Auxiliary Mixture Sampling for Age-Period-Cohort Models

Andrea Riebler & Leonhard Held

University of Zurich

Swiss Statistics Meeting

Lucerne, November 2007

Outline

1. Introduction

Goal

Detection of spatial and temporal patterns in epidemiological data

Methods:

Age-Period-Cohort model: to describe incidence or mortality rates

Auxiliary mixture sampling: to estimate the APC-model

Age-Period-Cohort Model

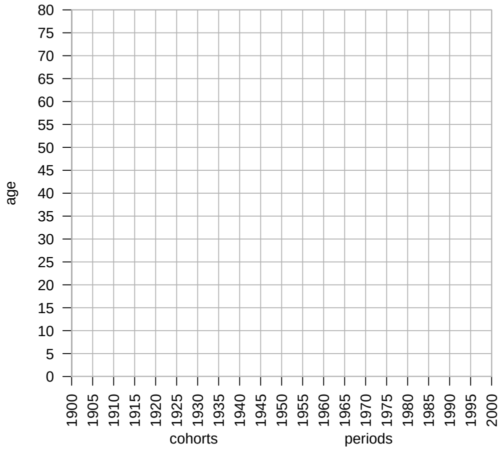
Problem

No or just limited data for possible disease factors available

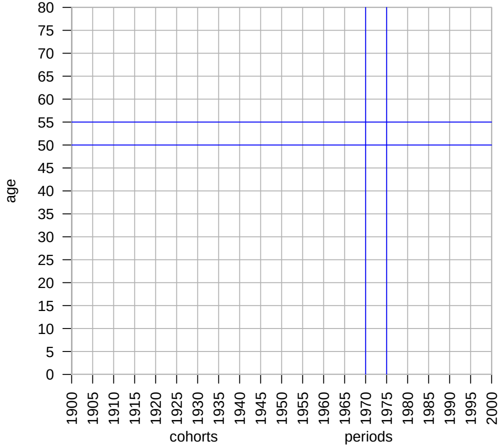
Analysis of incidence or mortality rates using three time scales

- **A**ge: age at diagnosis
- **P**eriod: date of diagnosis
- **C**ohort: date of birth

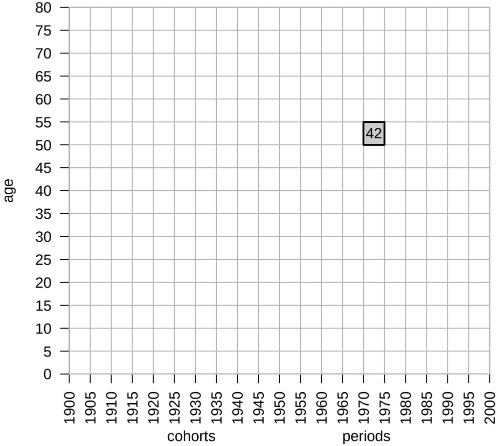
The Lexis-diagram



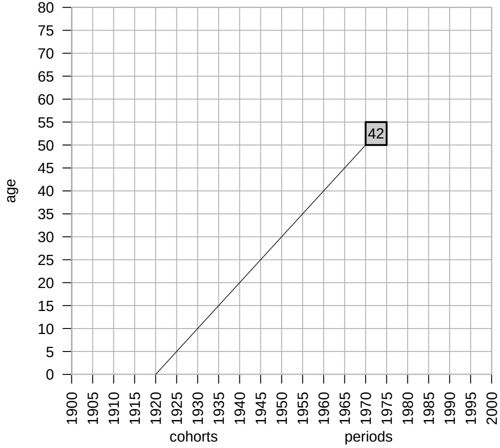
The Lexis-diagram



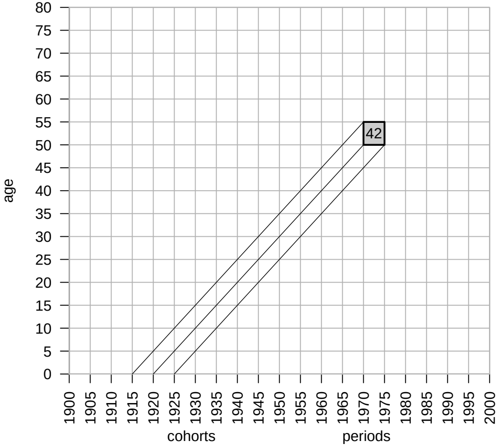
The Lexis-diagram



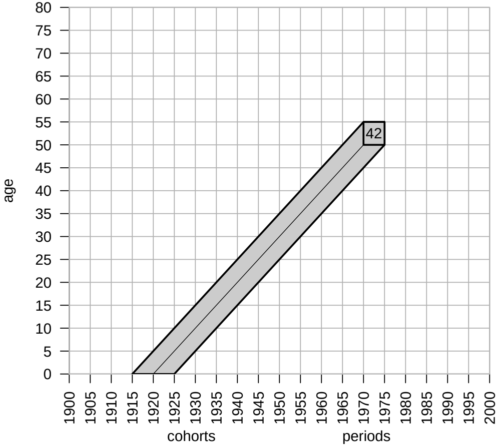
The Lexis-diagram



The Lexis-diagram



The Lexis-diagram



Age,- Period,- Cohort effects

Age effects

Consistent factors

Period effects

Factors that influence all persons under risk independent of the age, e.g. improvements of medical treatment

Cohort effects

Factors that influence persons of one generation, e.g. war

The cohort index

The cohort index k depends on the age group i and period j:

$$k = k(i,j) = (I - i) + j.$$

For different time grids

$$k = k(i,j) = G \cdot (I - i) + j,$$

where G is the grid factor.

$(i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, K$ and $K = G \cdot (I - 1) + J)$

The age-period-cohort model

y_{ij}: # counts in age group i at period j

n_{ij}: population in age group i at period j

The APC-model

$$y_{ij} \sim \mathrm{Po}(\underbrace{n_{ij} \cdot p_{ij}}_{\lambda_{ij}})$$

$$\eta_{ij} = \log(\lambda_{ij}) = \log(n_{ij}) + \mu + \alpha_i + \beta_j + \gamma_k$$

with age effect α_i, period effect β_j, cohort effect γ_k

Non-identifiability

To assure identifiability additional constraints are necessary

- Make the intercept μ identifiable

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_k \gamma_k = 0$$

- The APC parameters are still not identifiable:

$$\alpha_i \rightarrow \alpha_i + c \cdot \left(i - \frac{I+1}{2}\right);$$

$$\beta_j \rightarrow \beta_j - c \cdot \left(j - \frac{J+1}{2}\right);$$

$$\gamma_k \rightarrow \gamma_k + c \cdot \left(k - \frac{K+1}{2}\right)$$

$$\Rightarrow \eta_{ij} = \log(n_{ij}) + \mu + \alpha_i + \beta_j + \gamma_k$$

is left unchanged

Random walks (RW) as prior distributions

First-order random walk (RW1)

$$\alpha_i \sim \mathrm{N}(\alpha_{i-1}, \kappa^{-1}) \qquad i = 2, \ldots, I.$$

RW1 penalizes deviations from a model where all parameters are constant.

Second-order random walk (RW2)

$$\alpha_i \sim \mathrm{N}(2\alpha_{i-1} - \alpha_{i-2}, \kappa^{-1}) \qquad i = 3, \ldots, I$$

RW2 penalizes deviations from a linear trend $\alpha_i = 2\alpha_{i-1} - \alpha_{i-2}$.

Heterogeneity

To account for additional "unstructured" heterogeneity, an additional parameter $z_{ij} \sim \mathrm{N}(0, \delta^{-1})$ can be introduced

$$\xi_{ij} = \underbrace{\log(n_{ij}) + \mu + \alpha_i + \beta_j + \gamma_k}_{\eta_{ij}} + z_{ij}$$

Using a reparameterization, it follows

$$z_{ij} = \xi_{ij} - (\log n_{ij} + \mu + \alpha_i + \beta_j + \gamma_k) = \xi_{ij} - \eta_{ij}$$

The implied prior of the linear predictor is

$$f(\boldsymbol{\xi}|\boldsymbol{\eta}, \delta) \propto \delta^{IJ/2} \cdot \exp\left(-\frac{\delta}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}(\xi_{ij} - \eta_{ij})^2\right).$$

Full conditional distributions

For all parameters except ξ_{ij} Gibbs sampling is possible.

Problem

The full conditional distribution

$$f(\boldsymbol{\xi}|\mathbf{y}, \boldsymbol{\eta}, \kappa, \nu, \tau, \delta) \propto f(\mathbf{y}|\boldsymbol{\xi})f(\boldsymbol{\xi}|\boldsymbol{\eta}, \delta)$$

is a non-standard distribution.

$\Rightarrow$ **Auxiliary mixture sampling**

Auxiliary Mixture Sampling for APC models

Model specification:

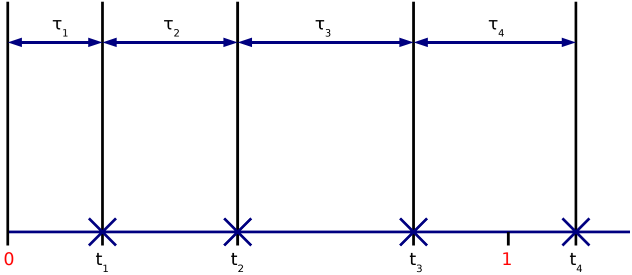
$$y_{ij} \sim \mathrm{Po}(\underbrace{n_{ij} \cdot p_{ij}}_{\lambda_{ij}})$$

$$\xi_{ij} = \log(\lambda_{ij}) = \underbrace{\log(n_{ij}) + \mu + \alpha_i + \beta_j + \gamma_k}_{\eta_{ij}} + z_{ij}$$

Data augmentation:

Introduce additional auxiliary variables to eliminate nonlinearity and non-normality.

Poisson Process



crosses mark the occurrence of an event,
$t_1, \ldots, t_4$ arrival times,
$\tau_1, \ldots, \tau_4$ inter-arrival times

Data Augmentation (1)

For each $y_{ij} > 0$ introduce $\boldsymbol{\tau}_{ij} = (\tau_{ij1}, \tau_{ij2})$, for $y_{ij} = 0$ just $\tau_{ij} = \tau_{ij1}$

- τ_{ij2} denotes the last jump before 1 and is $\mathrm{Ga}(y_{ij}, \lambda_{ij})$

$$\tau_{ij2} = \rho_{ij2}/\lambda_{ij}, \qquad \rho_{ij2} \sim \mathrm{Ga}(y_{ij}, 1)$$

Data Augmentation (1)

For each $y_{ij} > 0$ introduce $\boldsymbol{\tau}_{ij} = (\tau_{ij1}, \tau_{ij2})$, for $y_{ij} = 0$ just $\tau_{ij} = \tau_{ij1}$

- τ_{ij2} denotes the last jump before 1 and is $\mathrm{Ga}(y_{ij}, \lambda_{ij})$

$$\tau_{ij2} = \rho_{ij2}/\lambda_{ij}, \qquad \rho_{ij2} \sim \mathrm{Ga}(y_{ij}, 1)$$

- τ_{ij1} denotes the inter-arrival time between the last jump before and the first jump after 1 and is $\mathrm{Ex}(\lambda_{ij})$

$$\tau_{ij1} = \rho_{ij1}/\lambda_{ij}, \qquad \rho_{ij1} \sim \mathrm{Ex}(1).$$

Data Augmentation (2)

Reformulation:

$$-\log(\tau_{ij2}) = \underbrace{\log(\lambda_{ij})}_{\xi_{ij}} + \epsilon_{ij2} \qquad \epsilon_{ij2} \sim -\log \mathrm{Ga}(y_{ij}, 1)$$

$$-\log(\tau_{ij1}) = \underbrace{\log(\lambda_{ij})}_{\xi_{ij}} + \epsilon_{ij1} \qquad \epsilon_{ij1} \sim -\log \mathrm{Ex}(1)$$

where $\epsilon_{ij1} = -\log(\rho_{ij1})$ and $\epsilon_{ij2} = -\log(\rho_{ij2})$.

Mixture Approximation (1)

- ϵ_{ij1} follows the negative of a log $\mathrm{Ex}(1)$ distribution:

$$p(\epsilon_{ij}) = \exp(-\epsilon_{ij} - \exp(-\epsilon_{ij})) \approx \sum_{r=1}^{R} w_r \, \mathrm{N}(\epsilon_{ij}; m_r, s_r^2)$$

  with parameters m_r and s_r, and weight w_r for the rth component.

- ϵ_{ij2} follows a negative log Gamma distribution with integer shape parameter y_{ij}:

$$p(\epsilon_{ij}; y_{ij}) = \frac{\exp(-y_{ij}\epsilon_{ij} - \exp(-\epsilon_{ij}))}{\Gamma(y_{ij})} \approx \sum_{r=1}^{R(y_{ij})} w_r(y_{ij}) \, \mathrm{N}(\epsilon_{ij}; m_r(y_{ij}), s_r^2(y_{ij}))$$

Mixture Approximation (2)

Introduce the component indicators
$\mathbf{S} = \{r_{ij} = (r_{ij1}, r_{ij2}), i = 1, \ldots, I, j = 1, \ldots, J\}$ as missing data

$$-\log(\tau_{ij1}) = \xi_{ij} + m_{r_{ij1}}(1) + \epsilon_{ij1}, \qquad \epsilon_{ij1} \sim \mathrm{N}(0, s^2_{r_{ij1}}(1))$$

$$-\log(\tau_{ij2}) = \xi_{ij} + m_{r_{ij2}}(y_{ij}) + \epsilon_{ij2}, \qquad \epsilon_{ij2} \sim \mathrm{N}(0, s^2_{r_{ij2}}(y_{ij}))$$

For $y_{ij} = 0$ just the first equation is necessary.

Consequence

The full conditional distribution for ξ_{ij} now is a normal distribution
$\Rightarrow$ Gibbs-sampling possible

Sampling scheme

Select starting values for the hyperparameters, for $\boldsymbol{\tau}$, **S** and the main effects $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$.

1. Sample $\boldsymbol{\xi}$ conditional on $\boldsymbol{\tau}, \mathbf{S}, \mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \delta$ and **y**
2. Sample the inter-arrival times $\boldsymbol{\tau}$ and the component indicators **S** conditional on $\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$, $\boldsymbol{\xi}$ and **y** (exponential resp. discrete distributions)
3. Update the main effects and hyperparameters using Gibbs-sampling (multivariate Gaussian, gamma distributions)

Implementation

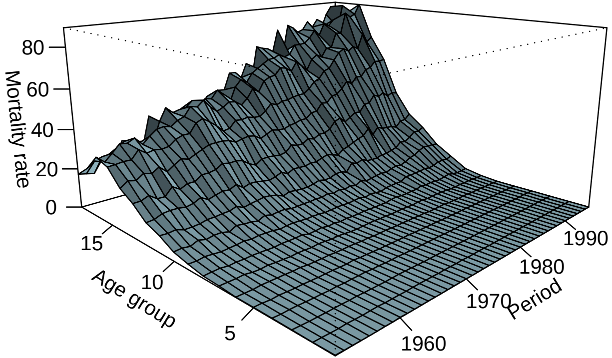
in C using the GMRFLib library by Håvard Rue
(`http://www.math.ntnu.no/~hrue/GMRFLib`)

GMRFLib includes functions:

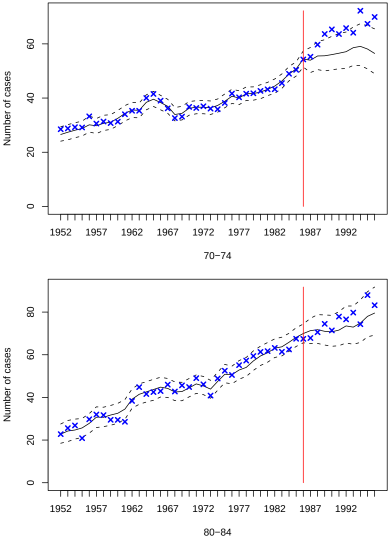
- for efficient sampling from age, period and cohort blocks
- for defining auxiliary mixture sampling for Poisson and binomial observational models

Application

Lung cancer mortality rates for females in West Germany



Retrospective prediction (RW1)



Summary and Outlook

Summary

Presentation of a Bayesian framework based on Gaussian Markov random fields to estimate APC models by drawing from standard distributions only.

Outlook

Application to multivariate datasets evaluating different APC models with joint or separated main effects.

References

Frühwirth-Schnatter, S., R. Frühwirth, L. Held, and H. Rue (2007).
Improved Auxiliary Mixture Sampling for Hierarchical Models of Non-Gaussian Data.
IFAS Research Report 2007-25, http://www.ifas.jku.at, Johannes Kepler University Linz.

Rue, H. and L. Held (2005).
Gaussian Markov Random Fields: Theory and Applications.
Volume 104 of *Monographs on Statistics and Applied Probability*, Boca Ratin, FL:Chapmann & Hall/CRC.

Thank you for your attention!

Any questions?