Using the HMD to assess longevity risk

Andres Barajas Paz, Andrew Cairns, Torsten Kleinow

Heriot-Watt University, Edinburgh

ab108@hw.ac.uk

House of Finance Days 2019 - HMD users conference,
Université Paris-Dauphine, March, 2019

Overview

- Motivation
- Main Objective
- MLE and Bayesian approaches
- Model and Notation
- Results
- Conclusions
- Forthcoming research

Overview on risk assessment

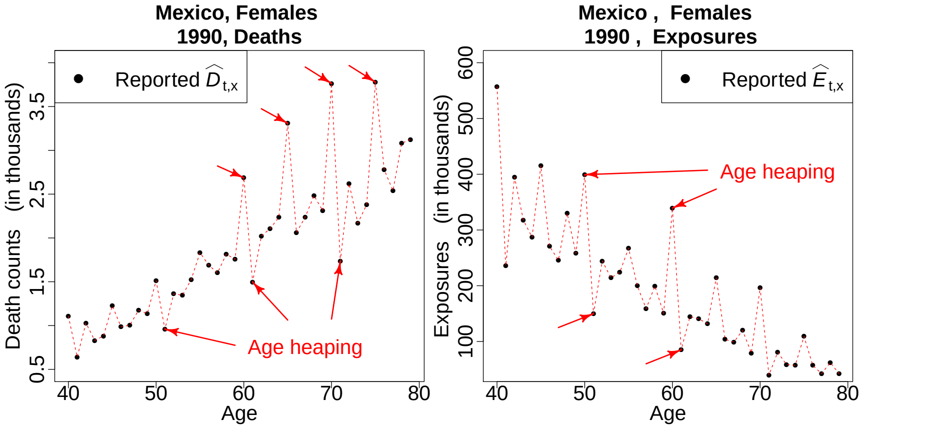
True force of mortality $\rightarrow$ Key in actuarial work.

- Underestimating mortality $\rightarrow$ life insurance companies have to pay the benefits sooner than expected.
- A major risk in pension plans is longevity risk $\rightarrow$ paying the benefits longer than expected.

HMD data quality → Key to giving us confidence in the risk assessment.

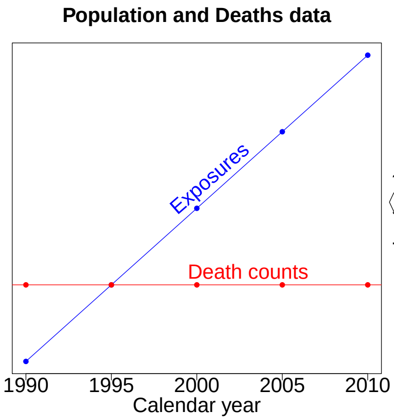
- The Dutch Actuarial Society → HMD for mortality projections.
- The Canadian Institute of Actuaries → HMD for mortality improvements.
- IFoA and SOA → HMD for mortality analyses.

Motivation.

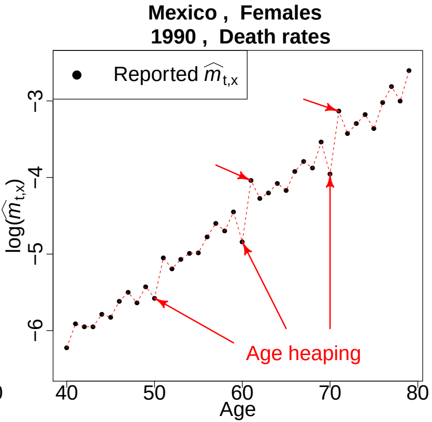


Age Heaping occurs when people misreport age.

Motivation.



Population and Deaths data
Exposures
Death counts
1990
1995
2000
2005
2010
Calendar year



Mexico , Females
1990 , Death rates
Reported $\widehat{m}_{t,x}$
$\log(\widehat{m}_{t,x})$
−3
−4
−5
−6
Age heaping
40
50
60
70
80
Age

Motivation.

- Mortality analyses $\xrightarrow{\text{Good quality mortality data}}$ HMD.

However, in many other countries population and deaths data can be somewhat unreliable.

- Population $\xrightarrow{\text{Misreporting of age}}$ census.
- Deaths $\xrightarrow{\text{Misreporting of age}}$ deaths data.

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Main Objective

- Develop mortality models for countries where their population data is affected by age heaping.

Application: Reported data $\rightarrow$ Smoothed HMD $\rightarrow$ International Reinsurance.

MLE and Bayesian approaches

We design a model taking into account two dimensional data. Hence, we consider the data by age x and across calendar years $t \in \mathcal{T}$ for deaths, and census years $t \in \mathcal{T}_c$.

- $\mathcal{T} = \{1990, \cdots, 2010\}$,
- $\mathcal{T}_c = \{1990, 1995, 2000, 2005, 2010\}$.

For any calendar year t we denote by $|t|$ the number of years, that is, $n_t = |t|$ is the length of the years t in our data set. Where x represents the reported age and y the true age.

$$E_{t,y} \xrightarrow{\text{Age heaping}} \widehat{E}_{t,x}$$

$$D_{t,y} \xrightarrow{\text{Age heaping}} \widehat{D}_{t,x}$$

Model and Notation

$$\widehat{D}(t,x)\Big|\underline{m},\underline{E},\theta \sim Poisson\Big(\sum_y m(t,y)E(t,y)g^D(t,x,y)\Big),\ \forall t \in \mathcal{T}$$

$$\widehat{E}(t,x)\Big|\underline{E},\theta \sim Poisson\Big(\sum_y E(t,y)g^E(t,x,y)\Big),\ \forall t \in \mathcal{T}_c.$$

$$m(t,y) = \exp\left[a(t) + b(t)\big(y-\bar{y}\big) + c(t)\Big(\big(y-\bar{y}\big)^2 - \sigma_y^2\Big)\right].$$

log-likelihood function

$$\ell_E(\theta) = \sum_x \sum_{t \in \mathcal{T}_c} \Big[\widehat{E}(t,x) \log \Big(\sum_y E(t,y) g^E(t,x,y) \Big) - \sum_y E(t,y) g^E(t,x,y) \Big] + \text{Const}$$

$g^* \rightarrow$ Probability for an individual aged y to report age x,

$\eta^* \rightarrow$ Captures the increasing improvements across years,

$H^* \rightarrow$ Describes what ages are "popular".

Bayesian approach

Where,

$$\theta = \{\underline{E}, \underline{\eta}^D, \underline{\eta}^E, \underline{H}^D, \underline{H}^E, \underline{a}, \underline{b}, \underline{c}, \delta_A, \mu_b, \mu_c\}$$

Prior distributions

$$E(t, y) \sim U(0, \infty) \textbf{ iid},$$

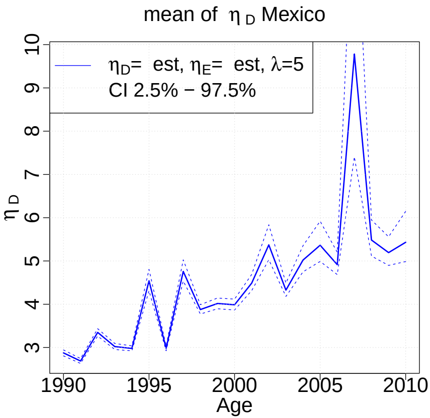
$$\eta^D(t) \sim \textbf{Exp}\left(1/2\right) \textbf{ iid}, \qquad \eta^E(t) \sim \textbf{Exp}\left(1/2\right) \textbf{ iid},$$

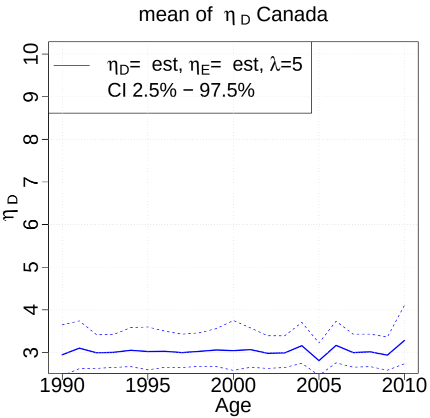
$$\left|H^D(x)\right| \sim \textbf{Exp}\left(\lambda\right) \textbf{ iid}, \qquad \left|H^E(x)\right| \sim \textbf{Exp}\left(\lambda\right) \textbf{ iid}.$$

Full log posterior $\log(\pi(\theta))$

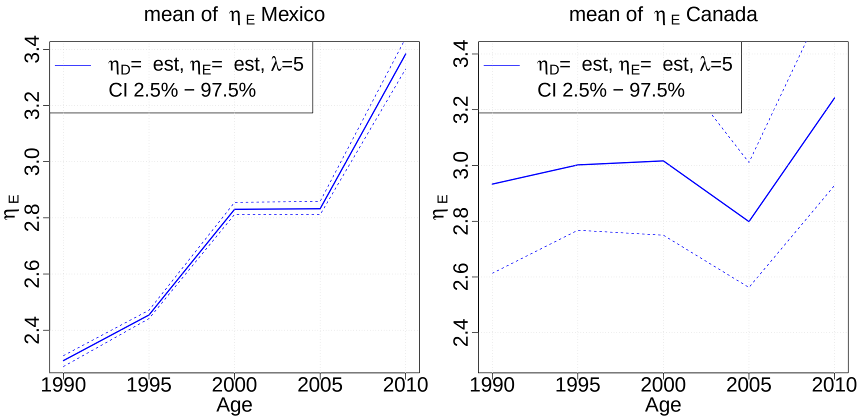
$$
\begin{aligned}
&\propto \sum_{x}\sum_{t\in\mathcal{T}}\Bigg[\widehat{D}(t,x)\log\Big(\sum_{y} m(t,y)E(t,y)g^{D}(t,x,y)\Big)\\
&-\sum_{y} m(t,y)E(t,y)g^{D}(t,x,y)\Bigg]\\
&+\sum_{x}\sum_{t\in\mathcal{T}_c}\Bigg[\widehat{E}(t,x)\log\Big(\sum_{y} E(t,y)g^{E}(t,x,y)\Big)-\sum_{y} E(t,y)g^{E}(t,x,y)\Bigg]\\
&-\Bigg[\sum_{t\in\mathcal{T}}\frac{1}{2}\eta^{D}(t)\Bigg]-\Bigg[\sum_{t\in\mathcal{T}_c}\frac{1}{2}\eta^{E}(t)\Bigg]-\Bigg[\sum_{x}\lambda\Big|H^{D}(x)\Big|\Bigg]-\Bigg[\sum_{x}\lambda\Big|H^{E}(x)\Big|\Bigg]
\end{aligned}
$$

Parameters η^D M-H, Mexico & Canada.

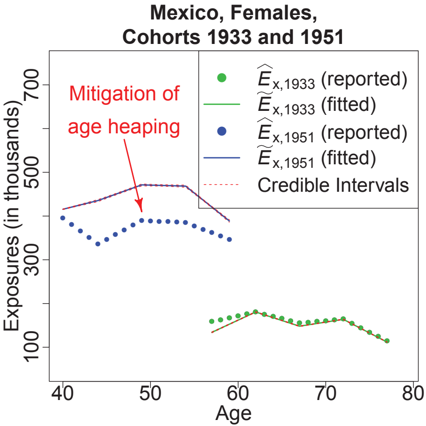


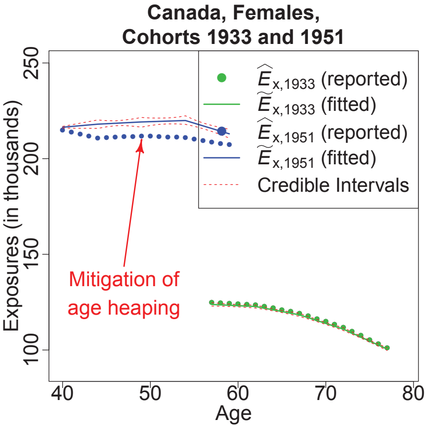


Parameters η^E M-H, Mexico & Canada.

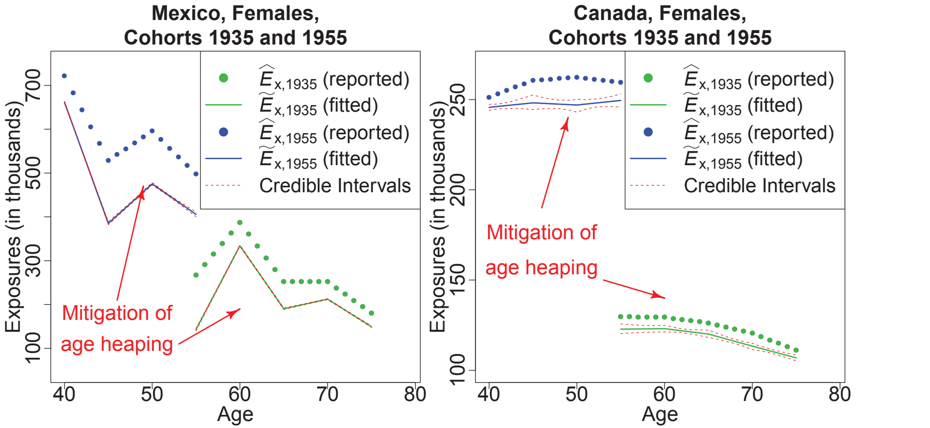


Fitted exposures by Cohort, Mexico & Canada.

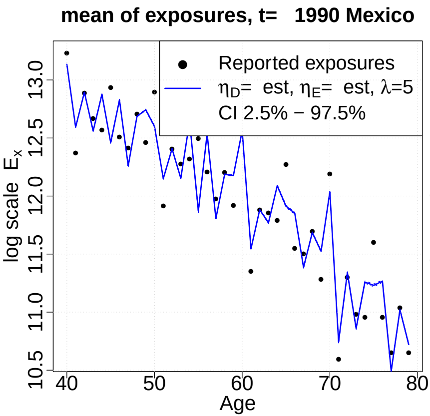


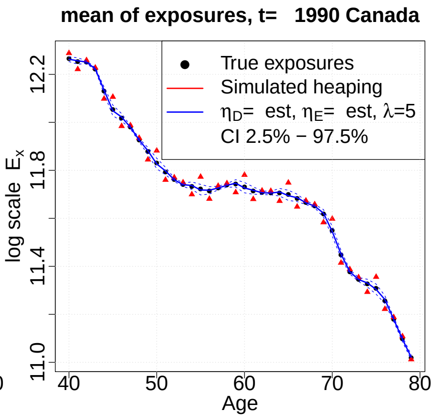


Fitted exposures by Cohort, Mexico & Canada.



Fitted exposures $E_{t,y}$, Mexico & Canada 1990.





Conclusions.

- Smooth time series $\underline{a}$, $\underline{b}$ and $\underline{c}$ $\xrightarrow{\text{Reduce age heaping}}$ $m(t,y)$ and population. However, we do not want to smooth too much because it would destroy the natural volatility from the data.
- This model improves the quality of the Mexican data by reducing age heaping across all calendar years.
- The remaining volatility in the fitted exposures comes from the death counts.

Conclusions

- Parameter η^* reflects the improvement in the quality of the data over time. Hence, we expect η^* to increase over time. In other words, there is less age heaping that there use to be, say in 1990.
- We will collaborate with HMD to see how their approach can be adapted to Mexican data for producing complete life table series, which is also relevant to international reinsurance.

Forthcoming research

- Include priors for parameters regarding the true force of mortality.
- Sensitivity of choosing different priors.
- We will collaborate with HMD to see how their approach can be adapted to Mexican data for producing complete life table series, which is also relevant to international reinsurance.

Thank You!

Questions?

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries