Recent Advances in the Statistical Analysis of Count and Survival Data, September 25-26, 2007

Auxiliary Mixture Sampling for Dynamic Survival Models

Helga Wagner
Johannes Kepler University Linz, Austria

Outline

- Motivating example
- Normal Dynamic Survival Model
- Data Augmentation for the exponential regression model
- Auxiliary mixture sampling for the dynamic survival model
- Applications
- Extensions

Example: Gastric Data

90 patients, randomly allocated to a therapy: chemotherapy (trt=0) or combination of chemotherapy and radiation (trt=1).
10 observation times are right censored.

Data: $(y_i, \delta_i, \mathbf{x}_i), i = 1, \ldots, n$

y_i	observation time
	independent censoring $y_i = \min(t_i, c_i)$
δ_i	failure indicator: $\delta_i = \begin{cases} 1 & \text{if } t_i < c_i \\ 0 & \text{otherwise} \end{cases}$
$\mathbf{x}_i$	covariate vector

The Cox Model

$$\lambda(t\,|\,\mathbf{x}) = \lambda(t) \cdot \exp(\mathbf{x}'\boldsymbol{\beta}).$$

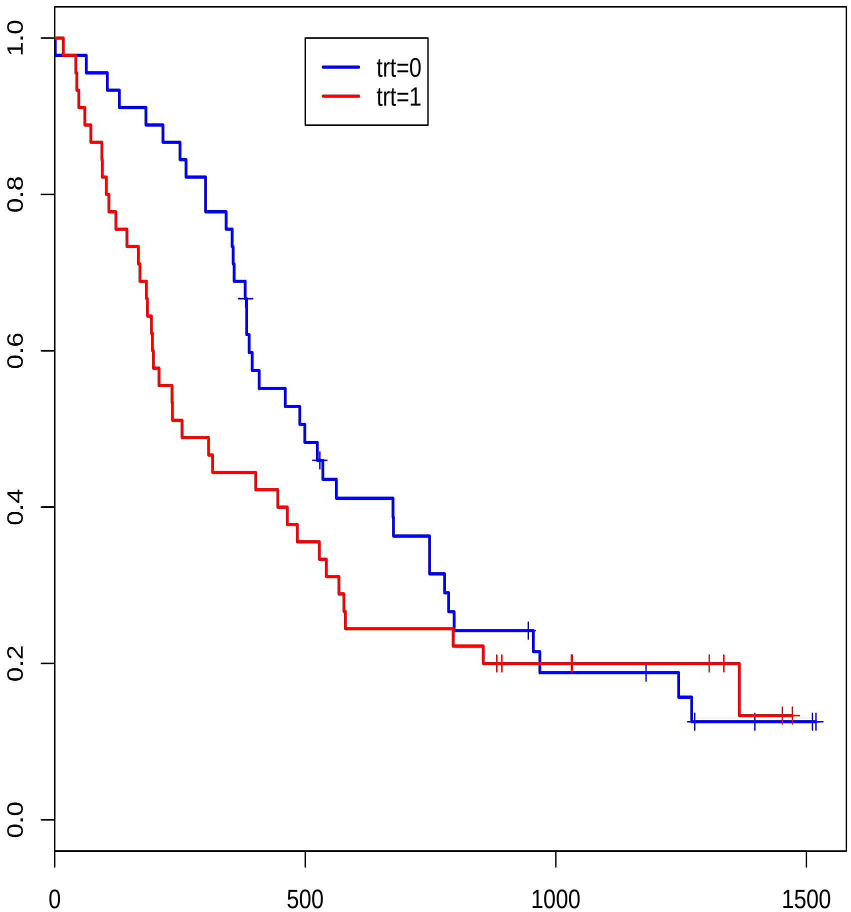
The hazard ratio for two individuals with covariate values $\mathbf{x}$ and $\mathbf{x}^*$ depends on the difference between their linear predictors

$$\frac{\lambda(t\,|\,\mathbf{x})}{\lambda(t\,|\,\mathbf{x}^*)} = \exp\left((\mathbf{x} - \mathbf{x}^*)'\boldsymbol{\beta}\right)$$

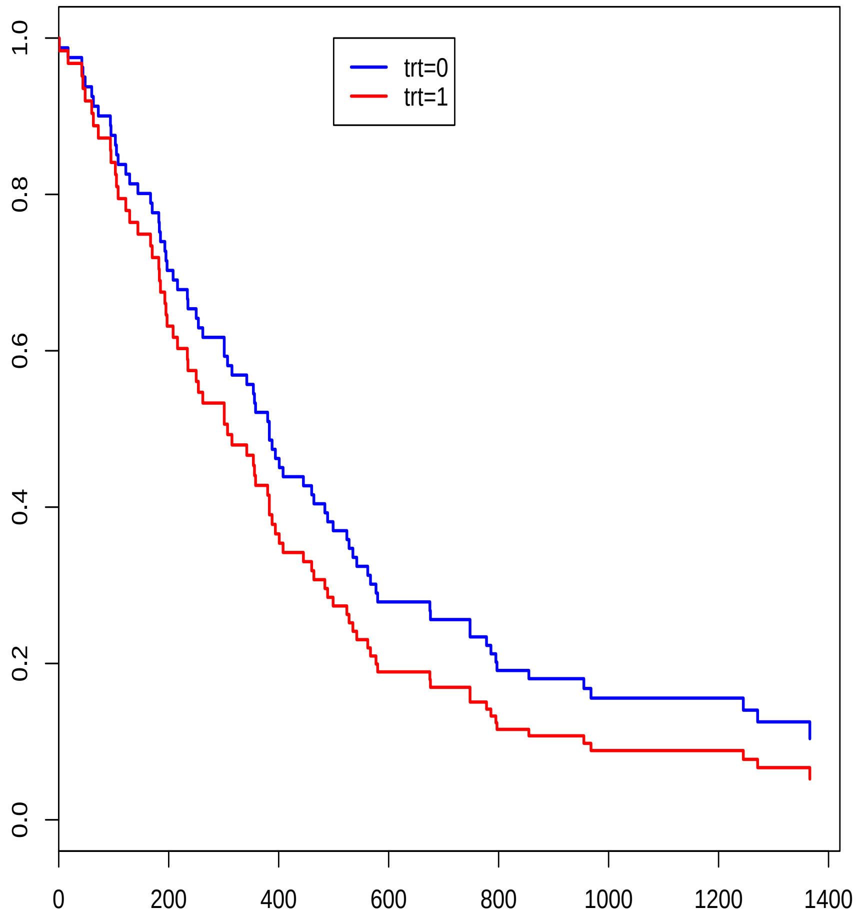
and is constant over time

Survival functions $S(t|\mathbf{x})$ and $S(t|\mathbf{x}^*)$ do not cross !

Example: Gastric Data



Kaplan-Meier Estimates



Proportional Hazards model

Dynamic Survival Model

Allow time-varying effects (Gamerman, 1991; Kneib and Fahrmeir, 2007; Hennerfeind et al., 2006)

$$\lambda(t|\mathbf{x}_i) = \exp\left(\beta_0(t) + \sum_{k=1}^{K} x_{ik}\beta_k(t)\right)$$

Dynamic survival model Gamerman (1991):

- piecewise exponential model for lifetimes
- correlated prior processes for the baseline log-hazard and the covariate effects.

Normal Dynamic Survival Model

Partition of the time axis: J intervals $(0, s_1], \ldots (s_{J-1}, s_J]$

- baseline log-hazard $\beta_0(t)$ and covariate effects $\beta_k(t)$ are constant in each interval

$$\beta_k(t) = \beta_{kj}, \qquad \text{for } t \in I_j = (s_{j-1}, s_j]$$

- each $\beta_k(t)$ follows a Gaussian random walk Hemming and Shaw (2002, 2005)

$$\beta_{kj} = \beta_{k,j-1} + w_{kj} \qquad w_{kj} \sim N\left(0, \theta_k\right).$$

Exponential regression model

Complete survival times: $t_1, \ldots, t_n$

$$t_i \sim \mathrm{Ex}\left(\lambda_i\right) \qquad \lambda_i = \exp(\mathbf{x}_i' \boldsymbol{\alpha})$$

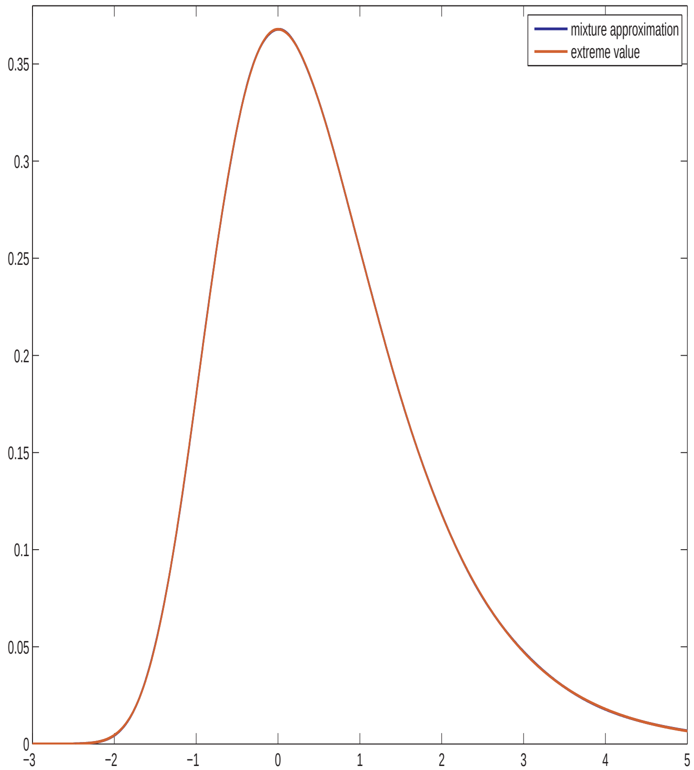
$\Longrightarrow$ linear model for transformed survival times:

$$-\ln(t_i|\mathbf{x}_i) = \mathbf{x}_i' \boldsymbol{\alpha} + \varepsilon_i$$

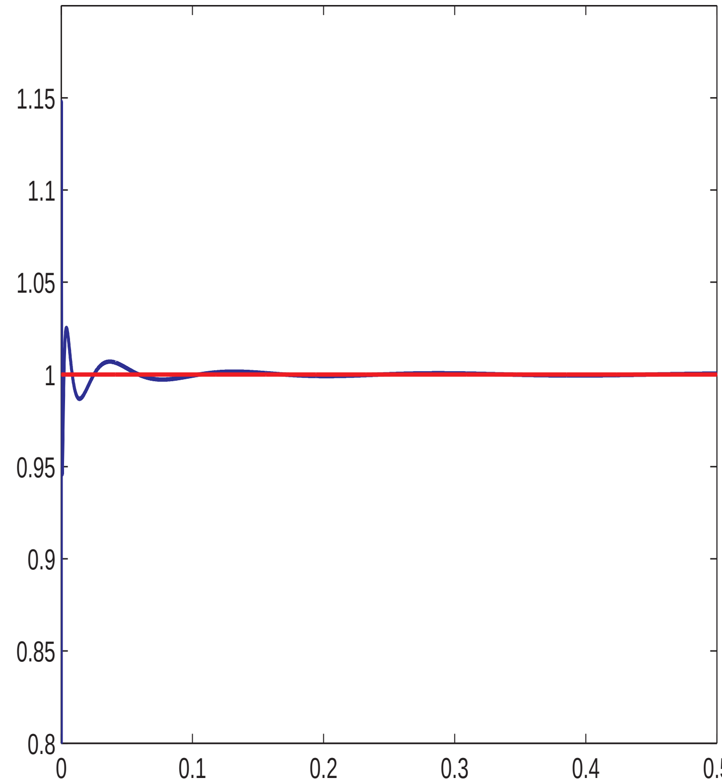
ε_i has a type I extreme value distribution

$\Longrightarrow$ mixture approximation can be used to estimate $\boldsymbol{\alpha}$

Mixture Approximation



Density of ε (exact and approx.)



Hazard of $\exp(\varepsilon)$ (exact and approx.)

Data Augmentation for Censored data

Censored survival times: $\mathbf{y} = (t_1, \delta_1), \dots, (t_n, \delta_n)$

Generate complete auxiliary survival times τ_i: Conditionally on $\{t_i > c_i\}$ the (unobserved) residual life time $\xi_i = t_i - c_i \sim \mathrm{Ex}\,(\lambda_i)$

Auxiliary survival times τ_i, i=1,. . . , n

$$\tau_i = \begin{cases} y_i & \text{if} \quad \delta_i = 1 \\ y_i + \xi_i & \text{if} \quad \delta_i = 0 \end{cases}$$

Data Augmentation for Censored data

Introduction of component indicators r_i (for the mixture approximation of $p(\varepsilon)$) leads to a linear normal regression model with heteroscedastic errors

$$\begin{aligned} -\ln \tau_i | \boldsymbol{\alpha}, r_i = & \quad \mathbf{x}_i' \boldsymbol{\alpha} + m_{r_i} + \varepsilon_i \\ \varepsilon_i \sim & \qquad N\left(0, s_{r_i}^2\right) \end{aligned}$$

Auxiliary variables: $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)$ where $\mathbf{z}_i = (\tau_i, r_i)$

Auxiliary Mixture Sampling

1. For each $i = 1, \ldots, n$ generate $\mathbf{z}_i = (\tau_i, r_i)$ from $p(r_i|\tau_i, \boldsymbol{\alpha}, \mathbf{y})p(\tau_i|\boldsymbol{\alpha}, \mathbf{y})$

2. Sample $\boldsymbol{\alpha}$ from $p(\alpha|\mathbf{y}, \mathbf{z}) \sim N\left(a_n, A_n\right)$

Implementation of Step 1:

- sample $\xi_i \sim \mathrm{Ex}\left(\lambda_i\right)$ if $\delta_i = 0$

- sample the component indicator from a discrete distribution with 10 categories

The normal dynamic survival model

In the normal dynamic survival model

- the latent process $\boldsymbol{\beta}_j, j = 1, \ldots, J$ (baseline log-hazard and covariate effects) is Gaussian

- observations follow the PE-Model and are subject to censoring

Aims of Data Augmentation

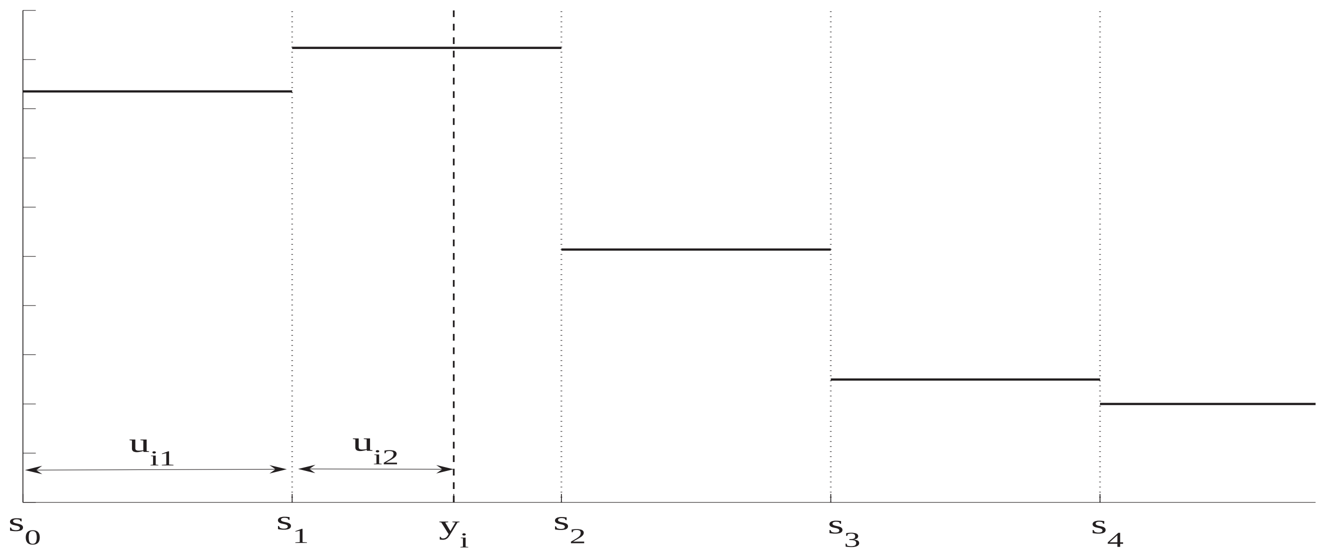
- Step I: Achieve a linear state space model

- Step II: Eliminate non-normality

Data Augmentation Step I

Split observation times $y_i \in I_j$ into episodes $u_{il}, l = 1, \ldots, j$, of constant hazard

$$\lambda_{il} = \lambda(t|\mathbf{x}_i, t \in I_l) = \exp\left((\mathbf{x}_i^f)'\boldsymbol{\alpha} + (\mathbf{x}_i^v)'\boldsymbol{\beta}_l\right)$$



Data Augmentation Step I

- For $y_i \in I_j$

$$u_{il} = \begin{cases} \Delta_l = s_l - s_{l-1} & l = 1, \ldots, j_i - 1 \\ y_i - s_{l-1} & l = j_i \end{cases}$$

- First $j_i - 1$ episodes u_{il} are censored by the interval length

$$u_{il} = \min(\tau_{il}, \Delta_l) \quad l = 1, \ldots, j_i - 1$$

$\Longrightarrow$ generate residual lifetimes $\xi_{il} \sim \mathrm{Ex}\left(\lambda_{il}\right)$

Data Augmentation Step I

- The auxiliary survival times

$$\tau_{il} = \begin{cases} u_{il} + \xi_{il} & l = 1, \ldots, j_i - 1 \\ u_{il} + (1 - \delta_i)\xi_{il} & l = j_i \end{cases}$$

follow the dynamic generalized linear model West et al. (1985)

$$\tau_{il} | \boldsymbol{\alpha}, \boldsymbol{\beta}_l \sim \mathrm{Ex}\left(\exp\left((\mathbf{x}_i^f)'\boldsymbol{\alpha} + (\mathbf{x}_i^v)'\boldsymbol{\beta}_l\right)\right),$$

$$\boldsymbol{\beta}_l = \boldsymbol{\beta}_{l-1} + \boldsymbol{\omega}_l, \quad \boldsymbol{\omega}_l \sim N\left(0, \mathrm{diag}(\boldsymbol{\theta})\right)$$

with $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_K)'$

Data Augmentation Step II

Introduction of a component number r_{il} for each $\tau_{il}, i = 1, \ldots, n; l = 1, \ldots j_i$ leads to a gaussian state space model with heteroscedastic errors

$$
\begin{aligned}
\ln \tau_{il} &= (\mathbf{x}_i^f)'\boldsymbol{\alpha} + (\mathbf{x}_i^v)'\boldsymbol{\beta}_l + m_{r_{il}} + \varepsilon_{il}, & \varepsilon_{il} \sim N\left(0, s^2_{r_{il}}\right) \\
\boldsymbol{\beta}_l &= \boldsymbol{\beta}_{l-1} + \boldsymbol{\omega}_l, & \boldsymbol{\omega}_l \sim N\left(0, \mathrm{diag}(\boldsymbol{\theta})\right)
\end{aligned}
$$

Auxiliary variables:

$\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)$ where $\mathbf{z}_i = (\tau_{il}, r_{il}, l = 1, \ldots, j_i)$

The Sampling Scheme

- Multi-move sampling for the whole sequence $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ by forward-filtering backward sampling from $p(\boldsymbol{\alpha}, \boldsymbol{\beta}|\boldsymbol{\theta}, \boldsymbol{\tau}, \mathcal{R})$ as in Frühwirth-Schnatter (1994), Carter and Kohn (1994), de Jong and Shephard (1995), or Durbin and Koopman (2002): recursive sampling from multivariate normals

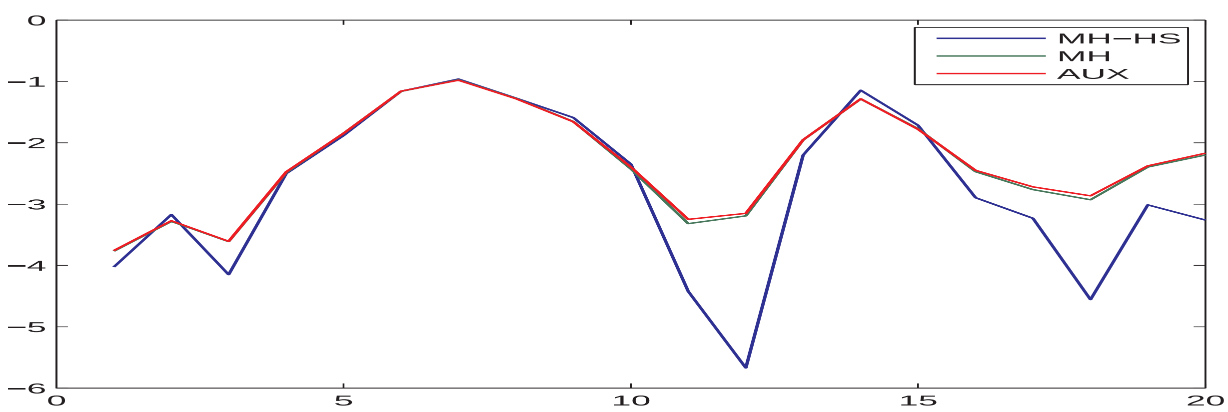
- Sample $\boldsymbol{\theta}$ from $p(\boldsymbol{\theta}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\tau}, \mathcal{R})$ (inverted Gamma distibutions)

- Sample $\boldsymbol{\tau}, \mathcal{R}$ from $p(\boldsymbol{\tau}, \mathcal{R}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, (\mathbf{y}, \boldsymbol{\delta}))$ (exponential resp. discrete distribution with 10 categories)

Applications: Simulated data

- $n = 200$ survival times y_i, no covariates, $\theta_0 = 0.3$
- 21 intervals: $(0, 1], \ldots (19, 20], (20, \max(y_i)]$
- Priors: $\beta_0 \sim N\,(0, 100)$

$$\theta_0 \sim \mathrm{IG}\,(c_0, C_0)$$



Application: Gastric Data

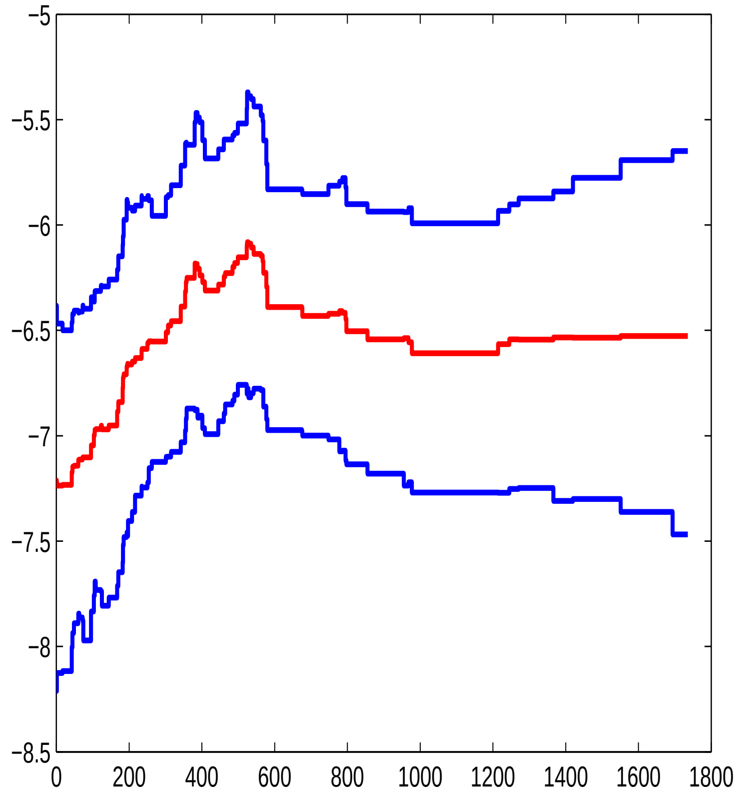
- 2 dimensional state vector
- division points of the time axis: observed failure times (80 time intervals)
  $\Longrightarrow$ monotone likelihood
- Priors:

$$\boldsymbol{\beta} \sim N\left(\mathbf{0}, 100\mathbf{I}\right)$$

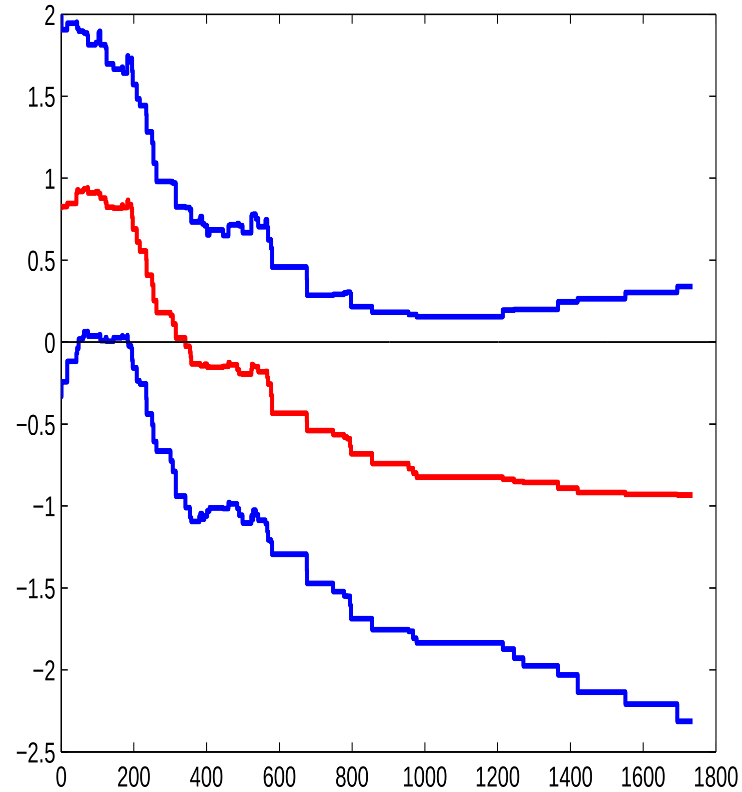
$$\theta_0 \sim \mathrm{IG}\left(0.01, 0.01\right)$$

- 3920 auxiliary survival times

Application: Gastric Data

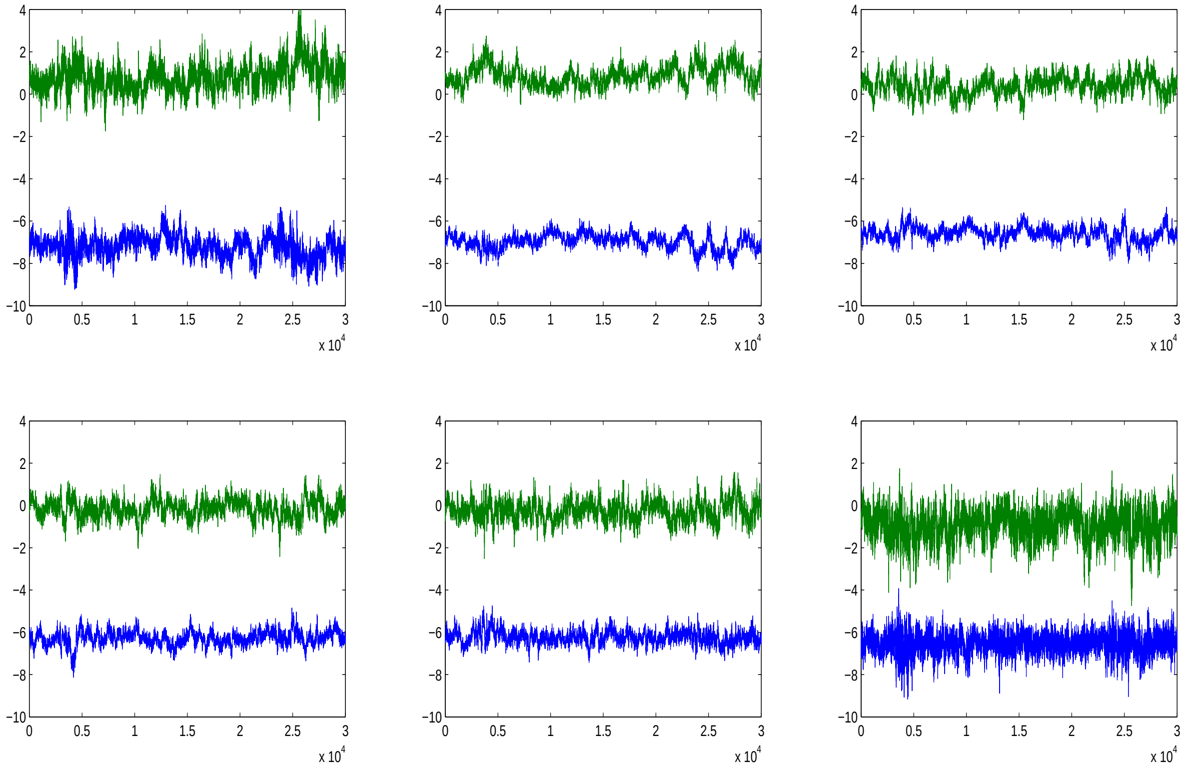


Baseline Log-Hazard



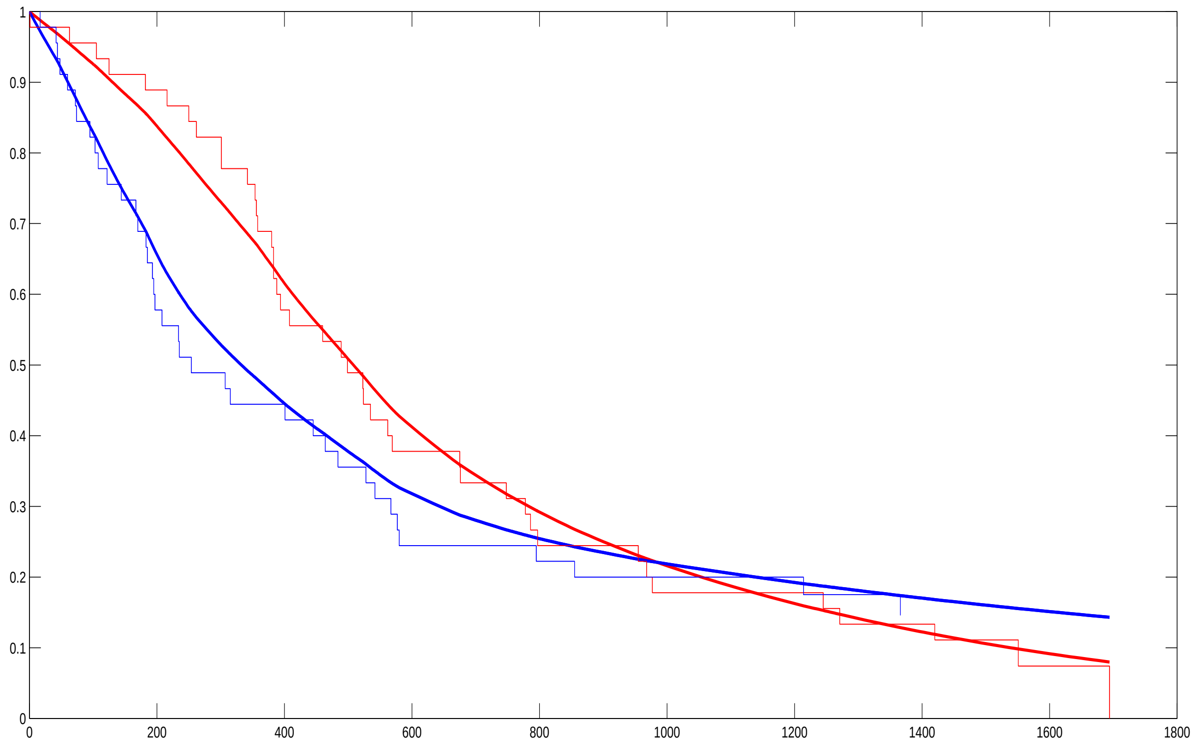
Treatment effect

Application: Gastric Data



MCMC plots for baseline log hazard β_{0j} (blue) and treatment effect β_{1j} (green) for $j = 1, 15, 30, 45, 60, 80$

Application: Gastric Data



Estimated Survival Functions in the Normal Dynamic Survival Model

Extensions

- Extensions of the linear predictor:
  - spatial effects
  - frailties
  - spline terms for nonlinear effects
  - time-varying covariates (step functions)

- Extensions to other types missing data:
  - left- and interval censoring
  - combinations of left-, right and interval censoring

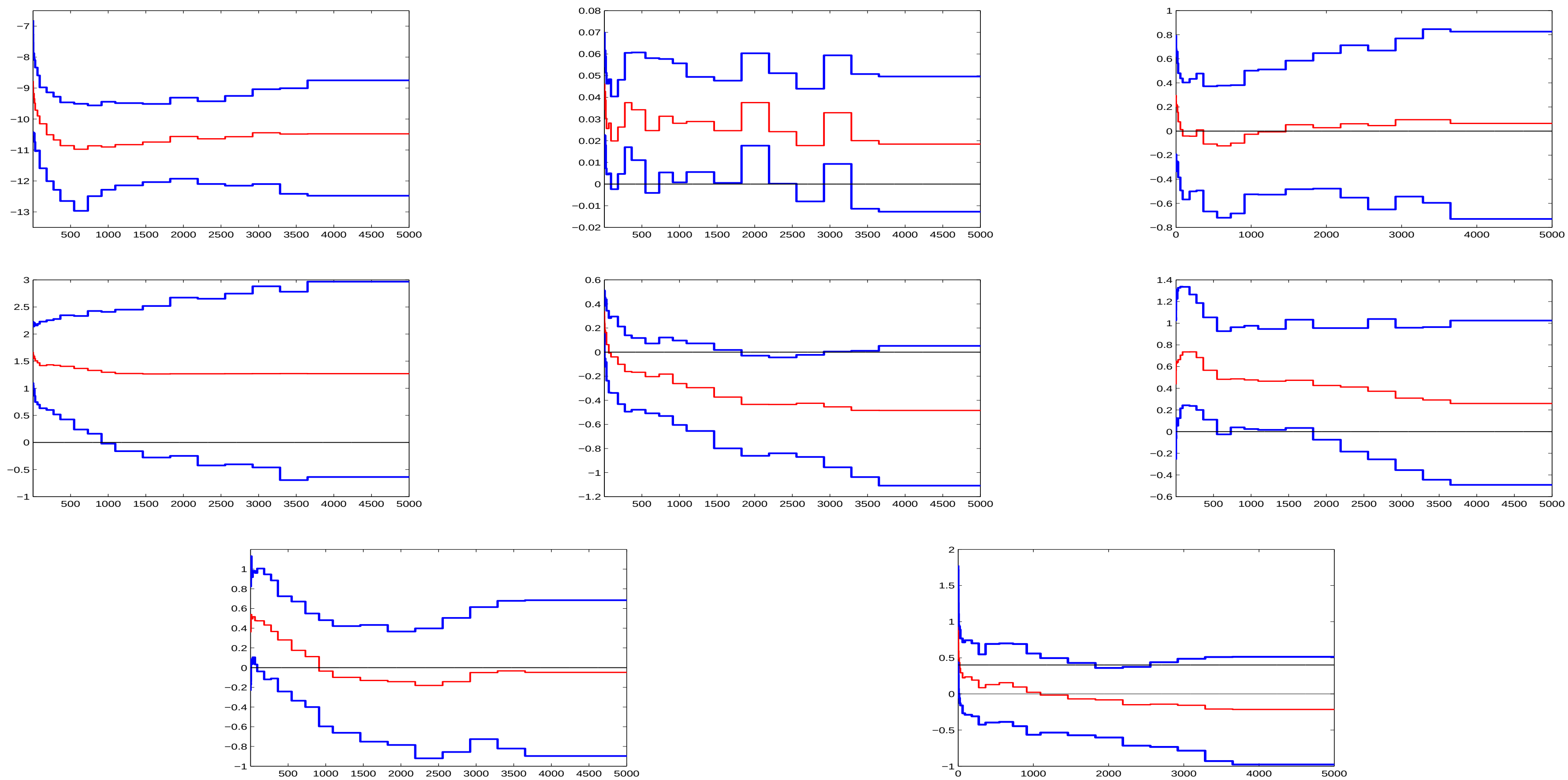
Worcester Heart Attack Study

481 patients with acute myocardial infarction

7 Covariates:

`age`	age at hospital admission in years
`sex`	0=male, 1=female
`sho`	cardiogenic shock complications (0=no,1=yes)
`cpk`	peak cardiac enzyme measured in international units (IU)
`chf`	left heart failure complications (0=no, 1=yes)
`miorder`	myocardial infection order (0=first, 1=recurrent)
`mitype`	myocardial infection type (0=Q wave)

Worcester Heart Attack Study



WHAS data: Posterior means and 95% credible regions for the baseline log-hazard and the effects of age, `sex` , `sho` , `cpk` , `chf` , `miorder` and `mitype`.

Variable selection for dynamic survival models

Variable selection for the dynamic survival model is based on the conditional normal state space model

$$\ln \tau_{il} = \beta_{0l} + \sum_{k=1}^{K} x_{ik}\beta_{kl} + m_{r_{il}} + \varepsilon_{il}, \quad \varepsilon_{il} \sim N\left(0, s^2_{r_{il}}\right)$$

$$\beta_{kl} = \beta_{k,l-1} + \omega_{kl}, \quad \omega_{kl} \sim N\left(0, \theta_k\right), \; k = 0, \ldots, K$$

Variable selection for dynamic survival models

Introduce indicators:

$$\eta_k = 0 \;\Leftrightarrow\; \beta_{k0} = 0, \qquad \text{for } k = 1, \ldots, K$$

$$\gamma_k = 0 \;\Leftrightarrow\; \theta_k = 0, \qquad \text{for } k = 0, \ldots, K$$

- $\gamma_0 = 1$ time varying baseline hazard
- $\gamma_0 = 0$ constant baseline hazard
- $\eta_k = 1, \gamma_k = 1$ time varying effect of regressor $x_k \;\; (k > 0)$
- $\eta_k = 1, \gamma_k = 0$ fixed effect of regressor $x_k \;\; (k > 0)$
- $\eta_k = 0, \gamma_k = 0$ no effect of regressor $x_k \;\; (k > 0)$

Variable selection for dynamic survival models

Select indicators and non-zero elements of $\boldsymbol{\beta}_0$ and $\boldsymbol{\theta}$ in the following representation of model

$$\ln \tau_{il} = \beta_{00} + \sum_{k=1}^{K} \eta_k x_{ik} \beta_{k0} + \sum_{k=0}^{K} \gamma_k (\pm\sqrt{\theta_k}) x_{ik} \tilde{\beta}_{kl} + m_{r_{il}} + \varepsilon_{il}$$

$$\tilde{\beta}_{kl} = \tilde{\beta}_{k,l-1} + \tilde{\omega}_{kl}, \qquad \tilde{\omega}_{kl} \sim N(0,1), \tilde{\beta}_{k0} = 0$$

where $\varepsilon_{il} \sim N\left(0, s^2_{r_{il}}\right)$

Summary

Dynamic survival models allow time-varying covariate effects for survival data.

Normal dynamic survival models may be represented as partially Gaussian state space models for auxiliary variables.

Gibbs type sampling with draws from well-known densities only (no need to use a Metropolis Hastings Algorithm) is feasible

Representations as a conditional normal model allows many extensions (inclusion of random effects, variable selection)

References

Carter, C. and Kohn, R. (1994). On Gibbs sampling for state space models. Biometrika, 81:541–553.

de Jong, P. and Shephard, N. (1995). The simulation smoother for time series models. Biometrika, 82:339–350.

Durbin, J. and Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. Biometrika, 89:603–15.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. Journal of Time Series Analysis, 15:183–202.

Gamerman, D. (1991). Dynamic Bayesian models for survival data. Applied Statistics, 40:63–79.

Hemming, K. and Shaw, J. E. H. (2002). A parametric survival

References

model applied to breast cancer survival times. Applied Statistics, 51:421–435.

Hemming, K. and Shaw, J. E. H. (2005). A class of parametric dynamic survival models. Lifetime Data Analysis, 11:81–98.

Hennerfeind, A., Brezger, A., and Fahrmeir, L. (2006). Geoadditive survival models. Journal of the American Statistical Association, 101:1065–1075.

Kneib, T. and Fahrmeir, L. (2007). A mixed model approach for geoadditive hazard regression. Scandinavian Journal of Statistics, 34:207–228.

West, M., Harrison, P. J., and Migon, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting (C/R: p84-

References

97). Journal of the American Statistical Association, 80:73–83.