Survival Analysis I (CHL5209H)

Olli Saarela

Dalla Lana School of Public Health
University of Toronto

olli.saarela@utoronto.ca

February 26, 2019

Breslow estimator

- As we have seen previously, the Cox partial likelihood is used for estimating the parameters β in the proportional hazards model

$$\lambda_i(t) = \lambda_0(t) \exp\{\beta' x_i\},$$

  while leaving the baseline hazard function $\lambda_0(t)$ unspecified.
- However, if we are interested in absolute risks or survival probabilities based on the fitted model, for example to compare these to the actual outcomes, we need a some kind of estimate for the cumulative baseline hazard $\Lambda_0(t)$.
- The survival probability is then given by the relationship

$$S_i(t) = \exp\{-\Lambda_0(t) \exp\{\beta' x_i\}\}.$$

Breslow estimator (2)

- An estimator for $\Lambda_0(t)$ is suggested by the earlier profile likelihood calculation (formula (2) in the 5th slides), and is known as the Breslow estimator:

$$\hat{\Lambda}_0(t) = \sum_{j:t_j \leq t} \frac{d_j}{\sum_{l=1}^{n} Y_l(t_j) \exp\{\hat{\beta}' x_l\}}.$$

- Here $\hat{\beta}$ are the partial likelihood estimates of the regression parameters, and the at-risk process $Y_i(t) = \mathbf{1}_{\{T_i \geq t\}}$ is used to check whether individual i is still at risk at time t.
- As before, t_j refer to the ordered observed event times, and d_j to the numbers of events at each time.

Motivation

- It is not that easy to define a counterpart of the residual in a linear regression model for censored time-to-event outcomes.
- Nevertheless, several types of residuals exist for the Cox model, and can be used as model diagnostics.
- Intuitively, the closest counterpart to the linear model residual is the martingale residual.
- For understanding this, we need to introduce the concept of a martingale.

Martingales

- Recall that we can understand the hazard function through the relationship

$$P(\mathrm{d}N_i(t) = 1 \mid \mathcal{F}_{t-}) = E[\mathrm{d}N_i(t) \mid \mathcal{F}_{t-}] = Y_i(t)\lambda_i(t)\,\mathrm{d}t.$$

- Since the quantity $Y_i(t)\lambda_i(t)\,\mathrm{d}t$ can be interpreted as the expected value of the counting process jump, this motivates consideration of the differences

$$\mathrm{d}M_i(t) = \mathrm{d}N_i(t) - Y_i(t)\lambda_i(t)\,\mathrm{d}t,$$

which have the usual kind of 'observed' minus 'expected' interpretation.

Martingales (2)

- Equivalently,

$$M_i(t) = N_i(t) - \int_0^t Y_i(u)\lambda_i(u)\,\mathrm{d}u,$$

or

$$N_i(t) = \int_0^t Y_i(u)\lambda_i(u)\,\mathrm{d}u + M_i(t).$$

- Here the process $M_i(t)$ has the property $E[\mathrm{d}M_i(t) \mid \mathcal{F}_{t-}] = 0$.
- A process with this property is called a martingale.
- The process $M_i(t)$ can thus be interpreted as 'noise', while the hazard function captures the systematic variation in the counting process $N_i(t)$.

Martingale residual

- For the Cox model we get

$$M_i(t) = N_i(t) - \int_0^t Y_i(u) \exp\{\beta' x_i\} \lambda_0(u) \, \mathrm{d}u,$$

or the estimated version

$$\hat{M}_i(t) = N_i(t) - \int_0^t Y_i(u) \exp\{\hat{\beta}' x_i\} \, \mathrm{d}\hat{\Lambda}_0(u),$$

where the cumulative baseline hazard increments $\mathrm{d}\hat{\Lambda}_0(t)$ are given by the Breslow estimator.

- The $\hat{M}_i$, evaluated at the end of the follow-up period of individual i, is the martingale residual.

Uses of martingale residuals

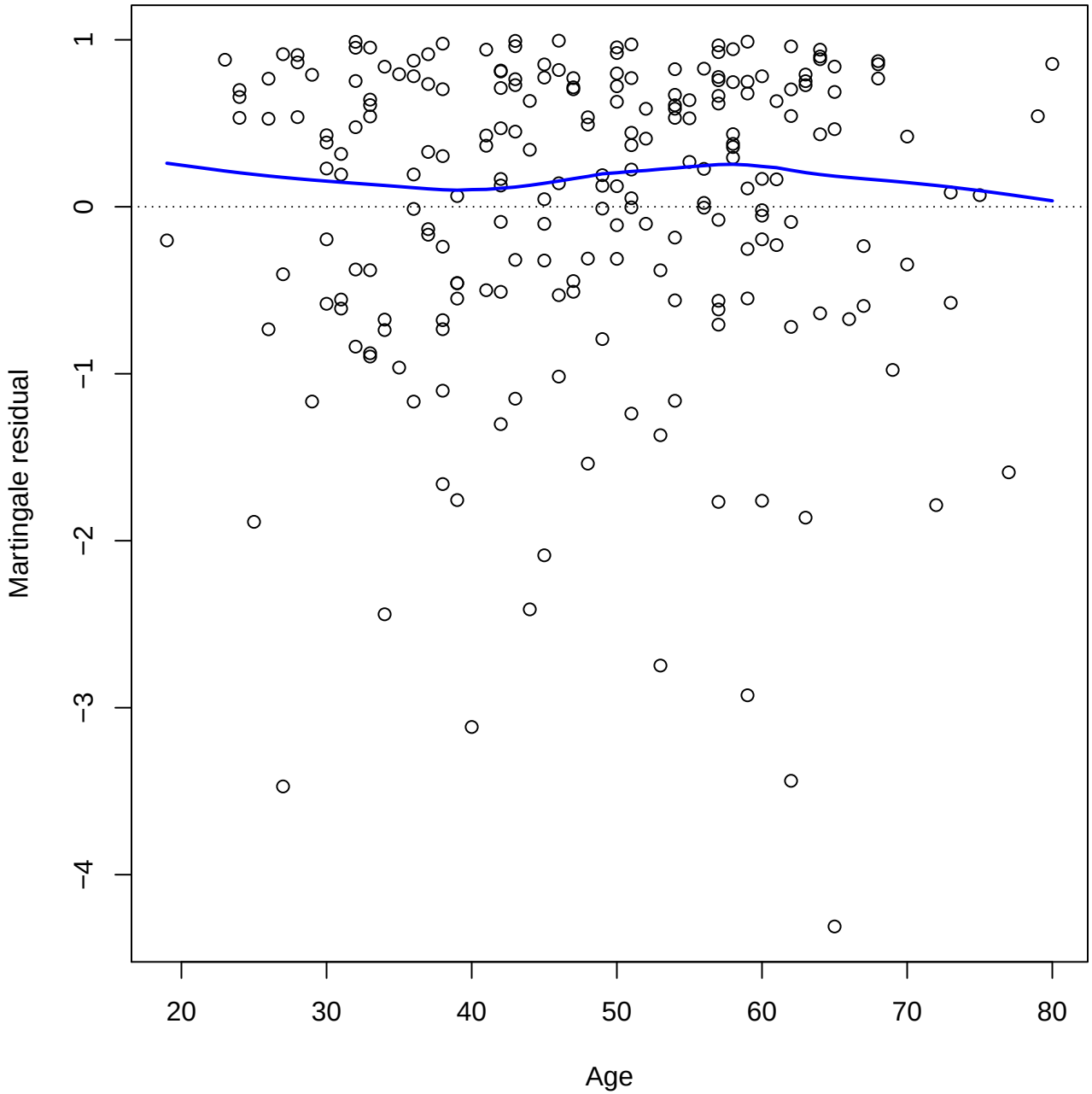
- The martingale residuals $\hat{M}_i$ for $i = 1, \ldots, n$ can be plotted for example against continuous covariates to check for non-linearity of the covariate effects.
- This is because any systematic deviation from linearity will show as systematic difference in the residuals.
- Note that $\hat{M}_i$ can receive values between $-\infty$ and one. (Why?)
- Thus, the residual plot will not be symmetric around zero.
- In R, the residuals are extracted from a fitted `coxph` object through the `residuals.coxph` function:

```
residuals(object,
  type=c("martingale", "deviance", "score", "schoenfeld",
         "dfbeta", "dfbetas", "scaledsch","partial"),
  collapse=FALSE, weighted=FALSE, ...)
```

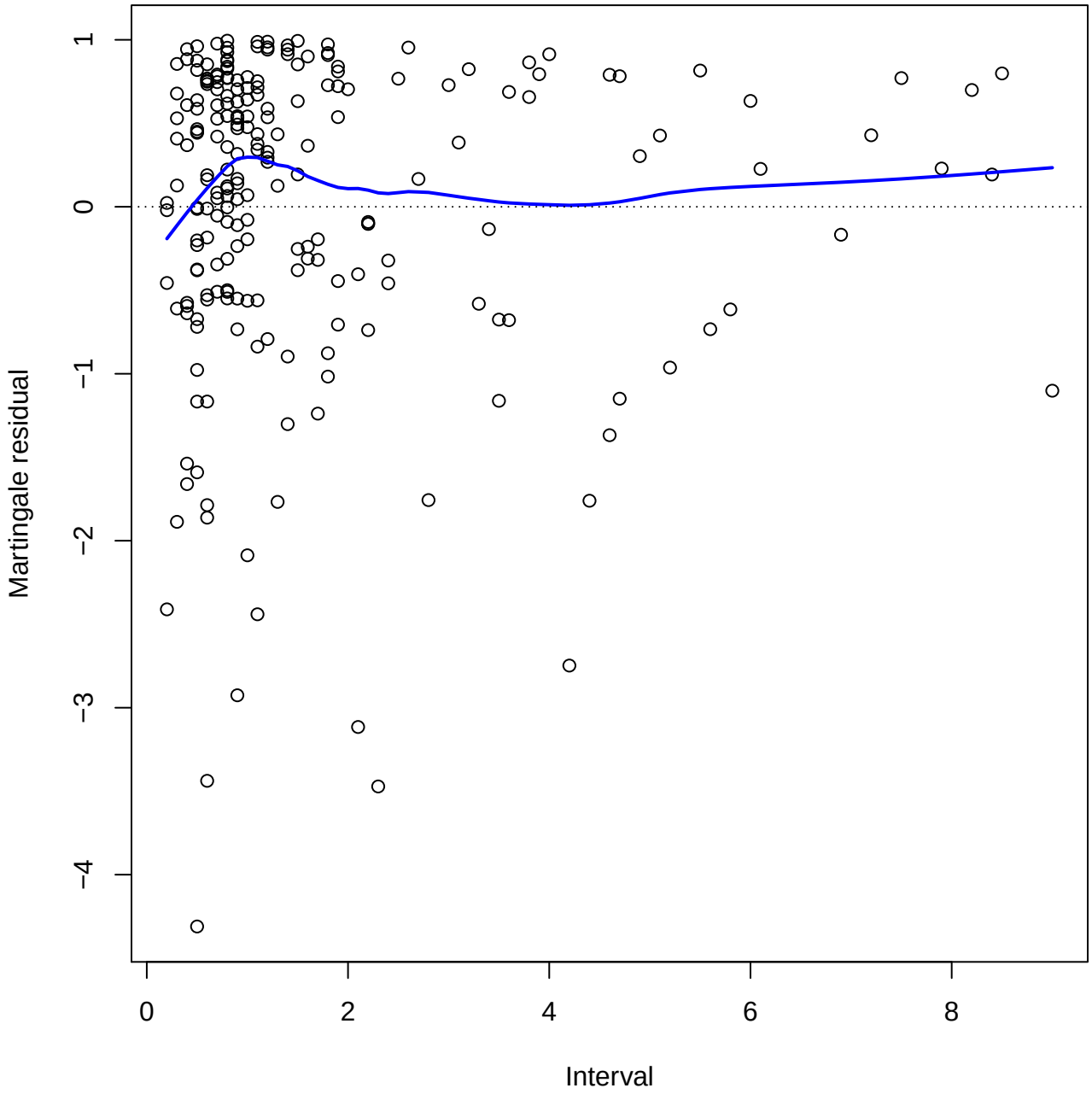
Example: brain tumor trial dataset

```
Call:
coxph(formula = Surv(weeks, event) ~ treat + resect75 + age +
    interval + karn + race + local + male + nitro + factor(path) +
    grade, data = brain)

  n= 221, number of events= 206

                    coef exp(coef)  se(coef)      z Pr(>|z|)
treat          -0.396796  0.672471  0.144512 -2.746 0.006037 **
resect75       -0.443613  0.641714  0.164990 -2.689 0.007172 **
age             0.017294  1.017445  0.006029  2.868 0.004127 **
interval       -0.139448  0.869838  0.047318 -2.947 0.003208 **
karn           -0.376704  0.686119  0.160759 -2.343 0.019115 *
race            0.592330  1.808197  0.270284  2.192 0.028415 *
local          -0.466002  0.627506  0.176343 -2.643 0.008228 **
male           -0.239337  0.787150  0.153227 -1.562 0.118294
nitro           0.480875  1.617490  0.154996  3.102 0.001919 **
factor(path)2 -0.640822  0.526859  0.217740 -2.943 0.003250 **
factor(path)3 -0.804242  0.447427  0.223189 -3.603 0.000314 ***
factor(path)4 -0.623663  0.535977  0.429548 -1.452 0.146528
grade          -0.857132  0.424377  0.293056 -2.925 0.003447 **
---
```

Martingale residuals by age



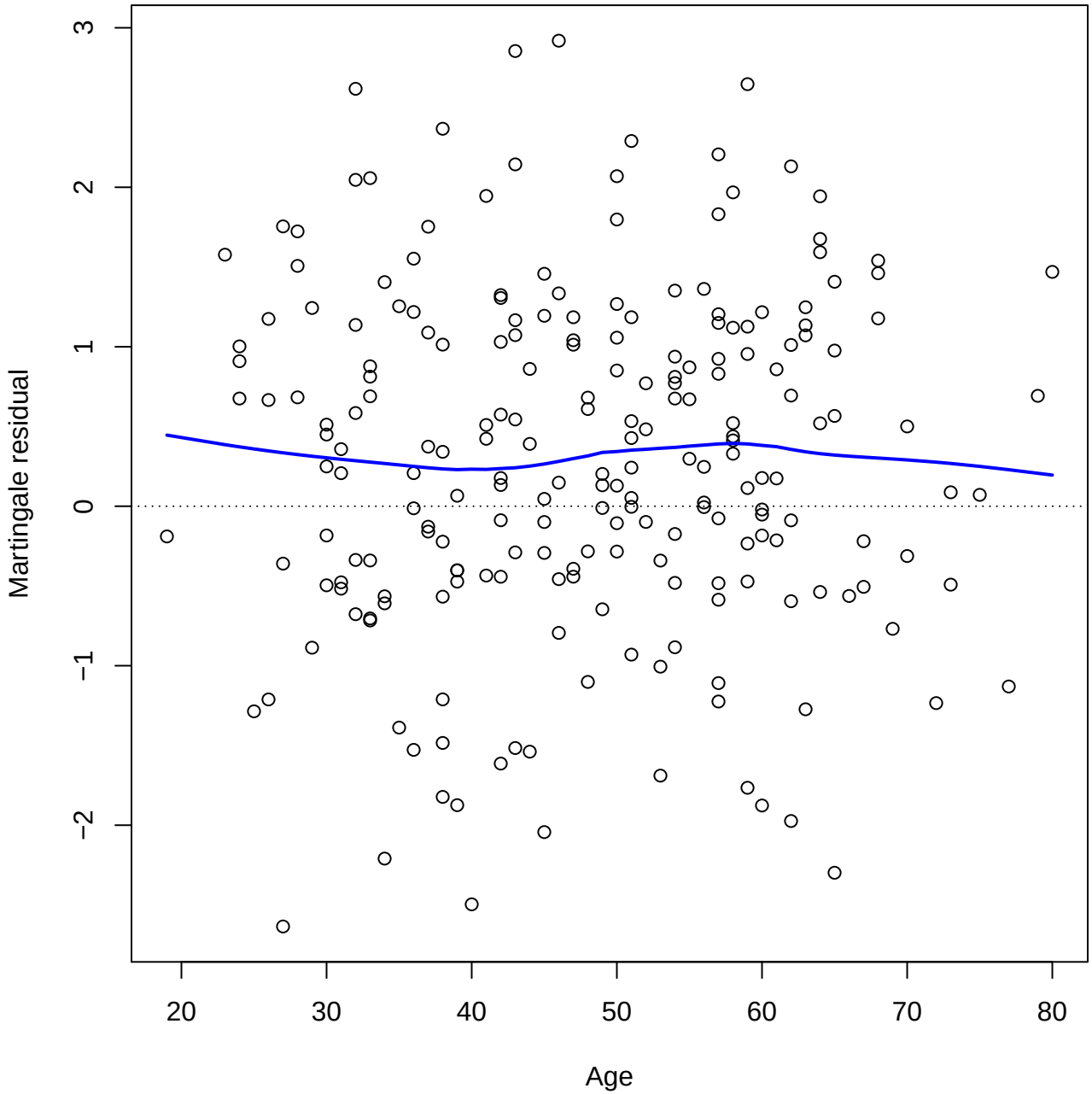
Martingale residuals by interval



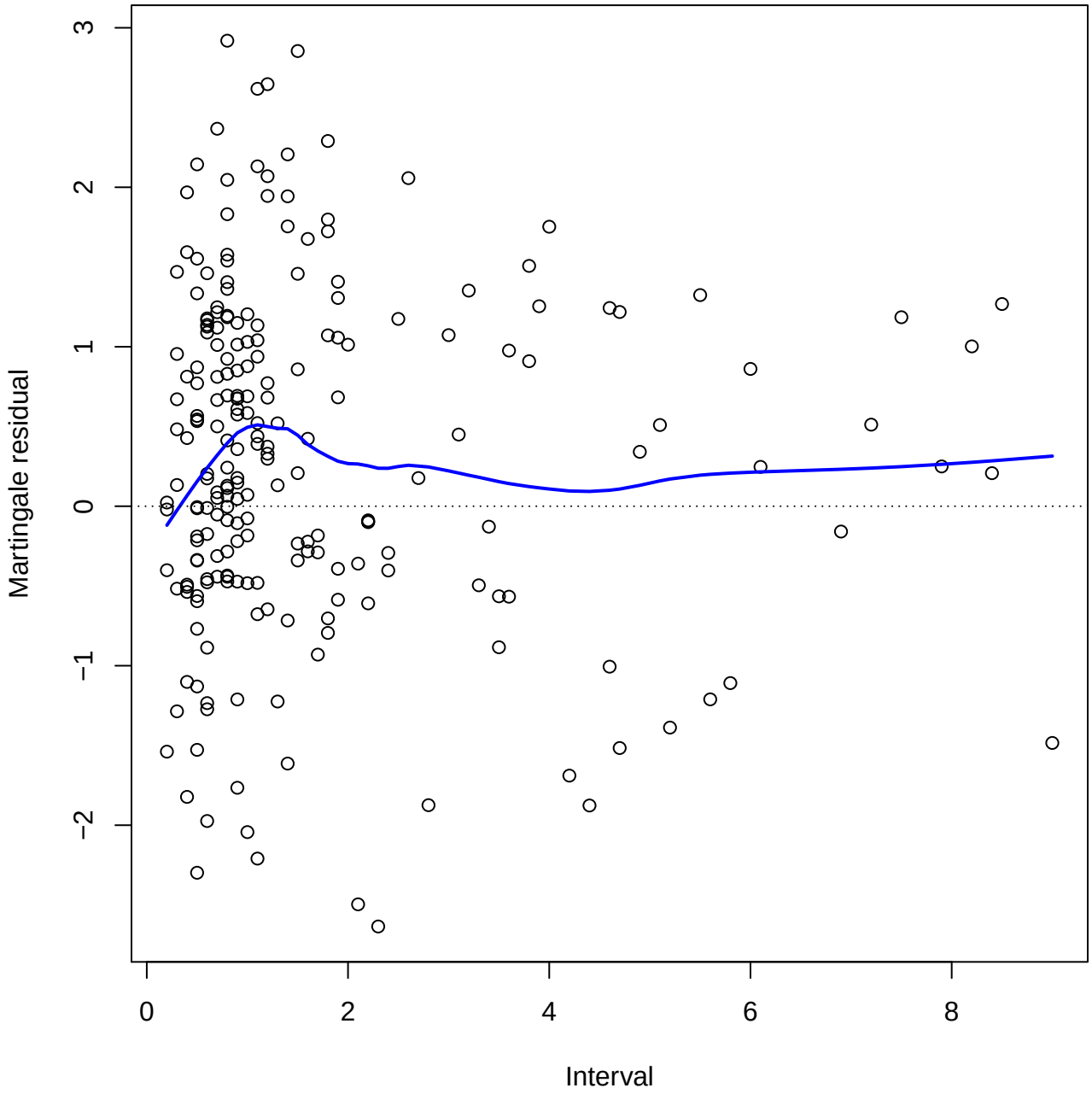
Other types of residuals

- Deviance residuals are rescaled versions of martingale residuals, to make them more symmetric around zero.
- Score residual and dfbeta measure the influence of the observation i on the parameter estimates when i is removed from the model fit. These are calculated separately for each regression coefficient.
- Schoenfeld residuals are defined only for non-censored observations, and can be interpreted as differences between observed and 'expected' (based on the model) covariate values.
- Since these 'expected' covariate values are based on the proportional hazards assumption, plotting the Schoenfeld residuals against time might be informative about violations of the proportionality assumption.

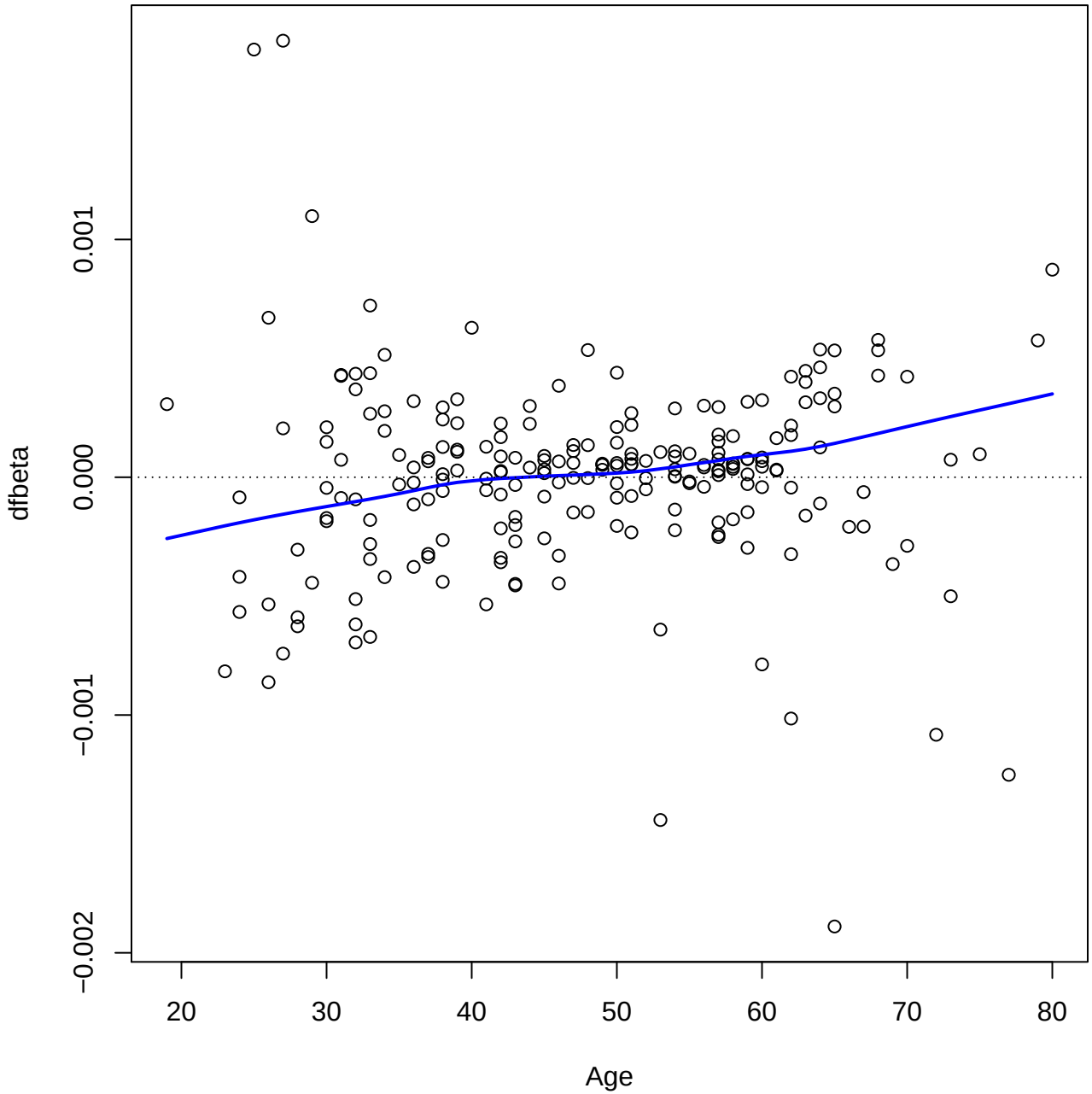
Deviance residuals by age



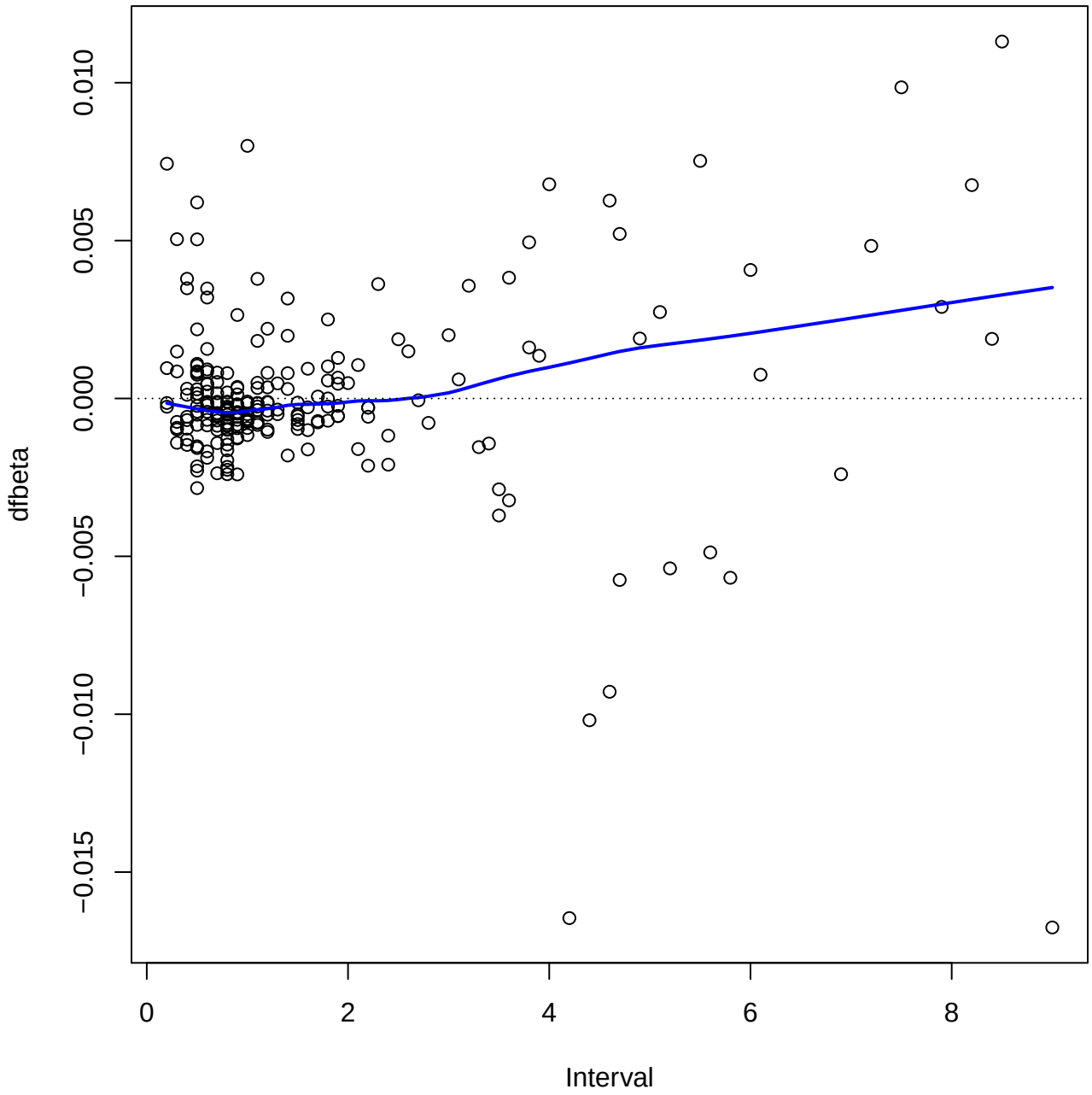
Deviance residuals by interval



dfbeta by age



dfbeta by interval



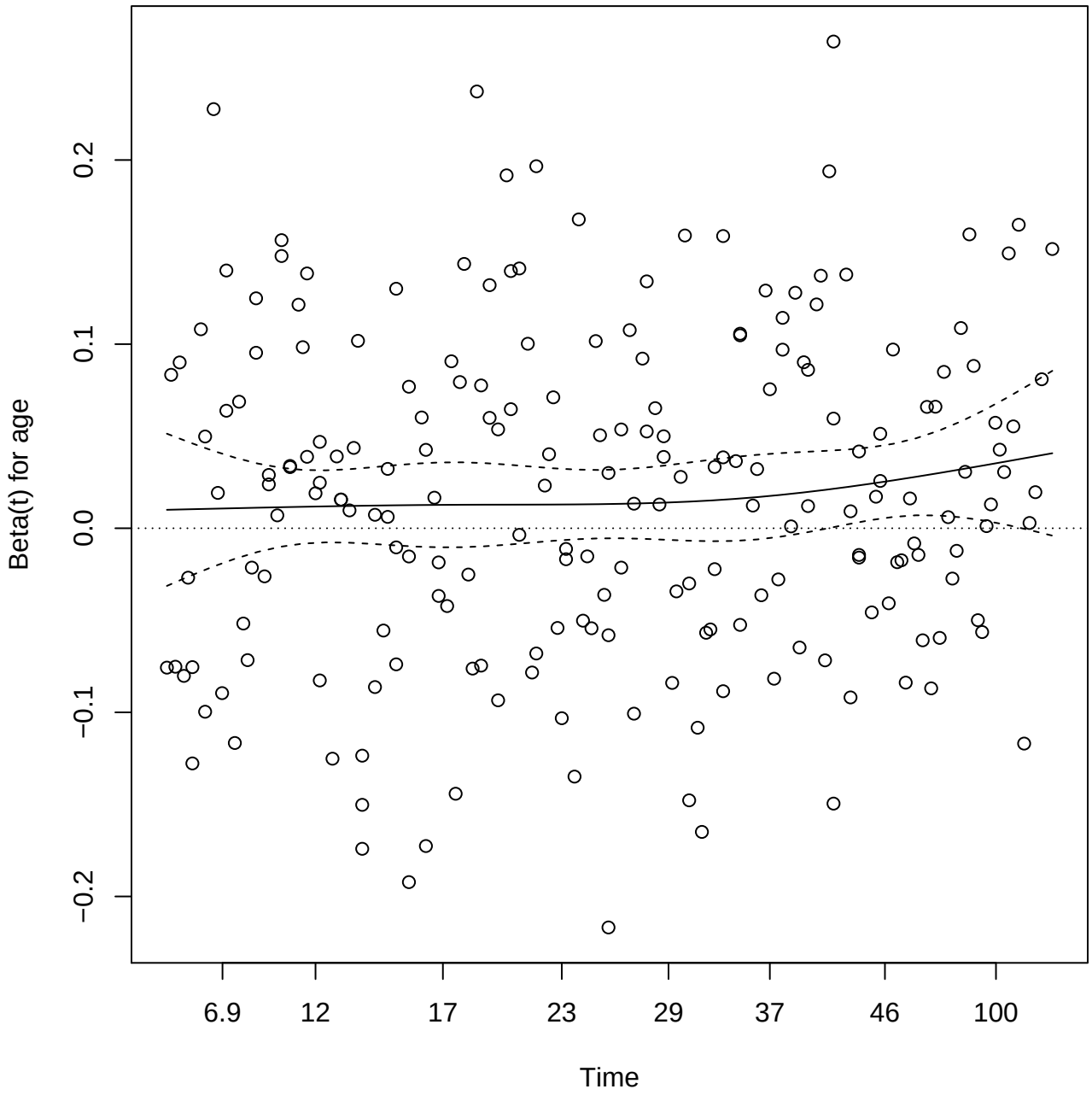
Testing proportionality

- In addition to plotting the Schoenfeld residuals against time, one could test whether there is a significant correlation with time.
- Such tests in R are calculated for each covariate by the `cox.zph` function.
- `plot.cox.zph` produces residual plots of scaled Schoenfeld residuals against time or some transformation.
- Another way to test proportionality would be to add covariate-time interaction terms into the model, and test whether these are significantly different from zero. In the R `coxph` function, such terms can be added using the `tt` argument.

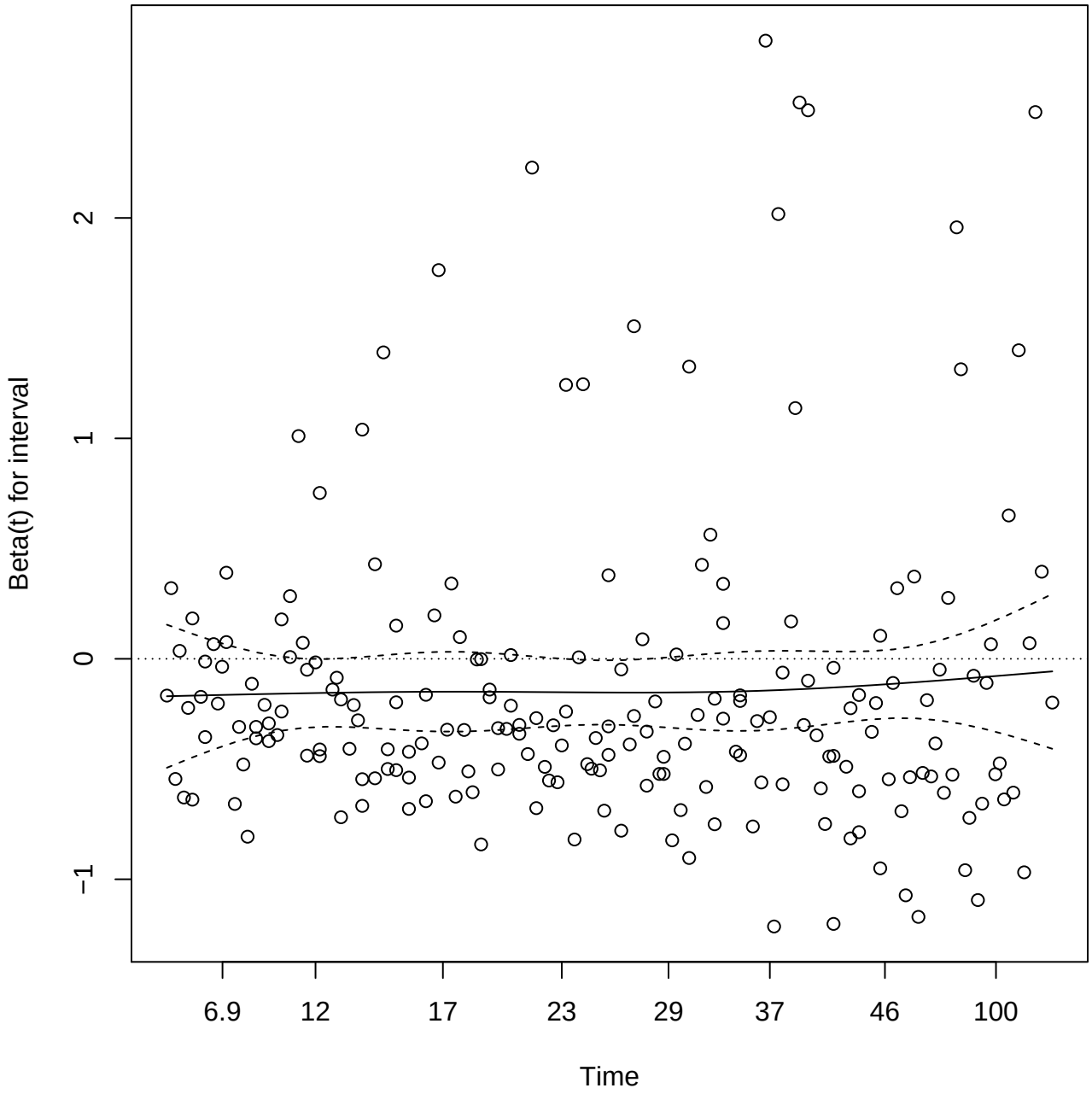
Tests for proportionality

```
1> cox.zph(model, global=FALSE)
                    rho   chisq      p
treat           0.05587 0.65886 0.4170
resect75       -0.13222 3.88410 0.0487
age             0.07592 1.26678 0.2604
interval        0.03134 0.20841 0.6480
karn            0.05952 0.87011 0.3509
race           -0.07737 1.24277 0.2649
local           0.01268 0.03645 0.8486
male           -0.15205 5.26865 0.0217
nitro          -0.03945 0.36563 0.5454
factor(path)2  -0.03441 0.25958 0.6104
factor(path)3  -0.07641 1.26805 0.2601
factor(path)4   0.06396 0.84754 0.3572
grade          -0.00283 0.00193 0.9649
```

Residual plot for age returned by `plot.cox.zph`



Plot for interval returned by `plot.cox.zph`



Plot for male returned by `plot.cox.zph`