Python*, Journal of Open Source Software, $4(40)$, 1317, *https://doi.org/10.21105/joss.01317*
- Lifelines package documentation, *https://lifelines.readthedocs.io*
- GitHub repository of the presentation, *https://github.com/Etamunu/CoxPresentation*