Test of Time-Varying Effects for Cox's Regression Model

Thomas H. Scheike
Department of Biostatistics
University of Copenhagen
email: ts@biostat.ku.dk
URL: www.biostat.ku.dk/~ts

OUTLINE :

- Proportional models.
- Test for proportionality.
- Semi-parametric modelling.

Framework

Given iid counting process data $N_i(t)$ observed over some period $[0, \tau]$.

Modelling of the intensity :

$$\lambda_i(t) = E(dN_i(t)|Y_i(t), F_i(t))$$

Y_i is a risk indicator and $F_i(t)$ contains the history of the ith subjects. With covariates, $x_i(t)$, $F_i(t)$ contains history of $Y_i(t)$, $x_i(t)$, and $N_i(t)$.

Standard analysis "Cox-regression"

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp(x_i'(t)\beta)$$

Time-varying Effects

Often effects will by time-varying.

In bio-medical applications: treatment effects often have effects that are diminishing over time.

Therefore important with test for time-varying effects and other model validation.

Regression modelling

Two regression models:

$$x_i^T(t)\beta \quad (1)$$

$$x_i^T(t)\beta(t) \quad (2)$$

- If all effects of (2) are constant then (1) will give the correct summary of effects.
- If all effects are not constant then (1) will give biased results.
- Model (2) always gives sensible description of data, but model may be too large.

The preferred model will often be the semi-parametric

$$x_i^T\beta(t) + z_i^T\gamma$$

where **both** $\beta(t)$ **and** γ **are multidimensional**.

Strategy

- Estimation and asymptotics for

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp(X_{i1}^T\beta_1(t) + Z_i^T\gamma)$$

- Test and simultaneous confidence intervals may be based on asymptotic distribution of estimators.
- Resampling of residuals (Lin, Wei, Ying) SM.

MSS, SM semiparametric kernel estimation technique.

Nonparametric model: Zucker and Karr, Pons, Sun and Cai, **Murphy and Murphy and Senn (sieve)**.

Testing for proportionality

With operational semi-parametric model we wish to compare

$$\lambda_i(t) = Y_i(t)h(X_{i1}^T\beta_1(t) + X_{i2}\beta_2(t) + Z_i^T\gamma)$$

and under $H_0 : \beta_2(t) \equiv \eta$

$$\lambda_i(t) = Y_i(t)h(X_{i1}^T\beta_1(t) + X_{i2}\eta + Z_i^T\gamma).$$

Successive tests !!!

Estimate $B(t) = \int_0^t \beta(s)ds$ and γ.

A simple test based on $\hat{B}_2(\cdot)$, only, is to compute

$$F_3(\hat{B}_2(\cdot)) = \sup_{t\in[0,\tau]} |\hat{B}_2(t) - \hat{B}_2(\tau)\frac{t}{\tau}|$$

or alternatively

$$\int_0^\tau (\hat{B}_2(t) - \hat{B}_2(\tau)\frac{t}{\tau})^2 dt.$$

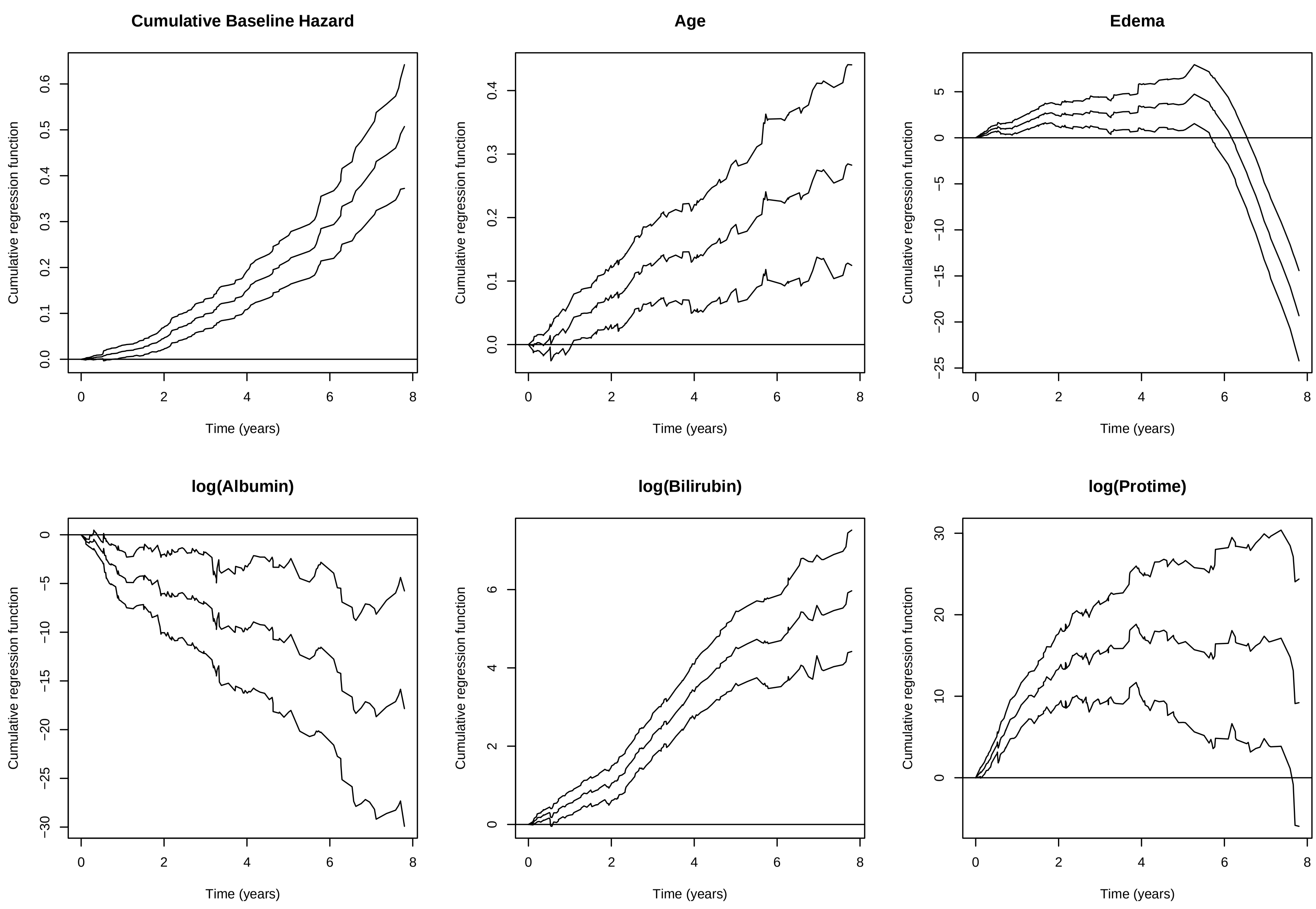
Asymptotic distribution or resampling to evaluate test statistics.

PBC data

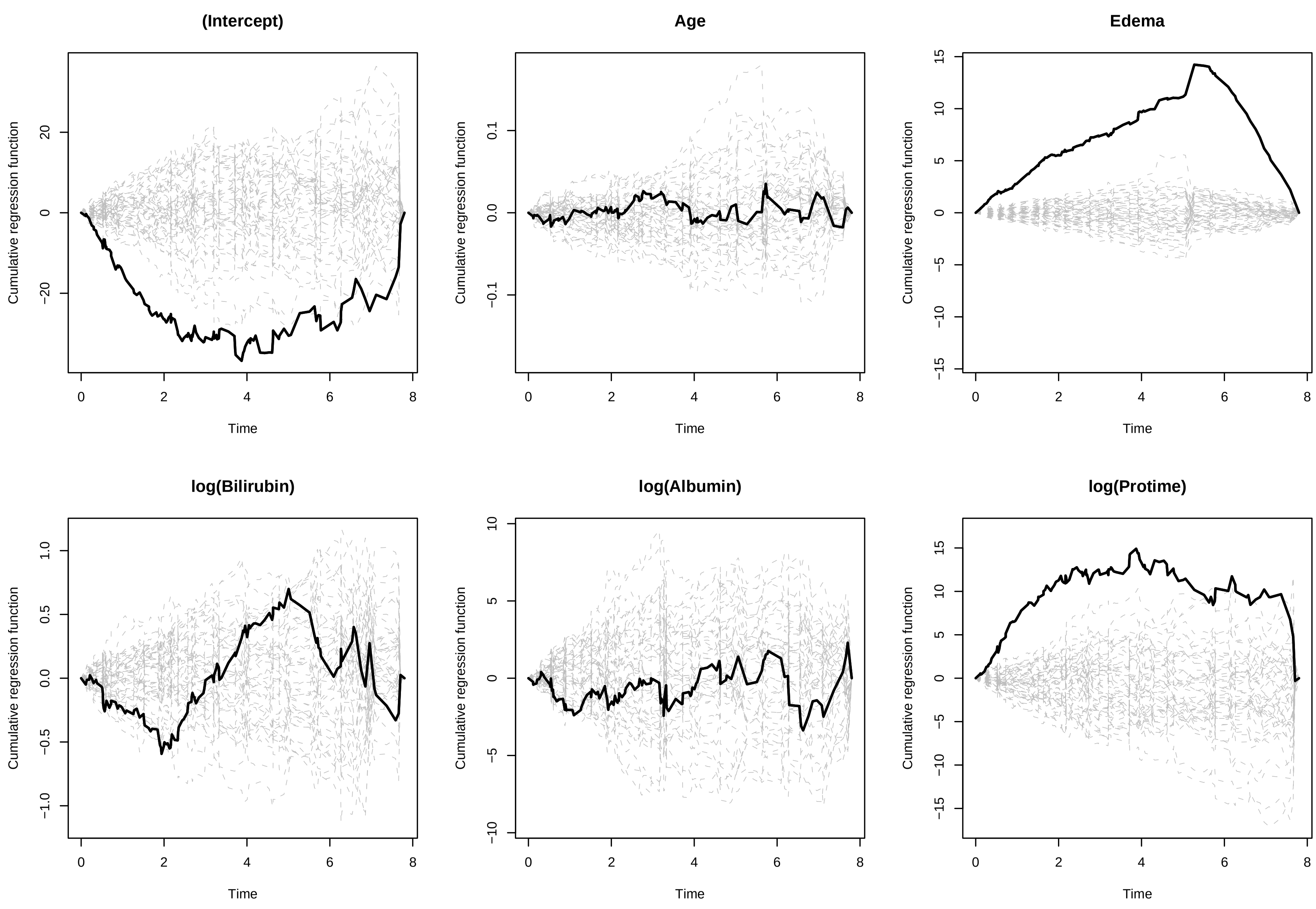
Fleming & Harington (1991):
Mayo Clinic trial in primary biliary cirrhosis (PBC) of the liver and the study was conducted between 1974 and 1984. A total of 418 patients are included in the data and followed until death or censoring.

We restrict attention to the first 3000 days of the study, where we censor all patients being at risk.

143 deaths.



Cumulative Baseline Hazard
Age
Edema
log(Albumin)
log(Bilirubin)
log(Protime)
Cumulative regression function
Time (years)


(Intercept)
Age
Edema
log(Bilirubin)
log(Albumin)
log(Protime)
Cumulative regression function
Time

Previous tests

Previous tests not successive !

Assess overall fit of Cox model. Formally (in the framework of this talk) they test

$$x_i'(t)\beta(t) \to x_i'(t)\beta.$$

Formal tests consider the model

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp(X_i^T\beta + (X_i f_i(t))^T\theta)$$

and use the test for $H : \theta = 0$.

Tempting to consider

$$\lambda_i(t) = Y_i(t)\alpha(t)\exp(X_i^T\beta + Z_i\gamma + Z_i\log(t)\theta)$$

and then use $H : \theta = 0$ as a test for time-varying effects of Z_i.

Schoenfeld residuals

The score with respect to θ (around Cox model $\hat{\beta}$) of

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp(X_i^T\beta + (X_i f_i(t))^T\theta)$$

is

$$\int G(t)\left(X_i(t)) - E(t,\hat{\beta})\right) dN_i(s)$$

with $G(t) = \text{diag}(f(t))$
$E(t,\beta) = \sum Y_i(t)X_i\exp(X_i^T\beta)/\sum Y_i(t)\exp(X_i^T\beta)$

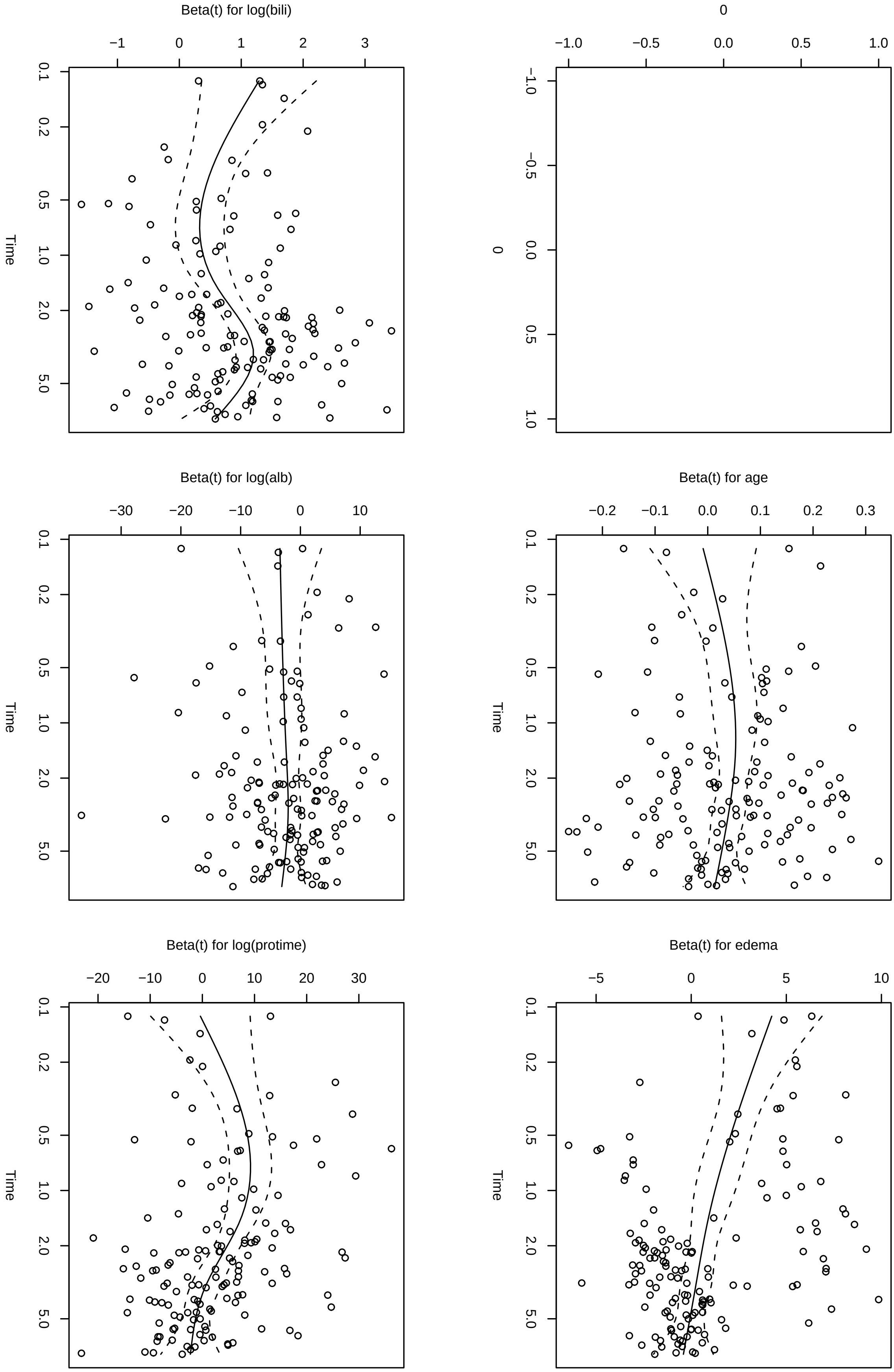
Therefore if Cox score for all t

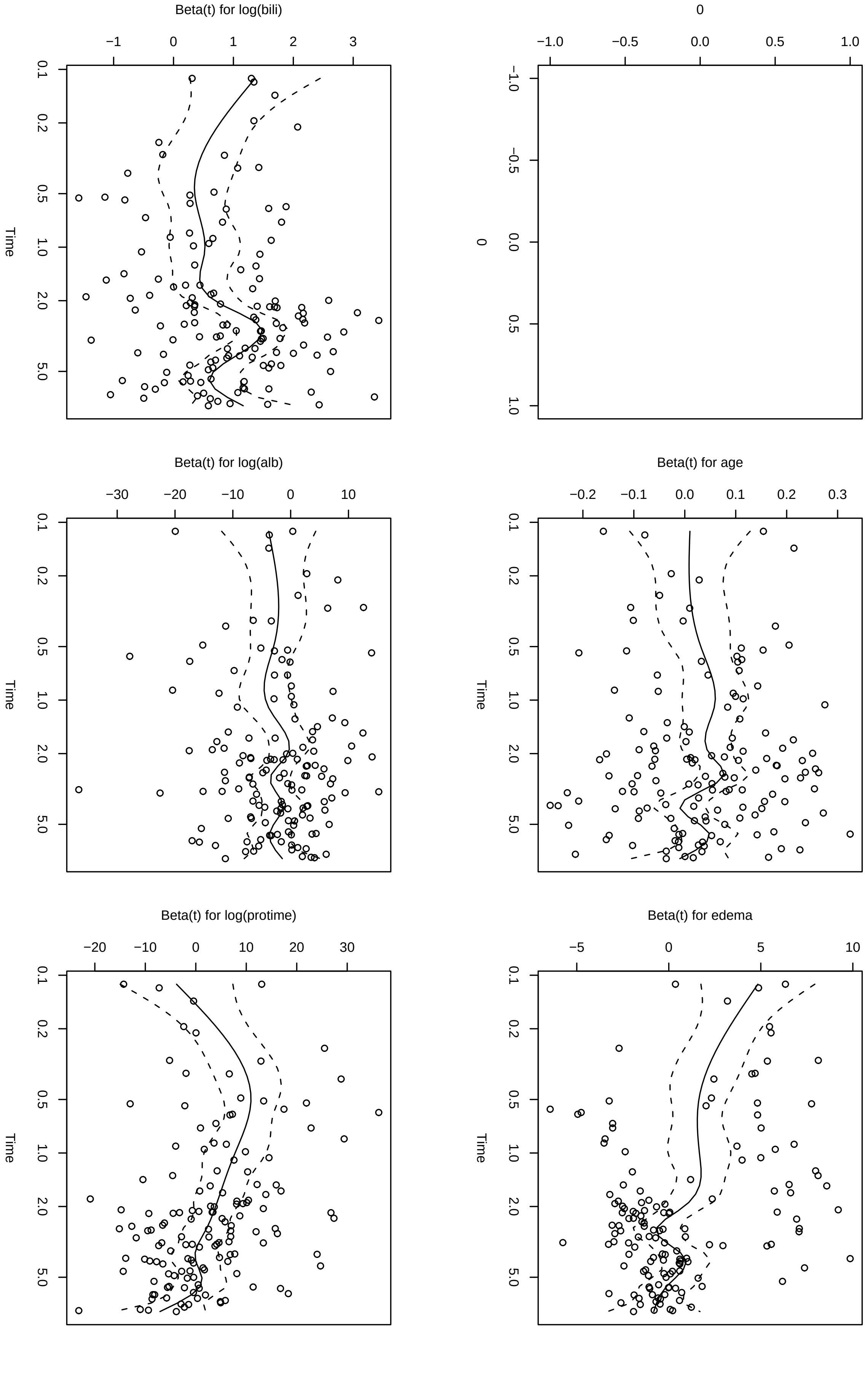
$$\int_0^t \left(X_i(t)) - E(t,\hat{\beta})\right) dN_i(s) = 0$$

or Schoenfeld residuals=0, then all effects are proportional.

- GT smooth Schoenfeld residuals to see if they are 0.
- Smoothing will give a first order estimate of $\beta(t)$ around $\hat{\beta}$.
- Iteration with $E(t,\hat{\beta})$ replaced with $E(t,\hat{\beta}^p(s))$ gives consistency (see MSS or WS).

```
> cox<-coxph(Surv(time/365,status)~age+edema+log(bili)+
+ log(alb)+log(protime),robust=T,pbc3000);
> summary(cox);
Call:
coxph(formula = Surv(time/365, status) ~ age + edema + log(bili) +
    log(alb) + log(protime), data = pbc3000, robust = T)

  n= 418

                coef exp(coef) se(coef) robust se     z       p
age           0.0362    1.0368  0.00806   0.00976  3.71 0.00021
edema         0.6841    1.9820  0.21491   0.25551  2.68 0.00740
log(bili)     0.8646    2.3740  0.08495   0.08957  9.65 0.00000
log(alb)     -2.4588    0.0855  0.67533   0.67195 -3.66 0.00025
log(protime)  2.6603   14.3002  0.85490   0.92193  2.89 0.00390

             exp(coef) exp(-coef) lower .95 upper .95
age             1.0368     0.9645    1.0172     1.057
edema           1.9820     0.5045    1.2012     3.270
log(bili)       2.3740     0.4212    1.9918     2.830
log(alb)        0.0855    11.6903    0.0229     0.319
log(protime)   14.3002     0.0699    2.3474    87.115

Rsquare= 0.403   (max possible= 0.979 )
Likelihood ratio test= 215  on 5 df,   p=0
Wald test            = 201  on 5 df,   p=0
Score (logrank) test = 276  on 5 df,   p=0,   Robust = 99  p=0

  (Note: the likelihood ratio and score tests assume independence of
     observations within a cluster, the Wald and robust score tests do not).
>
> time.test<- cox.zph(cox,transform="log")
> print(time.test)
                 rho   chisq        p
age          -0.0110  0.0203 8.87e-01
edema        -0.2833 15.0682 1.04e-04
log(bili)     0.1127  1.6262 2.02e-01
log(alb)      0.0210  0.0663 7.97e-01
log(protime) -0.2561  8.6255 3.31e-03
GLOBAL            NA 33.0293 3.71e-06
```


Beta(t) for log(bili)
0
Beta(t) for log(alb)
Beta(t) for age
Beta(t) for log(protime)
Beta(t) for edema
Time
0.1
0.2
0.5
1.0
2.0
5.0



Beta(t) for log(bili)
0
Beta(t) for log(alb)
Beta(t) for age
Beta(t) for log(protime)
Beta(t) for edema
Time

Schoenfeld residuals

To construct tests based on the Schoenfeld residuals one may cumulate them, this gives the cumulative score process.

Lin, Wei, and Ying :

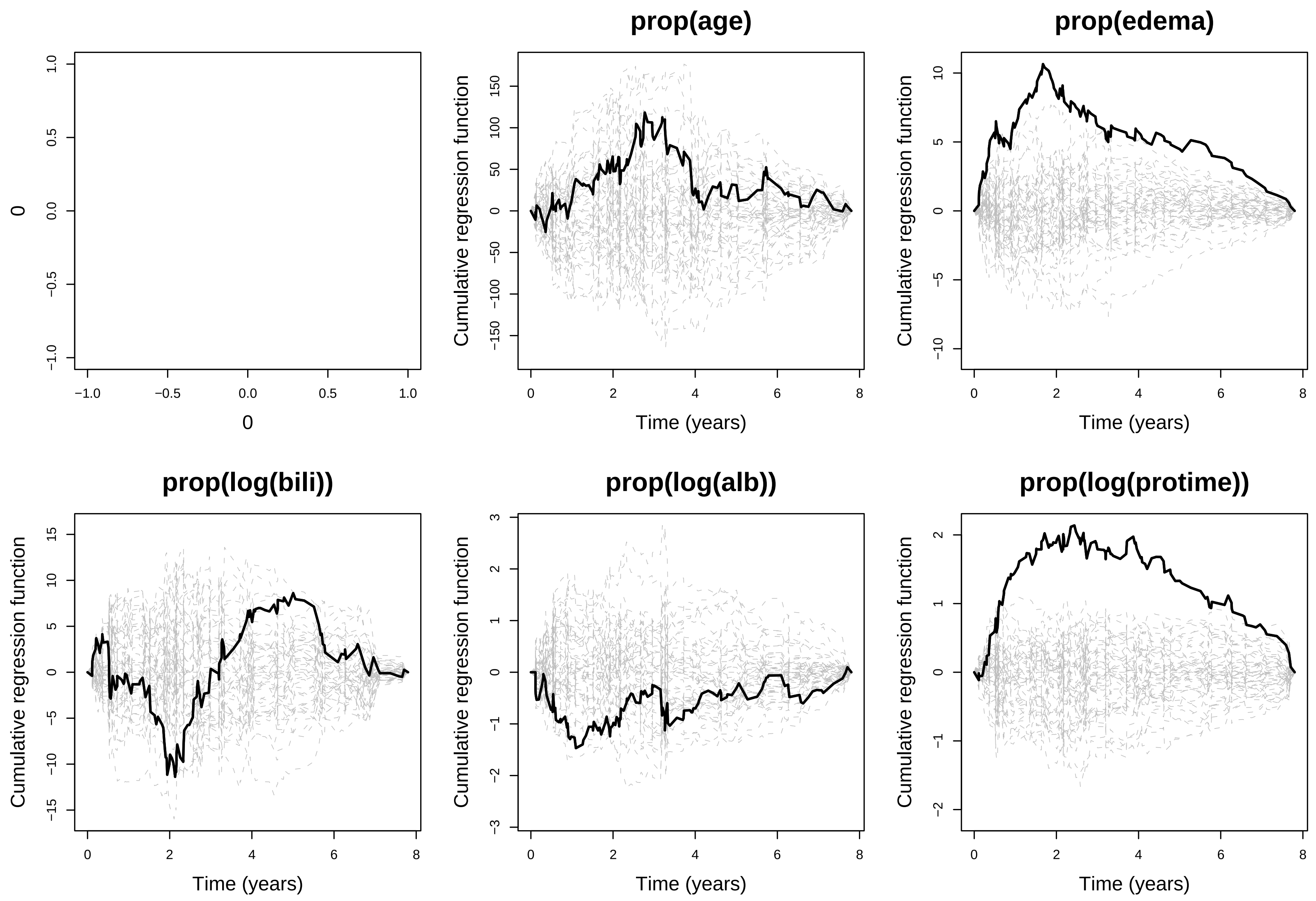
Observed score process under Cox model

$$S_j = \sup_t |U_j(\hat{\beta}, t)| \quad j = 1, .., p$$

whose percentiles can be simulated under the null.

LWY cumulative score processes:

```
> library(survival)
> #data(pbc);
> ourcoxw<-cox.aalen(Surv(time/365,status)~prop(age)+prop(edema)+
+ prop(log(bili))+prop(log(alb))+prop(log(protime)),pbc,
+ maxtime=3000/365,weighted.score=0)
Non unique survival times: Adds random noise Runif(0.0001)
May cause difficulties for counting process data
Right censored survival times
 Cox-Aalen Survival Model
Simulations starts N= 1000
> summary(ourcoxw)
Cox-Aalen Model

Test for Aalen terms
Test for non-siginificant effects
            sup| hat B(t)/SD(t) | p-value H_0: B(t)=0
(Intercept)               0.4193               0.845
Test for time invariant effects
            sup| B(t) - (t/tau)B(tau)| p-value H_0: B(t)=b t
(Intercept)                     3e-04                 0.944

Proportional Cox terms :
                     Coef. Std. Error (Sandwich) Robust Std.  Error
prop(age)           0.0362                0.0072            0.0098
prop(edema)         0.6832                0.2069            0.2559
prop(log(bili))     0.8639                0.0770            0.0895
prop(log(alb))     -2.4505                0.6850            0.6706
prop(log(protime))  2.6576                0.7454            0.9219
                   D2log(L)^-(1/2)
prop(age)                   0.0081
prop(edema)                 0.2150
prop(log(bili))             0.0849
prop(log(alb))              0.6751
prop(log(protime))          0.8552

Score Test for Proportionality
                   sup| hat U(t) | p-value H_0: U(t) Proportional
prop(age)                118.2769                           0.206
prop(edema)               10.6493                           0.000
prop(log(bili))           11.3826                           0.254
prop(log(alb))             1.4691                           0.315
prop(log(protime))         2.1379                           0.001


  Call: cox.aalen(Surv(time/365, status) ~ prop(age) + prop(edema) +
    prop(log(bili)) + prop(log(alb)) + prop(log(protime)), pbc,
    maxtime = 3000/365, weighted.score = 0)

> par(mfrow=c(2,3))
> plot(ourcoxw,score=T,xlab="Time (years)")
```


0
0
−1.0
−0.5
0.0
0.5
1.0
prop(age)
prop(edema)
prop(log(bili))
prop(log(alb))
prop(log(protime))
Cumulative regression function
Time (years)
−150
−100
−50
50
100
150
−10
−5
5
10
−15
15
−3
−2
−1
1
2
3
4
6
8

```
> ourcoxw<-cox.aalen(Surv(time/365,status)~prop(age)+prop(edema)+
+ prop(log(bili))+prop(log(alb))+prop(log(protime)),pbc,
+ maxtime=3000/365,weighted.score=1)
Non unique survival times: Adds random noise Runif(0.0001)
May cause difficulties for counting process data
Right censored survival times
 Cox-Aalen Survival Model
Simulations starts N= 1000
> summary(ourcoxw)
Cox-Aalen Model

Test for Aalen terms
Test for non-siginificant effects
            sup| hat B(t)/SD(t) | p-value H_0: B(t)=0
(Intercept)                0.4194              0.872
Test for time invariant effects
            sup| B(t) - (t/tau)B(tau)| p-value H_0: B(t)=b t
(Intercept)                      3e-04               0.956

Proportional Cox terms :
                     Coef. Std. Error (Sandwich) Robust Std.  Error
prop(age)           0.0361                0.0072            0.0097
prop(edema)         0.6861                0.2066            0.2556
prop(log(bili))     0.8647                0.0771            0.0896
prop(log(alb))     -2.4550                0.6854            0.6700
prop(log(protime))  2.6555                0.7452            0.9218
                   D2log(L)^-(1/2)
prop(age)                   0.0081
prop(edema)                 0.2150
prop(log(bili))             0.0850
prop(log(alb))              0.6751
prop(log(protime))          0.8551

Score Test for Proportionality
                   sup| hat U(t) | p-value H_0: U(t) Proportional
prop(age)                   1.9679                            0.450
prop(edema)                 4.3920                            0.000
prop(log(bili))             2.1117                            0.431
prop(log(alb))              2.2699                            0.289
prop(log(protime))          4.0079                            0.001


  Call: cox.aalen(Surv(time/365, status) ~ prop(age) + prop(edema) +
    prop(log(bili)) + prop(log(alb)) + prop(log(protime)), pbc,
    maxtime = 3000/365, weighted.score = 1)
```

The LWY score processes are compared to resampled processes under null.

$$\begin{aligned} M_i(t) &= N_i(t) - \int_0^t Y_i(s)\exp(X^T(t)\beta)\lambda_0(s)ds \\ &= N_i(t) - \int_0^t Y_i(s)\exp(X^T(t)\beta)d\Lambda_0(s) \end{aligned}$$

$$\begin{aligned} U(\hat{\beta}, t) &= \sum_{i=1}^{n} \int_0^t X_i(s)d\hat{M}_i(s) \\ &= \sum_{i=1}^{n} \int_0^t (X_i(s) - E(t, \hat{\beta}))d\hat{M}_i(s). \end{aligned}$$

Lin, Wei and Ying

$$\sum_{i=1}^{n} \int_0^t (X_i(s) - E(t, \hat{\beta}))dN_i(s)G_i$$

robust alternative

$$\sum_{i=1}^{n} \int_0^t (X_i(s) - E(t, \hat{\beta}))d\hat{M}_i(s)G_i$$

Simulation

Observed rejection probabilities for $H_0 : \beta_j(t) \equiv \gamma_j$.

Observed Power									
		$\theta g(t)$		LWY Test		G_{sup}		G_{is}	
n	Corr(X,Z)	$\beta_1(t)$	$\beta_2(t)$	$\beta_1(t)$	$\beta_2(t)$	$\beta_1(t)$	$\beta_2(t)$	$\beta_1(t)$	$\beta_2(t)$
200	0.0	82.1	6.8	100	5.0	76.2	5.9	81.5	5.4
200	0.1	85.5	7.1	100	6.8	76.0	5.8	81.7	5.6
200	0.3	88.3	4.9	100	19.5	77.2	7.0	80.5	7.6
200	0.7	94.3	6.6	100	89.3	73.1	6.6	75.2	6.2
400	0.0	98.3	10.0	100	6.0	97.2	6.6	98.9	5.9
400	0.1	99.4	9.1	100	100	97.6	6.6	99.2	5.3
400	0.3	99.7	9.0	100	100	96.4	5.3	98.6	5.9
400	0.7	100	8.0	100	100	98.1	6.7	98.8	6.7
800	0.0	100	14.0	100	6.8	100	5.7	100	5.6
800	0.1	100	13.3	100	100	100	5.3	100	5.2
800	0.3	100	11.0	100	83.7	100	5.6	100	5.4
800	0.7	100	14.0	100	100	100	5.9	100	5.8

Model

$$\lambda_i(t) = \exp(-1/2 + \beta_1(t)X_i + \beta_2(t)Z_i)$$

where (X_i, Z_i) are multivariate normal with mean 0, standard deviation 2, and with varying correlation (0,0.1, 0.3 and 0.7) and where $\beta_1(t) = 0.35 \cdot \sin(2t)$ for $t < \pi/2$ and $\beta_1(t) = \sin(2t)$ for $t \geq \pi/2$ and $\beta_2(t) \equiv 0.3$.

Censoring at time 3.

Graphical Procedure

A typical graphical procedure for investigating if Z_i has a time-varying effect is based on stratifying Z_i into S strata and then fitting the stratified model

$$\lambda_i(t) = Y_i(t)\alpha_s(t)\exp(X_i\beta) \quad Z_i \quad \in \text{"}s\text{"} \; s = 1, .., S.$$

A closer examination of the estimates of $\alpha_s(t)$ to see if these baseline intensities are consistent with the hypothesized baseline $\alpha(t)\exp(Z_i\gamma)$ is then used to check if the effect of Z_i depends on time.

Discussion

- Traditional test only access overall fit of model.
- Relatively simple to perform successive tests (?!).