Prediction of future mortality

Mathias Lindholm

Cramér, September, 2020

The presented material is based primarily on joint work with

Patrik Andersson, Uppsala University

in particular the paper

"Mortality forecasting using a Lexis-based state-space model"

To appear in Annals of Actuarial Science

Outline of the presentation

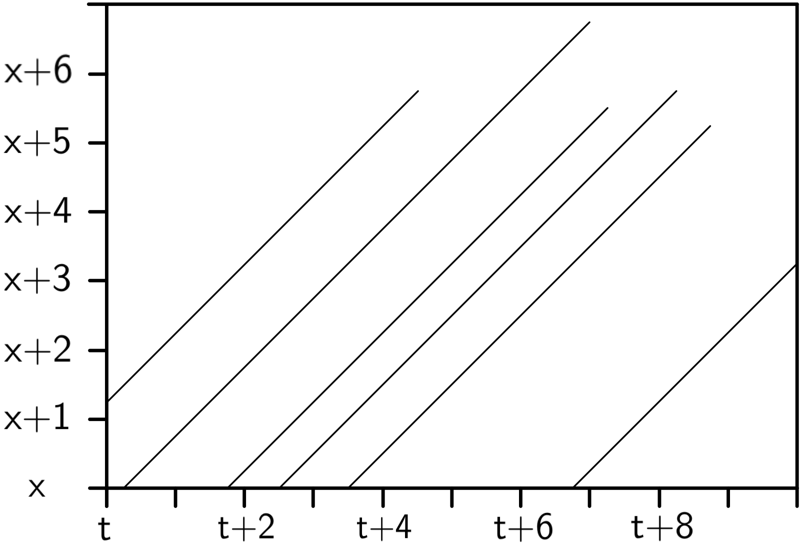
- What do we observe?
- Discuss how to go from observed mortality to predicting future mortality
- From Lee-Carter to state-space models
- Some words about ongoing work and extensions

Some notation:

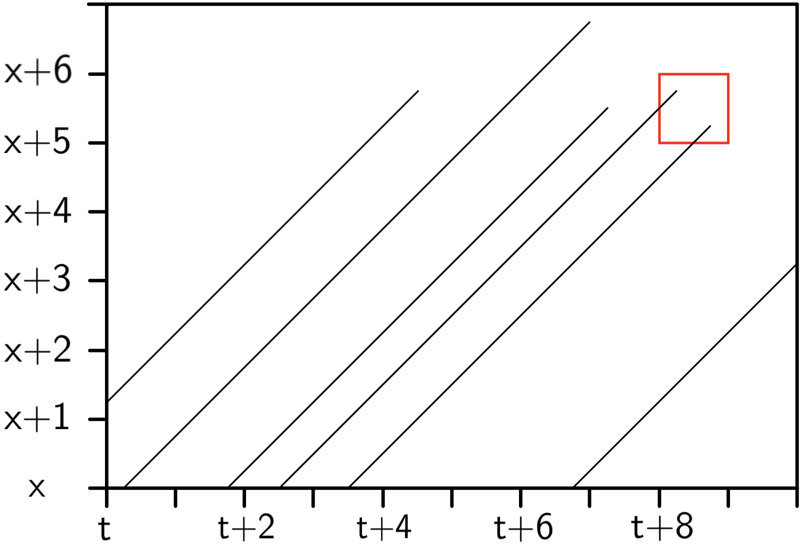
- all vectors are column vectors
- we will usually not define the dimension of vectors and matrices – should be clear from context
- X corresponds to a random quantity, x corresponds to an observation
- $x_{0:n} = (x_0, \ldots, x_n)$
- all Greek letters corresponds to model parameters

What we observe and structuring of data – Lexis diagrams

Lexis diagrams



Lexis diagrams



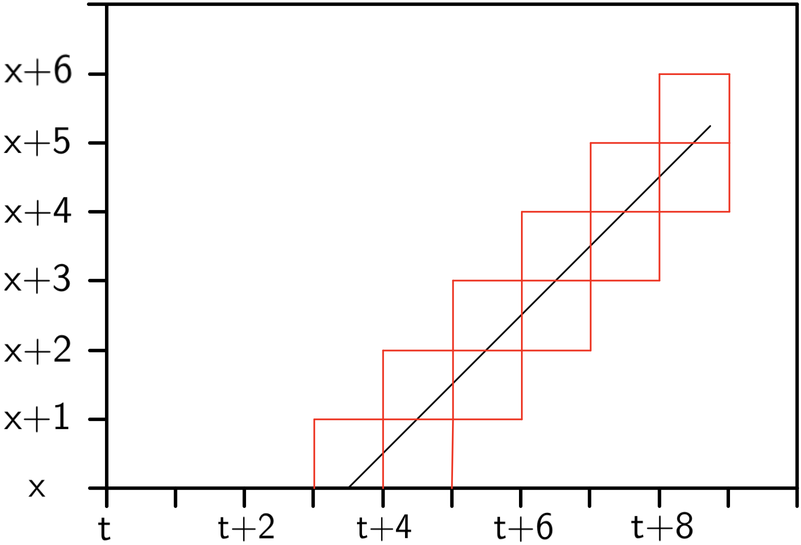
Lexis diagrams: models and likelihood

Assumptions:

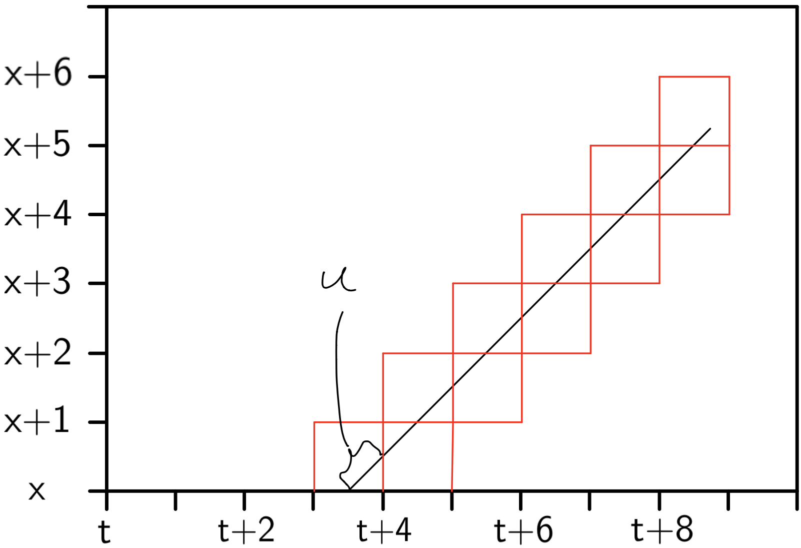
- all individuals are independent
- the mortality rate is constant within yearly Lexis squares

Start with a single individual!

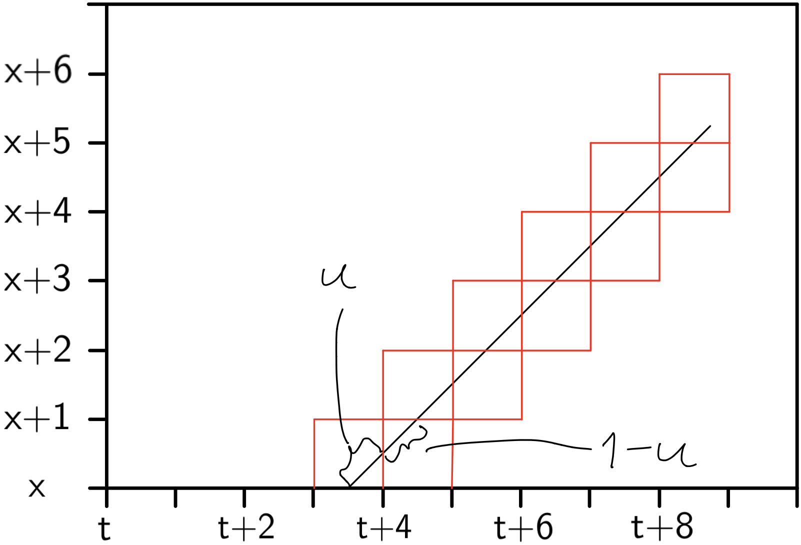
Lexis diagrams: models and likelihood – single individual



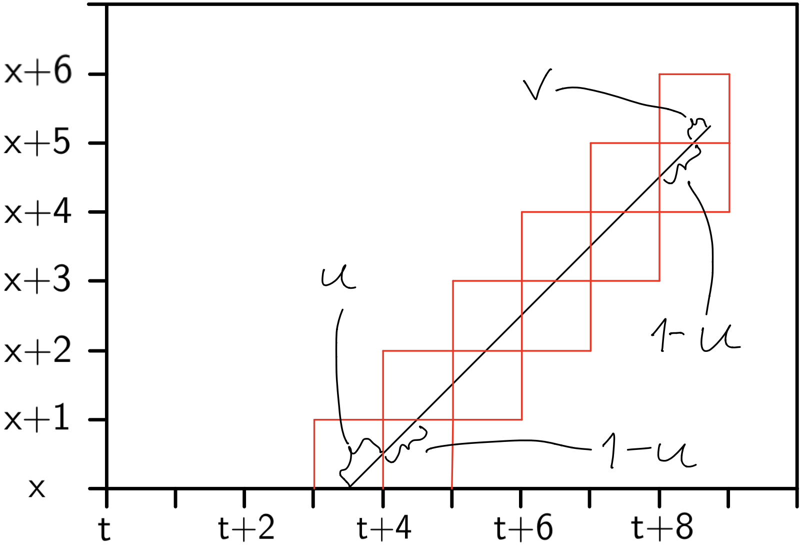
Lexis diagrams: models and likelihood – single individual



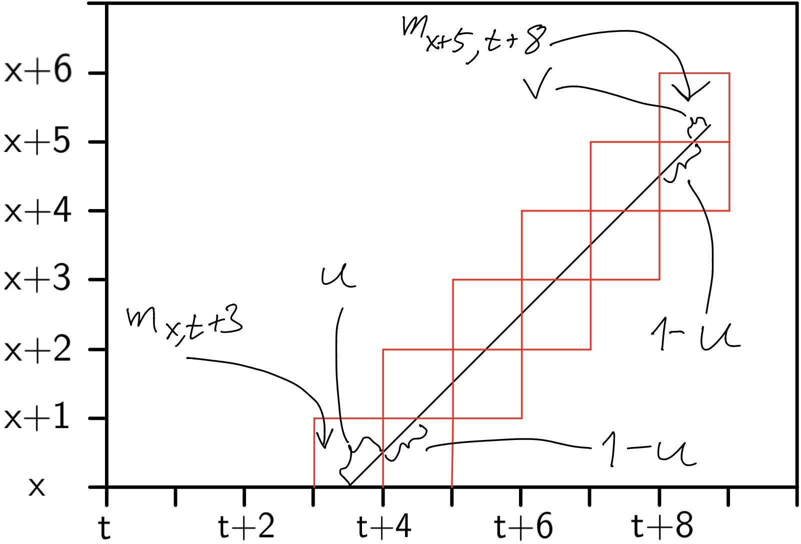
Lexis diagrams: models and likelihood – single individual



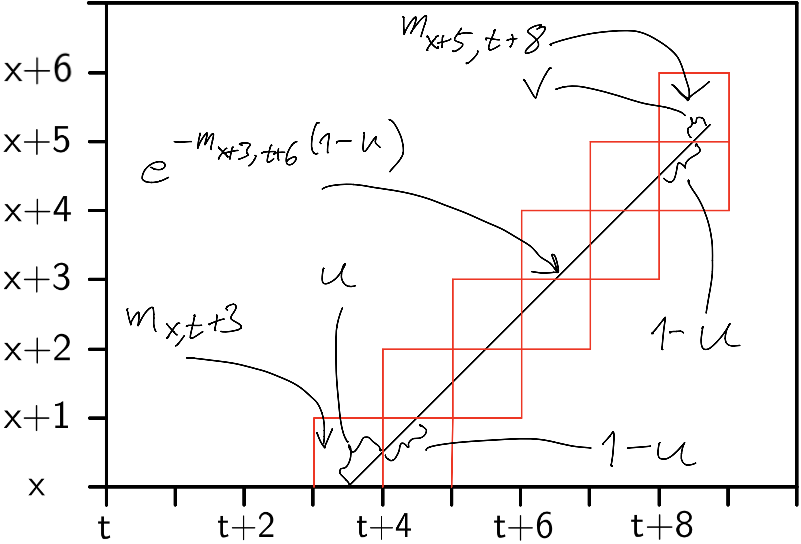
Lexis diagrams: models and likelihood – single individual



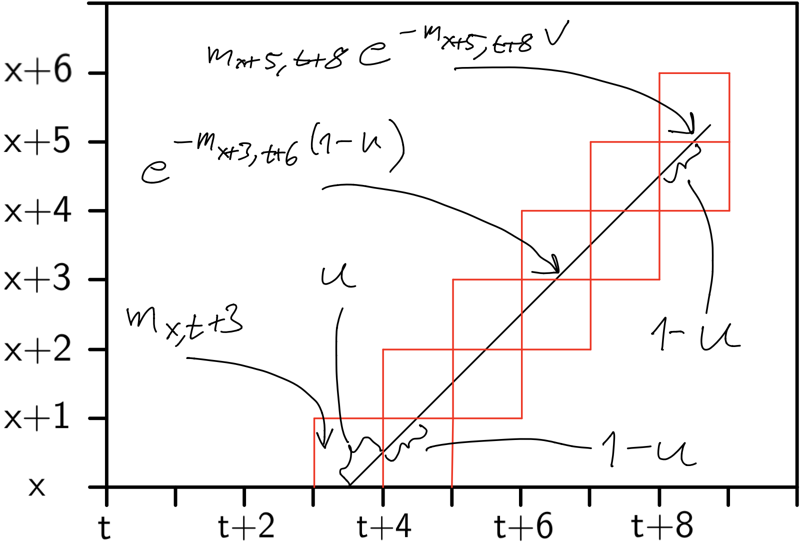
Lexis diagrams: models and likelihood – single individual



Lexis diagrams: models and likelihood – single individual



Lexis diagrams: models and likelihood – single individual

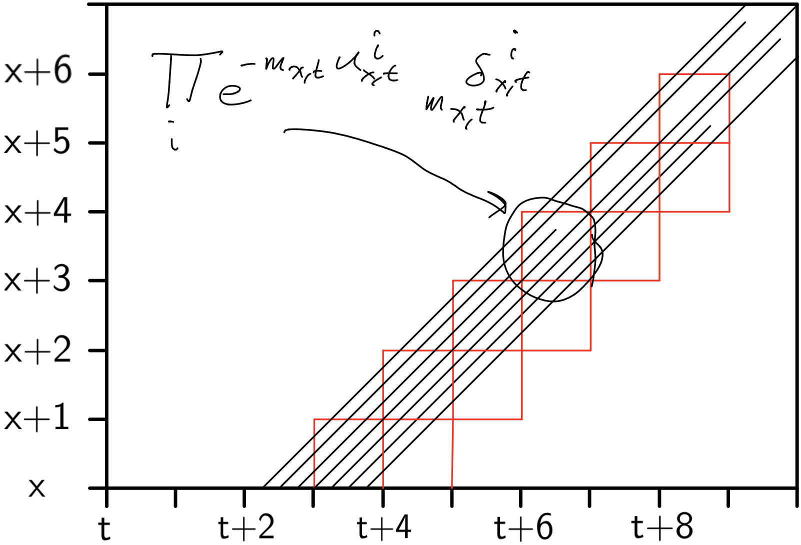


Lexis diagrams: models and likelihood – single individual

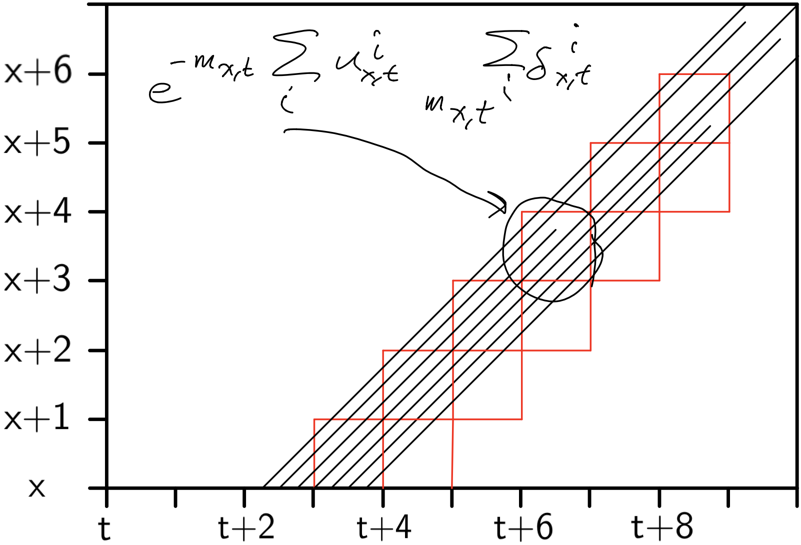
Single individual: multiply contribution from all Lexis squares visited by the individual

Single square, many individuals?

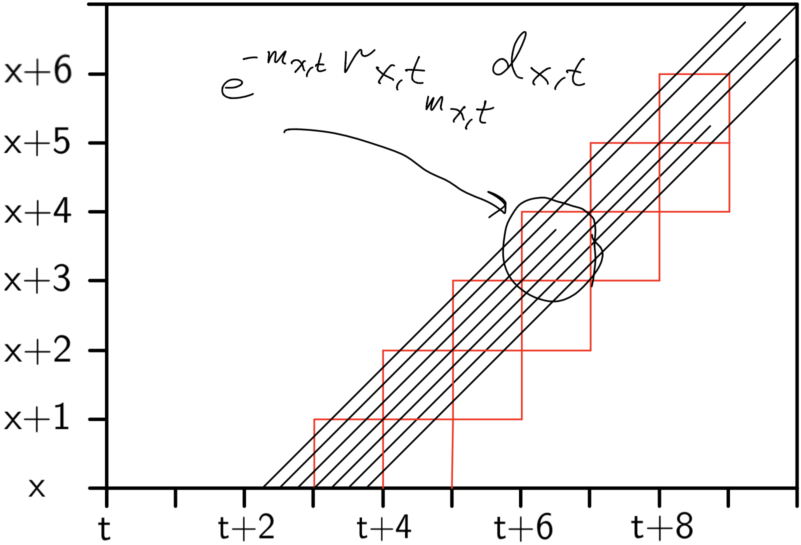
Lexis diagrams: models and likelihood – single square



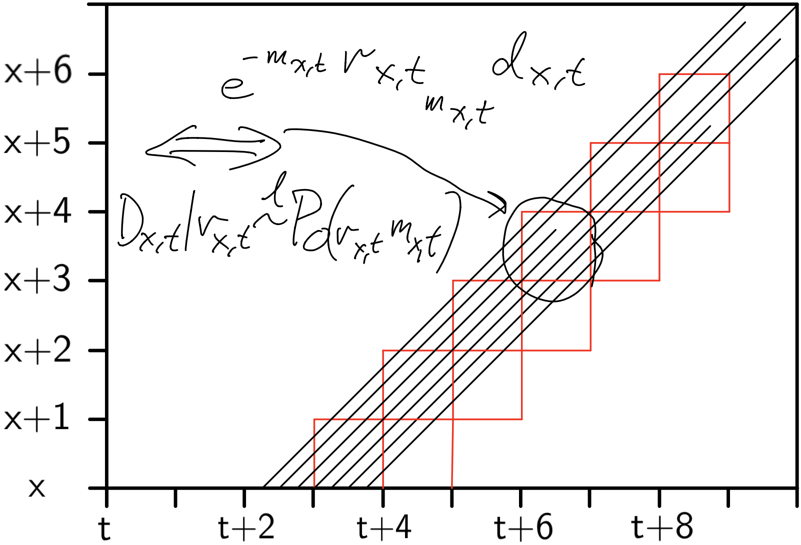
Lexis diagrams: models and likelihood – single square



Lexis diagrams: models and likelihood – single square



Lexis diagrams: models and likelihood – single square



Lexis diagrams: models and likelihood

Lemma 1

The log-likelihood for the total population, assuming independence between individuals and piecewise constant $m_{x,t}$, is,

$$l(\mathcal{M}) = \sum_{(x,t)\in\mathcal{S}} (d_{x,t} \log m_{x,t} - r_{x,t} m_{x,t}), \qquad (1)$$

where $\mathcal{S}$ denotes the set of all observed Lexis squares, $\mathcal{M} = \{m_{x,t} \mid (x,t) \in \mathcal{S}\}$, $d_{x,t}$ corresponds to the observed number of deaths in (x,t) and $r_{x,t}$ corresponds to the observed exposure-to-risk in (x,t).

Consequence of Lemma 1:

$$\widehat{m}_{x,t} = \frac{d_{x,t}}{r_{x,t}},$$

$\widehat{m}_{x,t}$ is often referred to as the "force of mortality" or "mortality rate".

Lexis diagrams: models and likelihood – prediction

...but we want to predict future mortality rates...

...and note that Lemma 1 **provides no structure between** $m_{x,t}$s

We need additional structure!

Can be done in a lot of different ways!

Lexis diagrams: models and likelihood

The "Lee-Carter approach", see e.g. [3]:

1. Estimate $\widehat{m}_{x,t}$ for all $(x, t) \in \mathcal{S}$
2. Given $\{\widehat{m}_{x,t}\}$, fit a stochastic process to the estimated $\widehat{m}_{x,t}$s

Sometimes a "post fitting adjustment" is done, due to problems caused by the two-stage procedure

Lexis diagrams: models and likelihood

In somewhat more detail:

1. Estimate $\widehat{m}_{x,t}$ for all $(x, t) \in \mathcal{S}$ and let Y be the $m \times n$ matrix defined as
$$Y_{x,t} = \log \widehat{m}_{x,t}$$
2. Let $\bar{y}$ denote the $m \times 1$ vector with age-wise estimated average mortality rates, and make the following SVD-approximation:
$$(Y - \bar{y})_{x,t} \approx \beta_x \kappa_t$$
3. Fit a (Gaussian) stochastic process to the estimated $\widehat{\kappa}_t$s

Note: 2. corresponds to **dimension reduction**

Lexis diagrams: models and likelihood

Note the close resemblance between the above procedures and the following state space model (see e.g. [1, 2]):

$$\begin{cases} Y_t & = \mu + \beta K_t + U_t, \quad U_t \sim N(0, \sigma_u^2 I) \\ K_t & = \alpha + K_{t-1} + V_t, \quad V_t \sim N(0, \sigma_v^2) \end{cases}$$

where $\mu, \beta, \alpha, \sigma_u$, and σ_v are unknown parameters and the K_ts correspond to **unobservable states**

Still, note that Y is a matrix based on **estimated** mortality rates!

Lexis diagrams: models and likelihood

Note that

- the Lee-Carter model is a model for **mortality rates**
- we **observe death counts**

Taking the above into account (Andersson & ML, [4]):

Use a Poisson state space model for **death counts**

Poisson state space models

The Poisson state space model

$$\begin{cases} D_{i,t} \mid M_{i,t}, r_{i,t} \sim \mathrm{Po}(r_{i,t} M_{i,t}) \\ M_{i,t} = \exp\{(\Upsilon X_t)_i\} \\ X_{t+1} = \Gamma X_t + \mu + V_t, \;\; V_t \sim \mathrm{N}(0, \Sigma) \\ X_0 \sim \mathrm{N}(\mu_0, \Sigma_0) \end{cases} \qquad (\star)$$

Note that

- Υ corresponds to the dimension reduction from m to p (a.k.a. "EPCA" or "GPCA"), where $X_t \in \mathbb{R}^p$, where m is the number of age groups
- $M_{i,t}$ is defined in terms of the standard link-function for a Poisson GLM and that $\log M_{i,t} \approx Y_{i,t}$

Poisson state space models

Note that, given estimated parameters and state vectors, $x_{0:t}$, it is easy to

- decompose the in-sample variation of

$$\widehat{M}_{i,t} := \frac{D_{i,t}}{r_{i,t}}$$

  in terms of variation from $M_{i,t} = \exp\{(\Upsilon X_t)_i\}$ and variation from $D_{i,t}$ conditional on $M_{i,t}$ and $r_{i,t}$
- project future mortality rates (and death counts) by first sampling a trajectory $x_{0:n}$

Estimation

Following [4], estimation of **parameters** and **state vectors** is done using Stochastic Approximation EM + particle filter techniques

- in a first step Υ is estimated using a Poisson likelihood
- given $\widehat{\Upsilon}$, the remaining model parameters, Γ, μ, and Σ, and state-vectors are estimated
- the SAEM algorithm makes use of that we have explicit sufficient statistics for the remaining model parameters

  these sufficient statistics are updated iteratively using a weighted averaging
- in each SAEM update state-vectors are sampled using the Forward Filtering Backward Smoothing algorithm given the latest parameter updates

Model evaluation and forecasting

We will split our data into two parts

- in-sample "training data"
- out-of-sample "validation data"

Recall that model ($\star$) allows us to

- predict future **mortality rates**
- (based on simulations) calculate

$$M_{i,t} = \frac{D_{i,t}}{r_{i,t}} \quad \text{(force of mortality)} \tag{2}$$

By using (2) for model validation we will capture both

(i) the population ("Poisson") variation
(ii) the variation from the mortality rates

Model evaluation and forecasting

Concerning model selection, Andersson & ML, [4], suggest an R^2-type measure that

- is based on using the likelihood
- works both in-sample and out-of-sample

Numerical illustrations

Numerical illustrations

We will use the following special case of model $(\star)$:

$$\begin{cases} D_{i,t} \mid M_{i,t}, r_{i,t} \sim \mathrm{Po}(r_{i,t} M_{i,t}) \\ M_{i,t} = \exp\{(\Upsilon X_t)_i\} \\ X_{t+1} = \Gamma_{xx} X_t + \Gamma_{xy} Y_t + \mu_x + V_t, \;\; V_t \sim \mathrm{N}(0, \Sigma_x) \\ Y_{t+1} = \Gamma_{yy} Y_t + \mu_y + U_t, \;\; U_t \sim \mathrm{N}(0, \Sigma_y) \\ X_0 \sim \mathrm{N}(\mu_0, \Sigma_0) \\ Y_0 \sim \mathrm{N}(\mu_0, \Sigma_0) \end{cases}$$

Numerical illustrations

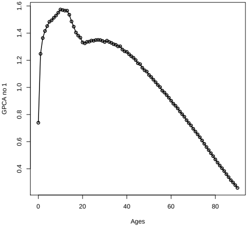
We will

- use Swedish male data from 1930–2016 with ages 0–90
- use 1–3 GPCA components
- illustrate the role of the Poisson part of the state space modelling

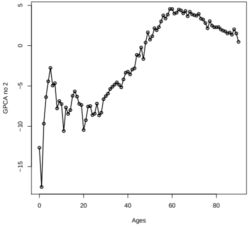
To start off: SWE male population, estimation 1930–1960, validation 1961–2016

(More examples, and more on model selection, can be found in Andersson & ML, [4])

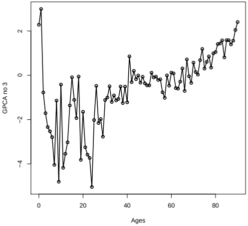
(a) 1st GPCA



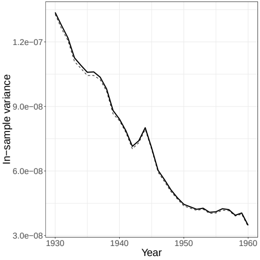
(b) 2nd GPCA



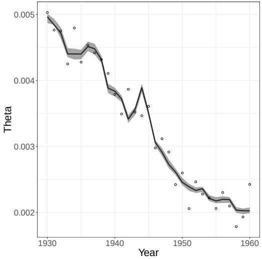
(c) 3rd GPCA



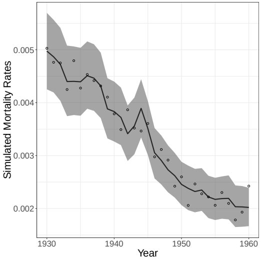
(d) Population var vs total var



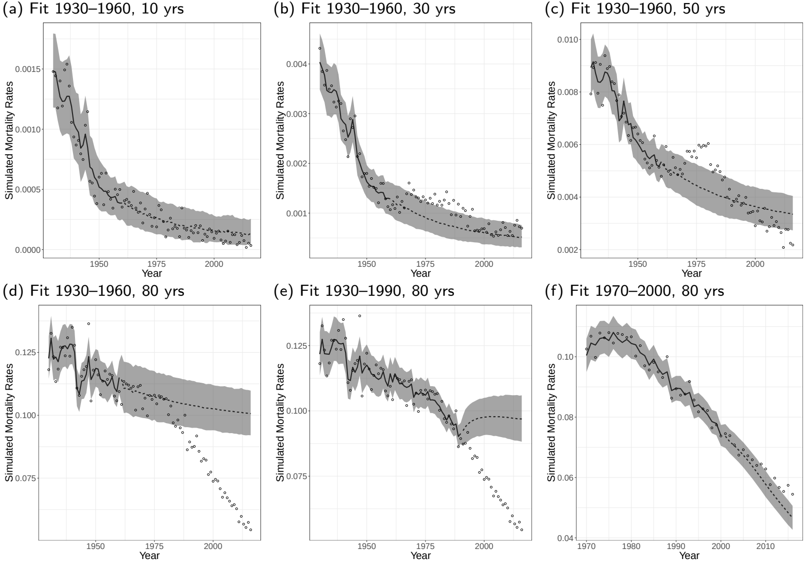
(e) $M_{i,t} = \exp(\Upsilon X_t)_i$



(f) $M^{\text{sim}}_{i,t} = D_{i,t}/r_{i,t}$



SWE males, 3 GPCA



Concluding remarks and extensions

We have

- introduced a flexible Poisson state space model which allows for dimension reduction...

  ...that avoids two-step estimation procedures
- illustrated the importance of explicitly capturing the Poisson part of the variation

...but

- we have used SAEM + particle filter techniques
  – other choices?
- one could consider allowing for even more flexibility
  – neural networks? ...estimation!?

...this is work in progress

References I

Piet De Jong and Leonie Tickle.
Extending lee–carter mortality forecasting.
Mathematical Population Studies, 13(1):1–18, 2006.

James Durbin and Siem Jan Koopman.
Time series analysis by state space methods.
Number 38. Oxford University Press, 2012.

Ronald D Lee and Lawrence R Carter.
Modeling and forecasting us mortality.
Journal of the American statistical association, 87(419):659–671, 1992.

Andersson Patrik and Mathias Lindholm.
Mortality forecasting using a lexis-based state-space model.
To appear in Annals of Actuarial Science, 2020.

Appendix

Female populations in SWE and US

Essentially similar results, will focus on ages 10, 40, and 80

- SWE females, fit 1950–1980, 2 GPCA
- US females, fit 1950–1980, 3 GPCA

Fit 1950–1980