Outline

- Cox's proportional hazards model.
- Goodness-of-fit tools
- More flexible models
- R-package `timereg`
- Forthcoming book, Martinussen and Scheike.

University of Copenhagen

http://www.biostat.ku.dk

Goodness-Of-Fit for Cox's Regression Model. Extensions of Cox's Regression Model.

Survival Analysis Fall 2004, Copenhagen

Torben Martinussen and Thomas Scheike

torbenm@dina.kvl.dk ts@biostat.ku.dk

Cox's proportional hazards model

In practice one has covariates: X_i (p-dimensional).

Hazard conditional on covariates: $\alpha_i(t, X_i)$.

The absolute dominant model is **Cox's proportional hazards model**:

$$\alpha_i(t) = \alpha_0(t) \exp(\beta^T X_i)$$

where $\alpha_0(t)$ is unspecified baseline hazard (hazard for $X_i = 0$).

Flexible model. Model is easily fitted using for example **SAS** or **R** (**S-Plus**).

Primary model for survival data because of its nice properties.

Suppose X is 1-dim. (or fix other covariates) then the relative risk

$$\frac{\alpha(t, X+1)}{\alpha(t, X)} = \exp(\beta)$$

is not depending on time (key assumption)!

This assumption is often violated !

Survival analysis

Standard setup for right-censored survival data. **IID** copies of (T, D) where

$$T = T^* \wedge C \quad D = I(T^* \leq C)$$

with T^* being the true survival time and C the (potential) censoring time and possibly covariates $X_i(t)$.

Hazard-function

$$\alpha(t) = \lim_{h \downarrow 0} \frac{1}{h} P(t \leq T^* < t + h \,|\, T^* \geq t, \mathcal{F}_{\sqcup -}).$$

Counting process

$$N_i(t) = I(T_i \leq t, D_i = 1)$$

Martingale

$$M_i(t) = N_i(t) - \Lambda_i(t)$$

where

$\Lambda_i(t) = \int_0^t Y_i(s)\alpha(s)\, ds$ (**compensator**), $\quad Y_i(t) = I(t \leq T_i)$ (**at risk process**).

```
> library(survival)
> data(pbc)
> attach(pbc)
> cbind(time,status,age,edema,bili,protime,alb)[1:5,]
     time status     age edema bili protime  alb
[1,]  400      1 58.7652     1 14.5    12.2 2.60
[2,] 4500      0 56.4463     0  1.1    10.6 4.14
[3,] 1012      1 70.0726     1  1.4    12.0 3.48
[4,] 1925      1 54.7406     1  1.8    10.3 2.54
[5,] 1504      0 38.1054     0  3.4    10.9 3.53
> sum(status)
[1] 161
> fit.pbc<-coxph(Surv(time/365, status) ~ age+edema+log(bili)+log(protime)+log(alb))
> fit.pbc
Call:
coxph(formula = Surv(time/365, status) ~ age + edema + log(bili) + log(protime) + log(alb))
                coef exp(coef) se(coef)     z       p
age           0.0382     1.039  0.00768  4.97 6.5e-07
edema         0.6613     1.937  0.20595  3.21 1.3e-03
log(bili)     0.8975     2.453  0.08271 10.85 0.0e+00
log(protime)  2.3458    10.442  0.77425  3.03 2.4e-03
log(alb)     -2.4524     0.086  0.65707 -3.73 1.9e-04
Likelihood ratio test=234  on 5 df, p=0  n= 418
```


PBC-data

PBC data (primary biliary cirrhosis): 418 patients are followed until death or censoring.

PBC is a fatal chronic liver disease.

Important explanatory variables:

- Age
- Albumin
- Bilirubin
- Edema
- Prothrombin time

Fitting Cox's model in R.

Graphical GOF for Cox's regression model

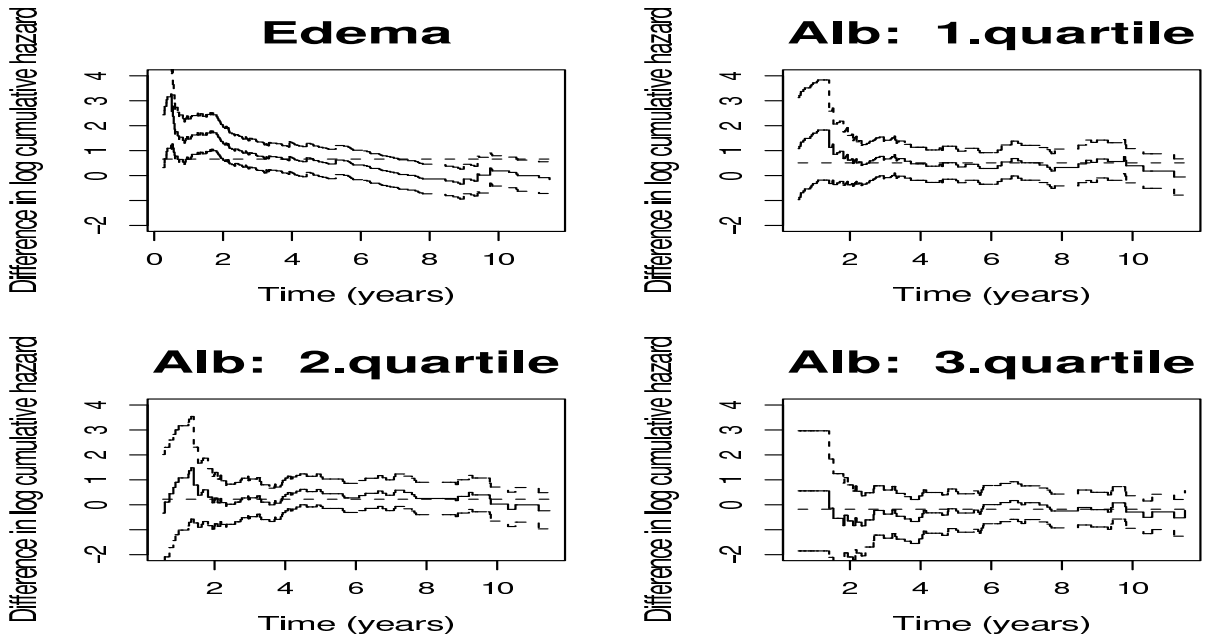


Figure 1: Estimated log-cumulative hazards difference along with 95% pointwise confidence intervals. The straight lines (dashed lines) are based on the Cox-model.

Cox's proportional hazards model

Traditional **goodness-of-fit tools**. Model:

$$\alpha_i(t) = \alpha_0(t)\exp(\beta_1 X_{i1} + ... + \beta_p X_{ip})$$

Investigate if each of the covariates are consistent with the proportional hazards assumption. Stratify based on a grouping (k=1,...,K) based on X_{i1}'s values:

$$\alpha_i(t) = \alpha_{0k}(t)\exp(\beta_2 X_{i2} + ... + \beta_p X_{ip}); \quad \text{if} X_{i1} \in A_k$$

Now, if the underlying full Cox-model is true the baseline estimates $\alpha_{0k}(t)$ should be proportional, as

$$\alpha_{0k}(t) = \alpha_0(t)\exp(\sum_{k=1}^{K} \beta_{1k} I(X_{i1} \in A_k)).$$

Graphical model-check of proportionality by making graphs of estimates of $\log(\int_0^t \alpha_{0k}(s)ds)$. Plotted against t they should be parallel.

Cumulative martingale residuals

Alternative: **Cumulative martingale residuals**, (Lin et al., 1993).

The martingales under the Cox regression model can be written as

$$M_i(t) = N_i(t) - \int_0^t Y_i(s)\exp(X^T\beta)d\Lambda_0(s);\ \hat{M}_i(t) = N_i(t) - \int_0^t Y_i(s)\exp(X^T\hat{\beta})d\hat{\Lambda}_0(s)$$

The score function, evaluated in the estimate $\hat{\beta}$, and seen as a function of time, can for example be written as

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

and is asymptotically equivalent to a Gaussian process (not a martingale) that can easily be simulated (LWY,93). Can now proceed to suggest some appropriate test statistic like

$$\sup_{t\in[0,\tau]} |U(\hat{\beta}, t)|$$

This is essentially a test for time-constant effects of all covariates !!

Cox's proportional hazards model

Traditional **goodness-of-fit tools**.

Make tests against specific deviations: Replace X_1 with $(X_1, X_1(\log(t)))$, say ($\beta_1 \to \beta_1 + \beta_{p+1} \cdot \log(t)$).

Test the null $\beta_{p+1} = 0$.

These methods are quite useful but also have some limitations :

- Graphical method:
  - Not parallel. What is acceptable?
  - What if a given covariate is continuous?
- Test: Ad hoc method. Which transformation to use?
- Both methods: They assume that model is ok for all the other covariates.

Resampling technique

The observed score process is given as $U(\hat{\beta}, t)$ and its asymptotic distribution is equivalent to the asymptotic distribution of

$$\sum_i \left(\hat{W}_{1i}(t) + \mathcal{I}(\hat{\beta}, t)\mathcal{I}^{-1}(\hat{\beta}, \tau)\hat{W}_{1i}(\tau) \right) G_i$$

where $G_1, ..., G_n$ are independent standard normals, and independent of the observed data.

Score process

The score process evaluated at β_0 can be written as

$$U(\beta_0, t) = \sum_i W_{1i}(t) \tag{1}$$

where

$$W_{1i}(t) = \int_0^t \left(Z_i - Z^T Y(\beta_0)(Y^T(\beta_0) W Y(\beta_0))^{-1} Y_i \right) dM_i(s),$$

with

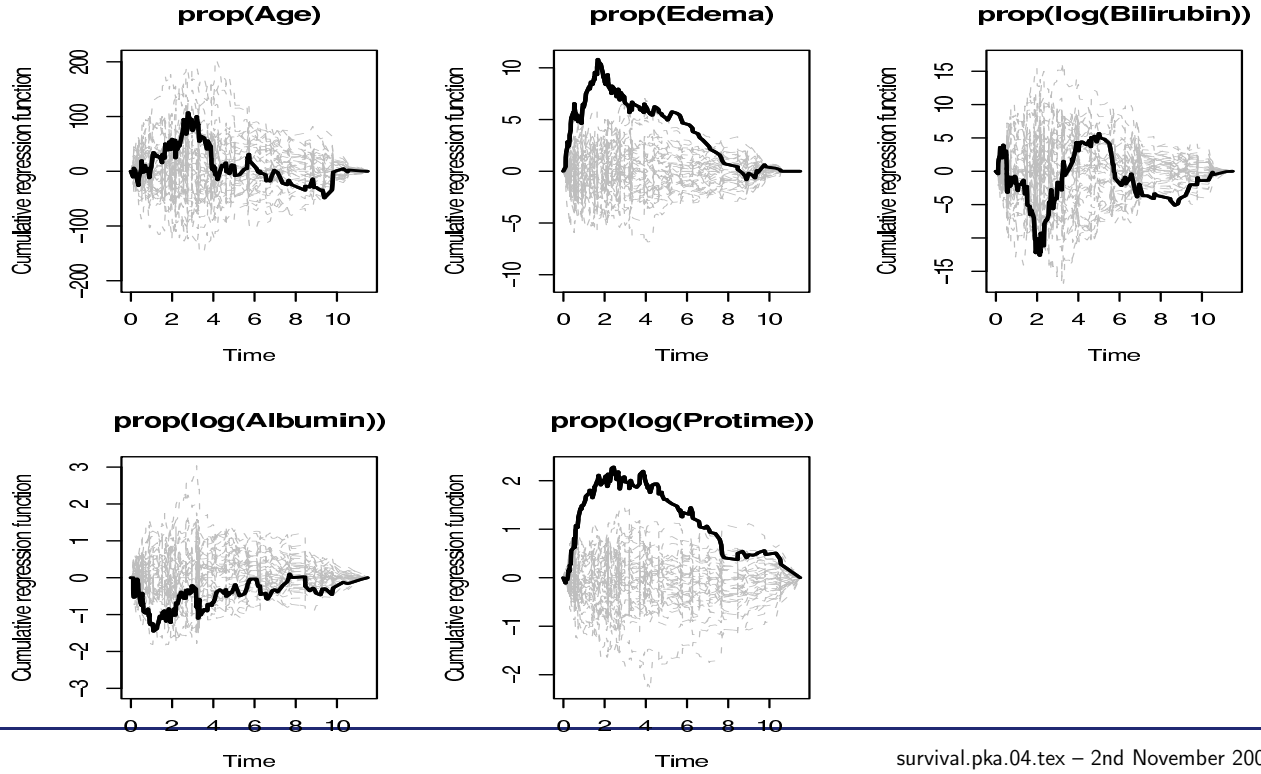
$$Y(\beta, t) = (Y_1 \exp(Z_1^T \beta), ..., Y_n \exp(Z_n^T \beta)) \quad W(t) = \text{diag}(Y_i \exp(-Z_i^T \beta))$$

Variance can then be estimated robustly by

$$\hat{\Sigma}_\beta = n\mathcal{I}^{-1}(\hat{\beta}, \tau) \left\{ \sum_i \hat{W}_{1i}(t)^{\otimes 2} \right\} \mathcal{I}^{-1}(\hat{\beta}, \tau).$$

where $\hat{W}_{1i}$ is obtained by replacing M_i by $\hat{M}_i$.

PBC-data



PBC-data

```
> library(timereg)
> fit<-cox.aalen(Surv(time/365,status)~prop(Age)+prop(Edema)+
+ prop(log(Bilirubin))+prop(log(Albumin))+prop(log(Protime)),
+ weighted.score=0,pbc);
Right censored survival times
Cox-Aalen Survival Model
Simulations starts N= 5000
> summary(ourcox)
Cox-Aalen Model

Score Test for Proportionality
                     sup| hat U(t) | p-value H_0: U(t) Proportional
prop(Age)                  106.9100                          0.3504
prop(Edema)                 10.8582                          0.0002
prop(log(Bilirubin))        12.5583                          0.1606
prop(log(Albumin))           1.4566                          0.3140
prop(log(Protime))           2.2851                          0.0018

> plot(fit,score=T,xlab="Time (years)")
```

Cumulative Residuals

Now,

$$\hat{M}(t) = \int_0^t G(\beta_0, s) dM(s) + \int_0^t \left\{ Y(\beta_0, s) Y^-(\beta_0, s) - Y(\hat{\beta}, s) Y^-(\hat{\beta}, s) \right\} dN(s).$$

The second term can be Taylor series expanded

$$\begin{aligned} & - \left[\int_0^t G(\beta^\star, s) \text{diag} \left\{ Y(\beta^\star, s) Y^-(\beta^\star, s) dN(s) \right\} Z(s) \right] (\hat{\beta} - \beta_0) \\ = & - \left[\int_0^t G(s) \text{diag} \left\{ Y(\beta^\star, s) Y^-(\beta^\star, s) dN(s) \right\} Z(s) \right] \mathcal{I}^{-1}(\beta^{\star\star}, \tau) U(\beta_0, \tau) \end{aligned}$$

where $\beta^\star$ and $\beta^{\star\star}$ are on line segment between $\hat{\beta}$ and β_0.

Therefore $\hat{M}(t)$ is asymptotically equivalent (see below) to

$$M(t) + B(\beta_0, t) M(\tau)$$

Cumulative Residuals

Model can be written as the $n \times 1$ vector

$$N(t) = Y(t) \lambda_0(t) + M(t)$$

where $M(t)$ is the mean-zero martingale, and then the martingales are estimated by

$$\hat{M}(t) = N(t) - \int_0^t Y(\hat{\beta}, s) Y^-(\hat{\beta}, s) dN(s) = \int_0^t G(s) dN(s).$$

where

$$G(\hat{\beta}, s) = I - Y(\hat{\beta}, s) Y^-(\hat{\beta}, s)$$

and

$$Y^-(\beta, s) = (Y^T W Y)^{-1} Y^T W.$$

Cumulative Residuals

$M_U(t)$ is asymptotically equivalent to

$$\sum_{i=1}^{n} \int_0^t U_i(t) - U^T(s)Y(\beta_0, s)\left\{Y^T(\beta_0, s)W(s)Y(\beta_0, s)\right\}^{-1} Y_i(s)dM_i(s)$$
$$- \left[\int_0^t U^T(s)G(s)\text{diag}\left\{Y(\beta_0, s)Y^-(\beta_0, s)dN(s)\right\} Z(s)\right] \sum_{i=1}^{n} W_{1,i} + o_p(n^{1/2}),$$
$$= \sum W_i(t) + o_p(n^{1/2}),$$

where $W_i(t)$ are i.i.d. an $W_{1,i}$ is an i.i.d. decomposition of $\hat{\beta} - \beta_0$.

Resample construction

$$\sum \hat{W}_i(t)G_i,$$

and $G_1, ..., G_n$ are standard normals have the same asymptotic distribution.

Where $\hat{W}_i(t)$ is obtained by using $\hat{M}_i$ instead of M_i.

Cumulative Residuals

A cumulative residual process is then defined by

$$M_U(t) = \int_0^t U^T(t)d\hat{M}(s)$$

and this process is asymptotically equivalent to

$$\int_0^t U^T(t)G(\beta_0, s)dM(s) - \left[\int_0^t U^T(s)G(s)\text{diag}\left\{Y(\beta_0, s)Y^-(\beta_0, s)dN(s)\right\} Z(s)\right](\hat{\beta} - \beta_0).$$

Denote the second integral in the latter display by $B_U(\beta_0, t)$. The variance of $M_U(t)$ can be estimated by the optional variation process

$$[M_U](t) = \int_0^t U^T(s)G(\hat{\beta}, s)\text{diag}(dN(s))U(s)G(\hat{\beta}, s) + B_U(\hat{\beta}, t)[U](\tau)B_U^T(\hat{\beta}, t)$$
$$- B_U(\hat{\beta}, t)[M_U, U](t) - [U, M_U](t)B_U^T(\hat{\beta}, t).$$

Cumulative residuals

```
> # our PBC version with no ties !!!!!!!!!!!!!!!!!!!!
> fit<-cox.aalen(Surv(time/365,status)~prop(Age)+ prop(Edema)+prop(log(Bilirubin))+prop(log(Albumin))+ prop(log(Protime)),pbc,
+ weighted.score=0,resid.mg=1);
Cox-Aalen Survival Model
Simulations starts N= 500
>
> X<-model.matrix(~-1+cut(Bilirubin,quantile(Bilirubin),include.lowest=T),pbc)
> colnames(X)<-c("1. quartile","2. quartile","3. quartile","4. quartile");
>
> resids<-mg.resids(fit,pbc,X,n.sim=1000,cum.resid=1)
> summary(resids)
Test for cumulative MG-residuals

Grouped Residuals consistent with model
            sup| hat B(t) | p-value H_0: B(t)=0
1. quartile           4.632               0.269
2. quartile           5.394               0.307
3. quartile           4.087               0.760
4. quartile           5.062               0.507
            int ( B(t) )^2 dt p-value H_0: B(t)=0
1. quartile            31.234               0.391
2. quartile            60.265               0.328
3. quartile            39.370               0.666
4. quartile            59.831               0.394

Residual versus covariates consistent with model
                     sup| hat B(t) | p-value H_0: B(t)=0
prop(Age)                   6.857               0.635
prop(log(Bilirubin))        9.030               0.146
prop(log(Albumin))          7.998               0.431
prop(log(Protime))          5.525               0.807
```


Cumulative Residuals

(Lin et al., 1993) suggest to cumulate the residuals over the covariate space as well as over time, and thus considers the double cumulative processes

$$M_j(x,t) = \int_0^t K^T(j,x,s)d\hat{M}(s)$$
$$= \int_0^t K^T(j,x,s)G(s)dM(s) \text{ for } j = 1,..,p$$

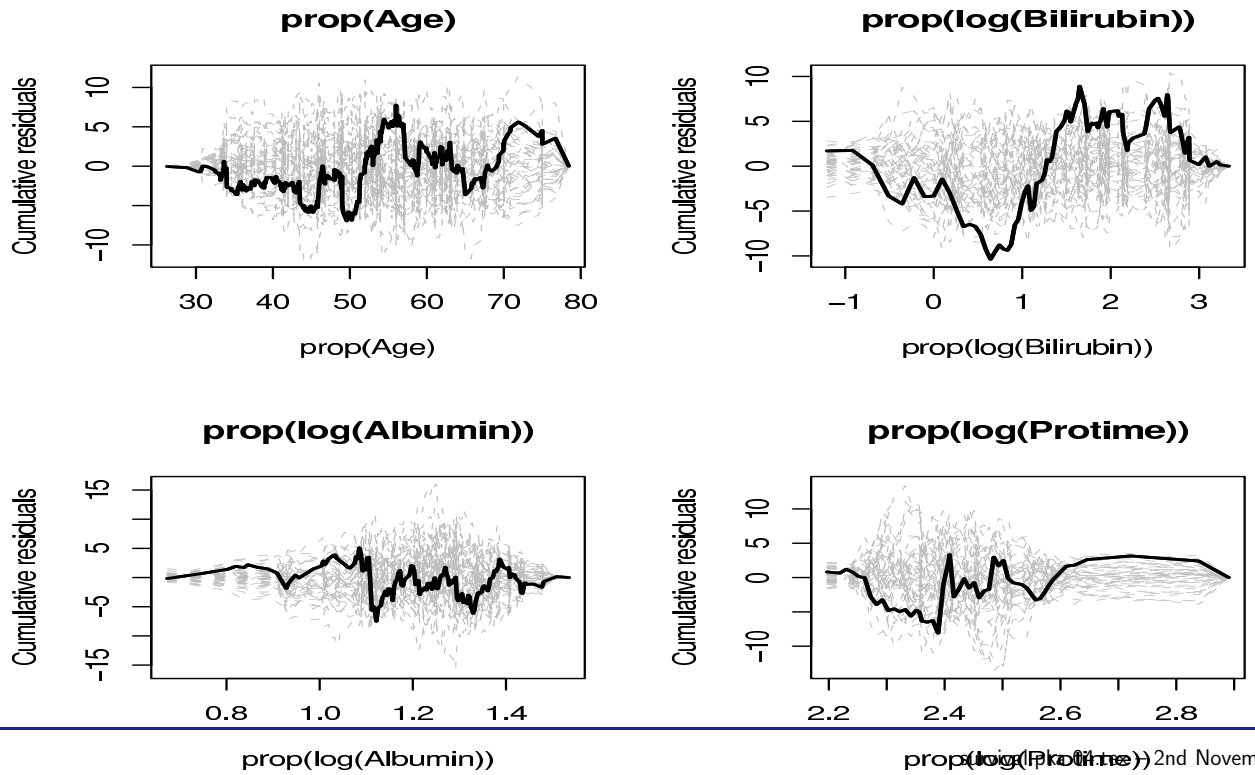
where $K(j,x,t)$ is an $n \times 1$ vector with ith

$$I(Z_{i,j}(t) \leq x) \text{ for } i = 1,..,n.$$

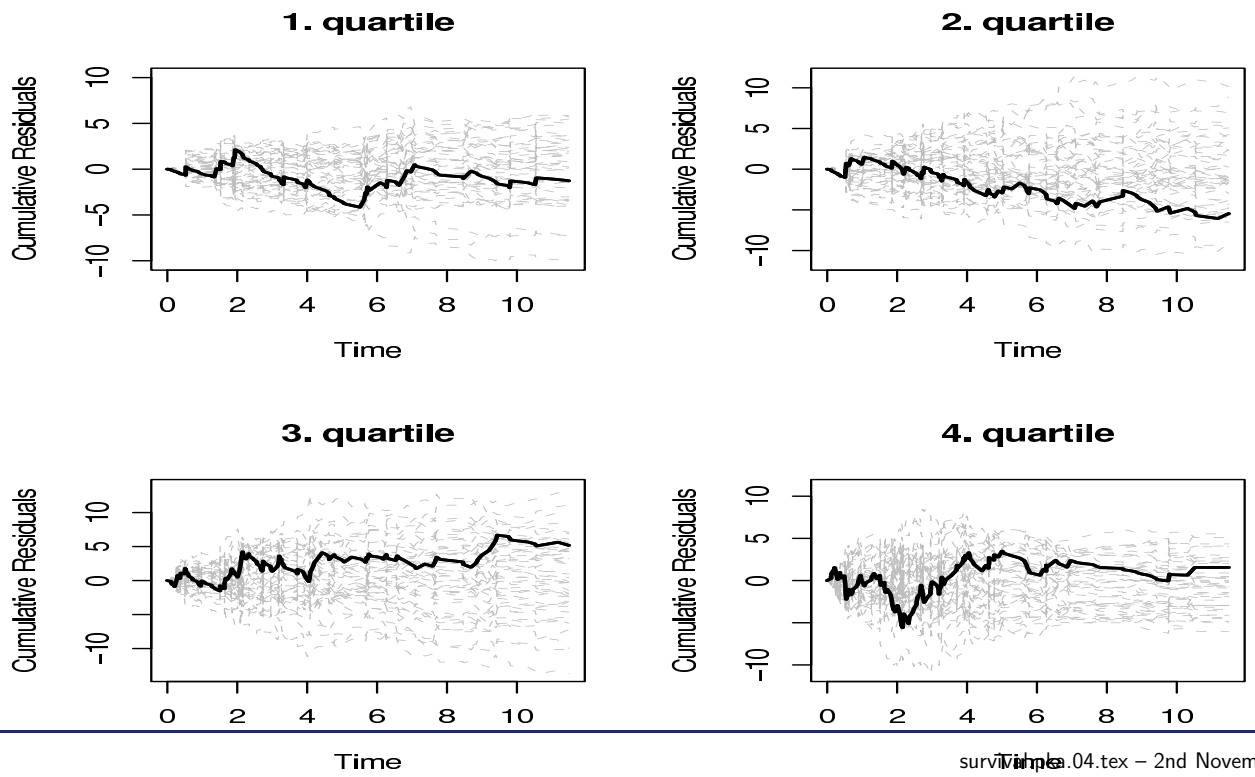
Integrating over time we get a process in x

$$M_j(x) = M_j(x,\tau) \tag{2}$$

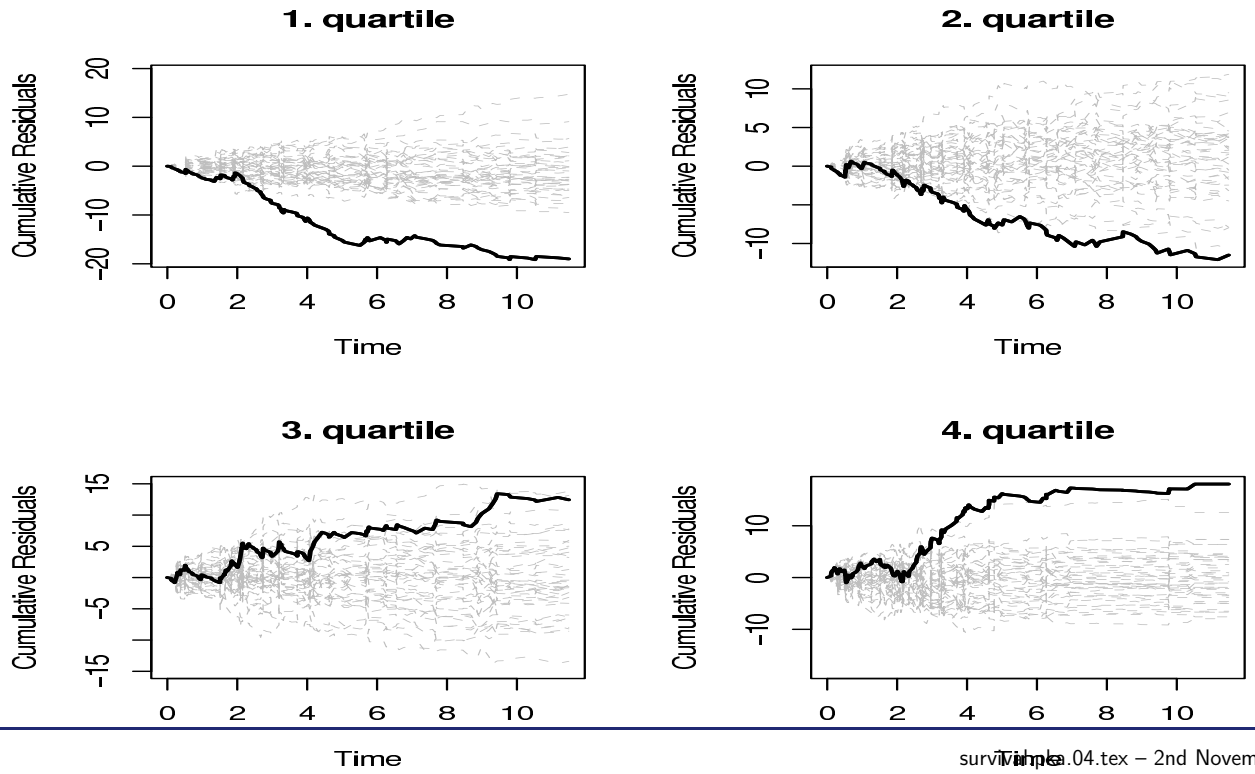
Cumulative martingale residuals



Cumulative martingale residuals



Cumulative martingale residuals



Cumulative residuals

```
> nfit<-cox.aalen(Surv(time/365,status)~prop(Age)+prop(Edema)+prop(Bilirubin)+prop(log(Albumin))+prop(log(Protime)),pbc,
weighted.score=0,resid.mg=1);
> nresids<-mg.resids(nfit,pbc,X,n.sim=1000,cum.resid=1)
> summary(nresids)
Test for cumulative MG-residuals

Grouped Residuals consistent with model
            sup| hat B(t) | p-value H_0: B(t)=0
1. quartile          16.360               0.000
2. quartile          10.585               0.039
3. quartile           8.909               0.123
4. quartile          17.911               0.000

            int ( B(t) )^2 dt p-value H_0: B(t)=0
1. quartile           968.735               0.000
2. quartile           309.346               0.039
3. quartile           228.277               0.122
4. quartile          1144.763               0.000

Residual versus covariates consistent with model
                    sup| hat B(t) | p-value H_0: B(t)=0
prop(Age)                    6.139               0.787
prop(Bilirubin)             27.530               0.000
prop(log(Albumin))           6.045               0.801
prop(log(Protime))           7.983               0.358
```

Cox's model with time-dependent effects

A typical deviation from Cox's model is **time-dependent covariate effects**.

- Treatment is effective for some time, but then effect levels off.
- Takes some time before treatment has an effect.

Model

$$\alpha_i(t) = \exp\left(\beta(t)^T X_i\right)$$

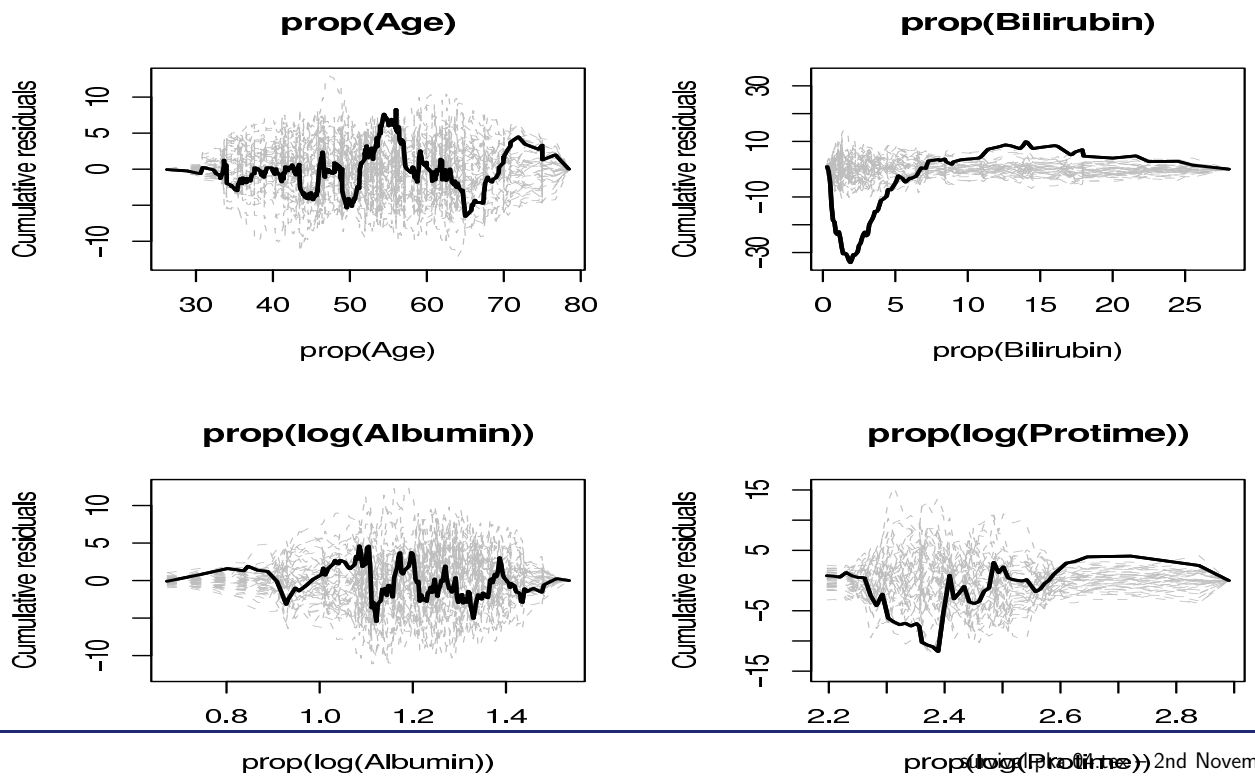
where coefficients $\beta(t)$ are now depending on time!

- Score-equation:

$$X(t)^T (dN(t) - \lambda(t)dt) = 0 \tag{3}$$

Cannot solve (3). Taylor expansion and integration of (3) yields an algorithm.

Cumulative martingale residuals



Cox's model with time-dependent effects

Algorithm:

$$g(\tilde{B})(t) = \int_0^t \tilde{\beta}(s)\,ds + \int_0^t \tilde{A}(s)^{-1}X(s)^T dN(s) - \int_0^t \tilde{A}(s)^{-1}X(s)^T\tilde{\lambda}(s)\,ds, \tag{4}$$

with $\tilde{A}(t) = A_{\tilde{\beta}}(t) = \sum_i Y_i(t)e^{x_i(t)^T\tilde{\beta}(t)}x_i(t)x_i(t)^T$ and $\tilde{\beta}(t)$ obtained from $\tilde{B}(t)$ by **smoothing**.

Theorem

- (4) has a solution $g(\hat{B}) = \hat{B}$
- $n^{1/2}(\hat{B} - B) \xrightarrow{\mathcal{D}} U$
- $\hat{\Sigma}(t) = n\int_0^t \hat{A}(s)^{-1}\,ds$
- $\hat{B}$ is efficient

Reference:(Martinussen et al., 2002).

Cox's model with time-dependent effects

If consistent estimate, $\tilde{\beta}(t)$, is present for estimating $\beta(t)$ Newton-Raphson suggests that

$$\lambda(t)dt = \tilde{A}(s)^{-1}X(s)^T dN(s) - \tilde{A}(s)^{-1}X(s)^T\tilde{\lambda}(s)\,ds,$$

where $\tilde{A}(t) = A_{\tilde{\beta}}(t) = \sum_i Y_i(t)e^{x_i(t)^T\tilde{\beta}(t)}x_i(t)x_i(t)^T$

PBC-data

```
> fit<-timecox(Surv(time/365,status)~Age+Edema+log(Bilirubin)+log(Albumin)+log(Protime),pbc,
+ maxtime=3000/365,band.width=0.5);
Right censored survival times
Nonparametric Multiplicative Hazard Model
Simulations starts N= 5000
> plot(fit,ylab="Cumulative coefficients",xlab="Time (years)");
> summary(fit)
Multiplicative Hazard Model

Test for nonparametric terms

Test for non-siginificant effects
               sup| hat B(t)/SD(t) | p-value H_0: B(t)=0
(Intercept)                    6.232               0.000
Age                            3.937               0.002
Edema                          5.069               0.000
log(Bilirubin)                 9.629               0.000
log(Albumin)                   5.895               0.000
log(Protime)                   6.519               0.000
Test for time invariant effects
               sup| B(t) - (t/tau)B(tau)| p-value H_0: B(t)=b t
(Intercept)                        36.796                 0.005
Age                                 0.035                 0.937
Edema                              14.230                 0.000
log(Bilirubin)                      0.699                 0.494
log(Albumin)                        3.378                 0.875
log(Protime)                       14.911                 0.010
```


Cox's model with time-dependent effects

Important: can also handle the semi-parametric model

$$\lambda(t) = Y(t)\lambda_0(t)\exp(X^T(t)\beta(t) + Z^T(t)\gamma) \tag{5}$$

- Can investigate the important $H_0 : \beta_p(t) \equiv \gamma_{q+1}$ of non-time-dependency;
- Notice that it may be done in a model allowing other covariates to have time-dependent effects!

PBC-data

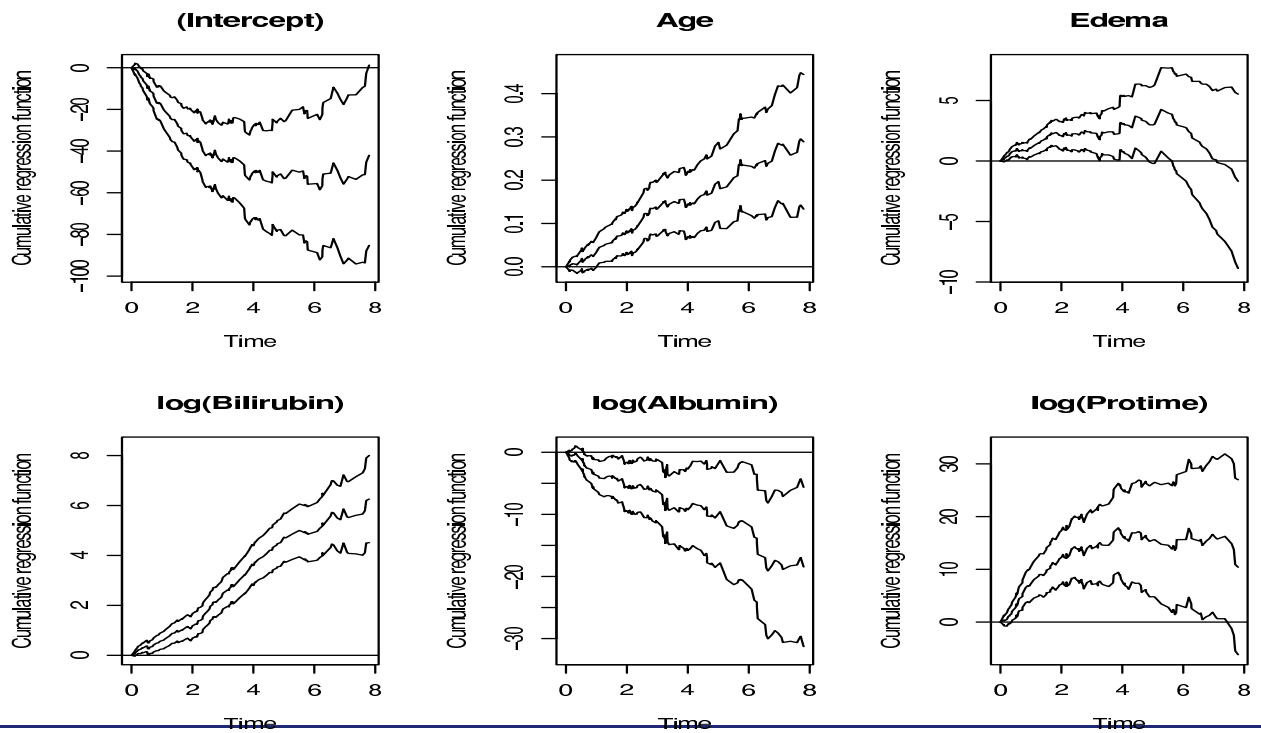
```
> fit.semi<-timecox(Surv(time/365,status)~semi(Age)+Edema+semi(log(Bilirubin))+semi(log(Albumin))+log(Protime),pbc,
maxtime=3000/365,band.width=0.5)
Right censored survival times
Semiparametric Multiplicative Risk Model
Simulations starts N= 5000
> summary(fit.semi)
Multiplicative Hazard Model

Test for nonparametric terms

Test for non-siginificant effects
             sup| hat B(t)/SD(t) | p-value H_0: B(t)=0
(Intercept)                  6.617                    0
Edema                        5.243                    0
log(Protime)                 5.673                    0
Test for time invariant effects
             sup| B(t) - (t/tau)B(tau)| p-value H_0: B(t)=b t
(Intercept)                      35.352                 0.001
Edema                            13.489                 0.000
log(Protime)                     14.044                 0.001


Parametric terms :
                      Coef. Std. Error Robust Std.  Error
semi(Age)             0.038      0.009              0.009
semi(log(Bilirubin))  0.827      0.098              0.086
semi(log(Albumin))   -2.417      0.673              0.643
```


PBC-data



Mix of Aalens and Cox's models

Cox-Aalen model

$$\alpha_i(t) = \alpha(t)^T X_i(t) \exp(\beta^T Z_i(t)),$$

- Gives a mix of Aalens and Cox's models
- Flexible modelling in additive part and multiplicative relative risk parameters for Z.

Reference : (Scheike and Zhang, 2002; Scheike and Zhang, 2003)

Model can also be fitted in `timereg`.

Gives excess risk interpretation of additive part and relative risk interpretation for multiplicative part.

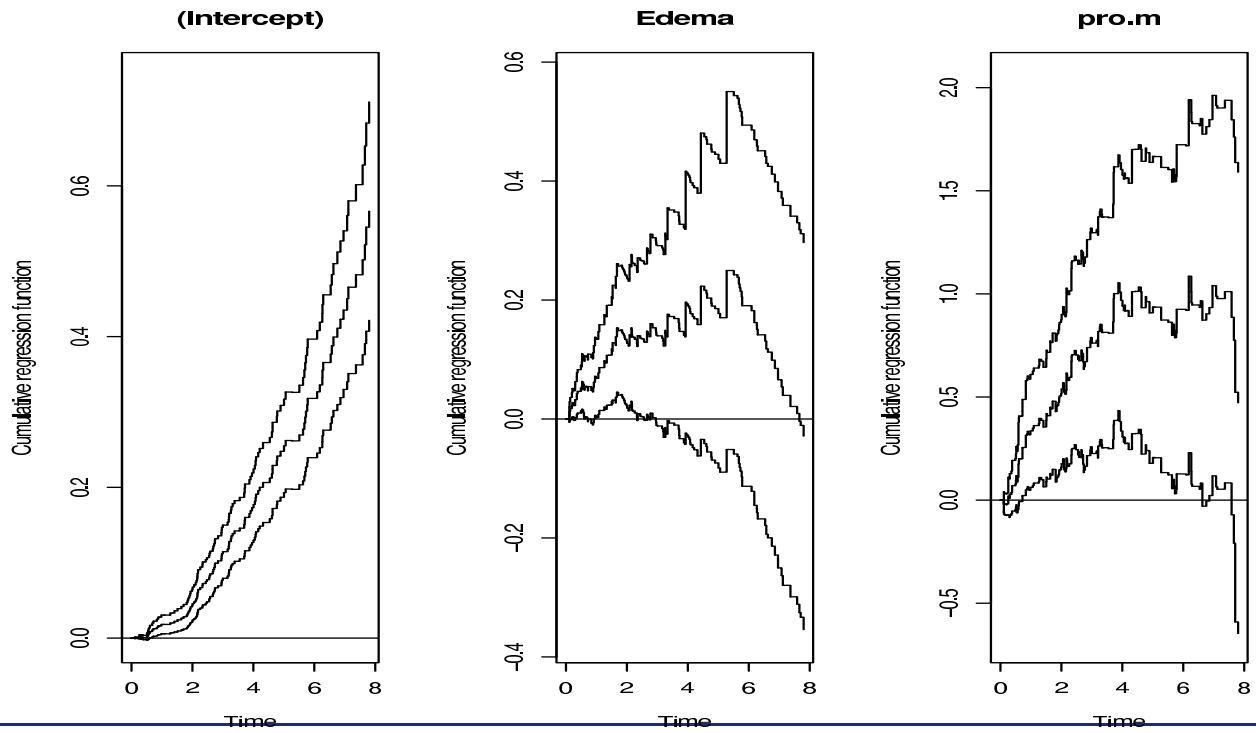
Cox's regression model and stratified versions of it is a special case of the model.

Cox's model with time-dependent effects

- Model with timedependent and constant effects is from a theoretically point of view much more satisfactory
- Model is available using the **R**-package **timereg**.
- Practical experience is needed.
- May need quite a bit of data to get reliable inference.
- Needs to choose a bandwidth. Optimal, cross-validation.
- Additional reference: (Scheike and Martinussen, 2004)
- Alternative models that allow for timedependent effects without the unpleasant bandwidth choice are for example :
  - **Aalens's additive hazards model**
  - **Cox-Aalen model**
  - **Proportional Excess hazard models**

  These can also be fitted using **timereg**!

Cox-Aalen model



PBC-data

```
> logbili.m<-log(pbc$Bilirubin)-mean(log(pbc$Bilirubin));
> logalb.m<-log(pbc$Albumin)-mean(log(pbc$Albumin));
> Age.m<-pbc$Age-mean(pbc$Age);
> fit<-cox.aalen(Surv(time/365,status)~prop(Age.m)+Edema+
+ prop(logbili.m)+prop(logalb.m)+log(Protime),resid.mg=1,
+ max.time=3000/365,pbc)
Cox-Aalen Survival Model
Simulations starts N= 500
> summary(fit)
Test for non-siginificant effects
             sup| hat B(t)/SD(t) | p-value H_0: B(t)=0
(Intercept)                  3.258               0.028
Edema                        2.653               0.096
log(Protime)                 3.439               0.016
Test for time invariant effects
             sup| B(t) - (t/tau)B(tau)| p-value H_0: B(t)=b t
(Intercept)                       2.060                 0.020
Edema                             0.269                 0.000
log(Protime)                      0.821                 0.032

Proportional Cox terms :
                 Coef. Std. Error Robust SE D2log(L)^-1
prop(Age.m)      0.035      0.007     0.010       0.008
prop(logbili.m)  0.800      0.078     0.087       0.087
prop(logalb.m)  -2.459      0.676     0.648       0.675

Score Tests for Proportionality
                sup| hat U(t) | p-value H_0
prop(Age.m)              75.739        0.640
prop(logbili.m)          17.331        0.006
prop(logalb.m)            0.524        0.994
```

Summary

- Cox's proportional hazards model.
- Are the relative risks really not depending on time? Check model carefully.
- More flexible models
  - Multiplicative model with timevarying covariate effects and also constant effects. Inference.
  - Other flexible models: Cox-Aalen model and excess risk models.
  - Aalens additive hazards model (and the semiparametric version).
- Can all be fitted in the R-package `timereg`: www.biostat.ku.dk/∼ts/timereg.html

Mix of Aalens and Cox's models

Excess-risk type model

$$\alpha_i(t) = \alpha(t)^T X_i(t) + \rho_i \lambda_0(t) \exp(\beta^T Z_i(t)),$$

- $\rho_i = 1$, all i, gives a mix of Aalens and Cox's models
- Model is perhaps most naturally seen as an excess risk model: ρ_i is excess indicator eg $I(d_i > 0)$ with d_i dosis for ith subject.

Has proven useful in cancer studies, see (Zahl, 2003).

Notice

$$\alpha_i(t) = \alpha(t)^T X_i(t) + \rho_i \lambda_0(t) \exp(\beta^T Z_i(t)) = \psi(t)^T \tilde{X}_i$$

where

$$\psi(t) = (\alpha(t), \lambda_0(t)),\ \tilde{X}_i^T = (X_i^T, \phi_i(\beta)),\ \phi_i(\beta) = \rho_i \exp(\beta^T Z_i).$$

May derive estimators of unknown parameters and also their large sample properties, see (Martinussen and Scheike, 2002) Model can also be fitted in `timereg`.

References

Lin, D.Y., Wei, L.J., and Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80**, 557–572.

Martinussen, T. and Scheike, T.H. (2002). A flexible additive multiplicative hazard model. *Biometrika* **89**, 283–298.

Martinussen, T., Scheike, T.H., and Skovgaard, I.M. (2002). Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models. *Scandinavian Journal of Statistics* **28**, 57–74.

Scheike, T.H. and Martinussen, T. (2004). On efficient estimation and tests of time-varying effects in the proportional hazards model. *Scandinavian Journal of Statistics* **31**, 51–62.

Scheike, T.H. and Zhang, M.J. (2002). An additive-multiplicative Cox-Aalen model. *Scandinavian Journal of Statistics* **28**, 75–88.

Scheike, T.H. and Zhang, M.J. (2003). Extensions and applications of the Cox-Aalen survival model. *Biometrics* **59**, 1033–1045.

Zahl, P. (2003). Regresion analysis with multiplicative and time-varying additive regression coefficients with examples from breast and colon cancer. *Statistics in Medicine* **22**, 1113–1127.