FORECASTING MORTALITY OF NOT EXTINCT COHORTS WITH THE PENALIZED COMPOSITE LINK MODEL

Silvia Rizzi, Søren Kjærgaard, Carlo Giovanni Camarda, Marie-Pier Bergeron-Boucher, Rune Lindahl-Jacobsen and James W. Vaupel

Interdisciplinary Center on Population Dynamics (CPOP)

University of Southern Denmark

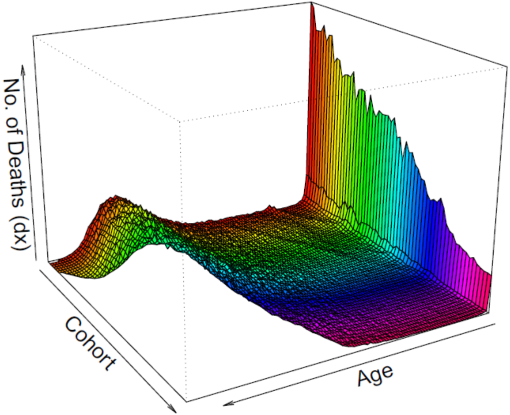
August 8, 2019

- Study mortality using birth cohort data vs. period data: advantages and limitations.

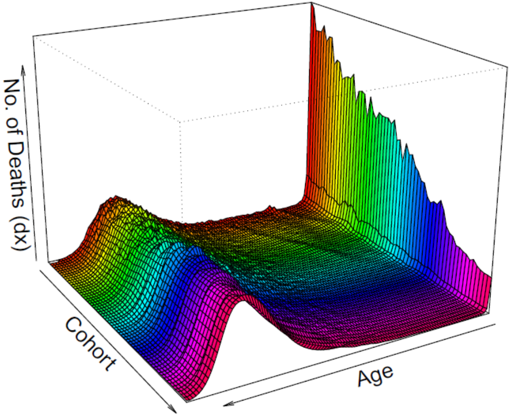
- Study mortality using birth cohort data vs. period data: advantages and limitations.
- Our aim: forecast cohort d_x, i.e. forecast age-at-death distribution for each cohort.

- Study mortality using birth cohort data vs. period data: advantages and limitations.
- Our aim: forecast cohort d_x, i.e. forecast age-at-death distribution for each cohort.
- Our target: Complete cohorts born up to 1960.

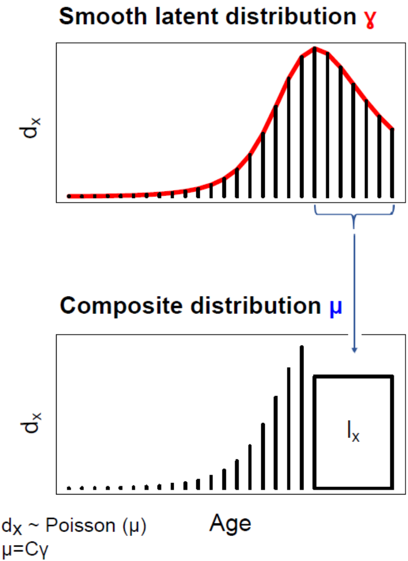
What we have



What we want



Our model


Smooth latent distribution γ
dx
Composite distribution μ
dx
lx
dx ~ Poisson (μ)
μ=Cγ
Age

- Cohort d_x observed counts, i.e. y_i, in I age steps.
- y_i assumed to be Poisson distributed with $E(y_i) = \mu_i$.
- Vector μ_i results from grouping the $J > I$ observed counts in the age steps $\mu_i = c_{ij}\gamma_j$.
- Vector γ_j assumed to be smooth.
- Model estimation by penalized maximum likelihood (Eilers, 2007).

Transition matrix

$$
C = \begin{array}{c} \\ d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_A \\ d_{A+} \\ 0 \end{array}
\begin{array}{c}
\begin{array}{ccccccccccc} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_A & \gamma_{A+1} & \cdots & \gamma_{120} & \gamma_{121} & \cdots & \gamma_{130} \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 0 & \cdots & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & \cdots & 1
\end{pmatrix}
\end{array}.
$$

- The PCLM method fails for cohorts where only the younger ages groups are observed.

The penalized composite link model (PCLM) with augmented data

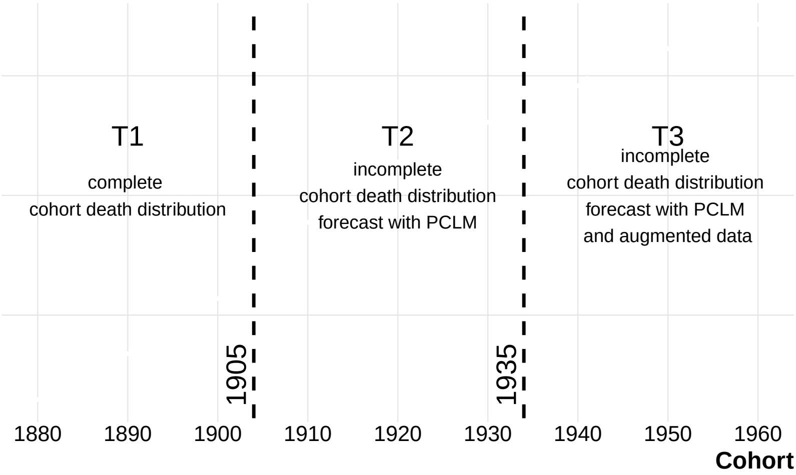
- The PCLM method fails for cohorts where only the younger ages groups are observed.
- Incorporate demographic forecasts when forecasting recent cohorts.

The penalized composite link model (PCLM) with augmented data

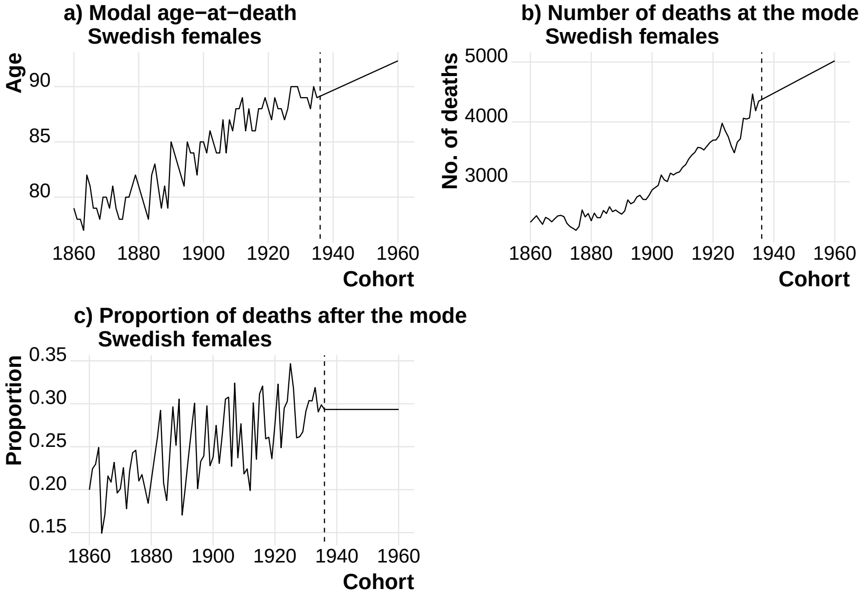
- The PCLM method fails for cohorts where only the younger ages groups are observed.
- Incorporate demographic forecasts when forecasting recent cohorts.
- Forecasts of the modal age at death.
- Forecasts of d_x at the mode.
- Forecasts of deaths after the modal age at death.

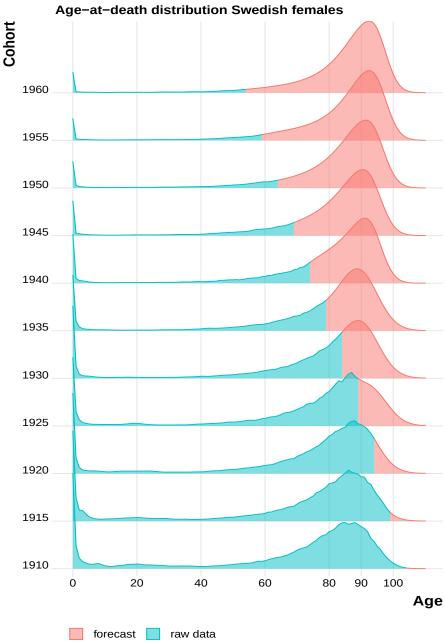
$$
C = \begin{array}{c} \\ d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_A \\ d_{M-} \\ d_M \\ d_{M+} \\ 0 \end{array}
\begin{array}{c}
\begin{array}{ccccccccccccccc}
\gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_A & \gamma_{A+1} & \cdots & \gamma_{M-1} & \gamma_M & \gamma_{M+1} & \cdots & \gamma_{120} & \gamma_{121} & \cdots & \gamma_{130}
\end{array} \\
\left(\begin{array}{ccccccccccccccc}
1 & 0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 0 & \cdots & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & 0 & 1 & \cdots & 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & 0 & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & \cdots & 1
\end{array}\right)
\end{array} .
$$

Data periods

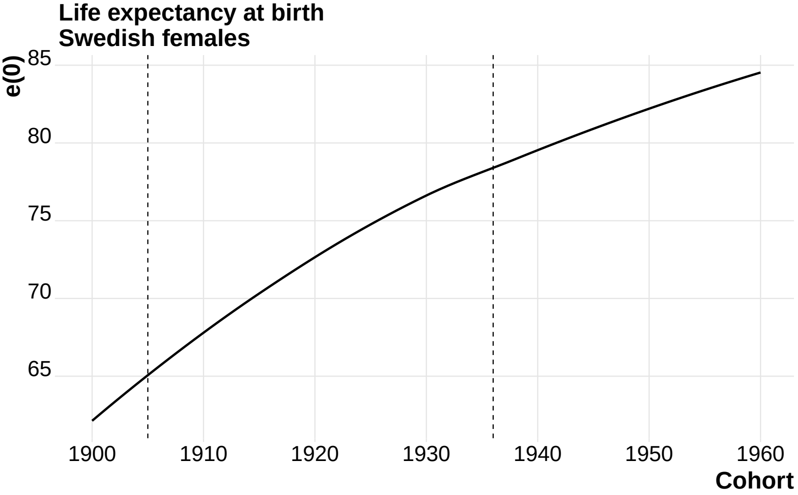


Forecasts of the modal age of death

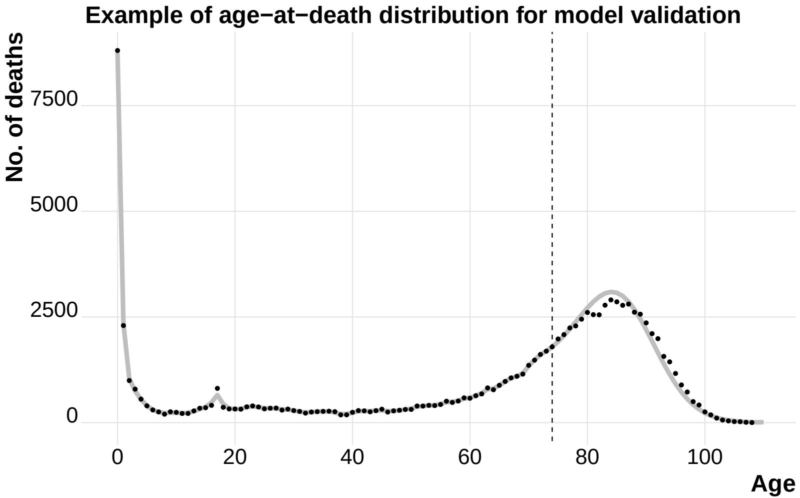



Age−at−death distribution Swedish females
Cohort
1960
1955
1950
1945
1940
1935
1930
1925
1920
1915
1910
0
20
40
60
80
90
100
Age
forecast
raw data

Life expectancy forecast



- (1) Deaths after age x are set missing using a completed cohort. x=75 in the example
- (2) Out-of-sample validation. A part of the data period is used for validation


Example of age−at−death distribution for model validation
No. of deaths
7500
5000
2500
0
0
20
40
60
80
100
Age

Years left out	10 years	15 years	20 years	25 years	30 years
Switzerland, females					
RMSE for $e_{0,i}$	0.0016	0.0046	0.0324	0.0389	0.3357
RMSE for $e_{50,i}$	0.0040	0.0081	0.0429	0.0520	0.2858
RMSE for $e_{65,i}$	0.0035	0.0074	0.0495	0.0595	0.2953
Sweden, females					
RMSE for $e_{0,i}$	0.0028	0.0043	0.0120	0.0378	0.3380
RMSE for $e_{50,i}$	0.0047	0.0069	0.0168	0.0503	0.3142
RMSE for $e_{65,i}$	0.0045	0.0063	0.0173	0.0558	0.3305
Switzerland, males					
RMSE for $e_{0,i}$	0.0007	0.0018	0.0098	0.0240	0.1731
RMSE for $e_{50,i}$	0.0037	0.0040	0.0143	0.0366	0.1305
RMSE for $e_{65,i}$	0.0020	0.0031	0.0175	0.0454	0.1312
Sweden, males					
RMSE for $e_{0,i}$	0.0010	0.0207	0.0564	0.1289	0.2381
RMSE for $e_{50,i}$	0.0031	0.0160	0.0420	0.0988	0.1775
RMSE for $e_{65,i}$	0.0019	0.0171	0.0420	0.0995	0.1785

Conclusions

- Forecast age-at-death distributions of not extinct cohort.
- Method that combines demographic information with data driven process.
- Computation of cohort life expectancy vs. period life expectancy.

- $y = (y_1, ..., y_I)'$ actually observed counts in I age intervals.
- y_i assumed Poisson distributed with $E(y_i) = \mu_i$.
- Let γ be the latent distribution defined on a grid of J narrow intervals (1 year age grid for the last age interval).
- Vector $\mu \in \mathbb{R}^I$ results from grouping the $J > I$ original (expected) counts γ_J into the open-ended age group: $\mu = C\gamma$
- C is the composition matrix.

- C is an indicator (0/1) matrix, such that the number of 1 in each row defines the number of elements of grouped into one age step.

$$C = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & \cdots & 1 \end{pmatrix}.$$

- To guarantee non-negative values of γ we write:

$$\mu = C\gamma = Ce^{\beta}$$