Workshop on Semiparametric Models and Applications
In Honor of Professor Catherine Huber-Carol
Mont Saint Michel, France, May 15-17, 2003

The Covariate Order Method for Nonparametric Exponential Regression and Some Applications in Other Lifetime Models

Bo Lindqvist

Department of Mathematical Sciences,
Norwegian University of Science and Technology,
Trondheim

bo@math.ntnu.no
http://www.math.ntnu.no/~bo/

Joint work with Jan Terje Kvaløy,
Stavanger University College, Norway

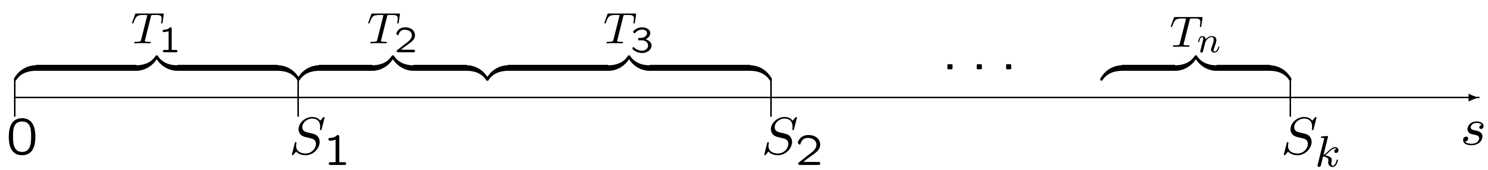
MOTIVATION

- Suppose $(T_1, \delta_1), (T_2, \delta_2), \ldots$ is a set of (possibly) censored *i.i.d. exponential lifetimes* from distribution $\mathrm{E}(\lambda)$.

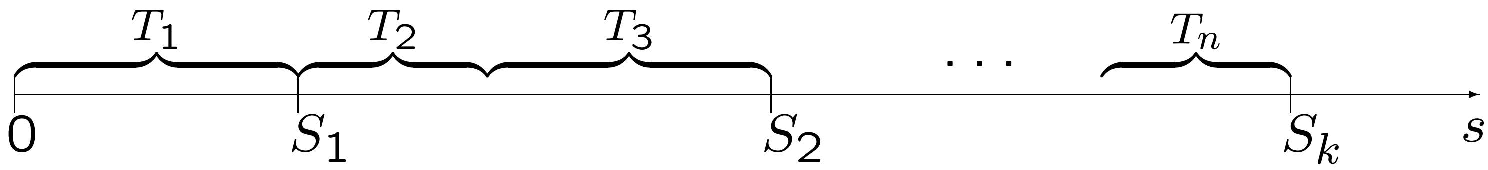
- Let δ be censoring indicator, independent censoring.

- For example, $\delta_1 = 1, \delta_2 = 0, \delta_3 = 1, \ldots$

- Then $S_1, S_2, \ldots$ is a homogeneous Poisson-process (HPP)



$S_1, S_2, \ldots$ is an HPP ...



SIMPLE APPLICATIONS

- Derive standard (χ^2-based) confidence interval for rate of exponential distribution, λ.

- Sort $(T_1, \delta_1), (T_2, \delta_2), \ldots$ according to some secondary variable (e.g. covariate) to see whether there is an unwanted/unexpected trend (implying deviations from HPP).

APPLICATION IN COX-REGRESSION

MODEL: Hazard function

$$\alpha(t|\boldsymbol{x}) = \alpha_0(t)\,\mathrm{exp}(\boldsymbol{\beta}\,'\boldsymbol{x})$$

BASIC RESULT for lifetime Z

$A_0(Z)\,\mathrm{exp}(\boldsymbol{\beta}\,'\boldsymbol{X}) \sim \mathrm{E}(1)$ given $\boldsymbol{X}$

$(A_0(t) = \int_0^t \alpha_0(u)du)$.

COX-SNELL RESIDUALS
for data $(\boldsymbol{T_1}, \boldsymbol{\delta_1}, \boldsymbol{X_1}), \cdots, (\boldsymbol{T_n}, \boldsymbol{\delta_n}, \boldsymbol{X_n})$

$\widehat{r}_i = \widehat{A}_0(T_i)\,\mathrm{exp}(\hat{\boldsymbol{\beta}}\,'\boldsymbol{X}_i)$, $i = 1, \ldots, n$

$(\widehat{r}_1, \delta_1), \ldots, (\widehat{r}_n, \delta_n)$ *should behave like a censored sample from E*(1).

EXPONENTIAL REGRESSION MODEL

- Z = lifetime, C = censoring time, $\boldsymbol{X}$ = covariate vector

ASSUMPTIONS

- $Z|\boldsymbol{X} = \boldsymbol{x}$ is $\mathrm{E}(\lambda(\boldsymbol{x}))$, i.e. has density $\lambda(\boldsymbol{x})e^{-\lambda(\boldsymbol{x})t}$

- $C|\boldsymbol{X} = \boldsymbol{x}$ has density $f_C(t|\boldsymbol{x})$

- Z, C are independent given $\boldsymbol{X}$

AIM

- Estimate $\lambda(\boldsymbol{x})$ from sample of $(T, \delta, \boldsymbol{X})$ where $T = \min(Z, C)$, $\delta = I(Z \leq C)$

COVARIATE ORDERING

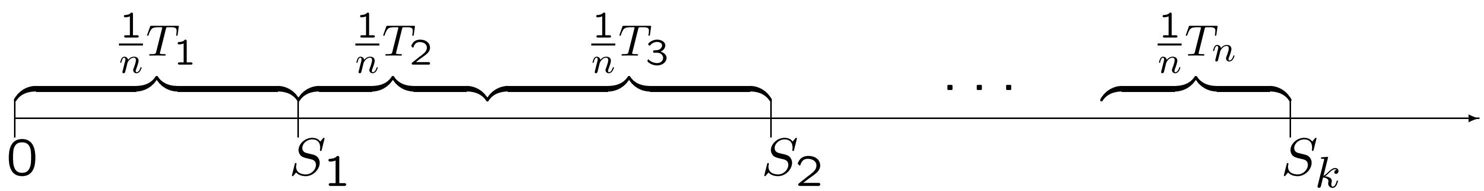
(single continuous covariate)

Order data $(T_1, \delta_1, X_1), \cdots, (T_n, \delta_n, X_n)$

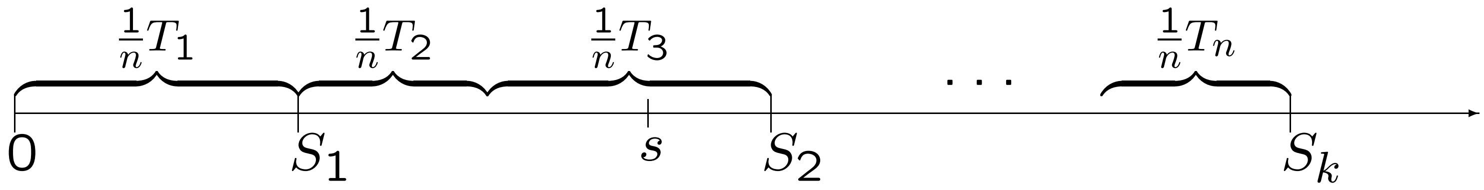
with $X_1 \leq X_2 \leq \cdots \leq X_n$

COVARIATE ORDER PROCESS

Point process $S_1, S_2, \ldots$ formed by successive $(1/n)T_i$, with *events* defined at end of uncensored $(1/n)T_i$.



COVARIATE ORDER ESTIMATOR



Conditional intensity of process $S_1, S_2, \ldots$ is

$$\rho(s) = n\lambda(X(s))$$

where $X(s)$ is the x corresponding to the time T_i under "risk" at s.

STEP 1: Estimate $\rho(s)$ by ordinary kernel estimator

$$\widehat{\rho}(s) \equiv \frac{1}{nh_s} \sum_{i=1}^{k} K\left(\frac{s - S_i}{h_s}\right)$$

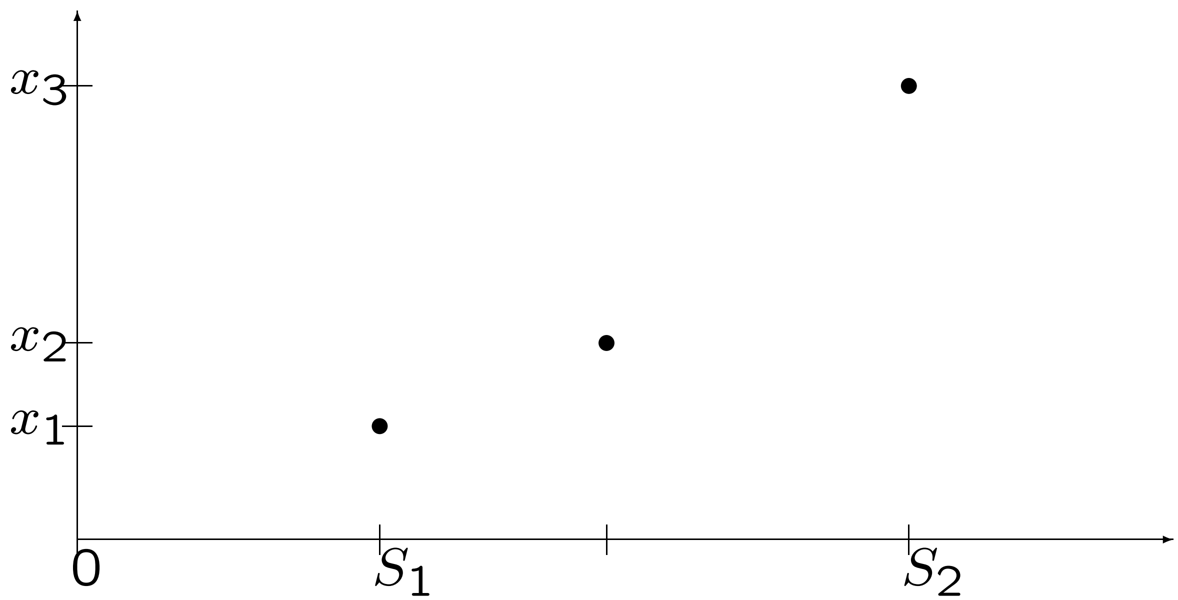
(or possibly another (sub)density estimator).

STEP 2: Invert relation $\widehat{\rho}(s) = n\widehat{\lambda}(X(s))$ to get

$$\widehat{\lambda}(x) = \widehat{\rho}(\widehat{s}(x))$$

where $\widehat{s}(x)$ is the s "corresponding to" x.

In practice: The *correspondence function* $\widehat{s}(x)$ is obtained by some smooth of the points $(\sum_{i=1}^{m} T_i, X_i)$, $i = 1, \ldots, n$.



THEOREM

Let $K(\cdot)$ be a positive kernel function and let h be a smoothing parameter which may depend on x. Then the estimator

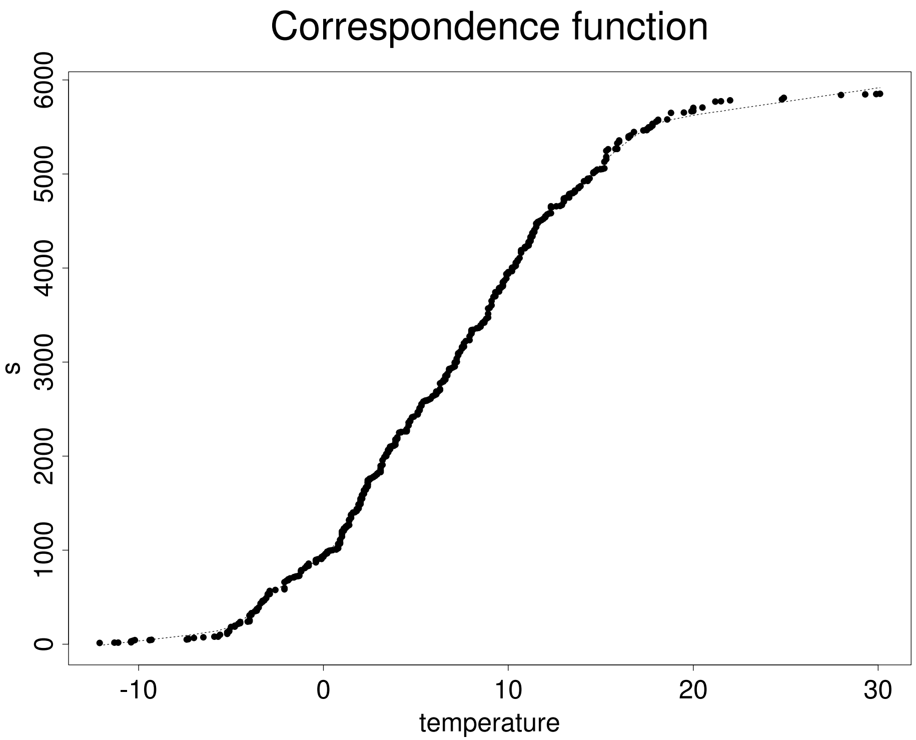
$$\hat{\lambda}(x) = \frac{1}{nh} \sum_{i=1}^{r} K\left(\frac{\widehat{s}(x) - S_i}{h}\right) \ ; \ \ x \in \mathcal{X}$$

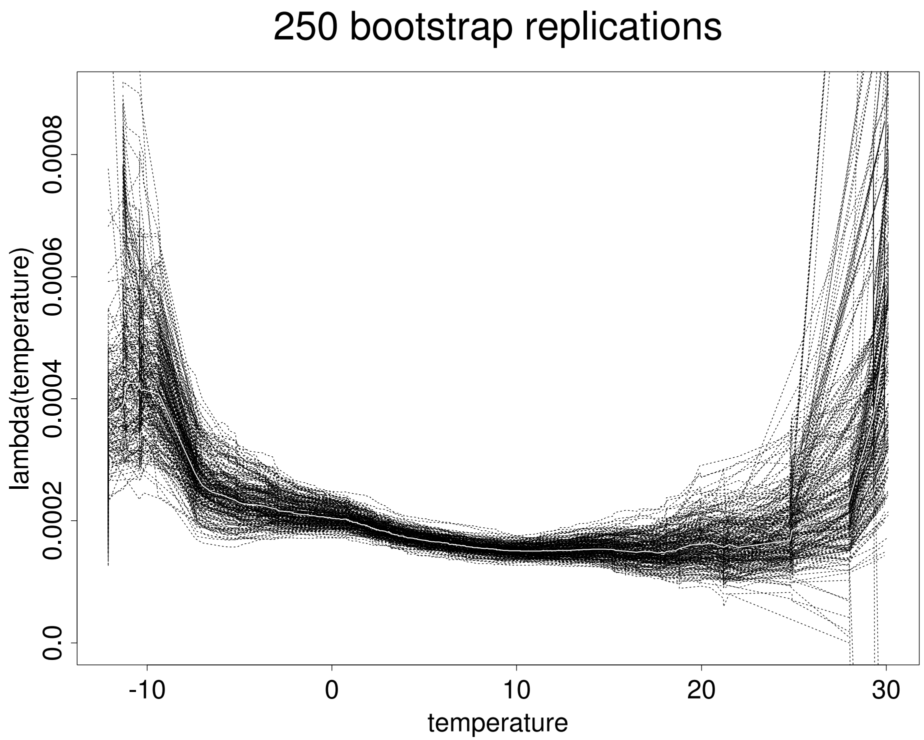
is a uniformly consistent estimator of $\lambda(x)$.

EXAMPLE – CARDIAC ARREST

Times of out-of-hospital cardiac arrests reported to a Norwegian hospital over a 5 years period.

$Z =$ inter-event times, $X =$ temperature.





COX-SNELL RESIDUALS
for data $(\boldsymbol{T_1}, \boldsymbol{\delta_1}, \boldsymbol{X_1}), \cdots, (\boldsymbol{T_n}, \boldsymbol{\delta_n}, \boldsymbol{X_n})$

$\widehat{r}_i = \widehat{A}_0(T_i)\exp(\hat{\boldsymbol{\beta}}'\boldsymbol{X}_i),\ i = 1, \ldots, n$

Under correct model: $(\widehat{r}_1, \delta_1), \ldots, (\widehat{r}_n, \delta_n)$ behave like censored sample from E(1).

RESIDUAL PLOTS

For each single covariate X_k, apply covariate order method to

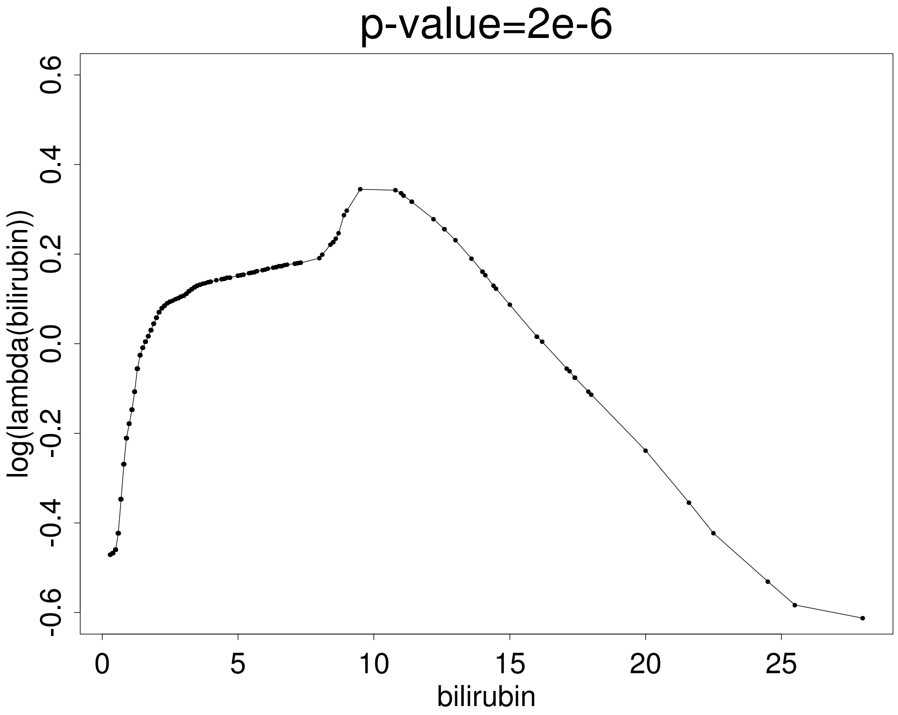
$$(\widehat{r}_1, \delta_1, X_{1k}), \ldots, (\widehat{r}_n, \delta_n, X_{nk}),$$

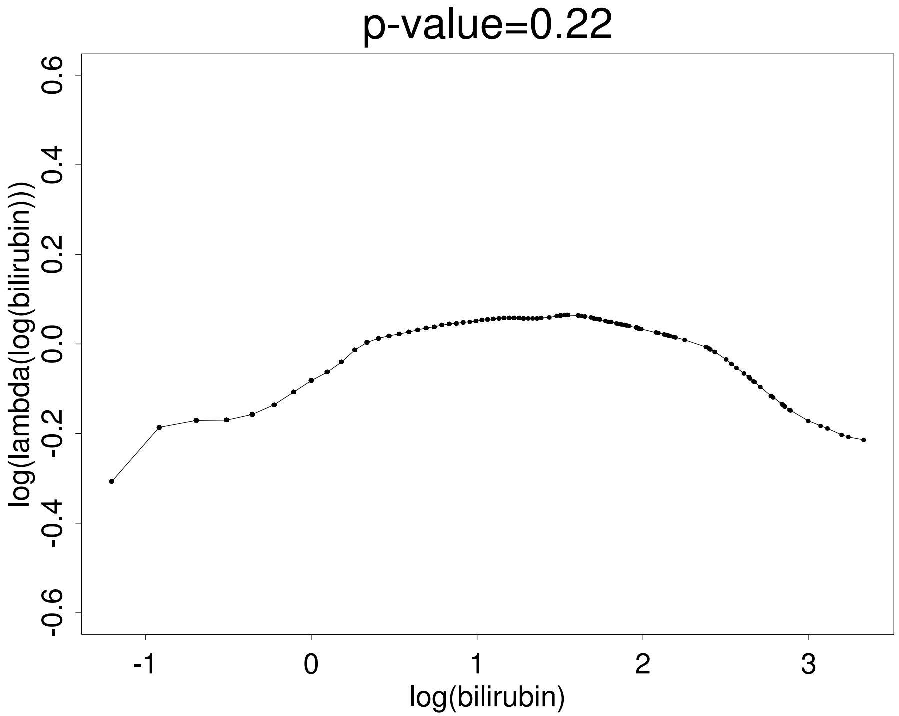
where X_{ik} is the kth covariate for the ith observation unit.

EXAMPLE - RESIDUAL PLOTS FOR PBC-DATA

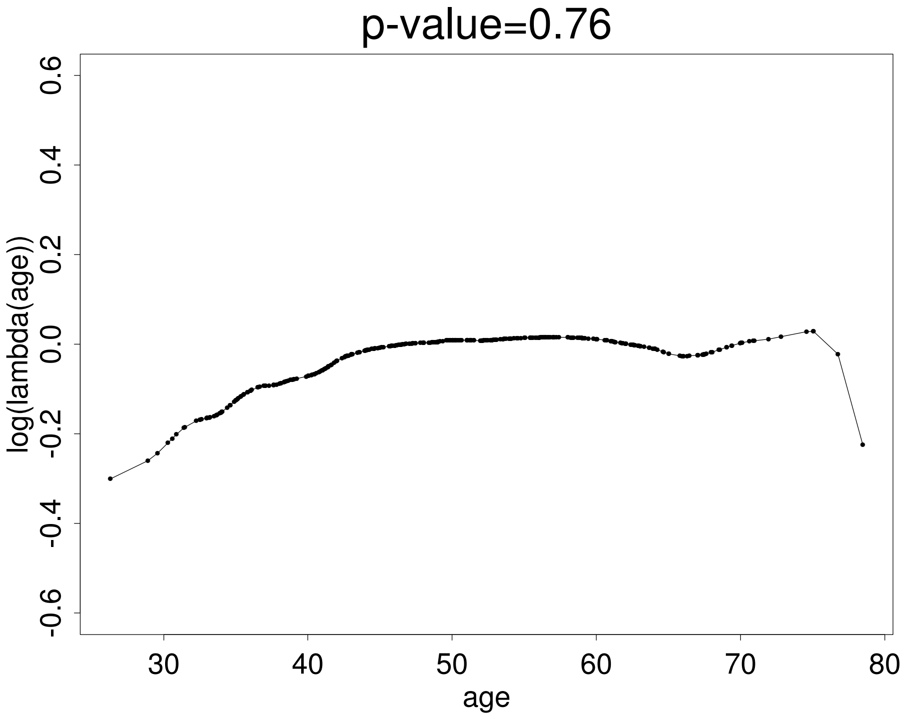
Estimated log($\lambda(x)$) for Cox-Snell residuals vs covariates $x =$*bilirubin* and $x =$ log*(bilirubin)*, respectively.

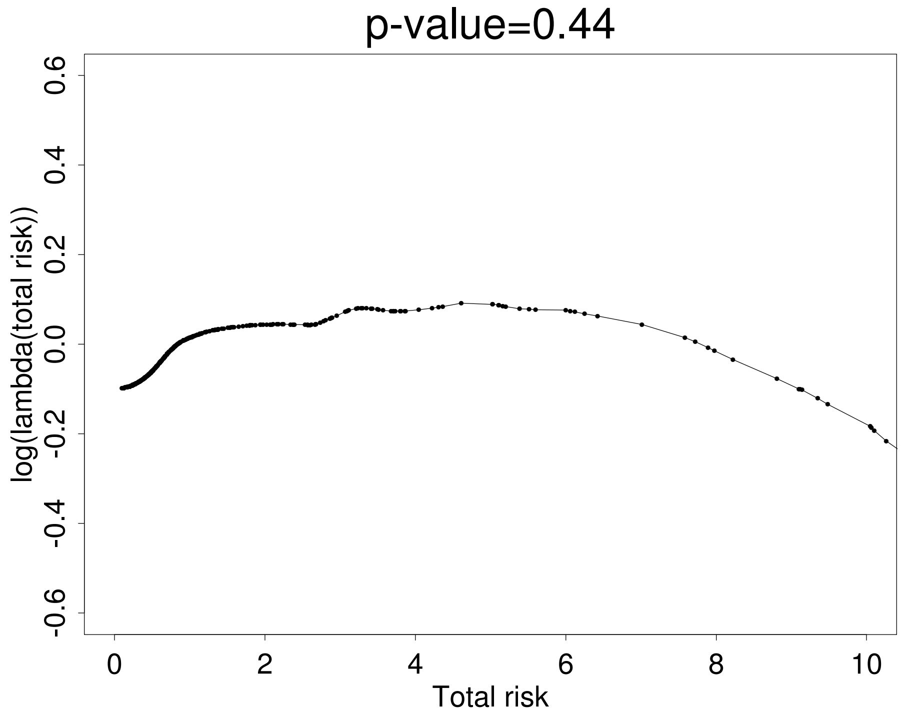
P-value from Anderson-Darling test for the null hypothesis of constant hazard function.





OTHER RESIDUAL PLOTS FOR PBC-DATA





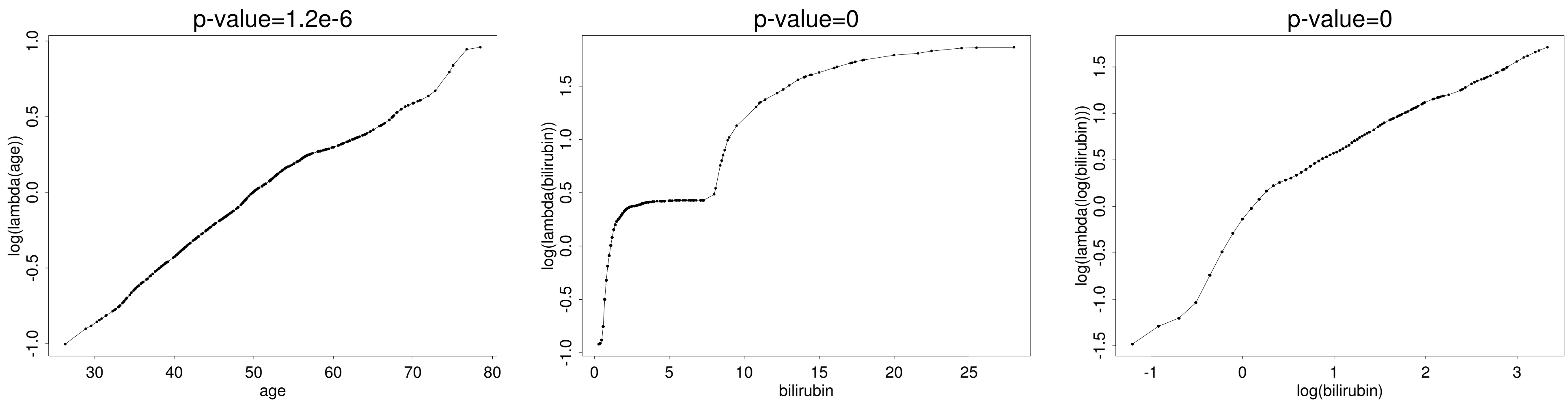
EXAMPLE - FUNCTIONAL FORM FOR COVARIATES IN PBC-DATA

BASIC RESULT (single covariate):
$A_0(Z)\exp(\beta X) \sim \mathrm{E}(1)$ given X.

This implies $A_0(Z) \sim \mathrm{E}(\exp(\beta X))$

Use covariate order method to data $(\hat{A}_0(T_i), \delta_i, X_{ik})$ to suggest functional form for kth covariate.

Plots of log-hazards - straight line implies functional form $\exp(\beta x)$:



SEVERAL COVARIATES

$\boldsymbol{X} = (X_1, \ldots, X_m)$

Generalization to several covariates is not immediate without imposing structure

GENERALIZED ADDITIVE MODEL

$$\lambda(\boldsymbol{x}) = e^{\alpha + g_1(x_1) + , \ldots , + g_m(x_m)}$$

$g_1(\cdot), \ldots, g_m(\cdot)$ estimated iteratively (backfitting)

Basic result:

$Z \sim \mathsf{E}(e^{\alpha + g_1(x_1) + \ldots + g_m(x_m)})$

$$\Downarrow$$

$Z e^{\alpha + g_1(x_1) + \ldots + g_{j-1}(x_{j-1}) + g_{j+1}(x_{j+1}) + \ldots + g_m(x_m)} \sim \mathsf{E}(e^{g_j(x_j)}).$

CONCLUSIONS - COVARIATE ORDER METHOD

- Covariate order method is simple and intuitive
- It is easy to use with existing software
- Method is flexible due to “free” choice of density estimation method
- Applicable to non-exponential lifetime models (e.g. Cox-regression) by transformation.
- Simulations indicate that method is competitive w.r.t. competing methods, for example local likelihood methods, in particular for high censoring and few observations.

LITERATURE

Nonparametric lifetime regression:

Hastie and Tibshirani (book, 90)

Clayton & Cuzick (JRSS 85)

Staniswalis (JASA 89)

Gentleman & Crowley (BIOMCS 91)

Diagnostic plots for model checking in PH models:

Arjas (JASA 88)

Grambsch, Therneau & Fleming (BIOMCS 95)

Therneau & Grambsch (book, 00)

Covariate order method

Kvaløy (LDA 02)

Kvaløy & Lindqvist (Comp Stat 03)

Kvaløy & Lindqvist (Workshop and book)

THEOREM 1

Assume that

$$0 < a \leq \inf_{x\in\mathcal{X}} \lambda(x) \leq \sup_{x\in\mathcal{X}} \lambda(x) \leq M < \infty$$

and that $\sup_{x\in\mathcal{X}} \lambda'(x) \leq D < \infty$. Further assume that the conditional distribution of C given x has finite first and second order moments and that $f_C(t|x)$ has bounded first derivative in x for all $x \in \mathcal{X}$. Then

$$\rho_n(s|\mathcal{F}_s^n)/n \xrightarrow{p} \lambda(\eta(s))$$

as $n \to \infty$ uniformly in s, where $\eta(s)$ is a deterministic function from the s-axis to the covariate axis, the inverse of which is given by

$$s(x) = \mathsf{E}(TI(X \leq x)).$$

THEOREM 2

Let $K(\cdot)$ be a positive kernel function which vanishes outside [-1,1] and has integral 1, and let h_s be a smoothing parameter which is either constant or varying along the s-axis. Assume that $h_s \to 0$ as $n \to \infty$ for all s. Further assume that there is a sequence h_n such that $h_s \geq h_n$ for all s, n where $nh_n \to \infty$ as $n \to \infty$. Then the estimator

$$\widehat{\lambda}(x) = \frac{1}{nh_s} \sum_{i=1}^{r} K\left(\frac{\widehat{s}(x) - S_i}{h_s}\right) \; ; \;\; x \in \mathcal{X}$$

is a uniformly consistent estimator of $\lambda(x)$.