A matter of life and death: mortality estimation and prediction

Jon Forster
Department of Statistics
University of Warwick, UK

Joint work with Jakub Bijak, Erengul Dodd, Jason Hilton and Peter Smith

Why mortality matters

Mortality estimates forecasts are vitally important

- Tax and Expenditure
- Healthcare
- Pensions and Insurance

- **Goal**: Forecast future mortality with a realistic quantification of uncertainty.

The life table

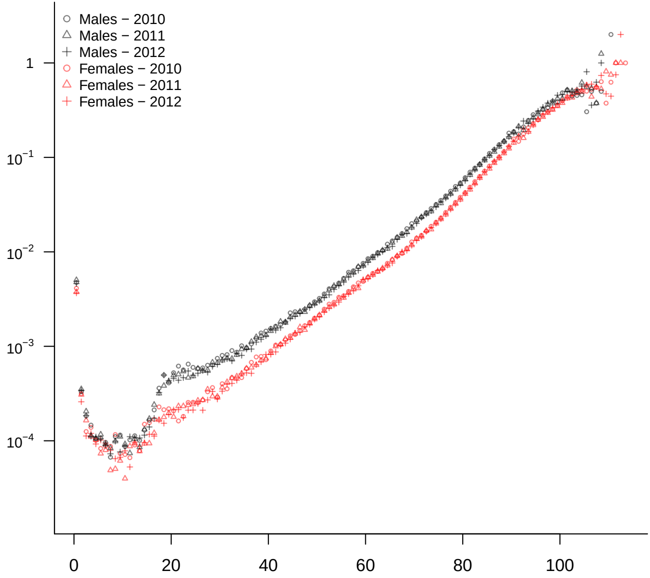
A static summary of the distribution of age at death for a population:

English Life Tables No 17

Period expectation of life
Based on data for England and Wales for the years 2010-2012

Age				Males				
x	m_x	q_x	l_x	d_x	L_x	T_x	μ_x	e_x
0	0.004757	0.004746	100000	475	99576.3	7896837		78.97
1	0.000306	0.000306	99525	30	99510.2	7797072	0.000369	78.34
2	0.000207	0.000207	99495	21	99484.6	7697562	0.000246	77.37
3	0.000147	0.000147	99474	14	99467.0	7598078	0.000172	76.38
4	0.000115	0.000115	99460	12	99453.9	7498612	0.000128	75.39
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
109	0.676172	0.491440	8	4	5.5	11	0.661588	1.43
110	0.701065	0.503943	4	2	2.8	5	0.685990	1.39
111	0.725677	0.516003	2	1	1.4	3	0.709972	1.34
112	0.750015	0.528125	1	1	0.7	1	0.733841	1.30

Crude mortality rates 2010-2012



A basic smoothing model

We estimate the *Central Mortality Rate*

$$m_x \equiv \frac{\text{Expected number of deaths aged x in 1 year}}{\text{Population aged x at risk}}$$

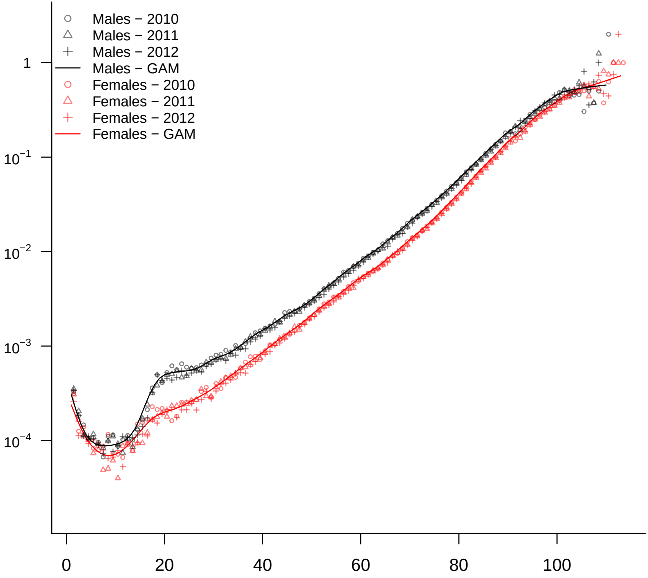
In a generalised additive (smooth) statistical model, we estimate the m_x by

$$\hat{m}_x = \exp s(x)$$

where s is a function which 'trades off' smoothness as a function of age with how closely the $\hat{m}_x$ are to the corresponding observed rates

$$\frac{\text{Observed number of deaths aged x in 1 year}}{\text{Population aged x at risk}}$$

Smooth mortality rates 2010-2012



Models for older ages and extrapolation

To obtain a more robust fit at older ages, and to extrapolate the mortality function m_x beyond the range of the observed data, one might use the log-linear Gompertz model

$$\log m_x = \beta_0 + \beta_1 x, \qquad x \geq x_0$$

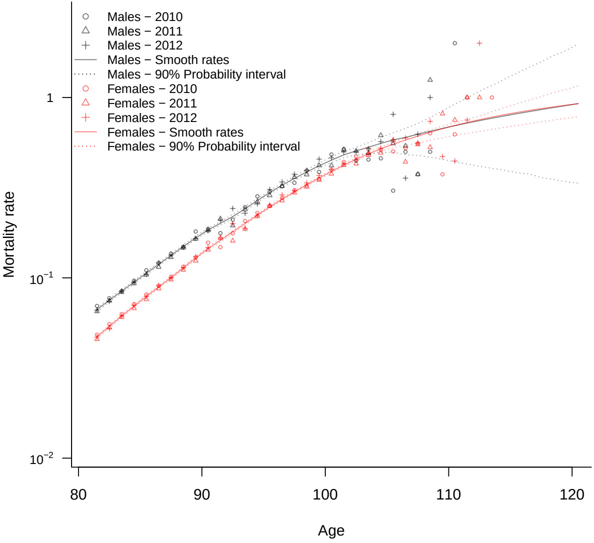
where x_0 is a suitable threshold

or

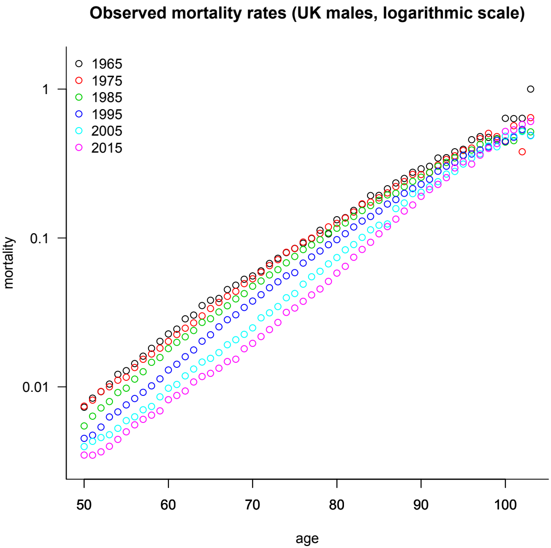
$$m_x = \frac{\beta_2 \exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}, \qquad x \geq x_0$$

where mortality rates flatten off, converging to the limit β_2 as $x \to \infty$.

ELT17 modelling at high ages



Mortality improvement



Dynamic models and projection

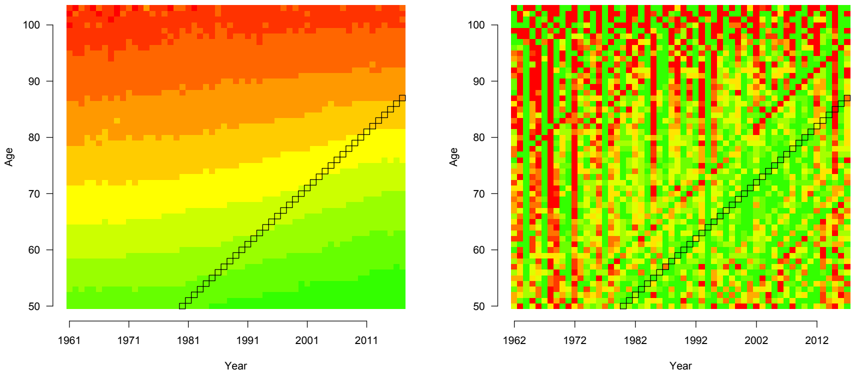
Now mortality varies not just by age, but over time as well.

We denote by m_{xt} the central mortality rate aged x, in year t, in population of interest, for $t = 1, \ldots, T =$present.

Statistical mortality models provide a framework for projecting m_{xt} etc for $t = T + 1, T + 2, \ldots$

Hence providing us with the information we need for planning.

The three time dimensions



A cohort is a subpopulation sharing a common birth-year. (1930 birth cohort identified above)

The model

Age-period-cohort (APC) GAM for mortality improvements

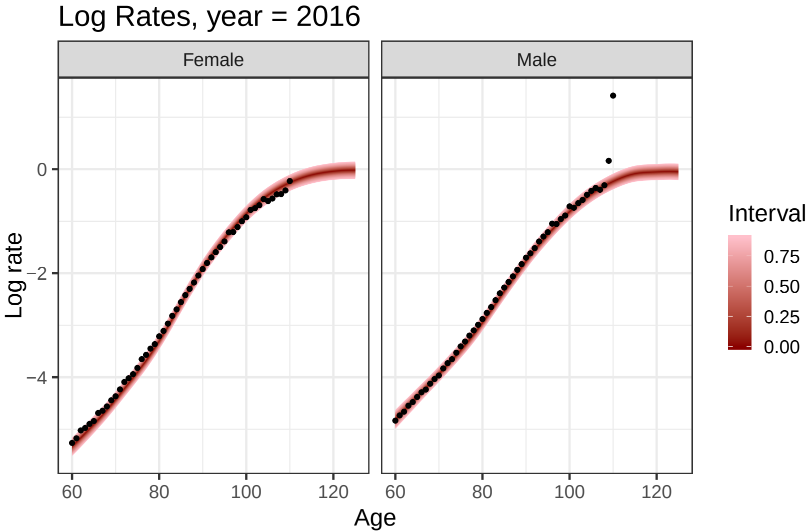
$$\log \frac{m_{xt}}{m_{x\,t-1}} = s_\alpha(x) + \kappa_t + s_\gamma(t-x)$$

For the highest ages $x > x_0$, use the structured model

$$m_{xt} = \frac{\beta \exp\left(\mu_0 + \mu_1 x + (\alpha_0 + \alpha_1 x)t\right)}{1 + \exp\left(\mu_0 + \mu_1 x + (\alpha_0 + \alpha_1 x)t\right)} \exp\left(\kappa_t + s_\gamma(t-x)\right)$$

Forecast (with uncertainty)

Fit on data up to 2006, 10 year projection.



Forecast life expectancy (with uncertainty)

Fit on data up to 2006.