A robust approach to stochastic mortality modelling

Helena Aro
Teemu Pennanen

Institute of Mathematics, Helsinki University of Technology

AFIR/LIFE Colloquium, Munich, September 10, 2009

Introduction

- Stochastic mortality modelling: capturing the uncertainty in future development of mortality
- We propose a general discrete-time framework
  - Flexible but relatively simple
  - Incorporates population-specific characteristics and user preferences
  - Robust in calibration
  - Allows for a choice of easily interpretable risk factors
- To illustrate, two versions of the model are fitted into Finnish mortality data
- Simulations for future Finnish mortality are provided

The model

- Let $E(x, t)$ be the size of population aged x (cohort) at the beginning of year t
- Denoting the number of deaths occurring in cohort x during year t by $D(x, t)$ we have

$$D(x, t) = E(x, t) - E(x + 1, t + 1)$$

- Assume the conditional distribution of $D(x, t)$ given $E(x, t)$ is binomial:

$$D(x, t) \sim \text{Bin}(E(x, t), q(x, t))$$

  where $q(x, t)$ is the probability that an individual aged x at the beginning of year t dies before the end of the year

The model

- Probabilities $q(x,t)$ modelled as a stochastic process
- *Systematic mortality risk*: uncertainty in future values of $q(x,t)$
- $D(x,t)$ obtained by sampling from $\mathrm{Bin}(E(x,t), q(x,t))$
- In large populations, dynamics are well described by $D(x,t) = E(x,t)q(x,t)$, uncertainty in $D(x,t)$ mainly arising from the systematic mortality risk

The model

- We model the logistic probabilities by

$$\operatorname{logit} q(x,t) := \ln\left(\frac{q(x,t)}{1-q(x,t)}\right) = \sum_{i=1}^{n} w_i(t)\phi_i(x),$$

  where $\phi_i(x)$ are user-defined *basis functions* across cohorts, and $w_i(t)$ stochastic *risk factors* that vary over time
- In other words, $q(x,t) = q_{w(t)}(x)$, where $w(t) = (w_1(t), \dots, w_n(t))$, and q_w is the parametric function defined for each $w \in \mathbb{R}^n$ by

$$q_w(x) = \frac{\exp\left(\sum_{i=1}^{n} w_i\phi_i(x)\right)}{1+\exp(\sum_{i=1}^{n} w_i\phi_i(x))}$$

- Modelling the logit transforms instead of $q(x,t)$ directly guarantees that $q(x,t) \in (0,1)$.

The model

- Selection of basis functions determines several characteristics of the model
- Certain desired properties of $q(x, t)$, e.g. continuity or smoothness across cohorts, are achieved by corresponding choices of $\phi_i(x)$
  - Incorporation of user preferences and/or population-specific characteristics
- Appropriate choice of basis functions assigns interpretations to risk factors
- Concrete interpretations facilitate the modelling of risk factors, which is advantageous in assessing estimation results and applications such as engineering of mortality-linked financial instruments

The model

Example

If the basis functions are such that $\phi_k(x) = 1$ but $\phi_i(x) = 0$ for $i \neq k$ for a certain cohort x, then the risk factor $w_k(t)$ equals the logistic death probability in cohort x in year t.

Example

In the model of Cairns et al. (2006) [1],

$$\operatorname{logit} q(x,t) = \kappa_1(t) + \kappa_2(t)(x - \bar{x}),$$

where $\bar{x}$ is the mean age over the cohorts, $w_i = \kappa_i$ for $i = 1, 2$, and the basis functions are $\phi_1(x) = 1$ and $\phi_2(x) = (x - \bar{x})$.

---

[1] A.J.G.Cairns, D.Blake, K.Dowd. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk & Insurance*, 73(4), 2006

The model

- After choosing the basis functions, the vector $w(t) = (w_1(t), \dots, w_n(t))$ of risk factors in modelled as a multivariate stochastic process in discrete time.
- Model can be based on expert views, historical data, or both
- The historical values of $w(t)$ can be constructed by maximum likelihood estimation
- Given the past values of $E(x,t)$ and $D(x,t)$, the log-likelihood function for annual values of $w(t)$ can be written for each t as

$$
\begin{aligned}
l_t(w) &= \ln \prod_x \binom{E(x,t)}{D(x,t)} q_w(x)^{D(x,t)} \left(1 - q_w(x)\right)^{E(x,t)-D(x,t)} \\
&= \sum_x \left[D(x,t) \sum_i w_i \phi_i(x) - E(x,t) \ln(1 + e^{\sum_i w_i \phi_i(x)}) + \ln \binom{E(x,t)}{D(x,t)} \right].
\end{aligned}
$$

The model

- The basis functions ϕ_i are *linearly independent* on a set A of cohorts if the only vector $w \in \mathbb{R}^n$ that satisfies

$$\sum_{i=1}^{n} w_i \phi_i(x) = 0 \quad \forall x \in A$$

is the zero vector $z = 0$.
- Violation of this condition means that at least one basis function can be removed without any effect on the properties of the model

Proposition

The log-likelihood function $l_t : \mathbb{R}^n \to \mathbb{R}$ is concave. If the basis functions ϕ_i are linearly independent on the set of cohorts

$$A(t) = \{x \mid E(x,t) > 0\},$$

then l_t is strictly concave.

The model

- Concavity implies that the local maxima of l_t are 'true' ML-estimators
- Strict concavity ensures uniqueness of the estimators
- Concavity enables the use of convex programming methods in the numerical maximization of l_t
- Assuming Poisson distributed deaths would not result in a concave log-likelihood function

The model

To summarize, the modelling proceeds as follows:

1. Choose a set $\{\phi_i\}_{i=1}^n$ of basis functions that describes the features of the death probability curve $q(x,t)$.
2. Construct historical values of $w(t)$ from data using maximum likelihood estimation
3. Model the future development of $w(t)$ as a stochastic process, using its historical values and/or expert information
4. The future death probabilities are given by

$$q(x,t) = \left[1 + \exp\left(-\sum_i w_i(t)\phi_i(x)\right)\right]^{-1}$$

5. The future deaths $D(x,t)$ are obtained by sampling from $\mathrm{Bin}(E(x,t), q(x,t))$ or simply by $D(x,t) = E(x,t)q(x,t)$, if we are only interested in systematic mortality risk

Modelling Finnish mortality

- We consider the mortality of Finnish males aged 18-100 years
- Data consists of annual values of $E(x, t)$ and $D(x, t)$, covering years 1900-2007 [2]
- Two models are fitted into the data:
  - A two-parameter model with linear basis functions
  - A three-parameter model with piecewise linear basis functions

[2] Source: Human mortality database, www.mortality.org

Modelling Finnish mortality: two-parameter model

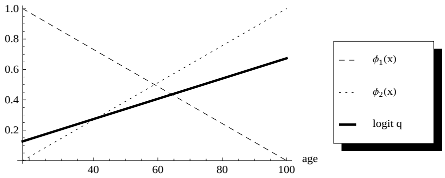
- Two-parameter model:

$$\operatorname{logit} q(x,t) = w_1(t)\phi_1(x) + w_2(t)\phi_2(x)$$

- Basis functions are linear:

$$\phi_1(x) = 1 - \frac{x-18}{82} \quad \text{and} \quad \phi_2(x) = \frac{x-18}{82}$$

- The linear combination $\operatorname{logit} q(x,t) = \sum_{i=1}^{2} w_i \phi_i(x)$ is also linear:



- Factor values for each t are two points on the line for $\operatorname{logit}\ q(x,t)$, hence having a natural interpretation

Modelling Finnish mortality: two-parameter model

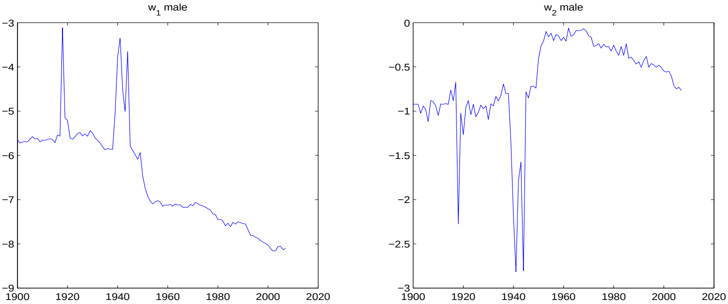


Figure: Estimated parameter values

Modelling Finnish mortality: two-parameter model

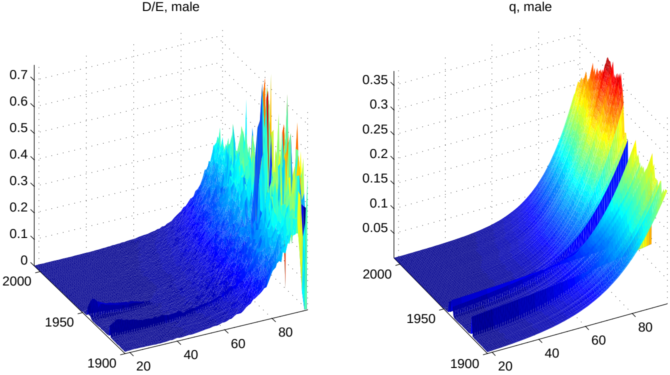


Figure: $\frac{D(x,t)}{E(x,t)}$ vs. estimated values of $q(x,t)$

Modelling Finnish mortality: three-parameter model

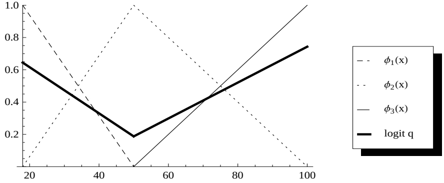
- In the three-parameter model

$$\text{logit}\ \ q(x,t) = w_1(t)\phi_1(x) + w_2(t)\phi_2(x) + w_3(t)\phi_3(x),$$

where the basis functions are piecewise linear:

$$\phi_1(x) = 1 - \frac{x-18}{32} \quad , \quad \phi_2(x) = \begin{cases} \frac{1}{32}x & x \le 50 \\ 2 - \frac{x}{50} & x > 50 \end{cases} \text{ and } \quad \phi_3(x) = \frac{x}{50} + 1.$$

- The linear combination now also piecewise linear:



- The factor values for each t again points on the line for $\text{logit}\ \ q(x,t)$

Modelling Finnish mortality: three-parameter model

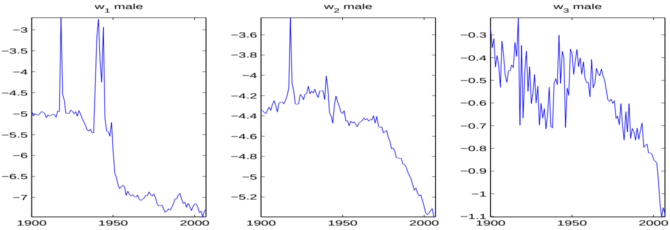


Figure: Estimated parameter values

Modelling Finnish mortality: three-parameter model

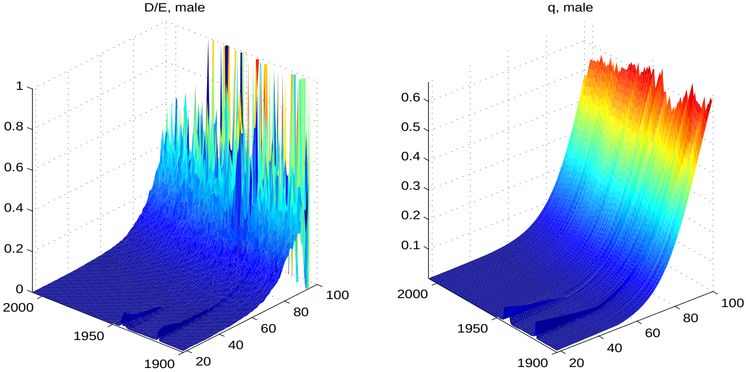


Figure: $\frac{D(x,t)}{E(x,t)}$ vs. estimated values of $q(x, t)$

Modelling the process w

- Estimated values for $w(t)$ in both models display drifting behaviour from year 1960 onwards
- Two- and three-dimensional random walks with a drift are fitted into the estimated values of $w(t)$ for the years 1960-2007
- General n-dimensional random walk with a drift is of the form

$$\Delta w(t) := w(t) - w(t-1) = \mu + CZ(t),$$

where $\mu \in \mathbb{R}^{n \times 1}$ is obtained as the mean vector of $\Delta w(t)$, $C \in \mathbb{R}^{n \times n}$ from the Cholesky decomposition of their covariance matrix $V = CC'$, and $Z(t)$ is a $n \times 1$ vector of independent standard Gaussian random variables: $Z_i(t) \sim N(0, 1)$

Simulations

- Death probabilities and cohort sizes $q(x,t)$ for Finnish males were simulated for 30 years into the future, using the two mortality models
- Process w was simulated with the Monte Carlo method (sample size 10000), and probabilities $q(x,t)$ were calculated as
$$q(x,t) = \left[1 + \exp\left(-\sum_i w_i(t)\phi_i(x)\right)\right]^{-1}$$
- The number of deaths $D(x,t)$ in each cohort was approximated by its expected value $E(x,t)q(x,t)$
- Cohorts aged 30 and 65 in the final observation year 2007 were chosen as reference cohorts

Simulations: two-parameter model

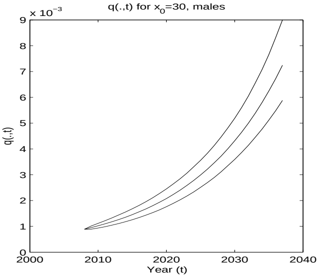


Figure: Medians and 90% confidence intervals for mortality rates $q(., t)$ and cohort sizes $E(., t)$. Cohort aged 30 in 2007.

Simulations: two-parameter model

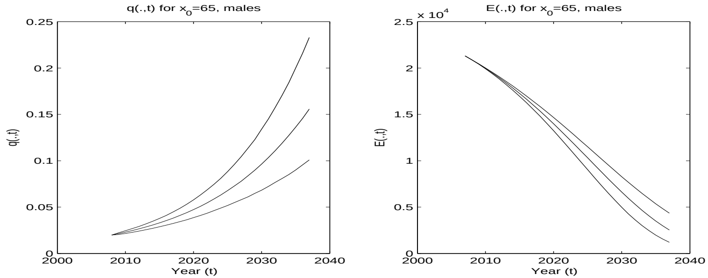
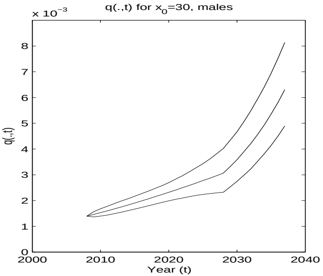


Figure: Medians and 90% confidence intervals for mortality rates $q(., t)$ and cohort sizes $E(., t)$. Cohort aged 65 in 2007

Simulations: three-parameter model



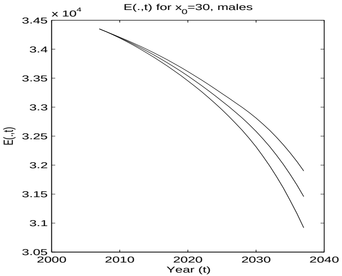


Figure: Medians and 90% confidence intervals for mortality rates $q(., t)$ and cohort sizes $E(., t)$. Cohort aged 30 in 2007.

Simulations: three-parameter model

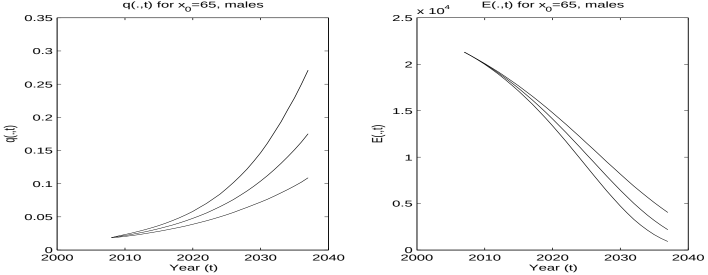


Figure: Medians and 90% confidence intervals for mortality rates $q(., t)$ and cohort sizes $E(., t)$. Cohort aged 65 in 2007.

Thank you!