Estimating and modelling mortality rates in the absence of population denominators

Andrés M. Villegas and Munir Hiabu

UNSW | Business School

AFRIC, Victorial Falls, Zimbabwe
Tuesday 25th July 2023

Agenda

Motivation

Calculating the (forward) mortality rate (the usual way – if population data is available)

The reversed mortality rate

- From reversed mortality rate to forward mortality rate
- Modelling the reversed mortality rate

Illustration using HMD data for England and Wales males

Conclusion

Motivation

Mortality rate (heuristic):

$$\frac{\#\ \text{occurred deaths}}{\text{population size}}$$

Problem: Denominator often poor quality or not known at all.

- Developing countries
- Subpopulations
- Old ages

Motivation - Quality of population data is sometimes doubtful

Colombia's population was overestimated by 5 million: Instead of the projected 50 millions population expected in 2018 in the Census 2005 projections, the population in 2018 was 45.5 million

Calculating the mortality rate (the usual way – if population data is available)

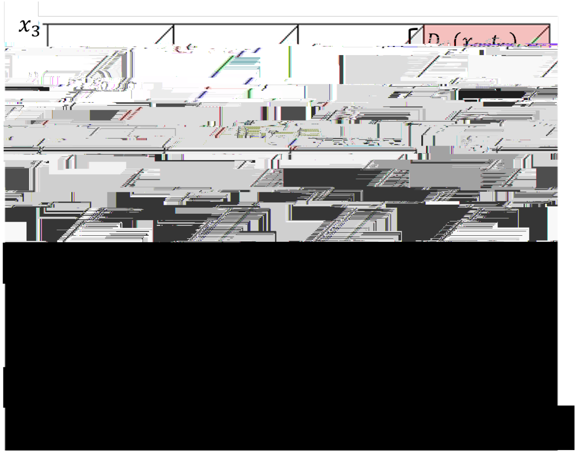
Calculating the mortality rate (the usual way if populati data is available)

We are interested in estimating the (forward) mortality rate

$$\mu(x|t) = \lim_{h \downarrow 0} h^{-1} \Pr\{\underbrace{X \in [x, x+h) | X \geq x}_{\text{Die in the next instant given survival to age } x}, T - X = t - x\}.$$

X is age of death
T is date (also called period) of death
$C = T - X$ is cohort; known before death

Calculating the mortality rate – The Lexis diagram



$$D_U(x_j, t_k) = \sum_i \mathbb{1}\{X_i \in [x_j, x_{j+1}), T_i - X_i \in [t_k - x_{j+1}, t_k - x_j)\}$$

$$D_L(x_j, t_k) = \sum_i \mathbb{1}\{X_i \in [x_j, x_{j+1}), T_i - X_i \in [t_k - x_j, t_{k+1} - x_j)\}$$

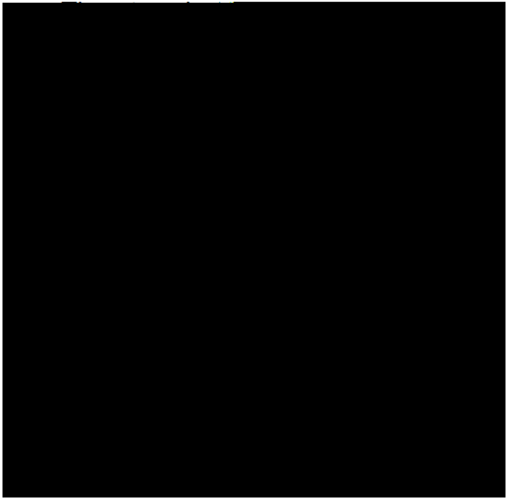
$$P(x_j, t_k) = \sum_i \mathbb{1}(T_i > t_k, T_i - X_i \in [t_k - x_{j+1}, t_k - x_j))$$

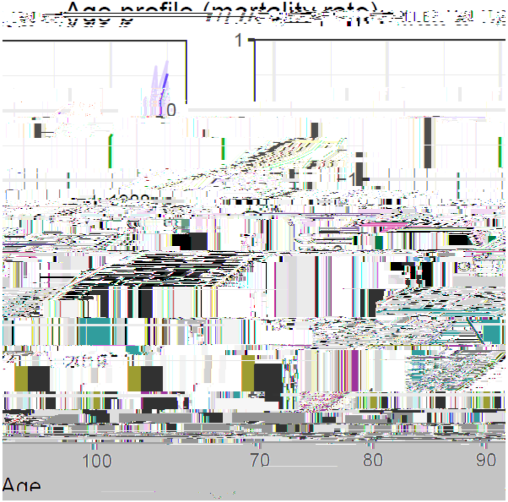
Calculating the mortality rate (the usual way – if population data is available) – central mortality rate

The central mortality rate is defined as

$$m(x_j | t_k) = \frac{D(x_j; t_k)}{E(x_j; t_k)} = \frac{D(x_j; t_k)}{\frac{1}{2}\{P(x_j; t_k) + P(x_j; t_{k+1})\} + \frac{1}{3}\{D_L(x_j; t_k)\}}$$

The central mortality rate (England and Wales males)





The reveresed mortality rate

We aim to estimate

$$^{R}(x|t) = \lim_{h \downarrow 0} h^{-1} \Pr\{ \underbrace{X \in (x \quad h; x] | X \quad x}_{\text{Die in the previous instant given dead by agex}}; C = c\}.$$

The reversed central mortality rate is given as

$$m^{R}(x_j|t_k) = \frac{D(x_j; t_k)}{E^{R}(x_j; t_k)}.$$

The deaths counts $D(x_j; t_k)$ are the same as before

Now: <u>How to calculate $E^{R}(x_j; t_k)$</u> ?

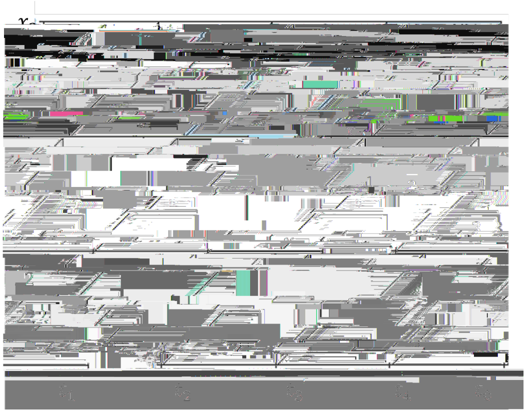
Exposed to risk (Under the assumption of closed population)

Forward rate: The number of people at risk of dying in next instant is the number of all future deaths. But this number is only known for extinct cohorts.

Reversed rate: The number of people at risk of having died in the previous instant is the number of all people who have already died. **This number can be counted from death data**.

The reversed mortality rate – how to calculate $E^R(x_j;t_k)$

The figures below show the weights of the deaths when calculating the exposure.



$$\hat{E}^R(x_j;t_k) = \frac{1}{3}\{D_U(x_j;t_k) + D_L(x_j;t_k)\}$$
$$+\frac{1}{3}\sum_{l=0}^{j} D_L(x_{j-l};t_{k-l-1}) + D_U(x_{j-l-1};t_{k-l})$$
$$+\frac{2}{3}\sum_{l=1}^{j} D_U(x_{j-l};t_{k-l}) + D_L(x_{j-l};t_{k-l})$$

The reversed mortality rate

Under appropriate assumption,

$$m^{R}(x_j j t_k) = \frac{D(x_j ; t_k)}{E^{R}(x_j ; t_k)} ;$$

is an unbiased estimator of the expected value of $^{R}(X j T)$ for $(X ; T)$ conditioned on the square $[x_j ; x_j + 1)$ $[t_k ; t_k + 1)$.

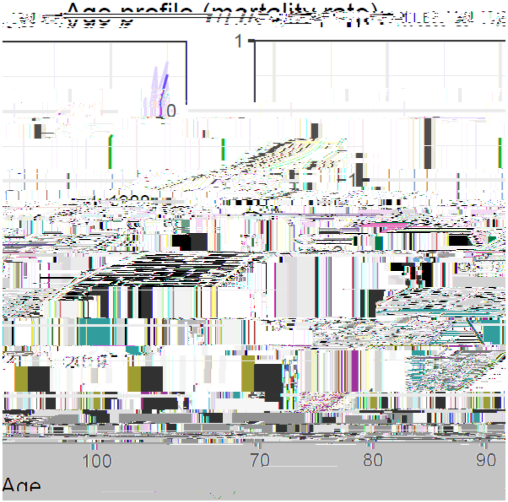
Is the reversed mortality rate useful?

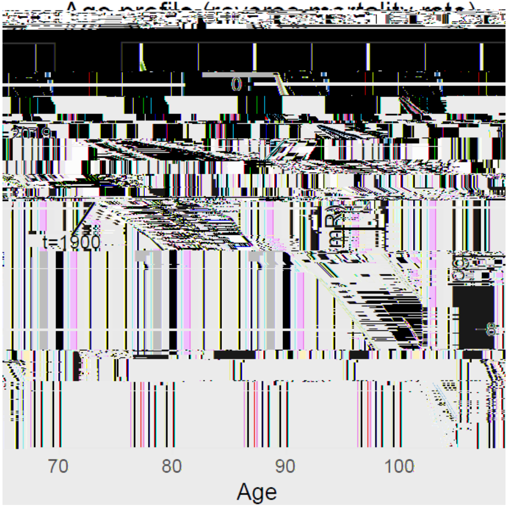
1. The reversed mortality rate can be interesting in its own right.
2. We can use the reversed mortality rate to estimate the forward mortality rate.
3. Modelling the reversed rate can give a new perspective on mortality forecasting.

Reversed mortality rate can be interesting in its own right

Reversed mortality rate can be interesting in its own right

Age pro le

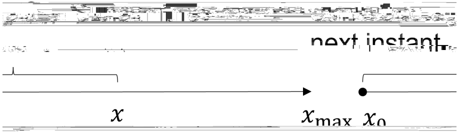




We can use the reversed mortality rate to estimate the forward mortality rate.

Relationship between reversed mortality rate and forward mortality rate

Forward time



$$(x) = \lim_{h\#0} h^{1} \mathrm{Prf}\, X \; 2 \; [x; x + h)g$$

$$f(x) = S(x)_{R} \;\; (x)$$

$$S(x) = e$$

From reversed mortality rate to forward mortality rate

We have then

$$(x|t) = \; {}^{R}(x|t)\frac{e^{\int_{x}^{x_{max}} {}^{R}(v|t \;\; x+v)dv}}{1 \;\; e^{\int_{x}^{x_{max}} {}^{R}(v|t \;\; x+v)dv}};$$

where

$$\frac{e^{\int_{x}^{x_{max}} {}^{R}(v|t \;\; x+v)dv}}{1 \;\; e^{\int_{x}^{x_{max}} {}^{R}(v|t \;\; x+v)dv}} = \frac{\text{Probability of dying before } x}{\text{Probability of dying after } x}$$

Problem: The integral runs over **unobserved ages for non-extinct cohorts.**

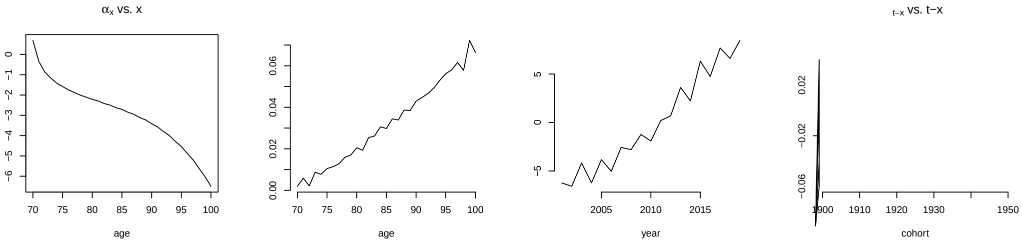
Solution: **Extrapolate the reversed mortality rate** to complete data for non-extinct cohort

From reversed mortality rate to forward mortality rate

$$q^R(x_l) = \frac{2m^R(x_l)}{2 + m^R(x_l)}; \quad m^F(x_j) = m^R(x_j)\frac{\prod_{l=j}^{J}\{1 - q^R(x_l)\}}{1 - \prod_{l=j}^{J}\{1 - q^R(x_l)\}}$$

Lee-Carter+Cohorts in Reverse

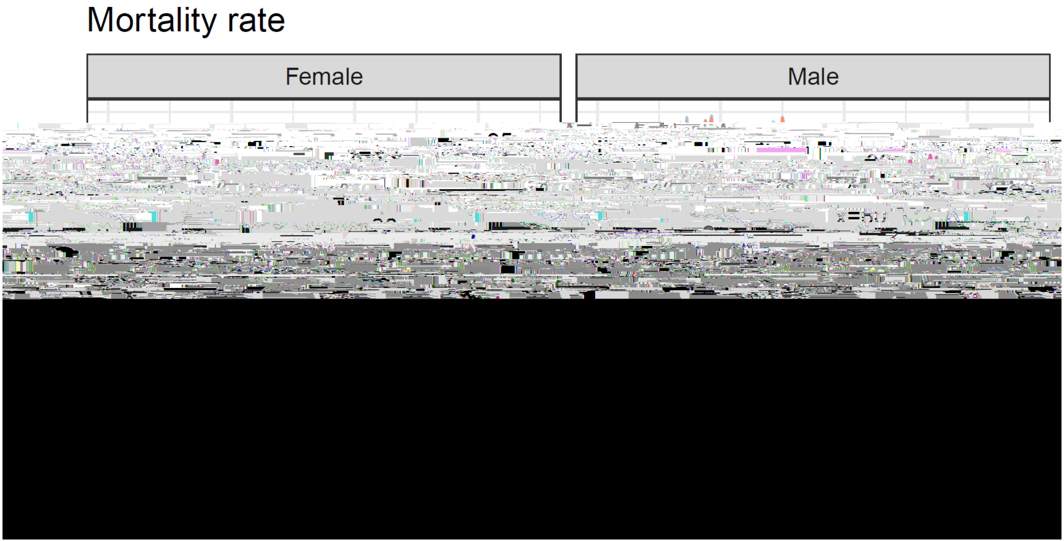
$$\log m^R(xjt) = \quad_x + \quad_x^{(1)} \quad_t^{(1)} + \quad_t \quad_x$$



Gompertz model in Reverse

How does the traditional central mortality rate $m(x_j; t_k)$ which uses population data, compare to $m^F(x_j; t_k)$ which only uses death counts?

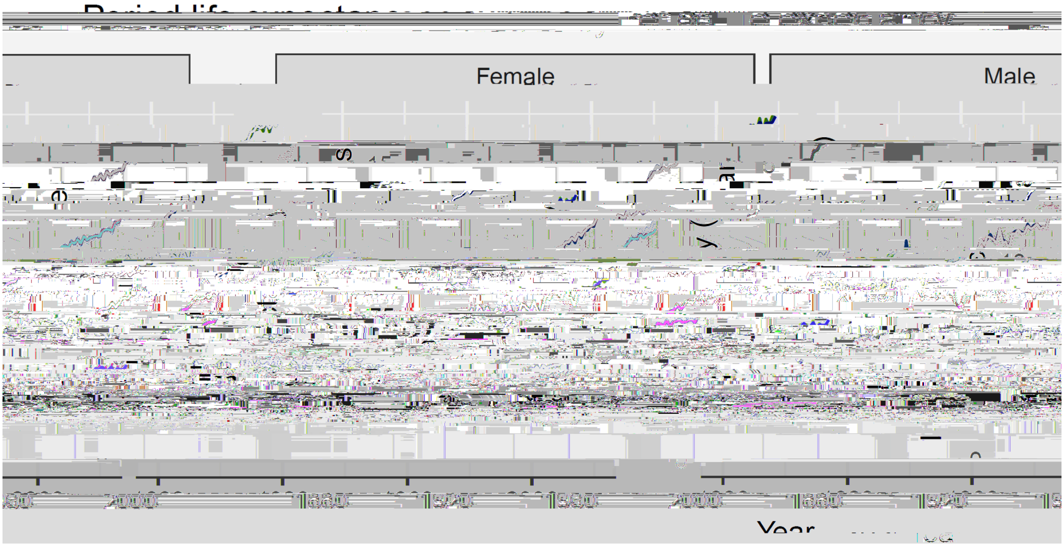
Mortality rates - England and Wales



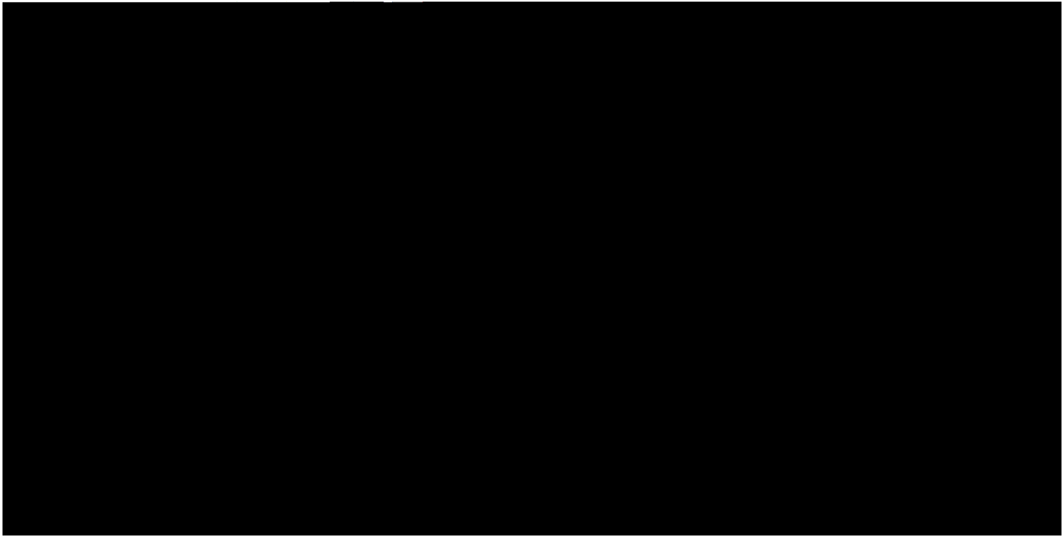
Mortality rates - England and Wales

Mean Absolute Percentage Error

Period Life Expectancy - England and Wales



Cohort Life Expectancy - England and Wales



Life Expectancy - England and Wales

Percentage Error

Conclusion

The actual size of the population of interest, if available at all, can often be poor quality

Propose a way to estimate mortality rates by using death counts only

The propose approach is reasonably accurate

- Good t of rates along both period and cohort
- Good estimates and projections of life expectancies

Useful new perspective for projection of mortality at older ages

- Explore out-of-sample forecast accuracy
- Check consistency of projections using population sizes
- Add diversity of projections ! model ensembles

Thank you!

a.villegas@unsw.edu.au