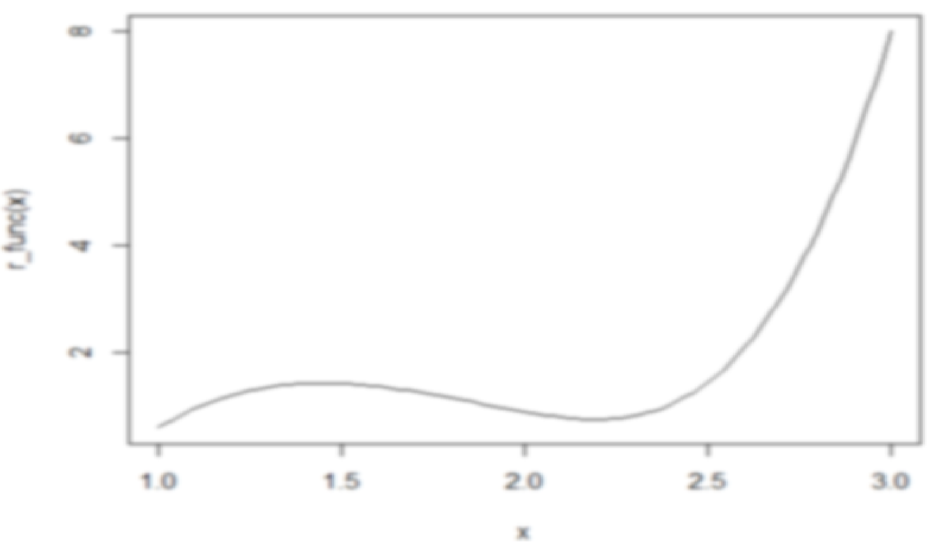
2019 6/3-6/7; MMR2019
11th Conference on Mathematical Methods in Reliability, Hong Kong

Programs for Semiparametric Cox Regression with Cubic M-spline

Emura T (National Central University, Taiwan)
Joint work with Shih JH



Semiparametric methods

- Cox regression (Cox 1972)

  - unspecified baseline hazard

- Cox model with splines (O'Sullivan 1998)

  - smooth polynomials for the baseline

  ↑ R programs have not been developed

Issues:

1. The best number of the basis functions
2. Any advantage over Cox regression?

- Baseline hazard function

$$\lambda_0(t) = \frac{\Pr(t \le X < t + dt \mid X \ge t, \mathbf{Z} = \mathbf{0})}{dt}$$

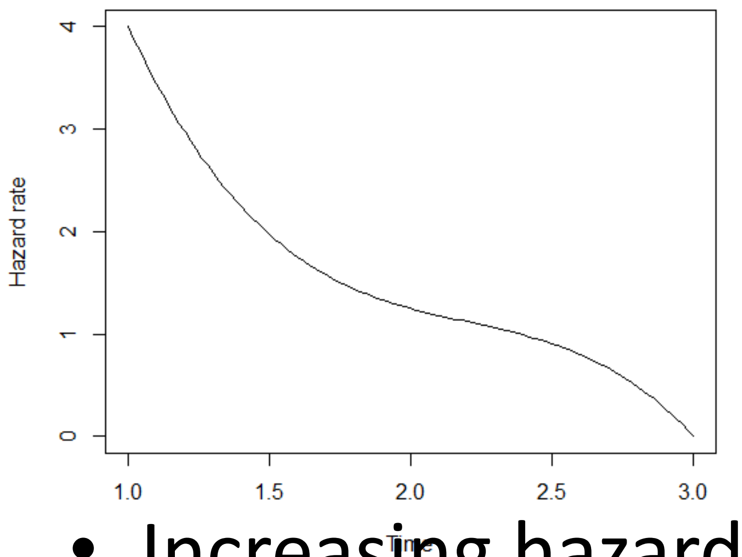
$X =$ time-to-event

$\mathbf{Z} =$ covariates

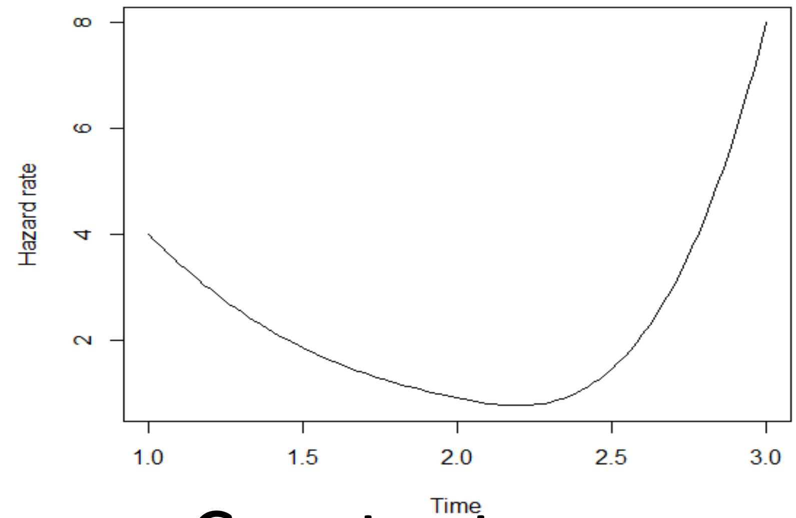
- The form $\lambda_0(\cdot)$ is unknown

  (decreasing/increasing, convex/concave, etc.)

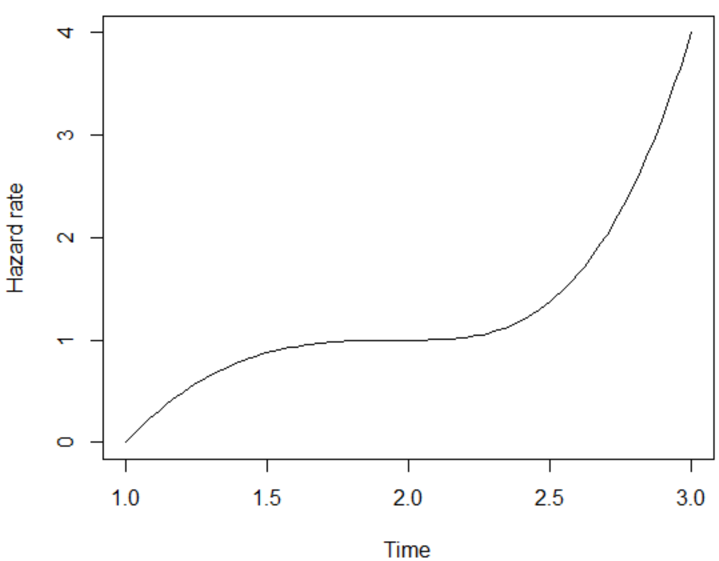
- Decreasing hazard



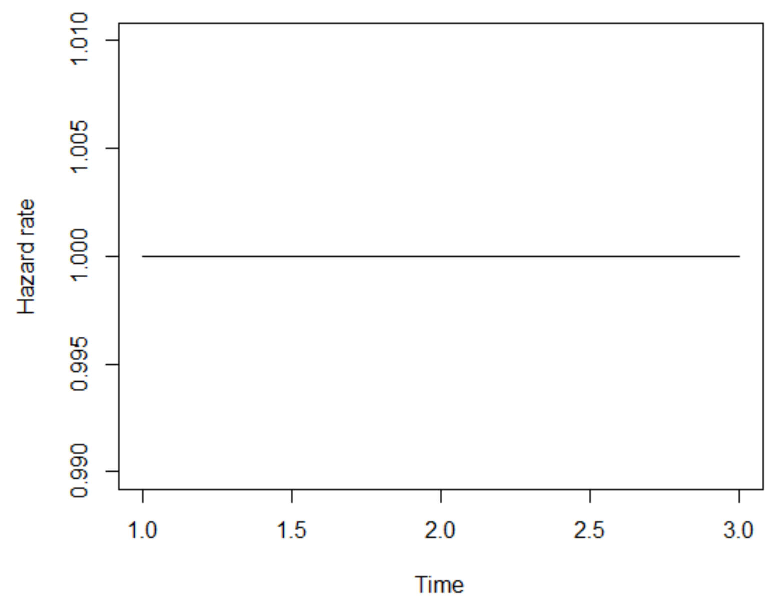
- Bus-tub shaped



- Increasing hazard



- Constant



Parametric hazard models

2-parameter models

- Weibull ➔ Increase or decrease
- Gamma ➔ Increase or decrease
- Lognormal ➔ Unimodal or decrease
- Log-logistic ➔ Unimodal or decrease
- Pareto ➔ Decrease

3-parameter models

- Burr (Burr 1942)
  ➔ Increase, decrease, or unimodal
- G- gamma (e.g. Balakrishnan and Pal 2015)
  ➔ Increase, decrease or unimodal.

4-parameter models

- Generalized F (G-gamma, Weibull, Lognormal, Log-logistic)

5-parameter spline model (Proposed)

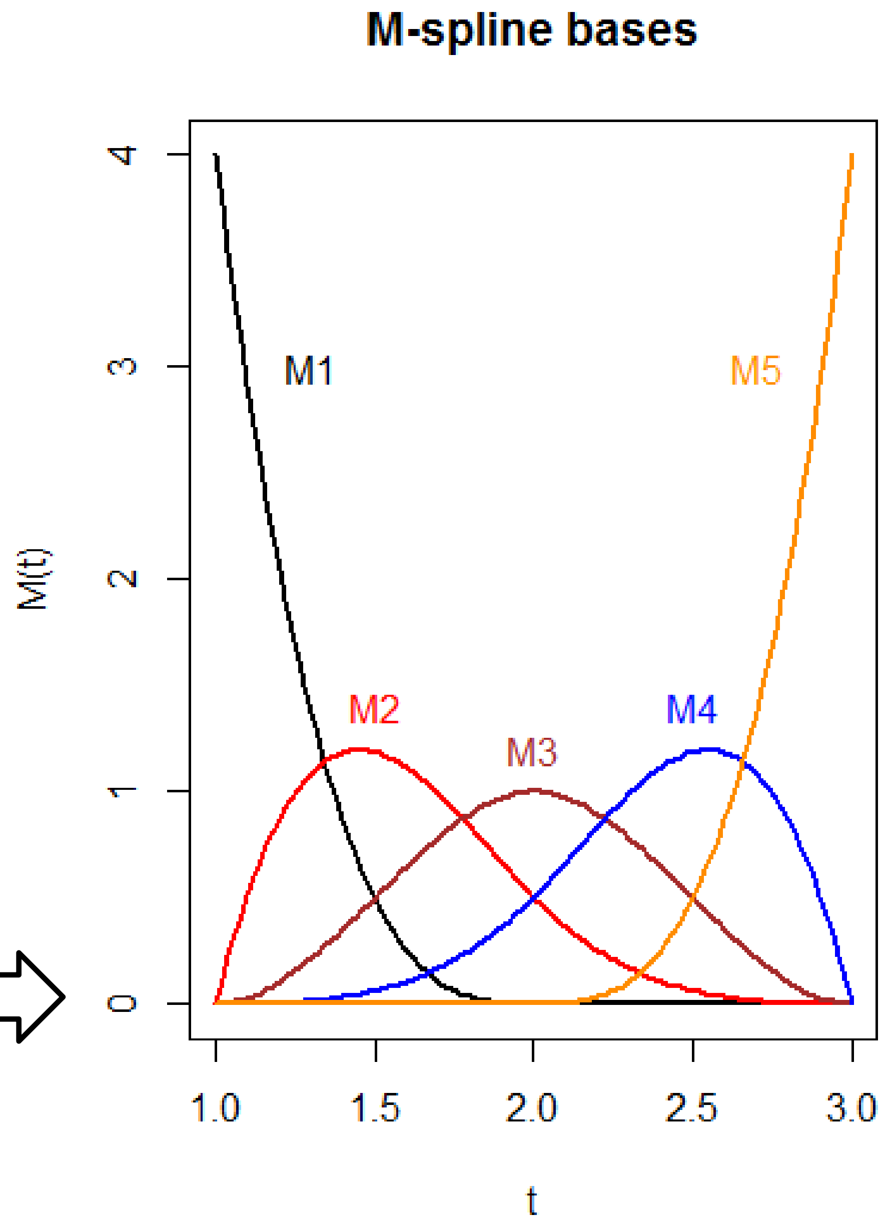
5-parameter M-spline (Ramsay 1988)

- Hazard function

$$\lambda_0(t;\mathbf{h}) = \sum_{\ell=1}^{5} h_\ell M_\ell(t)$$

- M-spline basis functions

$$\mathbf{M}(t) = (\, M_1(t),\ \ldots,\ M_5(t)\,)'$$



Define *M-spline basis* functions

$$M_1(t) = -\frac{4\mathbf{I}(\xi_1 \le t < \xi_2)}{\Delta} z_2(t)^3, \qquad M_5(t) = \frac{4\mathbf{I}(\xi_2 \le t < \xi_3)}{\Delta} z_2(t)^3,$$

$$M_2(t) = \frac{\mathbf{I}(\xi_1 \le t < \xi_2)}{2\Delta}\{\, 7z_1(t)^3 - 18z_1(t)^2 + 12z_1(t)\,\} - \frac{\mathbf{I}(\xi_2 \le t < \xi_3)}{2\Delta} z_3(t)^3,$$

$$M_3(t) = \frac{\mathbf{I}(\xi_1 \le t < \xi_2)}{\Delta}\{ -2z_1(t)^3 + 3z_1(t)^2 \,\} + \frac{\mathbf{I}(\xi_2 \le t < \xi_3)}{\Delta}\{\, 2z_2(t)^3 - 3z_2(t)^2 + 1\,\},$$

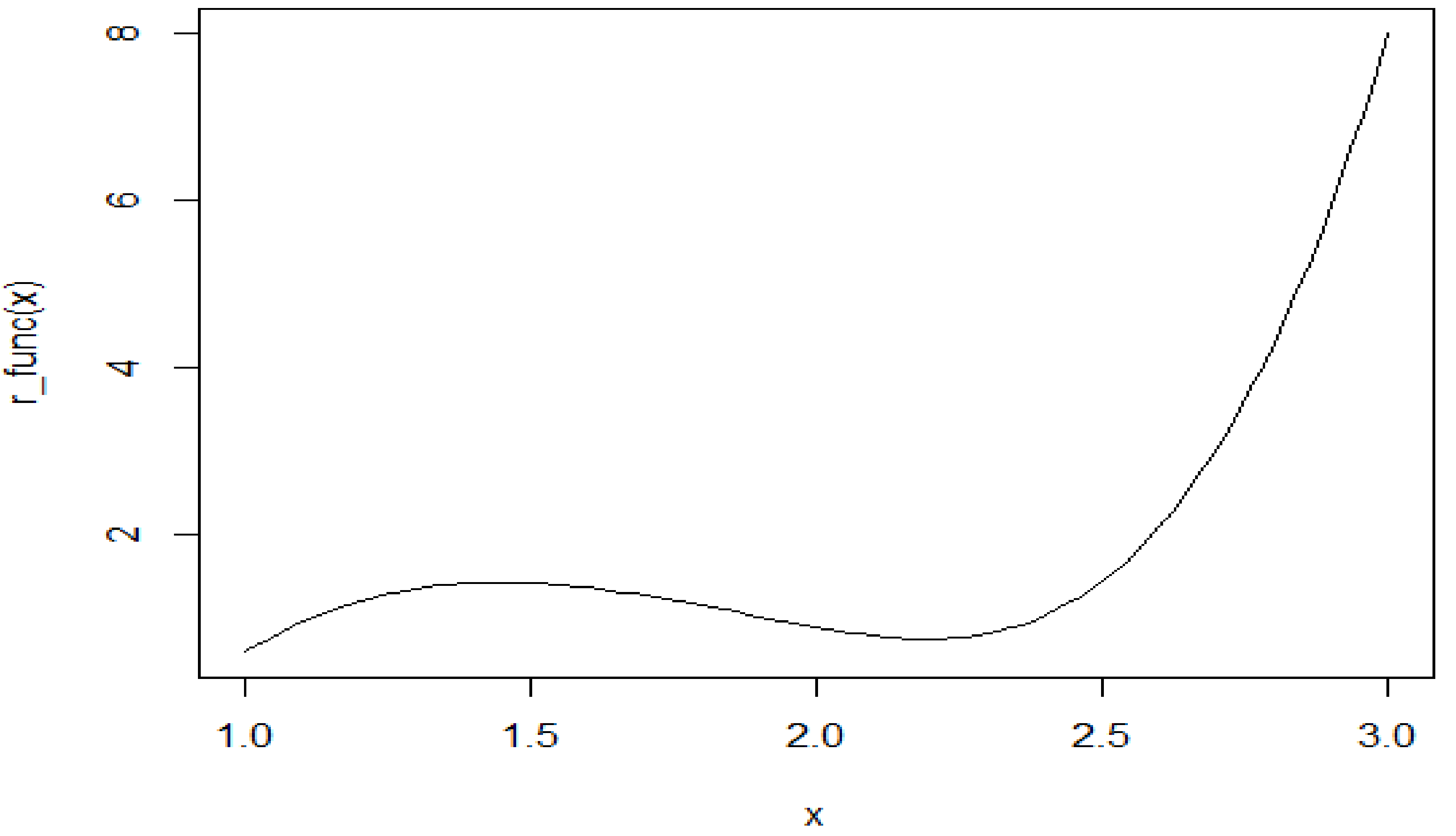
$$M_4(t) = \frac{\mathbf{I}(\xi_1 \le t < \xi_2)}{2\Delta} z_1(t)^3 + \frac{\mathbf{I}(\xi_2 \le t < \xi_3)}{2\Delta}\{ -7z_2(t)^3 + 3z_2(t)^2 + 3z_2(t) + 1\,\}.$$

↑ Appendix A of

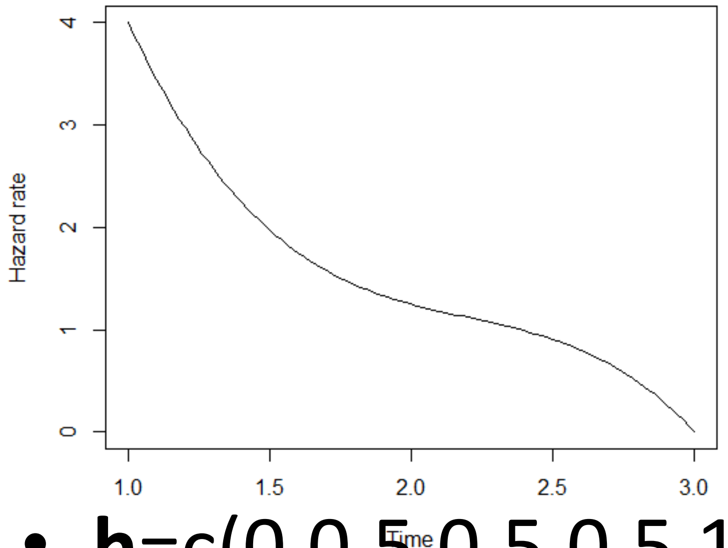
Emura T, Chen YH (2018), *Analysis of Survival Data with Dependent Censoring, Copula-Based Approaches* , JSS Research Series in Statistics, Springer,

Hazard function via cubic M-spline:

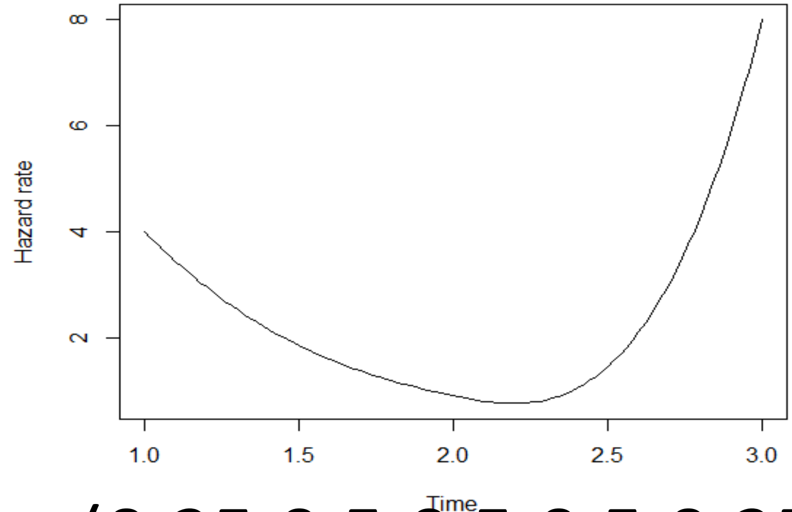
$$\lambda_0(t) = 0.15 \times M_1(t) + 1 \times M_2(t) + 0.3 \times M_3(t) + 0.2 \times M_4(t) + 2 \times M_5(t)$$



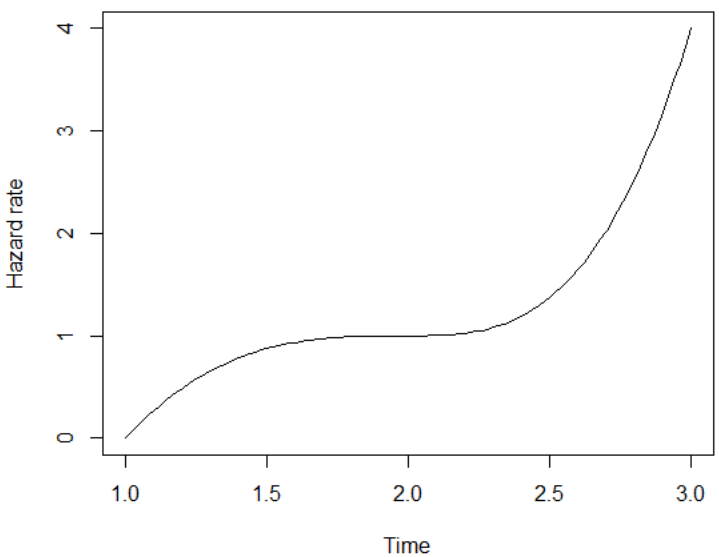
- **h**=c(1,1,0.5,0.5,0)



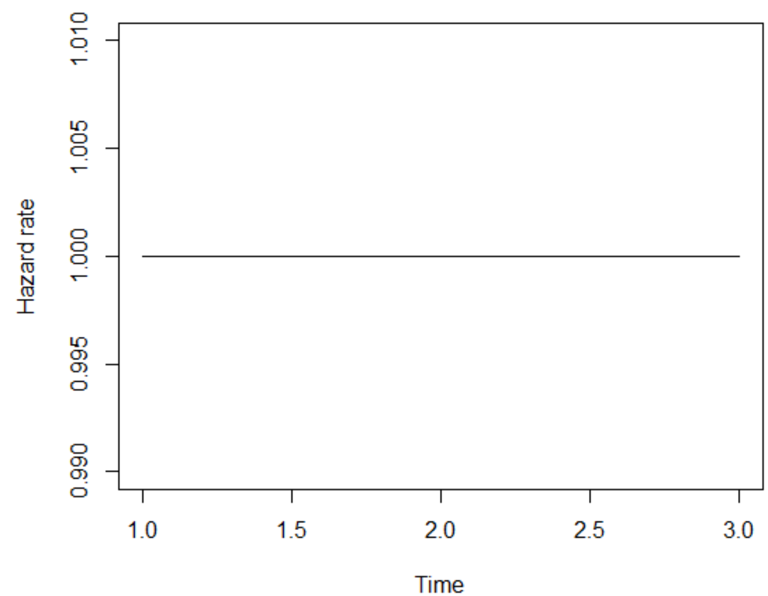
- **h**=c(1,1,0.3,0.2,2)



- **h**=c(0,0.5,0.5,0.5,1)

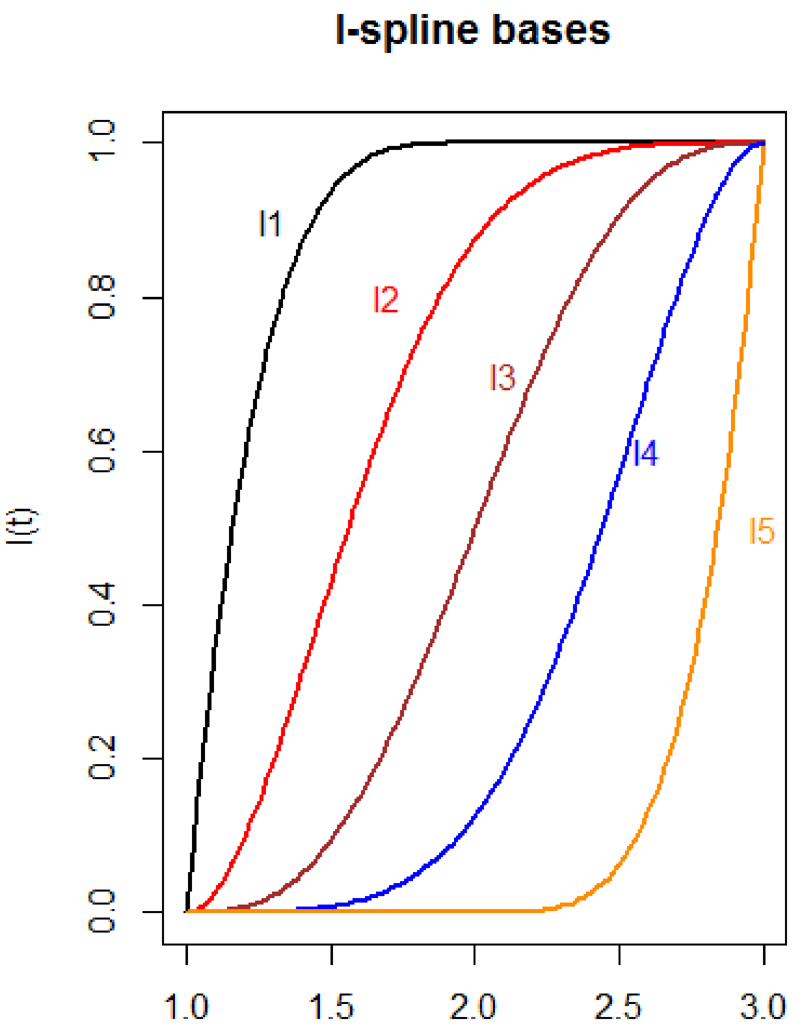


- **h**=c(0.25,0.5,0.5,0.5,0.25)



- Cumulative hazard: $\Lambda(t) = \sum_{\ell=1}^{5} h_\ell I_\ell(t)$

$$I_\ell(t) = \int_0^t M_\ell(u)du$$

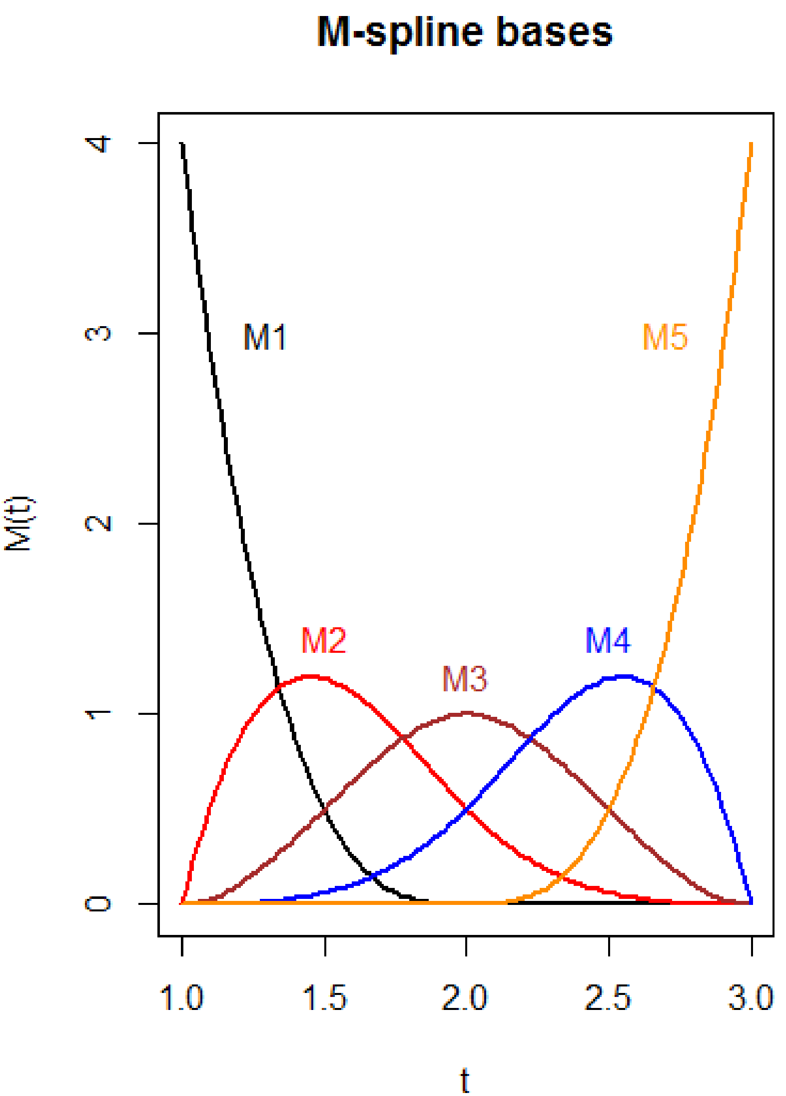


Explicit forms available (4-th order polynomial)

Computation

```
> library(joint.Cox)
> time=seq(1,3,length=10)
> M.spline(time,xi1=1,xi3=3)
```

	M1	M2	M3	M4	M5
[1,]	4.000	0.000	0.000	0.000	0.000
[2,]	1.882	0.927	0.123	0.005	0.000
[3,]	0.685	1.196	0.416	0.043	0.000
[4,]	0.148	1.037	0.747	0.148	0.000
[5,]	0.005	0.680	0.964	0.351	0.000
[6,]	0.000	0.351	0.964	0.680	0.005
[7,]	0.000	0.148	0.747	1.037	0.148
[8,]	0.000	0.043	0.416	1.196	0.685
[9,]	0.000	0.005	0.123	0.927	1.882
[10,]	0.00	0.000	0.000	0.000	4.00



Notations

- X_i : event time

- C_i : censoring time

$$\lambda(t \mid \mathbf{Z}_i) = -\frac{d \log S(t \mid \mathbf{Z}_i)}{dt}$$

$$S(t \mid \mathbf{Z}_i) = \Pr(X_i > t \mid \mathbf{Z}_i)$$

- $\mathbf{Z}_i = (Z_{i1}, ..., Z_{ip})'$: Covariates

  - $T_i = \min\{ X_i, C_i \}$
  - $\delta_i = \mathbf{I}(X_i \le C_i)$

Survival data consist of $\{ (T_i, \delta_i, \mathbf{Z}_i), \ i = 1, ..., n \}$

Cox model with splines

$$\lambda(t \mid \mathbf{Z}_i) = \lambda_0(t;\mathbf{h}) \exp(\boldsymbol{\beta}'\mathbf{Z}_i),$$

$$\lambda_0(t;\mathbf{h}) = \sum_{\ell=1}^{L} h_\ell M_\ell(t$$

$$\Lambda_0(t;\mathbf{h}) = \sum_{\ell=1}^{L} h_\ell I_\ell(t),$$

$$S_0(t;\mathbf{h}) = \exp\left[-\sum_{\ell=1}^{L} h_\ell I_\ell(t)\right],$$

$$I_\ell(t) = \int_0^t M_\ell(u)du$$

Log-likelihood

$$\ell(\boldsymbol{\varphi}) = \sum_{i=1}^{n} [\ \delta_i \{ \log \lambda_0(T_i; \mathbf{h}) + \boldsymbol{\beta}'\mathbf{Z}_i\ \} - \Lambda_0(T_i; \mathbf{h}) \exp(\boldsymbol{\beta}'\mathbf{Z}_i)\ ]$$

where $\boldsymbol{\varphi} = (\mathbf{h}, \boldsymbol{\beta})$

Penalized log-likelihood (O'Sullivan 1998)

$$\ell_{PL}(\boldsymbol{\varphi}) = \ell(\boldsymbol{\varphi}) - \kappa \int \ddot{\lambda}_0(t; \mathbf{h})^2 dt$$

Roughness

where $\ddot{f}(t) = d^2 f(t) / dt^2$.

Smoothing parameter;df=2(κ=Inf) ~ df=5(κ=0)

Simplification of the roughness

$$\int \ddot{\lambda}_0(t;\mathbf{h})^2 dt = \mathbf{h}'\Omega\mathbf{h}$$

$$\Omega = \begin{bmatrix} 192 & -132 & 24 & 12 & 0 \\ -132 & 96 & -24 & -12 & 12 \\ 24 & -24 & 24 & -24 & 24 \\ 12 & -12 & -24 & 96 & -132 \\ 0 & 12 & 24 & -132 & 192 \end{bmatrix}$$

How to choose $\hat{\kappa}$

➔ See the paper

Penalized ML estimator:

$$\hat{\boldsymbol{\varphi}} = (\hat{\boldsymbol{\beta}}^{\mathrm{PL}}, \hat{\mathbf{h}}^{\mathrm{PL}}) = \arg\max\left\{\ell(\boldsymbol{\beta},\mathbf{h}) - \hat{\kappa}\mathbf{h}'\Omega\mathbf{h}\right\}$$

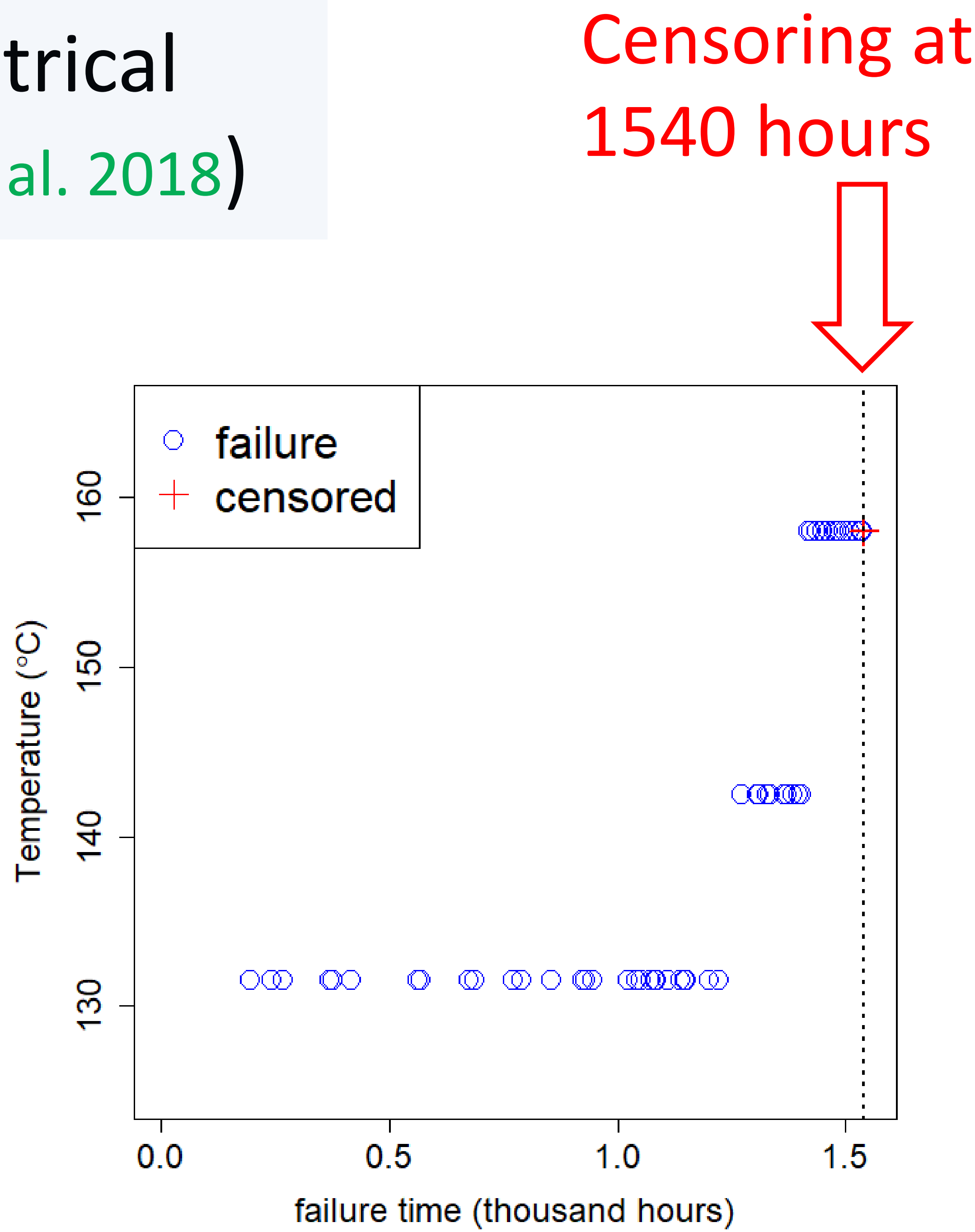
Life test data on electrical components (Zhou et al. 2018)

- *n*=64 samples
- Temperatures:

• **Z=0** for 158 ° C

• **Z=1** for 142.5 ° C

• **Z=2** for 131.5 ° C



Fitted results (regression coefficients)

- Cox regression (*coxph*) does not converge

```
> coxph(Surv(t.event,event)~Z)
Call: coxph(formula = Surv(t.event, event) ~ Z)
        coef        exp(coef)    se(coef)      z        p
Z    2.134e+01   1.859e+09   4.207e+03   0.005   0.996

Likelihood ratio test=120.1  on 1 df, p=< 2.2e-16
n= 64, number of events= 55
Warning message:
In fitter(X, Y, strats, offset, init, control, weights = weights,  :
  Loglik converged before variable  1 ; beta may be infinite.
```

Fitted results (regression coefficients)

- Cox regression

  $\hat{\boldsymbol{\beta}} = 21.34$ (Z-value=0.005): Un-converged value

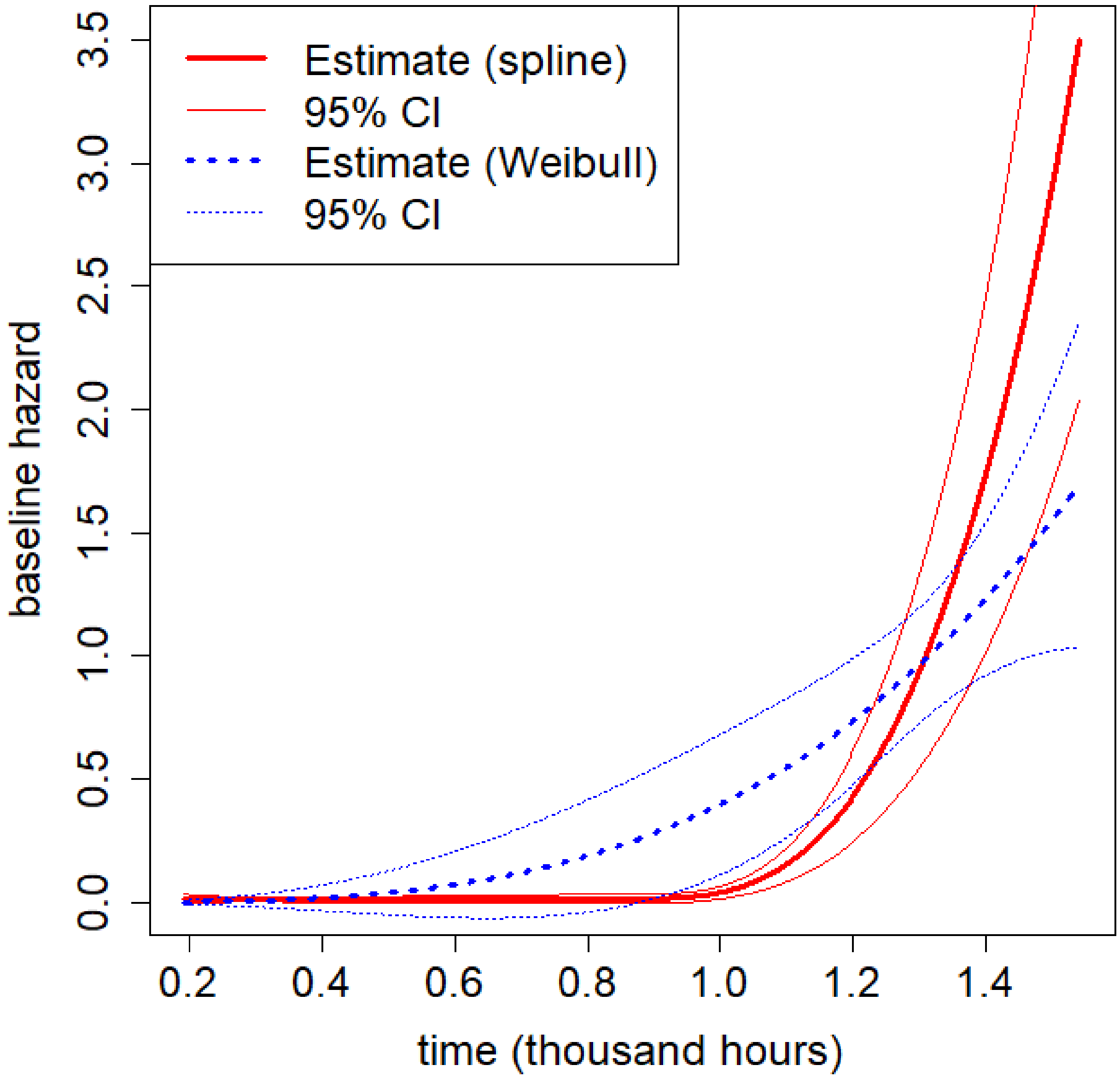
- The Weibull regression

  $\hat{\boldsymbol{\beta}} = 1.268$ (Z-value=6.651).

- The penalized likelihood with spline

  $\hat{\boldsymbol{\beta}}^{PL} = 2.092$ (Z-value=13.4) under $\kappa = 0$.

Fitted results (Baseline hazard)



Conclusion/Discussion

- **Cox regression does not converge** for small data
  - Partial likelihood may not have a maxima
  - Use "rank" to estimate **ß**
- **Cox regression cannot estimate baseline hazard** (only cumulative hazard by Breslow estimator)
- Spline models converges, as in parametric models
- Estimates **ß** are very close if both Cox regression and spline converges (results not shown)

Proposed methods available through an R package *joint.Cox*