Time-varying regression effects in the proportional odds model

CAS, OSLO

Joint with Andreas Andersen (Thomas Scheike)

Torben Martinussen
torbenm@dina.kvl.dk
Outline

- Time-varying regression effects: Aalen model and Cox model.
- \textit{R}-package \texttt{timereg} and forthcoming book.
- Proportional odds model. Basics.
- Modified partial likelihood estimation (\texttt{timereg}).
- Time-varying regression effects: Proportional odds model.
Survival analysis

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

\[ T = T^* \land C \quad D = I(T^* \leq C) \]

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

**Hazard-function**

\[ \alpha(t) = \lim_{h \downarrow 0} \frac{1}{h} P(t \leq T^* < t + h \mid T^* \geq t). \]

**Counting process**

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

**Martingale**

\[ M_i(t) = N_i(t) - \Lambda_i(t) \]

where

\[ \Lambda_i(t) = \int_0^t Y_i(s) \alpha(s) \, ds \text{ (compensator)}, \quad Y_i(t) = I(t \leq T_i) \text{ (at risk process)}. \]
Aalen’s additive hazard model

Intensity is given by

$$\lambda_i(t) = Y_i(t)(\beta_0(t) + \beta_1(t)X_{i1} \cdots \beta_p(t)X_{ip}) = X_i^T(t)\beta(t),$$


Written in vector form we have the decomposition

$$dN(t) = X(t)dB(t) + dM(t) \quad (1)$$

where \(i\)th row of \(X(t)\) is \(X_i^T(t)\) and \(B(t) = \int_0^t \beta(s) ds\).

Equation (1) suggest the estimator

$$\hat{B}(t) = \int_0^t X^-(s) dN(s),$$

where \(X^-(t)\) is a generalized inverse of \(X(t)\).

We see that:

$$\hat{B}(t) = B(t) + \int_0^t X^-(s) dM(s) + o_p(1/\sqrt{n}).$$
PBC-data

```r
> library(survival)
> fit.aalen<-aalen(Surv(time/365,status)~Age+Edema+
+ log(Bilirubin)+log(Albumin) +log(Protime),
+ pbc,maxtime=3000/365)
> plot(fit.aalen)
```
Proportional odds model

• Model:
  \[ S(t) = S(t|Z) = \frac{1}{1 + G(t) \exp(Z^T \beta)}; \quad \lambda(t) = \frac{g(t)}{\exp(-Z^T \beta) + G(t)} \cdot \]
  \[ \logit(1 - S_Z(t)) = \log(G(t)) + Z^T \beta, \quad (2) \]

  where \( G(t) \) is unknown with \( G(0) = 0, \ G(\infty) = \infty \) and \( g(t) = G'(t) \).

• Converging hazards: relative risk for two individual with covariates \( Z_1 \) and \( Z_2 \), respectively, is
  \[ \text{RR}(t) = \frac{\lambda(t, Z_2)}{\lambda(t, Z_1)} = \frac{\exp(-Z_1^T \beta) + G(t)}{\exp(-Z_2^T \beta) + G(t)} \]
  with \( \text{RR}(0) = \exp((Z_2 - Z_1)^T \beta) \) and \( \lim_{t \to \infty} \text{RR}(t) = 1 \)
Estimation in the proportional odds model

- Modified partial likelihood estimator.
- The martingale decomposition of $dN.(t)$ reads

$$dN.(t) = S_0(t, \beta, G)dG(t) + dM.(t),$$

where

$$S_0(t, \beta, G) = \sum_{i=1}^{n} Y_i(t) \exp(Z_i^T \beta) \lambda_0(\exp(Z_i^T \beta)G(t-))$$

writing $G(t-)$ to stress the needed predictability of the intensities. Breslow-type estimator (keeping $\beta$ fixed)

$$\tilde{G}(t, \beta) = \int_0^t \frac{1}{S_0(s, \tilde{G}, \beta)} dN.(s). \quad (3)$$
Estimation in the proportional odds model

- Estimation of $\beta$ may be based on the (partial) likelihood function

$$\prod_{i=1}^{n} \prod_{t \geq 0} \left[ Y_i(t) \exp(Z_i^T \beta) dG(t) \lambda_0(\exp(Z_i^T \beta) G(t-)) \right]^\Delta N_i(t)$$

$$\times \exp \left\{ - \int_0^\infty Y_i(t) \exp(Z_i^T \beta) \lambda_0(\exp(Z_i^T \beta) G(t-)) dG(t) \right\}.$$

- Replace $dG(t)$ with $d\tilde{G}(t, \beta)$ and $G(t)$ with $\tilde{G}(t, \beta)$:

$$\prod_{i=1}^{n} \prod_{t \geq 0} \left\{ Y_i(t) \exp(Z_i^T \beta) d\tilde{G}(t, \beta) \left( 1 + \exp(Z_i^T \beta) \tilde{G}(t-., \beta) \right)^{-1} \right\}^\Delta N_i(t),$$

Derivative with respect to $\beta$ of the log of this (pseudo) profile-likelihood is $\tilde{U}(\beta)$.

- $\hat{\beta}$ solves $\tilde{U}(\beta) = 0$; $\hat{G}(t) = \tilde{G}(t, \hat{\beta})$. 
Asymptotics for modified partial likelihood estimators

\[ n^{1/2}(\tilde{G}(t, \beta_0) - G_0(t)) = n^{-1/2} \int_0^t \frac{1}{n^{-1} S_0(s, \beta_0, \tilde{G})} dM_s(s) \]

\[ + n^{1/2} \int_0^t \frac{S_0(s, \beta_0, G_0) - S_0(s, \beta_0, \tilde{G})}{S_0(s, \beta_0, \tilde{G})} dG_0(s). \]

so converges to a process that solves a stochastic integral equation, (Bagdonavicius and Nikulin, 1999)

\[ n^{1/2}(\hat{\beta} - \beta_0) = \{n^{-1} I(\hat{\beta})\}^{-1} n^{-1/2} \tilde{U}(\beta_0), \]

\[ n^{1/2} \left( \hat{G}(t) - G_0(t) \right) = n^{1/2} \sum_{i=1}^n H_i(t, \beta_0) \]
Breast cancer data

```r
> cancer[1:3,]
  stain time status
1   1   23    1
2   1   47    1
3   1   69    1
> fit1<-prop.odds(Surv(time,status)~factor(stain),n.sim=100)
> summary(fit1)
Covariate effects
               Coef. Std. Error Robust SE D2log(L)^-1
factor(stain)2  1.32     0.597     0.589   0.634
```

![Graph showing Kaplan-Meier and PropOdds survival curves]
Test of the PO-model

- (Dauxois and Kirmani, 2003) considered test of the PO-model in a two-sample situation. With $\phi_j(t) = (1 - S_j(t))/S_j(t)$, we have $\phi_1(t) = e^{\beta} \phi_0(t)$. Let

$$
\psi_{jk} = \int_{\tau_1}^{\tau_2} k_j(t) \phi_k(t) \, dt, \quad j = 1, 2; \, k = 1, 2,
$$

and

$$
\gamma(k_1, k_2) = \psi_{11} \psi_{22} - \psi_{12} \psi_{21}.
$$

that is 0 if and only if the proportional odds model holds. Now use Kaplan-Meier estimates to calculate the test-statistic ($\hat{\phi}_j(t)$). For the breast cancer data one obtains, $p=0.92$.

- We will now consider the model

$$
\logit(1 - S_Z(t)) = \log(G(t)) + Z^T \beta(t), \quad (4)
$$

where the OR changes with time.
Time-varying effects in the PO-model

- Assuming

\[
\text{logit}(1 - S_Z(t)) = \log(G(t)) + Z^T \beta(t), \quad t > 0 \quad S_Z(t) = \frac{1}{1 + G(t)e^{Z^T \beta(t)}},
\]

\(G(0) = 0, \quad G(\infty) = \infty,\) corresponds to the hazard function

\[
\lambda_Z(t) = \frac{e^{Z^T \beta(t)}G'(t) + G(t)e^{Z^T \beta(t)}Z^T \beta'(t)}{1 + G(t)e^{Z^T \beta(t)}}, \quad (5)
\]

which we note in passing is different from

\[
\frac{e^{Z^T \beta(t)}G'(t)}{1 + G(t)e^{Z^T \beta(t)}}.
\]
Time-varying effects in the PO-model

• The hazard function may be written as

\[
\left( \frac{e^{Z^T\beta(t)}}{1 + G(t)e^{Z^T\beta(t)}} \right) \left( \frac{G(t)e^{Z^T\beta(t)}Z^T}{1 + G(t)e^{Z^T\beta(t)}} \right) \left( \begin{array}{c} G'(t) \\ \beta'(t) \end{array} \right),
\]

an Aalen model, so \( v(t) = (G(t), \beta(t))^T \) may be estimated by

\[
\hat{v}(t) = \int_0^t A^-(s, \hat{v}(s-))dN(s),
\]

where \( A(t, v(t-)) \) is the design matrix with \( i \)th row

\[
Y_i(t)\left( \frac{e^{Z_i^T\beta(t)}}{1 + G(t)e^{Z_i^T\beta(t)}} \right) \left( \frac{G(t)e^{Z_i^T\beta(t)}Z_i^T}{1 + G(t)e^{Z_i^T\beta(t)}} \right).
\]
Time-varying effects in the PO-model

Problems regarding estimation:

- Estimator of $v(t) = (G(t), \beta(t))^T$:

$$\hat{v}(t) = \int_0^t A_-(s, \hat{v}(s-))dN(s),$$

has a recursive structure. It needs to be started: $G(0) = 0$, but $\beta(0) = ?$.

- Design $Y_i(t)(\frac{e^{Z_i^T \beta(t)}}{1+G(t)e^{Z_i^T \beta(t)}}, \frac{G(t)e^{Z_i^T \beta(t)}Z_i^T}{1+G(t)e^{Z_i^T \beta(t)}})$ becomes singular when approaching zero.

- Consider for a moment test instead of estimation.
Test of time-varying effects in the PO-model

• Consider test of $H_0 : \beta(t) = k$

• Put $v_\delta(t) = \int_\delta^t dv(s)$ (with second component equal to zero under $H_0$) and

$$\hat{v}_\delta(t, \hat{\theta}) = \int_\delta^t A^{-}(s, \hat{v}(s-, \hat{\theta}))dN(s); \quad \hat{\theta} = (\hat{G}(\delta), \hat{\beta})^T (PO\text{-model})$$

for some prespecified $\delta$. Define $\hat{\beta}_\delta(t, \hat{\theta})$ as the second component in the latter display.

• Test-statistic: $\sup_{\delta \leq t \leq \tau} |n^{1/2} \beta_\delta(t, \hat{\theta}) - 0|$.

• $n^{1/2}(\hat{v}_\delta(t, \hat{\theta}) - v_\delta(t)) = n^{1/2}(\hat{v}_\delta(t, \theta) - v_\delta(t)) + n^{1/2}(\hat{v}_\delta(t, \hat{\theta}) - \hat{v}_\delta(t, \theta)) = (1) + (2)$. 
Test of time-varying effects in the PO-model

- $n^{1/2}(\hat{v}_\delta(t, \hat{\theta}) - v_\delta(t)) = n^{1/2}(\hat{v}_\delta(t, \theta) - v_\delta(t)) + n^{1/2}(\hat{v}_\delta(t, \hat{\theta}) - \hat{v}_\delta(t, \theta)) = (1) + (2)$.

- (1) converges in dist. to a process that solves a Volterra integral equation (under the null). Solution to this is a product intergral, see (Andersen et al, 1993).

- (2) is in the limit a functional of limit dist. of $n^{1/2}(\hat{\theta} - \theta)$ (under the null).

- Can obtain $n^{1/2}(\hat{v}_\delta(t, \hat{\theta}) - v_\delta(t)) = n^{1/2} \sum_{i=1}^{n} H_i(t)F_i$ with the $F_i$’s iid $N(0, 1)$, so resampling is possible.
Breast cancer data
Time-varying effects in the PO-model: Estimation

Want to use the estimator of \( v(t) = (G(t), \beta(t))^T \): \( \hat{v}(t) = \int_0^t A^-(s, \hat{v}(s-))dN(s) \).

Needs to be started. Follow here (Chen et al., 2002): Let \( \Lambda(x) = \log(1 + e^x) \), \( H(t) = \log(G(t)) \) with \( H(0) = -\infty \). Generic model then gives the martingale:

\[
M(t) = N(t) - \int_0^t Y(s)d\Lambda(Z^T\beta(s) + H(s))
\]

The increment at time \( t_1 \) of the compensator is:

\[
Y(t_1)\Lambda(Z^T\beta(t_1) + H(t_1)) - Y(t_1-)\Lambda(Z^T\beta(t_1-) + H(t_1-)) = Y(t_1)\Lambda(Z^T\beta(t_1) + H(t_1))
\]

Minimize

\[
\sum_{i=1}^n (dN_i(t_1) - Y(t_1)\Lambda(Z^T\beta(t_1) + H(t_1)))^2
\]

wrt \( \beta(t_1) \) and \( H(t_1) \) does not work.
Discussion

• Semiparametric model

\[
\text{logit}(1 - S_Z(t)) = \log(G(t)) + Z^T \beta(t) + X^T \gamma, \quad S_Z(t) = \frac{1}{1 + G(t)e^{Z^T \beta(t) + X^T \gamma}},
\]

It is possible to test \( H_0 : \beta(t) = k \), but how about estimation?

• Alternative model

\[
\text{logit}(1 - S_Z(t)) = \log(\beta_0(t) + Z^T \beta(t)) + X^T \gamma,
\]

\( \beta(0) = 0 \). Here, estimation is possible and implemented (\texttt{timereg}). This model, however, does not lead to OR that changes with time in a simple way:

\[
\text{OR}_{1:2} = \frac{e^{\log(\beta_0(t) + Z_1^T \beta(t)) + X^T \gamma}}{e^{\log(\beta_0(t) + Z_2^T \beta(t)) + X^T \gamma}} = \frac{\beta_0(t) + Z_1^T \beta(t)}{\beta_0(t) + Z_2^T \beta(t)} = \frac{\beta_0(t) + \beta_1(t) + Z_2 \beta_1(t)}{\beta_0(t) + Z_2 \beta_1(t)}
\]

if \( Z_1 = Z_2 + 1 \).
References


