Additive hazard regression models

STK4080 H16

1. Modifications of Cox-regression when prop.haz. does not hold
2. Aalen's additive hazards models
3. Some more additive hazards models

Repetition Cox-regression

With covariate x_i the survival time T_i has hazard

$$\alpha(t|x_i) = \exp(\beta' x_i)\alpha_0(t)$$

where $\alpha_0(t)$ is called basis (underlying) hazard and where β is a regressionsparameter.

With D_i indicator for death for individual i, $\tilde{T}_i$ right censoring time and $\mathcal{R}(t) = \{i : \tilde{T}_i \geq t\}$ = the risk set (right before) time t we estimate β by maximizing the Cox' partial likelihood

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{\exp(\beta' x_i)}{\sum_{k \in \mathcal{R}(\tilde{T}_i)} \exp(\beta' x_k)} \right]^{D_i} = \prod_{i=1}^{n} \left[\frac{\exp(\beta' x_i)}{S^{(0)}(\beta, t)} \right]^{D_i}$$

where $S^{(0)}(\beta, t) = \sum_{k \in \mathcal{R}(\tilde{T}_i)} \exp(\beta' x_k)$

Test for prop. assumption

can be based on Schoenfeld-residuals: The tests are directed to departures from the model given by

$$\alpha(t|x) = \exp((\beta_0 + \theta g(t))x)\alpha_0(t)$$

for specified functions $g(t)$. Actually they are score-tests for a new covariate $x'(t) = g(t)x$.

If the model is extended by a new time-dependent covariate $x_{p+1}(t) = x_p g(t)$ the score for the coefficient β_{p+1} is given

$$U_{p+1} = \sum_j g(\tilde{T}_j)[x_p - \bar{x}_p(\tilde{T}_j)]$$

as a weighted sum of the Schoenfeld-residuals.

Test for proportionality

Example: Melanoma data: KM transform.

Sex	$\chi^2_1 = 0.5$	p=0.46
Ulceration	$\chi^2_1 = 0.7$	p=0.40
Age	$\chi^2_1 = 2.8$	p=0.09
log(Tumor thickness)	$\chi^2_1 = 4.1$	p=0.04

Indication for departure for tumor thickness

R-syntax:

```
coxfit<-coxph(Surv(lifetime,dead)~sex+age+ulcer+logthick,data=mel)
cox.zph(coxfit)
             rho  chisq      p
     sex -0.095  0.536 0.4642
     age  0.200  2.828 0.0927
   ulcer  0.116  0.717 0.3972
logthick -0.299  4.079 0.0434
  GLOBAL     NA 10.450 0.0335
```

Strategies when proportional hazard fails

- Stratified Cox-regression
- Separate analyzes on disjoint time intervals
- Time-dependent covariates
- Alternative regression models
  - Accelerated failure time models
  - Additive models

Ex. Stratified Cox-regression

Weak departure wrt. thickness. Stratifies on `grthick`:

```
> coxstrat<-coxph(Surv(lifetime,dead)~sex+age+ulcer+strata(grthick),dat
> coxstrat

        coef exp(coef) se(coef)      z      p
  sex  0.4074     1.503  0.27351  1.490 0.1400
  age  0.0063     1.006  0.00837  0.753 0.4500
ulcer -0.9480     0.388  0.32572 -2.910 0.0036

Likelihood ratio test=13.2  on 3 df, p=0.00426  n= 205
> cox.zph(coxstrat)
           rho  chisq     p
   sex -0.0232 0.0313 0.860
   age  0.1178 1.0581 0.304
 ulcer  0.1037 0.5619 0.453
GLOBAL      NA 1.5924 0.661
```

But possibly the stratification changed other estimates somewhat?

Separate intervals

We may split the time interval i 2 and make separate

Cox-regressions within each interval:

Ex: Melanoma data. Half of death before $\tau = 3$
Analysis on $[0, 3 >$: Uses events only if `lifetime` < 3

```
coxph(Surv(lifetime,dead*(lifetime<3))~sex+age+ulcer+logthick,data=mel)
             coef exp(coef) se(coef)     z      p
     sex  0.4654     1.593    0.336  1.38 0.1700
     age  0.0106     1.011    0.010  1.06 0.2900
   ulcer -1.1979     0.302    0.449 -2.67 0.0076
logthick  0.6628     1.940    0.231  2.87 0.0041
```

Analysis on $[3, \infty >$: Uses only events with `lifetime` > 3

```
coxph(Surv(lifetime,dead*(lifetime>3))~sex+age+ulcer+logthick,data=mel)

             coef exp(coef) se(coef)      z      p
     sex  0.2204     1.247   0.3531  0.624 0.5300
     age  0.0332     1.034   0.0122  2.718 0.0066
   ulcer -0.5140     0.598   0.3833 -1.341 0.1800
logthick  0.2378     1.268   0.2148  1.107 0.2700
```

Time dependent covariats

If similar parameter estimates for age, sex and ulceration on these intervals we may fit a common model

$$\begin{aligned}\log\left(\frac{\alpha(t|x)}{\alpha_0(t)}\right) = \quad & \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \\ & + \beta_4 x_4 I(t < 3) + \beta_5 x_4 I(t \geq 3)\end{aligned}$$

i.e. Cox-regression with time dependent covariates

- $x_4 I(t < 3)$
- $x_4 I(t \geq 3)$

Cox-regression with time dependent covariates

Need to set up data-frame with time dependent covariates:

```
mel2 <- data.frame(intime = c(rep(0,205),rep(3,167)))
mel2$outtime <- c(pmin(mel$lifetime,3),mel$lifetime[mel$lifetime>3])
mel2$indi    <- c(mel$dead*(mel$lifetime<3),mel$dead[mel$lifetime>3])

mel2$sex    <- c(mel$sex,mel$sex[mel$lifetime>3])
mel2$ulcer    <- c(mel$ulcer,mel$ulcer[mel$lifetime>3])
mel2$age    <- c(mel$age,mel$age[mel$lifetime>3])

mel2$logtha <- c(mel$logthick,rep(0,167))
mel2$logthb <- c(rep(0,205),mel$logthick[mel$lifetime>3])

coxph(Surv(intime,outtime,indi)~sex+ulcer+age+logtha+logthb,data=mel2)
          coef exp(coef) se(coef)      z       p
   sex  0.3813     1.464  0.26901  1.417 0.16000
 ulcer -0.9845     0.374  0.32646 -3.016 0.00260
   age  0.0102     1.010  0.00823  1.239 0.22000
logtha  0.8985     2.456  0.24757  3.629 0.00028
logthb  0.2130     1.237  0.23346  0.912 0.36000

Likelihood ratio test=49  on 5 df, p=2.17e-09  n= 372
```

Advantages/disadvantages with strategies

1. Stratification
   - Easy
   - More difficult to show effect of stratification variable
   - Allows for only a few problem covariates
2. Separate intervals
   - Relatively easy
   - Choice of interval difficult/arbitrary
   - Looses power for covariates where the assumption is OK
   - Many parameter estimates

Advantages/disadvantages with strategies

3. Time dependent covariates

- Somewhat awkward to arrange (in R(?))
- Difficult choice of interval
- Only helpful when prop.haz. OK for most covar.

Consequences of departure from proportionality

- biased estimates of coefficients
- both for covariates where the assumption hold and fail
- biased survival estimates

Aalen's additive hazards model

$$\alpha(t|x) = \beta_0(t) + \beta(t)'x = \beta_0(t) + \sum_{j=1}^{p} \beta_j(t) x_j$$

allows for time-dependent covariate effects $\beta_j(t)$.

We will study Aalen's estimator in this model - not requiring any restrictions on the regression functions $\beta_j(t)$, but for which we typically estimate cumulative regression functions

$$B_j(t) = \int_0^t \beta_j(s) \mathrm{d}s$$

Simple example

Two groups: $x = 0, 1$ in group 0 and 1 resp.
Calculates Nelson-Aalen estimates $\hat{A}_x(t)$ for both groups.
Under the additive model $\alpha_x(t) = \beta_0(t) + \beta_1(t)x$ a special case of the Aalen additive hazards estimator becomes

$$\hat{B}_0(t) = \hat{A}_0(t) \qquad \text{and} \qquad \hat{B}_1(t) = \hat{A}_1(t) - \hat{A}_0(t)$$

and we may also estimate $\text{Var}[\hat{B}_1(t)]$ by

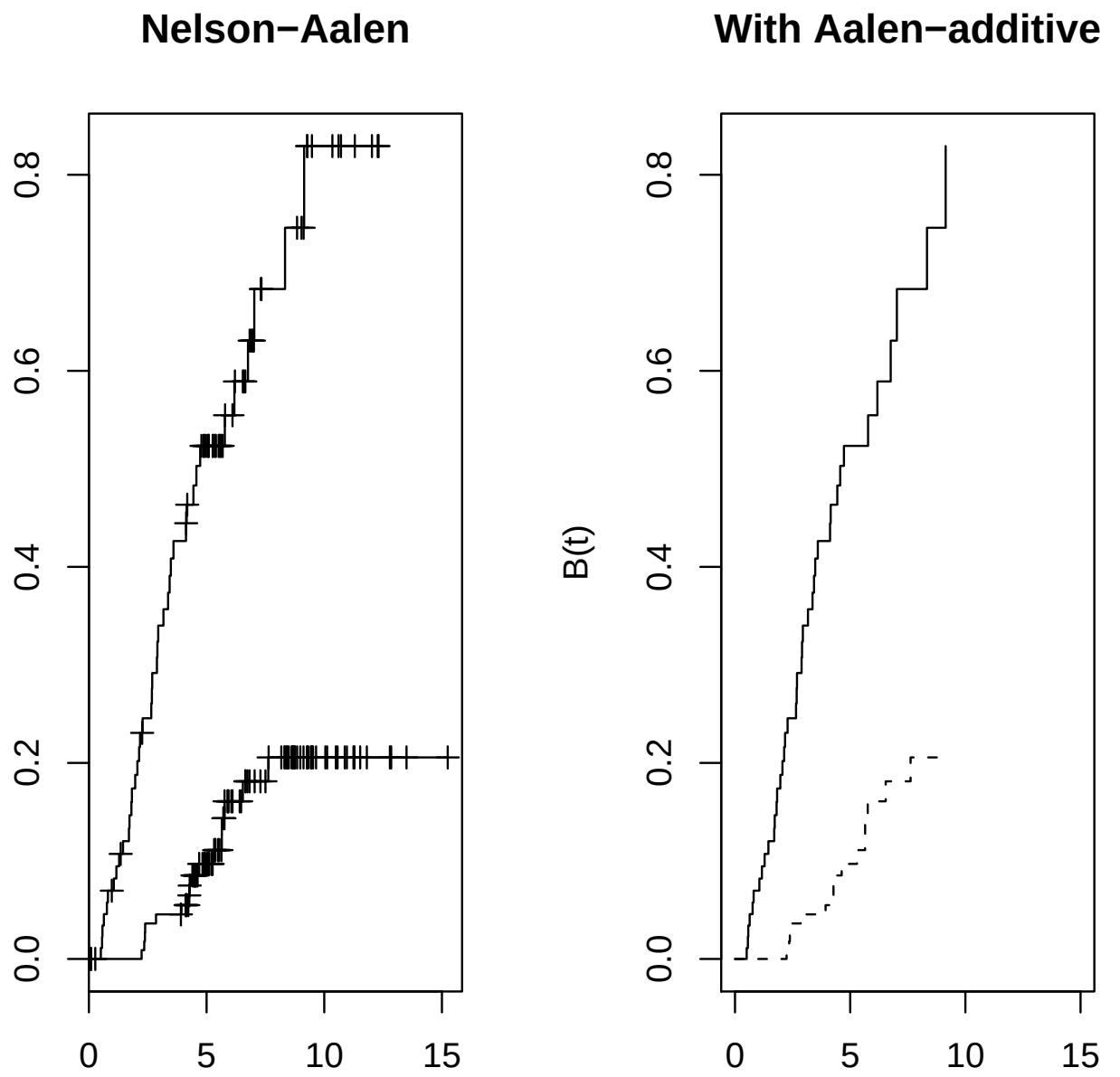
$$\widehat{\text{Var}}[\hat{B}_0(t)] = \int_0^t \frac{dN_{\bullet 0}(t)}{Y_{\bullet 0}(t)^2}$$

and

$$\widehat{\text{Var}}[\hat{B}_1(t)] = \int_0^t \frac{dN_{\bullet 0}(t)}{Y_{\bullet 0}(t)^2} + \int_0^t \frac{dN_{\bullet 1}(t)}{Y_{\bullet 1}(t)^2}$$

where $N_{\bullet x}(t)$ and $Y_{\bullet x}(t)$ are counting processes and number under risk in group $x = 0, 1$.

On the melanoma-data with x indicator for ulceration



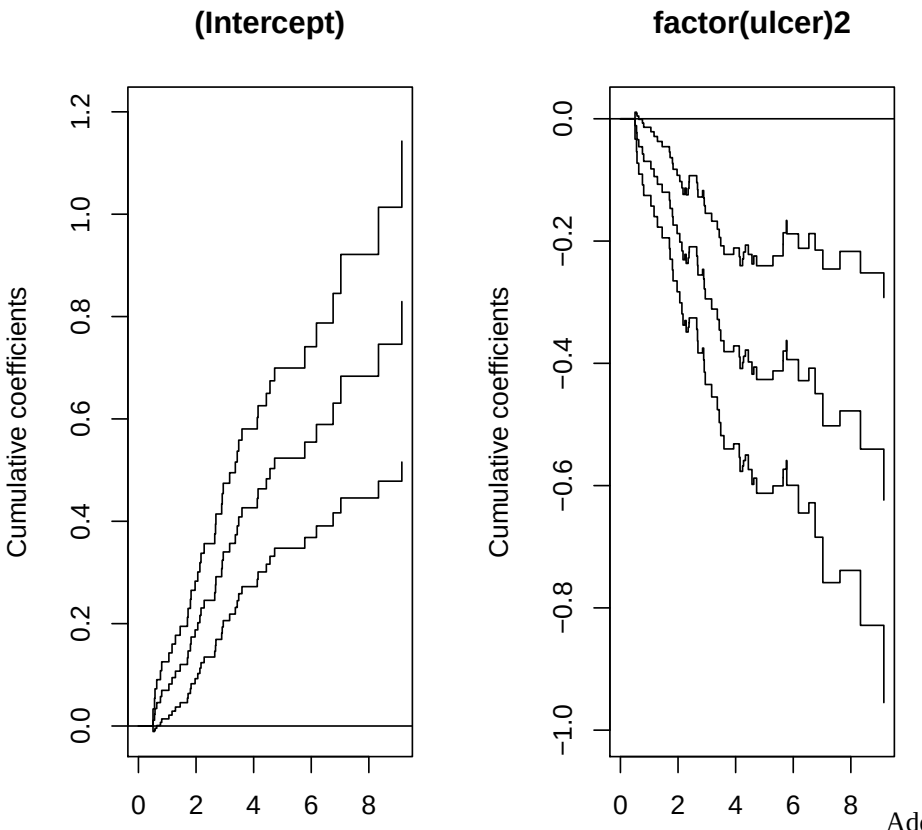
We use the `timereg` library:

```
library(timereg)
aalenfit<-aalen(Surv(lifetime,status==1)~factor(ulcer))
names(aalenfit)

 [1] "cum"               "var.cum"           "robvar.cum"        "resid
 [7] "pval.testBeqC"     "pval.testBeqC.is"  "obs.testBeqC"      "obs.t
[13] "sim.testBeqC.is"   "conf.band"         "test.procBeqC"     "sim.t
[19] "deviance"          "call"

NAafit<-survfit(Surv(lifetime,status==1)~ulcer,type="fh2")

par(mfrow=c(1,2))
plot(NAafit,fun="cumhaz")
title("Nelson-Aalen")

plot(stepfun(aalenfit$cum[,1],aalenfit$cum[,2]),type="l",xlim=c(0,15),x
lines(stepfun(aalenfit$cum[,1],aalenfit$cum[,2]+aalenfit$cum[,3]),lty=2
title("With Aalen-additive")
```

Standard plot from `timereg`

The plot on the previous pages were intended to show that the estimator in additive model and the Nelson-Aalen does the same thing. The regular plot from `timereg` produces $\hat{B}_x(t)$ with 95% CI.

```
plot(aalenfit)
```



Estimation in Aalen's additive model - in general

With counting process notation

$$dN_i(t) = Y_i(t)\beta_0(t)dt + Y_i(t)\sum_{j=1}^{p}\beta_j(t)x_{ij}dt + dM_i(t)$$

where $dN_i(t)$ indicates event for individual i at time t, $Y_i(t)$ the at risk indicator for individual i at time $t-$ and $dM_i(t)$ the martingale increment.

This may be interpreted as a linear regression model at each time t with

- $dN_i(t)$ as responses
- $\beta_j(t)dt$ as regression coefficients
- $Y_i(t)x_{ij}$ as covariates

Aalen regression - in general

Instead of $\beta_j(t)dt$ we may write $dB_j(t) = B_j(t) - B_j(t-)$.

Will then estimate changes (increments)
$dB(t)^\top = (dB_0(t), dB_1(t), \ldots, dB_p(t))$ by

- least squares estimator
- in regression problem given on previous page

With $d\hat{B}_j(t)$ estimate for $dB_j(t)$ in this regression we get estimate for cumulative regression-functions $B_j(t)$ as

$$\hat{B}_j(t) = \sum_{t_i \leq t} d\hat{B}_j(t_i) = \int_0^t d\hat{B}_j(s)$$

where this last integral is wrp to a counting-process (and events at exactly $t_1 < t_2 < \ldots$.)

Least squares estimators

With responses Y_i and covariate(-vectors) x_i the model $Y_i = \beta' x_i + \varepsilon_i$ is most often fit by minimizing $Q = \sum_{i=1}^{n} (Y_i - \beta' x_i)^2$

The solution to this problem may be written on matrix form

$$\hat{\beta} = (X^\top X)^{-1} X^\top Y$$

where $Y^\top = (Y_1, Y_2, \ldots, Y_n)$ is a vector of responser and

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

is the design matrix (matrix of covariates)

Estimator i Aalen additive model, explicitly

Let the responses at time t be given by

$$d\tilde{N}(t)^\top = (dN_1(t), dN_2(t), \ldots, dN_n(t))$$

and the design-matrix at time t

$$\tilde{X}(t) = \begin{bmatrix} Y_1(t) & Y_1(t)x_{11} & Y_1(t)x_{12} & \cdots & Y_1(t)x_{1p} \\ Y_2(t) & Y_2(t)x_{21} & Y_2(t)x_{22} & \cdots & Y_2(t)x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_n(t) & Y_n(t)x_{n1} & Y_n(t)x_{n2} & \cdots & Y_n(t)x_{np} \end{bmatrix}$$

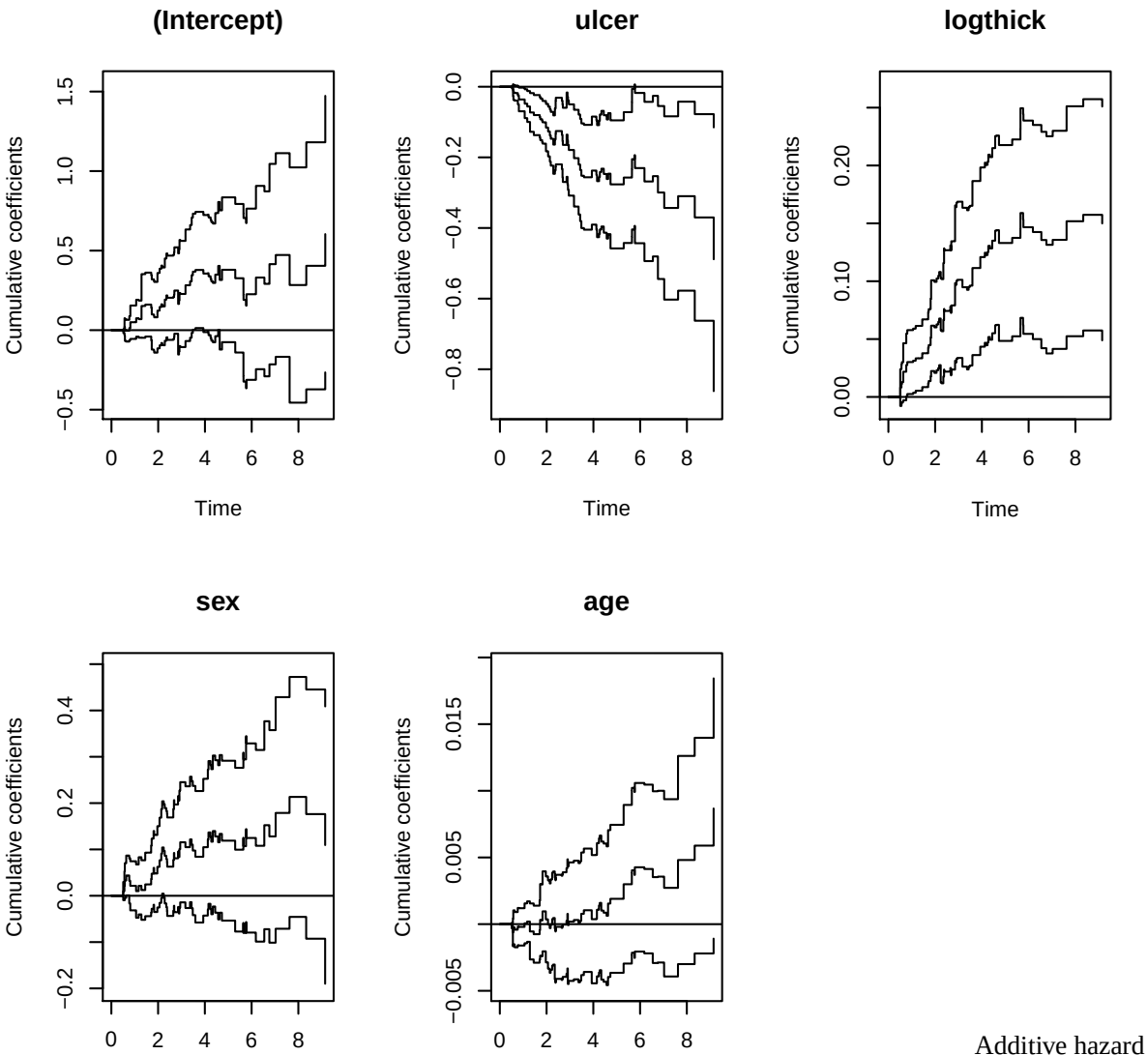
The increments $dB(t)$ are (usually) estimated by

$$d\hat{B}(t) = (\tilde{X}(t)^\top \tilde{X}(t))^{-1} \tilde{X}(t)^\top d\tilde{N}(t)$$

for t such that $\tilde{X}(t)$ has full rank.

Example: Melanoma data, all covariates

```
aalenfit<-aalen(Surv(lifetime,status==1)~ulcer+logthick+sex+age)
par(mfrow=c(2,3))
plot(aalenfit)
```



Interpretation of cumulative regression functions

If

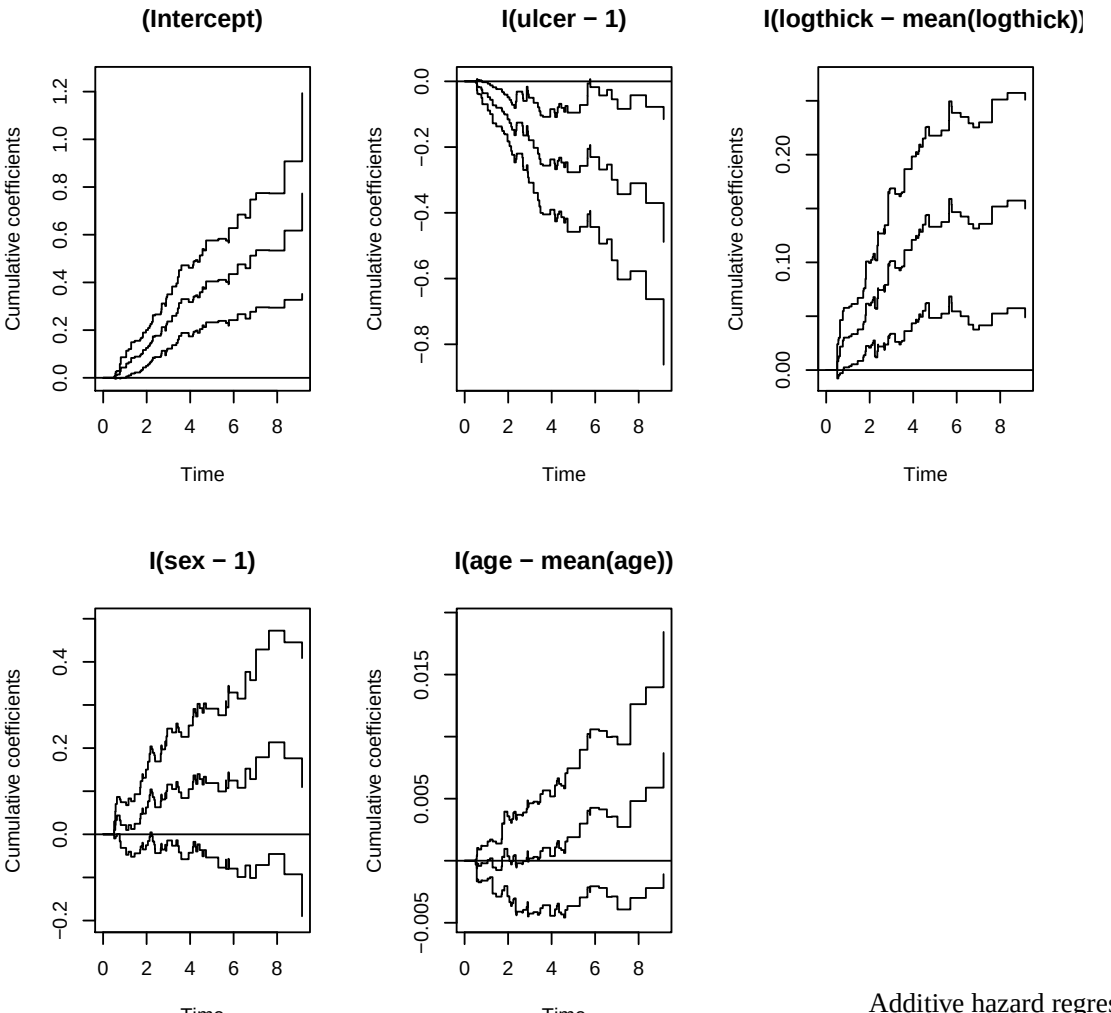
- $\hat{B}_j(t)$ is increasing in an interval:
  Higher risk with high value of x_j
- $\hat{B}_j(t)$ decreases in an interval:
  Lower risk with high value of x_j
- $\hat{B}_j(t)$ roughly constant in an interval:
  Little effect of x_j in the interval

It is thus possible to read off from the figures how the effect changes over time.

Interpretation of baseline $B_0(t)$

depends on coding of covariates. Better:

```
aalenfit<-aalen(Surv(lifetime,status==1)~I(ulcer-1)
+I(logthick-mean(logthick))+I(sex-1)+I(age-mean(age)))
par(mfrow=c(2,3))
plot(aalenfit)
```



$\hat{B}(t)$ is almost unbiased

Let $J(t)$ be the indicator that $\tilde{X}(t)^\top \tilde{X}(t)$ can be inverted (have full rank) at time t. Then

$$\hat{B}(t) - \int_0^t J(s) dB(s)$$

is a martingale and have expectation 0, i.e. $\hat{B}(t)$ is practically unbiased.

(NB. With this presentation it is only possible to estimate $\hat{B}(t)$ for t such that $J(t) = 1$.)

This resultat is of course closely related to the LSE is unbiased as long as the linear structure holds (even with dependent and heteroscedastic responses).

$\hat{B}(t)$ is almost unbiased

With $M_i(t) = dN_i(t) - Y_i(t)\beta_0(t)dt + Y_i(t)\sum_{j=1}^{p}\beta_j(t)x_{ij}dt$, $\beta(t) = (\beta_0(t), \ldots, \beta_p(t)$ and $\tilde{M}(t) = (M_1(t), \cdots, M_n(t))^\top$ we have

$$d\tilde{N}(t) = \tilde{X}(t)\beta(t)dt + d\tilde{M}(t)$$

Inserted into $d\hat{B}(t)$ this gives

$$\begin{aligned} d\hat{B}(t) = & \quad (\tilde{X}(t)^\top \tilde{X}(t))^{-1}\tilde{X}(t)^\top d\tilde{N}(t) \\ = & \quad (\tilde{X}(t)^\top \tilde{X}(t))^{-1}\tilde{X}(t)^\top \tilde{X}(t)\beta(t)dt \\ & \quad +(\tilde{X}(t)^\top \tilde{X}(t))^{-1}\tilde{X}(t)^\top d\tilde{M}(t) \\ = & \quad J(t)\beta(t)dt + (\tilde{X}(t)^\top \tilde{X}(t))^{-1}\tilde{X}(t)^\top d\tilde{M}(t) \end{aligned}$$

But the last term is a vector of a sum of uncorrelated martingales. Integrating up, the result on the previous page follows.

Variance-estimation, LSE $\hat{\beta}$

If $\mathrm{Var}(Y_i) = \sigma^2$ and Y_i-s are independent then the covariance matrix of $\hat{\beta}$ equals

$$\Sigma(\hat{\beta}) = (X^\top X)^{-1}\sigma^2,$$

but in general we have

$$\Sigma(\hat{\beta}) = (X^\top X)^{-1}X^\top \Sigma(Y) X (X^\top X)^{-1}$$

where $\Sigma(Y)$ is the covariance matrix of response vector Y.

(This kind of expression is often referred to as a "sandwich-estimator")

Variances of $\hat{B}(t)$

The responses in Aalen-regression $dN_i(t)$ have cond. "variances"

$$\text{Var}(dN_i(t)|\mathcal{F}_{t-}) = \text{E}[dN_i(t)|\mathcal{F}_{t-}] = Y_i(t)[\beta_0(t) + \sum_{j=1}^{p} \beta_j(t)x_{ij}]dt$$

which are unequal due to different $\sum_{j=1}^{p} \beta_j(t)x_{ij}$.

However, we may use $dN_i(t)$ as estimates of these variances.

Then, with independent individuals (uncorrelated martingale increments) we get a estimated conditional covariance matrix for $d\tilde{N}(t)$:

$$\Sigma_{d\tilde{N}(t)|\mathcal{F}_{t-}} = \text{diag}(d\tilde{N}(t))$$

where $\text{diag}(v)$ denotes the diagonal matrix of vector v.

Variance formula for $\hat{B}(t)$

By the sandwich formula we get that an estimate of the covariance-matrix of $d\hat{B}(t)$ is given as $\widehat{\Sigma}(d\hat{B}(t)) =$

$$(\tilde{X}(t)^\top \tilde{X}(t))^{-1} \tilde{X}(t)^\top \text{diag}(d\tilde{N}(t)) \tilde{X}(t) (\tilde{X}(t)^\top \tilde{X}(t))^{-1}$$

Furthermore, by the martingale property for $\hat{B}(t) - \int_0^t J(s) dB(s)$ we get uncorrelated increments $d\hat{B}(t)$.

Thus an estimator of the covariance matrix of $\hat{B}(u)$ is given by

$$\widehat{\Sigma}(\hat{B}(u)) = \int_0^u X^-(t) \text{diag}(d\tilde{N}(t)) X(t)^{-\top}$$

where $X^-(t) = (\tilde{X}(t)^\top \tilde{X}(t))^{-1} \tilde{X}(t)^\top$.

Estimation of regression function $\beta_j(t)$

may be carried out kernel smoothing

$$\hat{\beta}_j(t) = \frac{1}{b} \int K(\frac{t-s}{b}) d\hat{B}_j(s)$$

where $K(t)$ is kernel function (density with center in 0).

For $\hat{\beta}(t)$ the vector of kernel smoothed regression functions the covariance matrix becomes

$$\hat{\Sigma}_{\hat{\beta}(t)} = \frac{1}{b^2} \int K(\frac{t-s}{b})^2 X^-(s) \text{diag}(d\tilde{N}(s)) X(s)^{-\top}$$

Tests for $H_0 : \beta_q(t) = 0$ for all t

A (somewhat naive) idea might be based on a test statistic

$$\frac{\hat{B}_q(t)}{\text{se}(\hat{B}_q(t))} \sim \text{N}(0,1)$$

under H_0 (the denominator being the standard error of the numerator).

Disadvantages

- Choice of t arbitrary
- Might use $t = t_0 = \max(s : J(s) = 1)$, but for large s the increments may have large uncertainty $\rightarrow$ low power.
- Low power if $\beta_q(t)$ changes sign

Better tests for $H_0 : \beta_q(t) = 0$

With predictable weight functions $L_q(t)$ consider statistic

$$Z_q(t_0) = \int L_q(t) d\hat{B}_q(t)$$

which under the null is a martingale and with sensible choice of $L_q(t)$ approximately normal (more discussion in ABG)

The `timereg`-test is however based on a statistic

$$\sup_{0 \leq t \leq t_0} |\frac{\hat{B}_q(t)}{se(\hat{B}_q(t))}|$$

where $se(\hat{B}_q(t))$ is the standard error of $\hat{B}_q(t)$.

Output from Aalen-regression

```
> aalenfit<-aalen(Surv(lifetime,status==1)~ulcer+logthick+sex+age)
> summary(aalenfit)
Additive Aalen Model

Test for nonparametric terms

Test for non-significant effects
             Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                           1.98                0.357
ulcer                                 3.65                0.010
logthick                              3.47                0.005
sex                                   2.09                0.309
age                                   1.86                0.444
```

Tests for $H_0 : \beta_j(t) = \gamma_j$ for all t

`timereg` also allows for testing whether the regression functions $\beta_j(t)$ are constant and equal to some parameter γ by two different tests.

1. Kolomogorov-Smirnov (Supremum) test

$$\sup_{0 \leq t \leq t_0} |\hat{B}(t) - \frac{t}{t_0} \hat{B}(t_0)|$$

2. Cramer von-Mises test

$$\int_0^{t_0} |\hat{B}(t) - \frac{t}{t_0} \hat{B}(t_0)|^2 dt$$

For details on calculating the p-values see Martinussen & Scheike (2006): Dynamic Regression Models for Survival Data

Output from Aalen-regression, contd.

```
Test for time invariant effects
                  Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                       0.22700                       0.787
ulcer                             0.11400                       0.628
logthick                          0.06880                       0.046
sex                               0.12200                       0.356
age                               0.00394                       0.330

                    Cramer von Mises test p-value H_0:constant effect
(Intercept)                       0.085700                      0.801
ulcer                             0.022100                      0.662
logthick                          0.011700                      0.028
sex                               0.033100                      0.301
age                               0.000046                      0.247
```

Other additive hazards regression

Lin & Ying (1994) additive hazards regression: Model

$$\alpha(t|x) = \beta_0(t) + \gamma' x$$

is a special case of Aalen's hazards model with all regression functions constant, $\beta_j(t) = \gamma_j$, except for baseline $\beta_0(t)$.

McKeague & Sasieni (1994) is an in between case between the Aalen and Ling & Ying models, allowing for some time-dependent regression functions $\beta_j(t)$ and some time-constant $\beta_k(t) = \gamma_k$.

The Lin & Ying and McKeague & Sasieni estimators were both be developed within a counting process - martingale framework, but we omit details. These estimators are all implemented in `timereg` using the `aalen`-function.

Ex: McKeague & Sasieni model with time-dependent

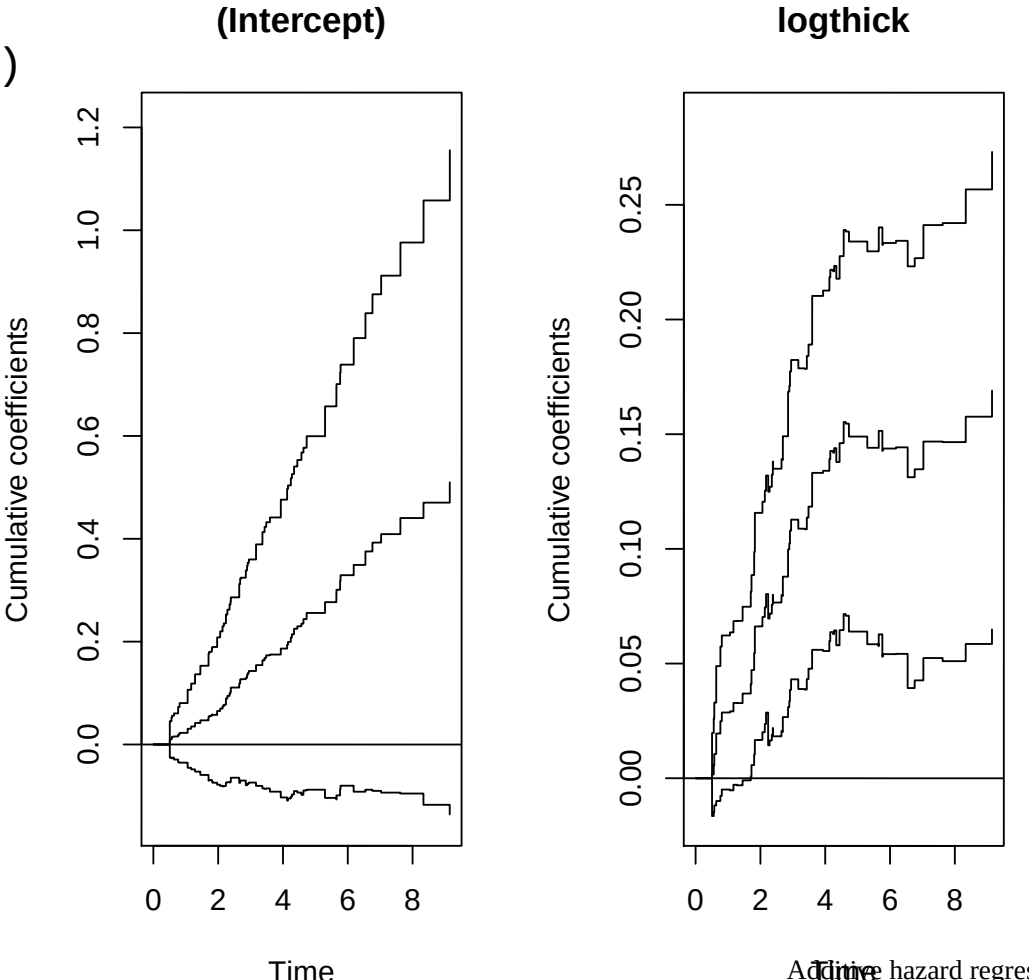
effect of ulceration and other covariate effects time-constant

```
> McKSasfit<-aalen(Surv(lifetime,status==1)~const(ulcer)+logthick
                                          +const(sex)+const(age))
> summary(McKSasfit)
Additive Aalen Model

Test for nonparametric terms

Test for non-significant effects
            Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                          1.65                0.143
logthick                             3.71                0.002

Test for time invariant effects
                  Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                        0.0451                       0.336
logthick                           0.0709                       0.045
                    Cramer von Mises test p-value H_0:constant effect
(Intercept)                       0.00482                       0.214
logthick                          0.01060                       0.091
```

Ex: McKeague & Sasieni model contd.

```
Parametric terms :
                  Coef.        SE Robust SE      z   P-val
const(ulcer) -0.045800 0.014500  0.014600 -3.15 0.00165
const(sex)    0.021500 0.014300  0.014400  1.50 0.13400
const(age)    0.000582 0.000456  0.000449  1.27 0.20200
```

```
> par(mfrow=c(1,2))
> plot(McKSasfit)
```



Ex: Lin & Ying model

```
> LinYing<-aalen(Surv(lifetime,status==1)~const(ulcer)+const(logthick)
                                        +const(sex)+const(age))
> summary(LinYing)
Additive Aalen Model

Test for nonparametric terms
Test for non-significant effects
            Supremum-test of significance p-value H_0: B(t)=0
(Intercept)                           1.6                0.237
Test for time invariant effects
                  Kolmogorov-Smirnov test p-value H_0:constant effect
(Intercept)                        0.0377                       0.767
                    Cramer von Mises test p-value H_0:constant effect
(Intercept)                       0.00255                       0.746

Parametric terms :
                     Coef.       SE Robust SE     z    P-val
const(ulcer)    -0.047700 0.015100   0.01540 -3.16 0.001600
const(logthick)  0.021500 0.006480   0.00716  3.32 0.000898
const(sex)       0.021800 0.014500   0.01480  1.51 0.132000
const(age)       0.000616 0.000458   0.00046  1.34 0.179000
```

Some more models

1. Special case of Lin & Ying model with $\beta_0(t)$ step-function

$$\alpha(t|x) = \beta_0(t) + \gamma' x$$

This generates a parametric model. As will be discussed next week the model can be fit by Poisson regression.

2. "Cox-Aalen" model: Covariate vectors x and z. Model

$$\alpha(t|x,z) = \alpha(t)^\top x \exp(\beta^\top z)$$

where $\alpha(t)$ is a vector of regression functions and β a parameter.

3. Proportional excess hazards model: Covariate vectors x and z.

$$\alpha(t|x,z) = \alpha(t)^\top x + \rho_i(t)\lambda_0(t) \exp(\beta^\top z)$$

where $\alpha(t)$ and β as in 2., $\rho_i(t)$ a known function (population mortality) and $\lambda_0(t)$ an unknown function. Both 2. and 3. may be fitted with `timereg`