Nonparametric Bayesian Intensity Model: Exploring Time-to-Event Data on Two Time Scales

Tommi Härkänen[1a] Anna But[b] Jari Haukka[b]

[a] Department of Public Health Solutions, National Institute for Health and Welfare, Finland
[b] Department of Public Health, University of Helsinki, Finland

EMS in Helsinki, July 24, 2017

NATIONAL INSTITUTE FOR HEALTH AND WELFARE
FINLAND

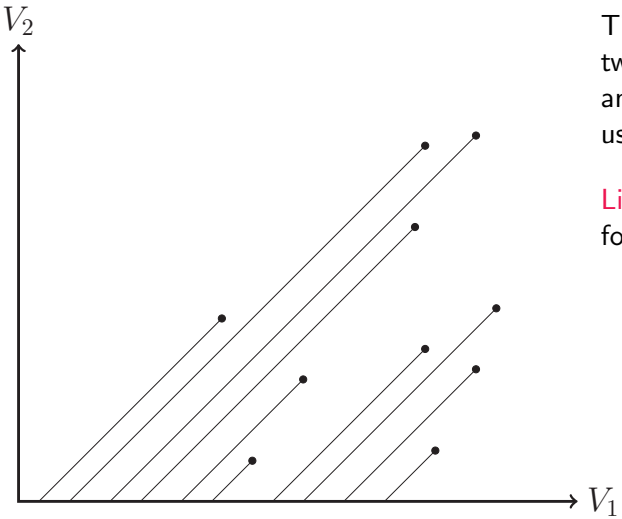
[1] E-mail: Tommi.Harkanen@thl.fi

Contents

1. Lexis diagram
2. Intensity models
3. Reversible jump MCMC
4. Simulation study
5. Empirical study on breast cancer

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Lexis diagram and multiple time scales



Time to event data with respect to two time scales V_1 and V_2 e.g. age and calendar time can be illustrated using Lexis diagram.

Life lines correspond to individual follow-up times.

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Intensity process for multiple time scales

For survival times T_i the intensity process is defined as

$$\hbar_i(t) = \lambda\big(V_{1i}(t), V_{2i}(t)\big)\,\mathbb{1}\{T_i > t\}$$

where

hazard surface $\lambda : \mathbb{R}^2 \to [0, \infty)$ and

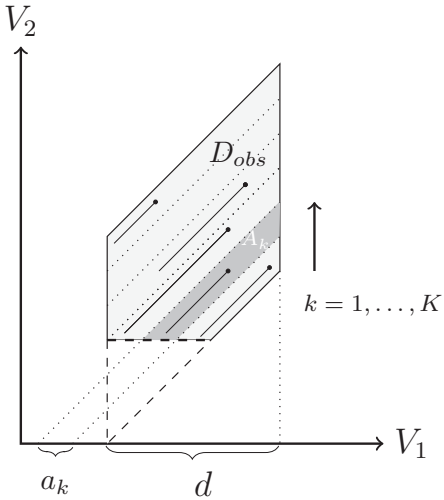
transformations of time t to time scales $V_{1i}(t) := t - a_i$ and $V_{2i}(t) := t - b_i$.

Approximation of hazard surface using 1-dimensional functions

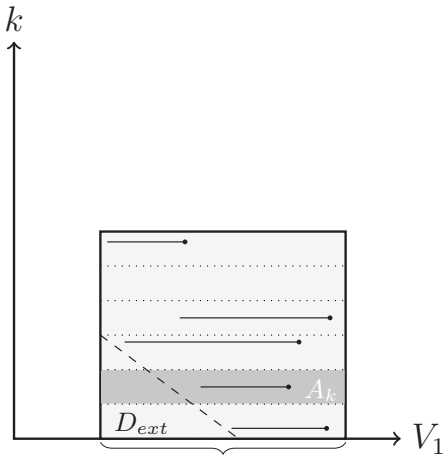
Divide Lexis diagram into diagonal strips A_k for $k = 1, 2, \ldots$.
Define the hazard surface using 1-dimensional hazard functions λ_k as

$$\lambda(v_1, v_2) = \lambda_k(v_1) \text{ for all } (v_1, v_2) \in A_k.$$

Transformation of the Lexis diagram



(a) Lexis diagram



(b) Isomorphic presentation of Lexis

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Piecewise constant (PC) hazard function

We approximate nonnegative hazard functions using PC functions

$$f(t) := \sum_{j \geq 0} h_j \, \mathbb{1}\{t \in (\varepsilon_j, \varepsilon_{j+1}]\}$$

where

- $(0 =:) \, \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_J < \varepsilon_{\text{Max}}$ are jump points and
- $h_j \geq 0$ corresponding levels.

Benefits:

- Calculation of Poisson likelihood for event times $t_{i,\ell}$ and follow-up intervals $(c_i, d_i]$ $\prod_i \exp\{-\int_{c_i}^{d_i} \lambda_i(s)\, ds\} \prod_\ell \lambda_i(t_{i,\ell})$ is simple.
- Addition and/or multiplication of PC functions result in a PC function.

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Prior distribution for one-dimensional PC hazard function

Arjas and Gasbarra (1994) in Statistica Sinica

We apply Bayesian inference, thus model parameters defining the PC functions are given prior distributions:

Jump points follow homogenous Poisson process $(\varepsilon_j) \sim \text{PoissonProcess}(\mu)$, $\mu > 0$, and

Levels are defined sequentially

$$\begin{aligned} h_0 &\sim \text{Gamma}(\alpha_0, \beta_0) \\ h_j &\sim \text{Gamma}(\alpha, \alpha/h_{j-1}),\ j > 0 \end{aligned}$$

Note that $\mathbb{E}[h_j \mid \alpha, h_{j-1}] = h_{j-1}$.

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Prior distribution for levels of two-dimensional PC hazard function

Härkänen, But and Haukka, to appear in Scandinavian Journal of Statistics

Start with the first strip A_1:
$$\begin{aligned} h_{1,0} &\sim \text{Gamma}(\alpha_0, \beta_0) \\ h_{1,j} &\sim \text{Gamma}(\alpha, \alpha/h_{1,j-1}),\ j > 0. \end{aligned}$$

Define average hazard level over interval $(a, b]$ as $\tau_k(a,b) := \int_a^b \lambda_k(t)\,dt/(b-a)$.

Define for A_k, $k > 1$ weighted average of previous level $h_{k,j-1}$ and $\tau_k(\varepsilon_{k,j}, \varepsilon_{k,j+1})$ with weight $\phi \geq 0$:

$$\bar{h}_{k,j} = \begin{cases} \left[h_{k,j-1} + \phi\,\tau_{k-1}(\varepsilon_{k,j}, \varepsilon_{k,j+1})\right]/(1+\phi) & k > 1,\ j > 0 \\ \tau_{k-1}(\varepsilon_{k,j}, \varepsilon_{k,j+1}) & k > 1,\ j = 0. \end{cases}$$

Define prior distributions of the levels $h_{k,j} \sim \text{Gamma}(\alpha, \alpha/\bar{h}_{k,j})$, $k > 1$.

- If $\phi = 0$, then hazard functions are *a priori* independent.
- If $\phi = 1$, then $h_{k,j-1}$ and $\tau_{k-1}(\varepsilon_{k,j}, \varepsilon_{k,j+1})$ have equal weight.

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Markov Chain Monte Carlo (MCMC)

Posterior distribution of the parameters is not possible to obtain analytically, thus we used numerical integration.
The posterior distribution of the parameters are approximated using Markov chain Monte Carlo (MCMC) methods:

- Jump points $\varepsilon_{k,j}$ and levels $h_{k,j}$ of the PC functions are updated one by one using the Metropolis-Hastings algorithm.
- The number of jump points is not fixed, thus at every iteration of the MCMC simulation the Reversible jump MCMC algorithm tries (with equal probability 0.5) to
  - add a new jump point (and corresponding level) to a PC function or
  - remove an existing jump point (and corresponding level) from a PC function

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Reversible jump MCMC

Algorithm to add a new jump point

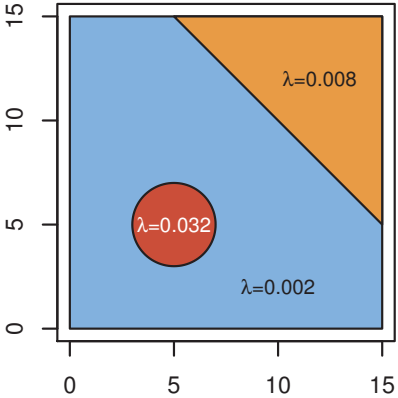
1. Select an interval j^* from $\text{Unif}\big(\{1, 2, \ldots, J+1\}\big)$.
2. Generate proposal jump point from that interval:
   - $\varepsilon^* \sim \text{Unif}(\varepsilon_{j^*-1}, \varepsilon_{j^*})$ and
   - level h^* using the Gamma prior.
3. Calculate the Hastings ratio, and determine whether to accept or reject the addition.

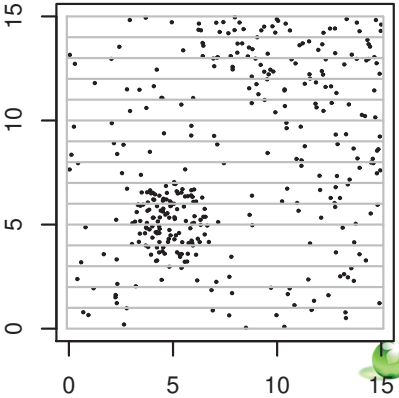
Similarly for deleting a jump point and corresponding level.

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Simulation study

500 simulated data sets containing each 5000 individuals and on average 357 events using hazard surface (and one simulated data set):





NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Results

Table 1: Median MSE (interquartile range IQR), and the number (percentage) of 1,485 examined points with 95% coverage probability within [0.94, 0.96] obtained for the different models.

Model	Number (%)	median MSE (IQR)
NPB, Lexis: $\alpha = 0.1$, $\phi = 0.1$	478 (32 %)	2.07e-06 (1.06e-06,5.68e-06)
NPB, Lexis: $\alpha = 0.1$, $\phi = 0.5$	581 (39 %)	2.09e-06 (1.10e-06,5.71e-06)
NPB, Lexis: $\alpha = 0.1$, $\phi = 1$	546 (37 %)	2.11e-06 (1.08e-06,5.73e-06)
NPB, Lexis: $\alpha = 0.1$, $\phi \sim \gamma(2, 2)$	536 (36 %)	2.02e-06 (1.08e-06,5.72e-06)
NPB, Lexis: $\alpha = 0.1$, $\phi = 2$	431 (29 %)	2.11e-06 (1.07e-06,5.76e-06)
NPB, Lexis: $\alpha = 1$, $\phi = 0.5$	150 (10 %)	1.86e-06 (8.69e-07,5.57e-06)
NPB, Lexis: $\alpha = 1$, $\phi = 1$	147 (10 %)	1.87e-06 (8.45e-07,5.35e-06)
NPB, Lexis: $\alpha = 1$, $\phi = 2$	134 (9 %)	1.88e-06 (8.29e-07,5.32e-06)
NPB, A-G: $\alpha = 0.1$	94 (6 %)	8.14e-06 (3.81e-06,2.72e-05)
NPB, A-G: $\alpha = 1$	65 (4 %)	2.10e-05 (8.01e-06,5.34e-05)
Poisson regression with splines	380 (26 %)	2.34e-06 (1.23e-06,1.19e-05)

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Empirical study on breast cancer

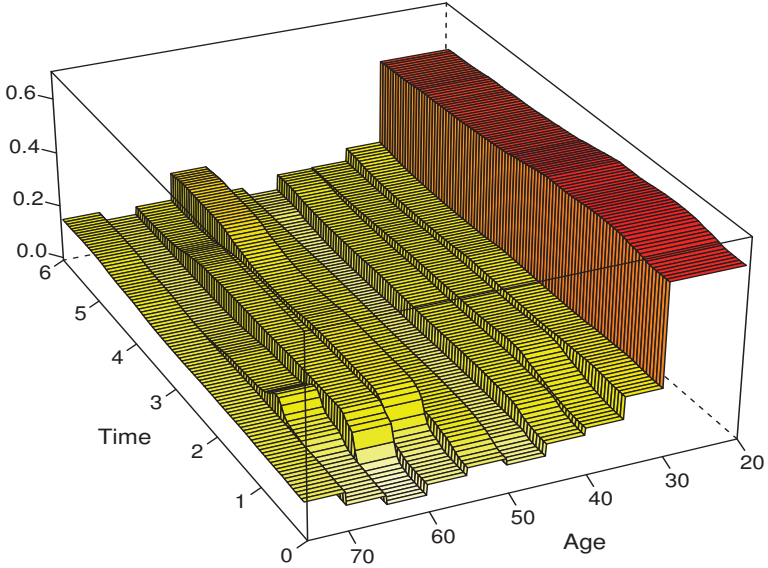
The gbcs[2] data set used here is based on a trial started in 1984 and conducted by the German Breast Cancer Study Group.

- 686 women diagnosed with primary, node-positive breast cancer
- Follow-up since the mastectomy until recurrence (N=299), death (N=171) or end of follow-up
- Seven prognostic factors: age at operation, menopausal status, hormone therapy, tumor size and grade, number of positive lymph nodes, number of progesterone and estrogen receptors
- Two time scales: age and time since mastectomy
- Ten strips according to age categories 20-30, 30-35, 35-40, . . . , 65-70, 70-80 years
- Patients within strips varied from 6 to 137, and events from 6 to 58

NATIONAL INSTITUTE FOR HEALTH AND WELFARE

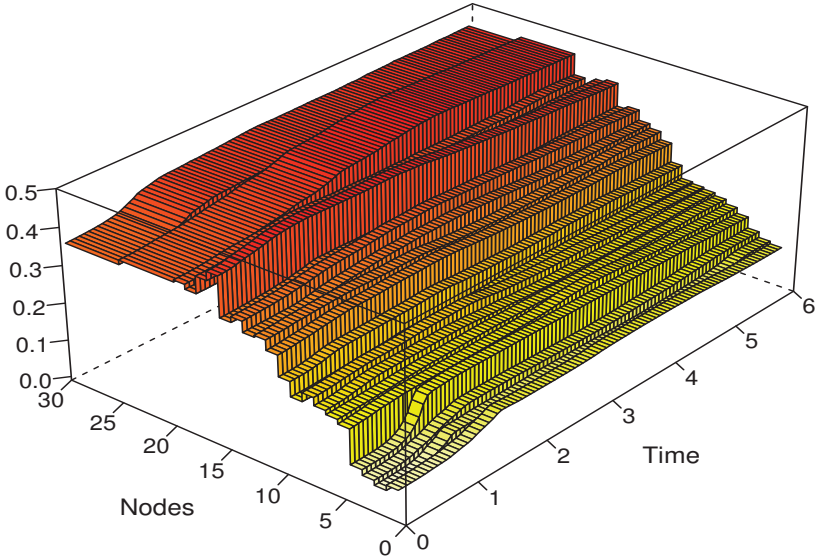
[2] https://www.umass.edu/statdata/statdata/data/

Hazard surface for age at operation and years since operation



NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Hazard functions for number of positive lymph nodes



NATIONAL INSTITUTE FOR HEALTH AND WELFARE

Conclusion

- Few methods currently available for Lexis diagrams and hazard surface models
- Two-dimensional smoothing provided more accurate results especially in the case with strict changepoint in the hazard
- Method can be applied also in case of ordinal covariates
- Our method works also in more than two dimensions
- Software and other material available at

`http://blogs.helsinki.fi/bayesian-intensity/`

NATIONAL INSTITUTE FOR HEALTH AND WELFARE