Model Choice in Cox-Type Additive Hazard Regression Models with Time-Varying Effects

Benjamin Hofner[1]
Wolfgang Hartl   Helmut Küchenhoff   Thomas Kneib

Department of Statistics
Statistical Consulting Unit
Ludwig-Maximilians-Universität München

11.03.2008

[1]benjamin.hofner@stat.uni-muenchen.de

Introduction

Cox PH model:

$$\lambda_i(t) = \lambda(t, \mathbf{z}_i) = \lambda_0(t) \exp(\mathbf{z}_i' \gamma)$$

with

- $\lambda_i(t)$ hazard rate of observation i $[i = 1, \dots, n]$
- $\lambda_0(t)$ baseline hazard rate
- $\mathbf{z}_i$ vector of covariates for observation i $[i = 1, \dots, n]$
- γ vector of regression coefficients

Problem: restrictive model, not allowing for

- non-proportional hazards (i.e. time-varying effects)
- non-linear effects

Motivation from Application

- **Why** do we need time-varying and non-linear effects?
- **Why** do we need variable selection?

Answer: Data at hand

- **Question:** treatment benefit in terms of 90-day survival
- **retrospective study** $\Rightarrow$ sensible confounder model needed allowing for
  - variable selection (which variables)
  - model choice (how to model these)

Motivation from Application

- **Why** do we need time-varying and non-linear effects?
- **Why** do we need variable selection?

Answer: Data at hand

- **Question:** treatment benefit in terms of 90-day survival
- **retrospective study** ⇒ sensible confounder model needed allowing for
  - variable selection (which variables)
  - model choice (how to model these)

Semiparametric Representation

Generalisation: Additive Hazard Regression [Kneib & Fahrmeir, 2007]

$$\lambda_i(t) = \exp(\eta_i(t))$$

with

$$\eta_i(t) = g_0(t) + \sum_{l=1}^{L} g_l(t) u_{il} + \sum_{j=1}^{J} f_j(x_{ij}) + \mathbf{z}_i' \gamma$$

where

- $g_o(t) = \log(\lambda_0(t))$ log-baseline ($\Rightarrow$ full likelihood available)
- $g_l(t)$ time-varying effects of covariates u_{il} $[l = 1, \ldots, L]$
- $f_j(x_{ij})$ smooth effects of covariates x_{ij} $[j = 1, \ldots, J]$
- $\mathbf{z}_i' \gamma$ as before

P-splines

flexible terms can be represented using P-splines [Eilers & Marx, 1996]

- model term:

$$f_j(x) = \sum_{m=1}^{M} \beta_{jm} B_{jm}(x) \quad \text{(analogous for } g_0 \text{ and } g_l\text{)}$$

- penalty:

$$\text{pen}(\boldsymbol{\beta_j}) = \kappa_j \, \boldsymbol{\beta_j}' \mathbf{K} \boldsymbol{\beta_j} \quad \text{(analogous for } g_0 \text{ and } g_l\text{)}$$

with

- $\mathbf{K} = \mathbf{D}'\mathbf{D}$ (i.e. cross product of difference matrix $\mathbf{D}$)
- κ_j smoothing parameter

Inference

Estimation based on Penalised Likelihood Criterion:

(NB: this is the **full** log-likelihood)

$$l = \sum_{i=1}^{n} \left[\delta_i \eta_i(t_i) - \int_0^{t_i} \lambda_i(t) dt \right] - \sum_{l=0}^{L} \text{pen}(\beta_l) - \sum_{j=1}^{J} \text{pen}(\beta_j)$$

Estimates for coefficients **and** smoothing parameters:
using mixed model based inference [Kneib & Fahrmeir, 2007]
(implemented in BayesX)

- T_i true survival time
- C_i censoring time
- $t_i = \min(T_i, C_i)$ observed survival time (right censoring)
- $\delta_i = \mathbb{1}(T_i \leq C_i)$ indicator for non-censoring

Model Choice

First Conclusion

- Estimation possible (given model structure)
- Variable selection (what to include) and model choice (how to include) not straight forward
- $\Rightarrow$ **Two-Stage Stepwise Procedure** [Hofner et al., 2008]

Side Note on Information Criterion

Remember: Estimation in a mixed model framework
Penalty represented by Gaussian random effects
most frequently used in this context: marginal AIC (not suitable here)
$\Rightarrow$ **use conditional AIC** instead:

$$\mathrm{AIC}_c = -2l + 2df$$

Model Choice

First Conclusion

- Estimation possible (given model structure)
- Variable selection (what to include) and model choice (how to include) not straight forward
- $\Rightarrow$ **Two-Stage Stepwise Procedure** [Hofner et al., 2008]

Side Note on Information Criterion

Remember: Estimation in a mixed model framework
Penalty represented by Gaussian random effects
most frequently used in this context: marginal AIC (not suitable here)
$\Rightarrow$ **use conditional AIC** instead:

$$\mathrm{AIC}_c = -2l + 2df$$

Two-Stage Stepwise Procedure

Starting Model: typically: empty model
(i.e. only baseline hazard rate)

Initial Choice Set: covariates not already included in the starting model

(i) **Modelling Alternatives:**
for each covariate in the choice set
- categorical: fixed vs. time-varying effect
- continuous: fixed vs. nonparametric vs. time-varying effect

(ii) **Estimation of Models:**
for each covariate and each modelling possibility:
- add effect to current model
- estimate hazard regression model
- store conditional AIC

Two-Stage Stepwise Procedure (ctd.)

(iii) **Selection Step** *with stopping criterion:*

- *Improvement of* AIC_c*:*
  - current model := best-fitting model (i.e. with $\min(\text{AIC}_c)$)
  - delete corresponding covariate from choice set
  - continue with step (iv)
- *Otherwise:*
  - terminate the algorithm

(iv) **Backward Deletion:**

- perform (classical) backward deletion step on current model
- *Improvement* $\Rightarrow$ add deleted covariate to choice set
- continue with step (i)

Toy Example

Variable **(stage 1)**	Modelling Alternative **(stage 2)**	AIC_c in step 1		2		3
Apache II score (continuous)	linear	3188.09	backward deletion: not applicable	–	backward deletion: no improvement	–
	smooth	**3186.21**		–		–
	time-varying	3188.37		–		–
palliative operation (categorical)	linear	3530.43		**3176.31**		–
	time-varying	3532.26		3177.98		–
age (continuous)	linear	3524.45		3178.18		**3168.55**
	smooth	3525.74		3178.37		3168.58
	time-varying	4073.94		3697.34		3685.98

Question / Data

Detailed Question:

Do surgical patients with severe sepsis have a treatment benefit in terms of 90-day survival from an activity-guided antithrombin III (AT 3) therapy?

Some more details on data

- **response:** 90-day survival
- **predictors:** 14 categorical predictors, 6 continous predictors
- **origin:** local database
  (Department of Surgery, Campus Großhadern, LMU Munich)
- **period of observation:** March 1st, 1993 – February 28th, 2005
- **N:** 545 septic patients [Moubarak et al., 2008]
  (462 complete cases used, 180 observations right-censored)

Some Changes in Fitting Procedure

Two-Stage Stepwise Procedure used for the Großhadern dataset:

- Starting Model **not** empty: 6 preset variables (age, sex, . . .)
  - modelling alternatives not fixed
  - ⇒ **Two-Stage Stepwise Procedure** without stopping criterion (i.e. model choice without variable selection)
- Build confounder model with starting model (NB: variables from starting model are not subject to backward deletion)
- **Last step:** add "AT 3"

Results: Confounder Model

Confounder model consists of

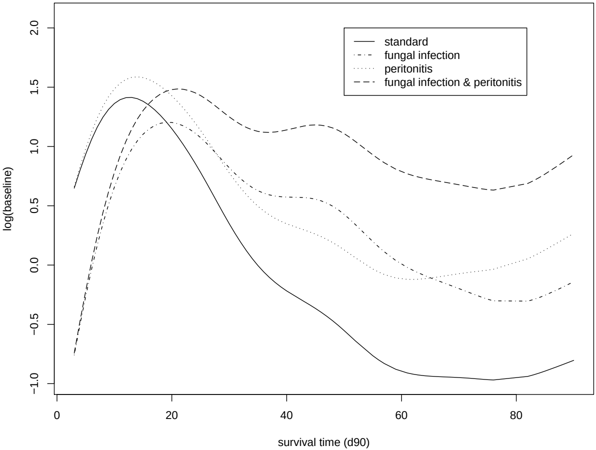
- 6 preset variables and
- 8 additional variables

with

- 3 smooth terms and
- 2 time-varying terms (only chosen for binary variables)

Results: Confounder Model (ctd.)

Time-Varing Effects (shown as log(baseline) in subgroups)



Results for AT3

Adding AT 3 as linear term leads to:

β_{AT3}	0.0385
Std. Dev.	0.1473
95% CI	$[-0.250, 0.327]$

$exp\{\beta_{AT3}\}$	1.0393
95% CI	$[0.779, 1.387]$

p-value	0.7937

Adding AT 3 as time-varying term leads to:

Results for AT3

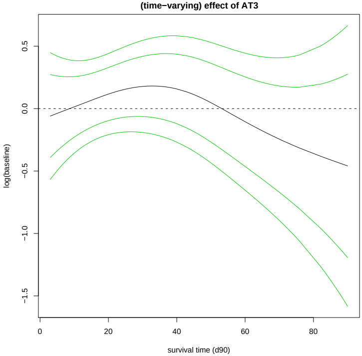
Adding AT 3 as linear term leads to:

β_{AT3}	0.0385
Std. Dev.	0.1473
95% CI	$[-0.250, 0.327]$

$exp\{\beta_{AT3}\}$	1.0393
95% CI	$[0.779, 1.387]$

p-value	0.7937

Adding AT 3 as time-varying term leads to:



Summary & Outlook

Two-Stage Stepwise Procedure...

- ...allows variable selection and model choice.
- ...allows flexible modelling (e.g. non-proportional hazard models).
- ...is not only applicable in survival models but in any type of flexible regression model.
- ...is expandable to interactions, spatial effects, ...
- ...could be used with fractional polynomials and other approaches.

Literature

- EILERS, P.H.C. & MARX, B.D. (1996). Flexible smoothing using B-splines and penalties. *Statistical Science*, **11**, 89–121.
- HOFNER, B., KNEIB, T., HARTL, W., KÜCHENHOFF, H. (2008). Model Choice in Cox-Type Additive Hazard Regression Models with Time-Varying Effects. *In Preparation*.
- KNEIB, T. & FAHRMEIR, L. (2007) A mixed model approach for geoadditive hazard regression. *Scand. J. Statist.*, **34**, 207–228.
- MOUBARAK, P., ZILKER, S., WOLF, H., HOFNER, B., KNEIB, T., KÜCHENHOFF, H., JAUCH, K.-W. & HARTL. W. H. (2008). Activity-guided antithrombin III therapy in severe surgical sepsis: Efficacy and safety according to a retrospective data analysis. *Submitted to Shock*.