

Time-dependent coefficients

Patrick Breheny

December 1

Introduction

- As we discussed previously, stratification allows one to handle variables that do not satisfy the proportional hazards assumption, but this comes at a cost
- Namely, the variables no longer act upon the likelihood in a regular way, and we cannot use any of the methods we have developed in this course to carry out inference or model selection concerning them
- An alternative approach is to directly model the dependence of those variables upon time

Schoenfeld residuals

- We will start with exploratory methods for examining the effect of a dependence on time; the most widely used residuals for this purpose are Schoenfeld residuals
- The Schoenfeld residual s_{ij} is, in a sense, the contribution that subject i makes to the j th score function:

$$\begin{aligned} s_{ij} &= d_i(x_{ij} - \boldsymbol{\pi}_{\cdot i}^T \mathbf{X}_j) \\ &= d_i(x_{ij} - \mathbb{E}_i X_j); \end{aligned}$$

note that $s_{ij} = 0$ for all censored observations

- In general, the Schoenfeld residuals have similar uses as delta-beta residuals in terms of identifying influential observations with respect to estimating $\hat{\beta}_j$

Schoenfeld residuals and time-varying effects

- However, Schoenfeld residuals have a very useful property in terms of investigating the dependence of an effect on time
- Suppose that the effect of covariate j is not necessarily constant, but varies with time:

$$\eta_i(t) = x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{ij}\beta_j(t)$$

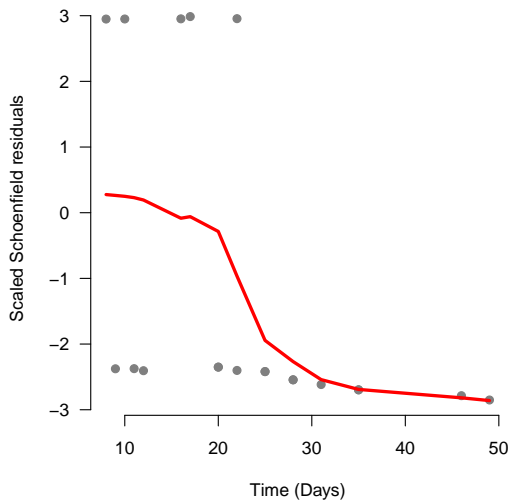
- Then one can define a scaled Schoenfeld residual, $\tilde{s}_{ij} = \hat{\beta}_j + s_{ij}/a$, where a is an appropriate scaling factor, such that the scaled residuals possess the following property:

$$\mathbb{E}\tilde{s}_{ij} = \beta_j(t)$$

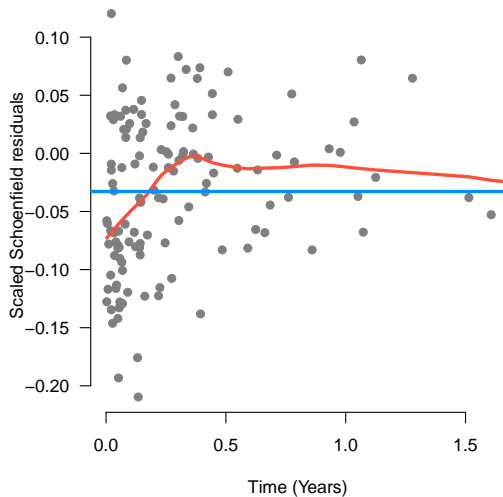
Schoenfeld residuals and time-varying effects (cont'd)

- In the interest of time, we'll have to skip the details of this derivation, but the details are in Grambsch and Therneau (1994), "Proportional hazards tests and diagnostics based on weighted residuals"
- Both Schoenfeld residuals (`type='schoenfeld'`) and scaled Schoenfeld residuals (`type='scaledsch'`) are available as options from `residuals.coxph`

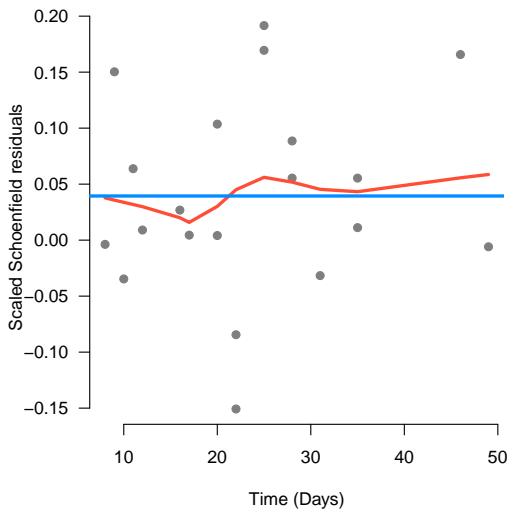
GVHD data: Treatment



VA Lung: Karnofsky



GVHD data: Age



Testing; R code

- One can also carry out a formal test for non-proportionality by testing whether the correlation between time and Schoenfeld residuals is zero
- These methods (the exploratory plot and the test) are available from the `cox.zph` function:

```
> fit <- coxph(Surv(Time, Status) ~ Group, GVHD)
> cox.zph(fit)
```

	rho	chisq	p
GroupMTX+CSP	-0.537	6.08	0.0137

```
> plot(cox.zph(fit))
```

Results

- For the VA lung data, Karnofsky status is highly non-proportional ($p = 0.0003$)
- For the GVHD data, the proportional hazards assumption is violated for treatment ($p = 0.01$) but not for age ($p = 0.6$)
- This would suggest that while a constant effect β is adequate for some terms, like age, we might wish to consider allowing time-dependent effects, $\beta(t)$, for other variables

Time-varying coefficients: Main idea

- To illustrate the basic idea, let's consider a simple linear time dependence:

$$\beta_j(t) = \beta_{j0} + \beta_{j1}t;$$

$\beta_{j1} = 0$ would correspond to the usual proportional hazards structure

- In principle, this poses no great challenge to estimation: the Cox partial likelihood revolves around calculating π_{ij} values, the probability of subject i failing at time t_j

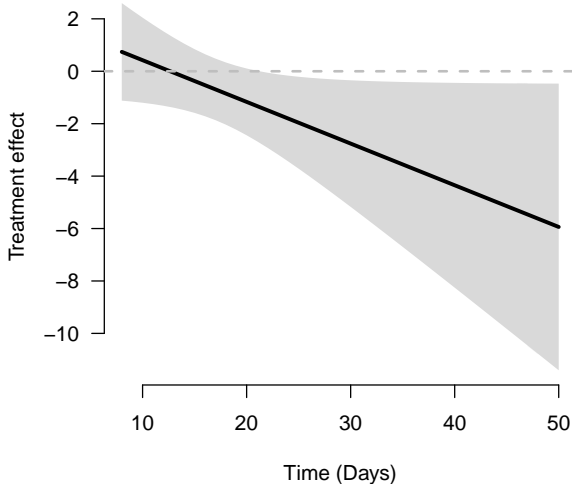
Specifics

- With a time-dependent effect, we simply need to plug the appropriate t_j into $\beta_j(t)$ when calculating π_{ij}
- Let's assume for the sake of simplicity that we only have one term in the model; we then have

$$\pi_{ij} = \frac{\exp(x_i\beta_0 + x_i\beta_1t_j)}{\sum_k \exp(x_k\beta_0 + x_k\beta_1t_j)}$$

- The score and information are simply functions of $\{\pi_{ij}\}$, so this is all we need to modify in order to estimate β_0 and β_1 and perform the usual inferential procedures

GVHD data: Treatment effect



What not to do

- On the surface, the idea of a time-varying effect may seem equivalent to an interaction between treatment and time
- In particular, you may be tempted to code it as

```
coxph(S ~ Trt + Trt:Time, GVHD)
```

- This is completely wrong, though; it results in

$$\eta_i = x_i\beta_0 + x_i t_i \beta_1;$$

where η_i is static across time, the effect being frozen at the point where subject i leaves the study

The tt function

- Instead, we need some way to tell `coxph` to recalculate the linear predictors at each distinct time t_j
- This is accomplished through the `tt` function (for “time transform”):

```
coxph(S ~ Trt + tt(Trt), GVHD,  
      tt=function(x, t, ...) {x*t})
```

- `tt` can be more complicated than this; in particular, it can return a matrix to represent a multiparameter function of time

Summary of tt models

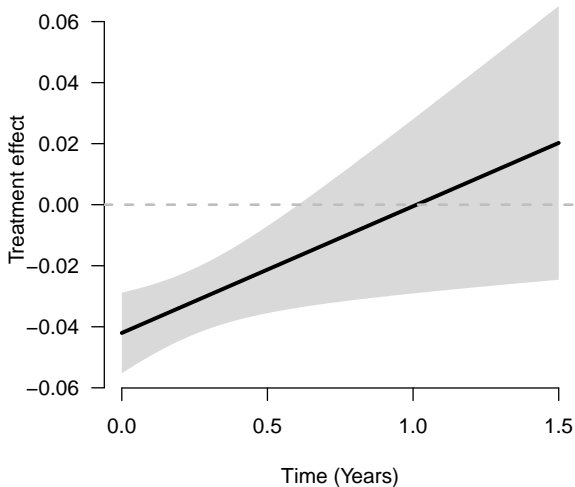
`summary` will provide the usual summary of the coefficients, for both the proportional hazard coefficients and the coefficients that represent the time-varying effect:

```
> summary(fit)
n= 64, number of events= 20

              coef exp(coef) se(coef)      z Pr(>|z|)
Trt           2.01256   7.48242  1.51909  1.325  0.1852
tt(Trt)    -0.15904   0.85296  0.08242 -1.930  0.0537 .
...
Concordance= 0.594 (se = 0.221 )
Rsquare= 0.171 (max possible= 0.915 )
Likelihood ratio test= 12.03 on 2 df, p=0.002437
```

In this case, the coefficient for the departure from proportionality is only borderline significant, but the overall test is highly significant

VA Lung data: Karnofsky effect (linear)



Exponential decay

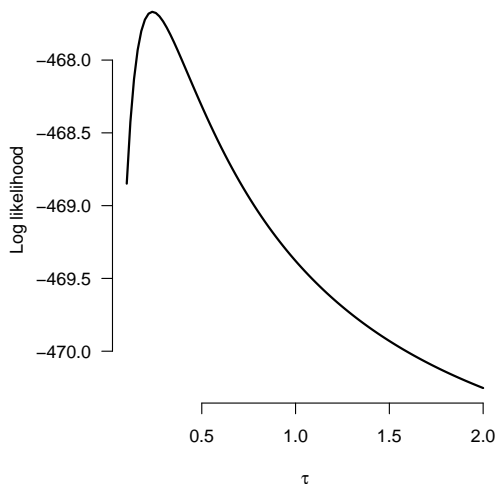
- Our exploratory plots, however, indicated that a linear term would not accurately capture the time-dependent effect of Karnofsky status
- A more appropriate model might be something like exponential decay:

$$\beta(t) = \beta_0 + \beta_1 \exp(-t/\tau),$$

where $\beta(t)$ gradually decays to a constant β_0 in the long run, and τ controls the rate of that decay

- Now, the τ parameter is not a typical linear coefficient, so the standard Cox Newton-Raphson method will not be able to estimate it; we will have to employ a grid search

Exponential decay rate



For comparison, the log likelihood of the linear model is -471.4

VA Lung data: Karnofsky effect (exponential decay)

