Yorkshire Actuarial Society

Agenda

14/03/2019

16:00 – 16:50	Machine Learning at Aviva – Steven Bacon
16:55 – 17:40	Yes, we CANN! – Mario Wüthrich
17:45 – 18:05	Tea & Coffee
18:05 – 19:00	Data Science meets Mortality Projection – Gareth Peters
19:00 – 20:00	Food & Drink

Aviva Public

Machine Learning at Aviva

Steven Bacon PhD

YAS 03.2019

Unfortunately Quantum declined YAS' request to share their slide content.

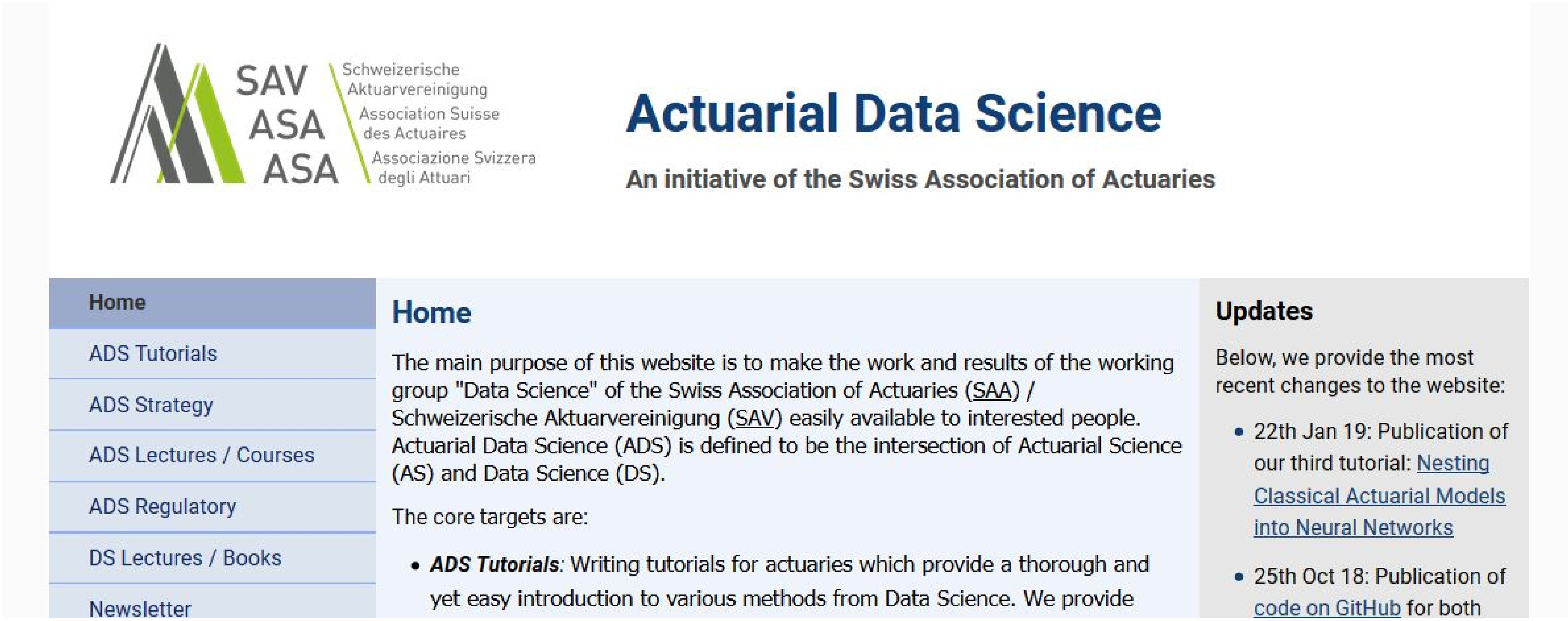
Yes, we CANN!

Mario V. Wüthrich
RiskLab, ETH Zurich

Yorkshire Actuarial Society
Aviva Offices, York, March 14, 2019

Overview

- **Yes, we CANN!** Editorial of ASTIN Bulletin 49/1, 2019
- **Actuarial Data Science** Initiative of the Swiss Association of Actuaries:
  - ⋆ Case study: French motor third-party liability claims, SSRN 2018
  - ⋆ Insights from inside neural networks, SSRN 2018
  - ⋆ Nesting classical actuarial models into neural networks, SSRN 2019
- **Data Analytics for Non-Life Insurance Pricing**
  Lecture Notes, ETH Zurich, SSRN 2019
- **Insurance Data Science Conference**, June 14, 2019, ETH Zurich
  `www.insurancedatascience.org`


SAV ASA ASA
Schweizerische Aktuarvereinigung
Association Suisse des Actuaires
Associazione Svizzera degli Attuari

Actuarial Data Science

An initiative of the Swiss Association of Actuaries

- **Home**
- ADS Tutorials
- ADS Strategy
- ADS Lectures / Courses
- ADS Regulatory
- DS Lectures / Books
- Newsletter

Home

The main purpose of this website is to make the work and results of the working group "Data Science" of the Swiss Association of Actuaries (SAA) / Schweizerische Aktuarvereinigung (SAV) easily available to interested people. Actuarial Data Science (ADS) is defined to be the intersection of Actuarial Science (AS) and Data Science (DS).

The core targets are:

- ***ADS Tutorials***: Writing tutorials for actuaries which provide a thorough and yet easy introduction to various methods from Data Science. We provide

Updates

Below, we provide the most recent changes to the website:

- 22th Jan 19: Publication of our third tutorial: Nesting Classical Actuarial Models into Neural Networks
- 25th Oct 18: Publication of code on GitHub for both

https://actuarialdatascience.org/

Yes, we CANN!

Actuarial pricing problem

- Determine from data $\mathcal{D} = ((Y_1, \boldsymbol{x}_1), \ldots, (Y_n, \boldsymbol{x}_n))$ an unknown regression function

$$\boldsymbol{x} \;\mapsto\; \mu(\boldsymbol{x}) = \mathbb{E}[Y].$$

- Variable $\boldsymbol{x}$ describes the covariates, for instance, in car insurance:
  - ⋆ **driver**: age, gender, nationality, size of household, marital status, date of driving test, occupation, medical conditions, credit record, etc.
  - ⋆ **car**: type, brand, size, weight, horse power, type of engine, cubic capacity, price, equipment, number of seats, age of car, leasing, etc.
  - ⋆ **contract**: type, duration, sales channel, deductible, other products, etc.
  - ⋆ **geographic**: province, zip code, city-rural area, type of flat, garage, etc.
  - ⋆ **driving**: annual distance, vehicle use, bonus level, claims experience, etc.

- Random variable $Y = Y_{\boldsymbol{x}}$ describes the claim.

Classical actuarial regression modeling

- Determine from data $\mathcal{D} = ((Y_1, \boldsymbol{x}_1), \ldots, (Y_n, \boldsymbol{x}_n))$ an unknown regression function

$$\boldsymbol{x} \;\mapsto\; \mu(\boldsymbol{x}) = \mathbb{E}[Y].$$

- Example of a generalized linear model (GLM):

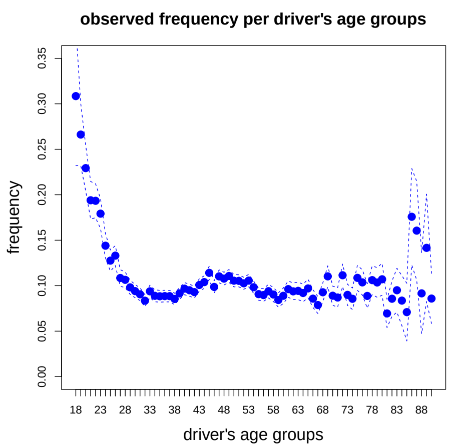
$$\boldsymbol{x} \;\mapsto\; \mu_{\boldsymbol{\beta}}^{\mathrm{GLM}}(\boldsymbol{x}) \;=\; \exp\langle \boldsymbol{\beta}, \boldsymbol{x}\rangle \;=\; \exp\left\{\sum_j \beta_j x_j\right\}.$$

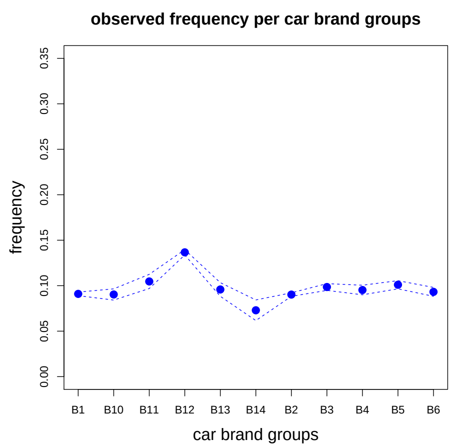
- Estimate the regression parameter $\boldsymbol{\beta}$ with maximum likelihood estimator $\widehat{\boldsymbol{\beta}}^{\mathrm{MLE}}$, received by minimizing the corresponding deviance loss

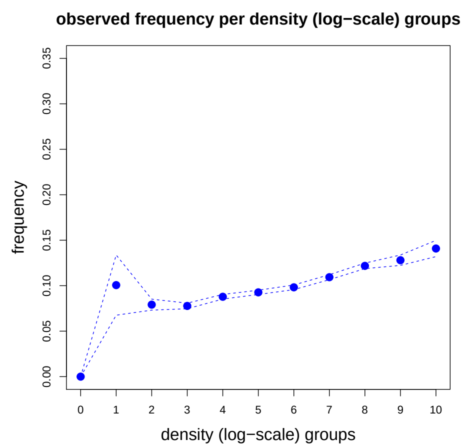
$$\boldsymbol{\beta} \;\mapsto\; \mathcal{L}_{\mathcal{D}}(\boldsymbol{\beta}).$$

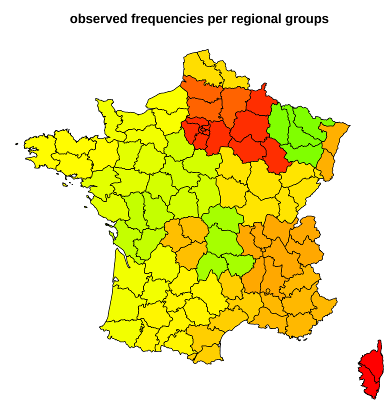
Example: car insurance frequencies

```
> str(freMTPL2freq)        #source R package CASdatasets
'data.frame':    678013 obs. of  12 variables:
 $ IDpol      : num  1 3 5 10 11 13 15 17 18 21 ...
 $ ClaimNb    : num 1 1 1 1 1 1 1 1 1 1 ...
 $ Exposure   : num  0.1 0.77 0.75 0.09 0.84 0.52 0.45 0.27 0.71 0.15 ...
 $ Area       : Factor w/ 6 levels "A","B","C","D",..: 4 4 2 2 2 5 5 3 3 2 ...
 $ VehPower   : int  5 5 6 7 7 6 6 7 7 7 ...
 $ VehAge     : int  0 0 2 0 0 2 2 0 0 0 ...
 $ DrivAge    : int  55 55 52 46 46 38 38 33 33 41 ...
 $ BonusMalus: int  50 50 50 50 50 50 50 68 68 50 ...
 $ VehBrand   : Factor w/ 11 levels "B1","B10","B11",..: 4 4 4 4 4 4 4 4 4 4 ...
 $ VehGas     : Factor w/ 2 levels "Diesel","Regular": 2 2 1 1 1 2 2 1 1 1 ...
 $ Density    : int  1217 1217 54 76 76 3003 3003 137 137 60 ...
 $ Region     : Factor w/ 22 levels "R11","R21","R22",..: 18 18 3 15 15 8 8 20 20 12 ...
```









Example: Poisson claims frequency model

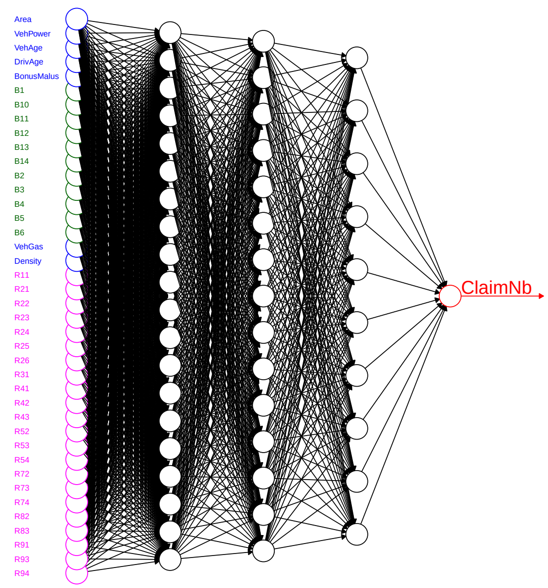
```
glm(formula = ClaimNb ~ VehPowerGLM + VehAgeGLM + DrivAgeGLM + BonusMalusGLM + VehBrand + VehGas +
                DensityGLM + AreaGLM + Region, family = poisson(), data = dat, offset = log(Exposure))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7036  -0.3777  -0.2886  -0.1626   6.9026

                Estimate Std. Error  z value Pr(>!z!)
(Intercept)   -4.0237251  0.0387027 -103.965  < 2e-16 ***
VehPowerGLM5   0.1994946  0.0192540   10.361  < 2e-16 ***
VehPowerGLM6   0.2281708  0.0191569   11.911  < 2e-16 ***
.
RegionR93     -0.0917983  0.0219777   -4.177 2.96e-05 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1  1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 200974  on 610211  degrees of freedom
Residual deviance: 190732  on 610164  degrees of freedom
```

	run time	# param.	in-sample loss	out-of-sample loss
homogeneous ($\mu \equiv \mathrm{const.}$)	0.1s	1	32.935	33.861
Model GLM	17s	48	31.257	32.149

Neural network regression model

- Choose network of depth $d \in \mathbb{N}$ with network parameter θ

$$\boldsymbol{x} \;\mapsto\; \mu_\theta^{\mathrm{NN}}(\boldsymbol{x}) = \exp\left\langle \theta^{(d+1)}, \boldsymbol{z}^{(d)} \circ \cdots \circ \boldsymbol{z}^{(1)}(\boldsymbol{x}) \right\rangle.$$

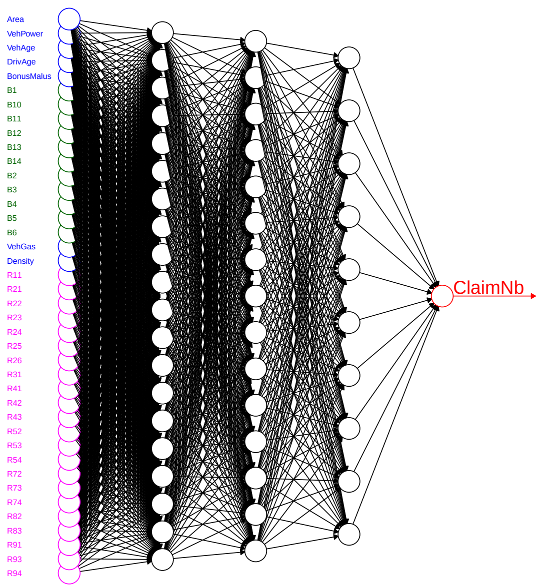


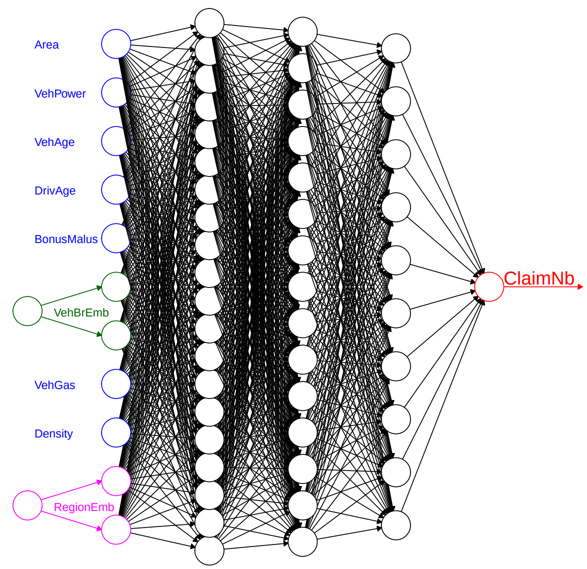
- Recall GLM: $\boldsymbol{x} \;\mapsto\; \mu_{\boldsymbol{\beta}}^{\mathrm{GLM}}(\boldsymbol{x}) = \exp\langle \boldsymbol{\beta}, \boldsymbol{x} \rangle$.

Neural network embedding layers

- Choose network of depth $d \in \mathbb{N}$ with network parameter θ

$$\boldsymbol{x} \;\mapsto\; \mu_\theta^{\mathrm{NN}}(\boldsymbol{x}) = \exp\left\langle \theta^{(d+1)}, \boldsymbol{z}^{(d)} \circ \cdots \circ \boldsymbol{z}^{(1)}(\boldsymbol{x}) \right\rangle.$$

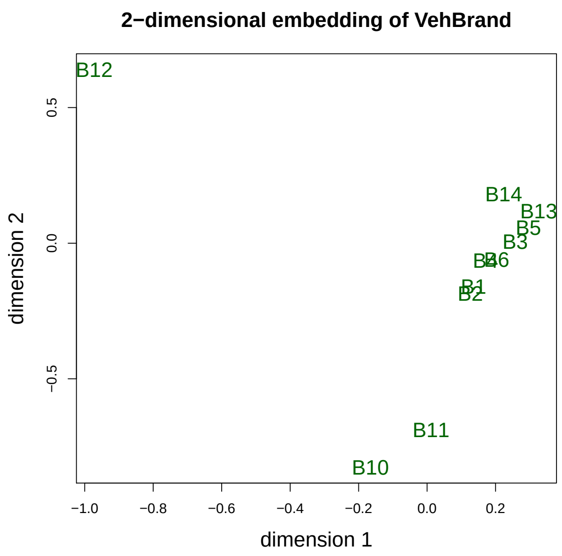


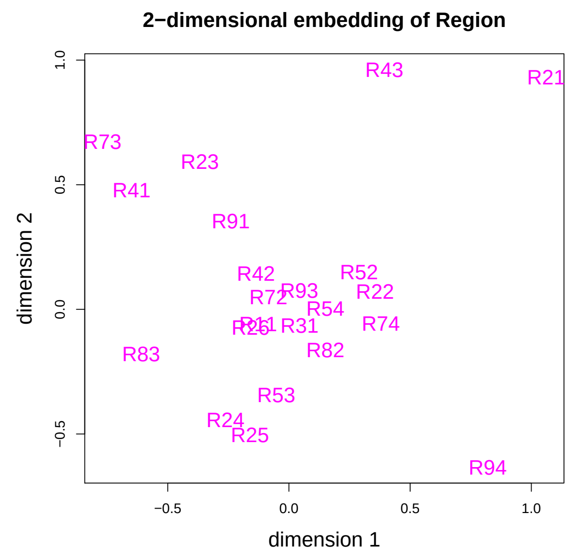


- Gradient descent method provides $\widehat{\theta}$ w.r.t. deviance loss $\theta \mapsto \mathcal{L}_{\mathcal{D}}(\theta)$.

NN example: car insurance frequencies

	run time	# param.	in-sample loss	out-of-sample loss
homogeneous ($\mu \equiv \text{const.}$)	0.1s	1	32.935	33.861
Model GLM	17s	48	31.257	32.149
NN (2-dim. embeddings)	365s	792	30.165	31.453

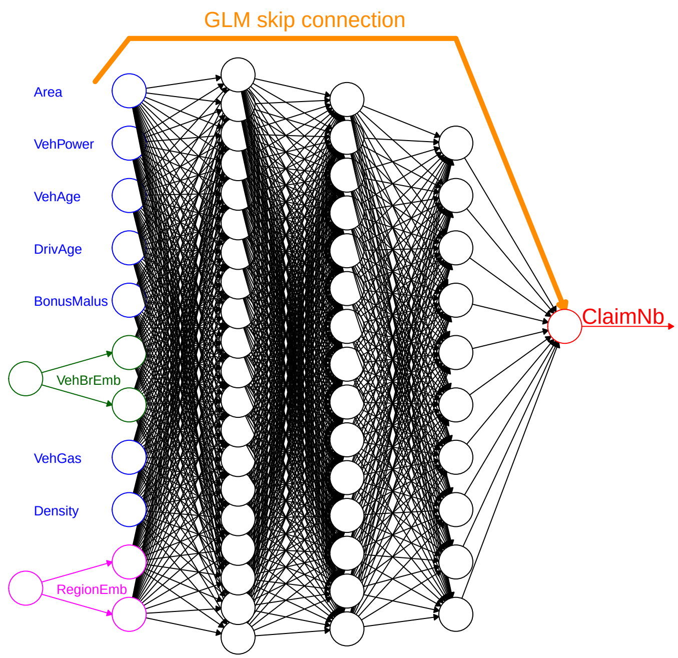




Combined Actuarial Neural Network

- Choose regression function with parameter $(\boldsymbol{\beta}, \theta)$

$$\boldsymbol{x} \;\mapsto\; \mu^{\rm CANN}_{(\boldsymbol{\beta},\theta)}(\boldsymbol{x}) = \exp\left\{\langle \boldsymbol{\beta}, \boldsymbol{x}\rangle + \left\langle \theta^{(d+1)}, \boldsymbol{z}^{(d)} \circ \cdots \circ \boldsymbol{z}^{(1)}(\boldsymbol{x})\right\rangle\right\}.$$

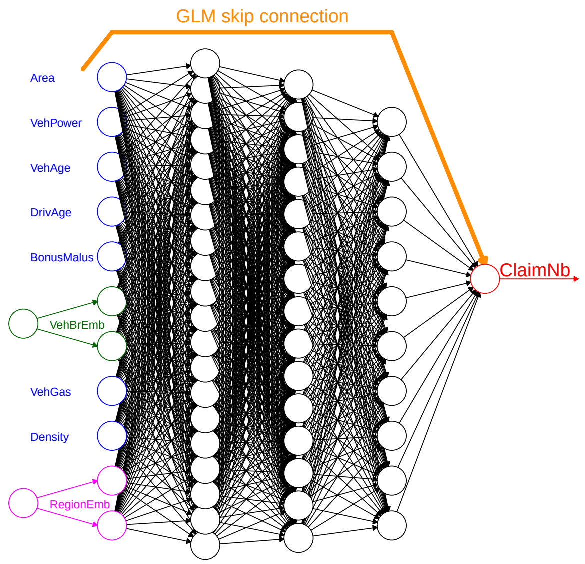


- Gradient descent method provides $(\widehat{\boldsymbol{\beta}}, \widehat{\theta})$ w.r.t. deviance loss $(\boldsymbol{\beta}, \theta) \mapsto \mathcal{L}_{\mathcal{D}}(\boldsymbol{\beta}, \theta)$.

Combined Actuarial Neural Network

- Choose regression function with parameter $(\boldsymbol{\beta}, \theta)$

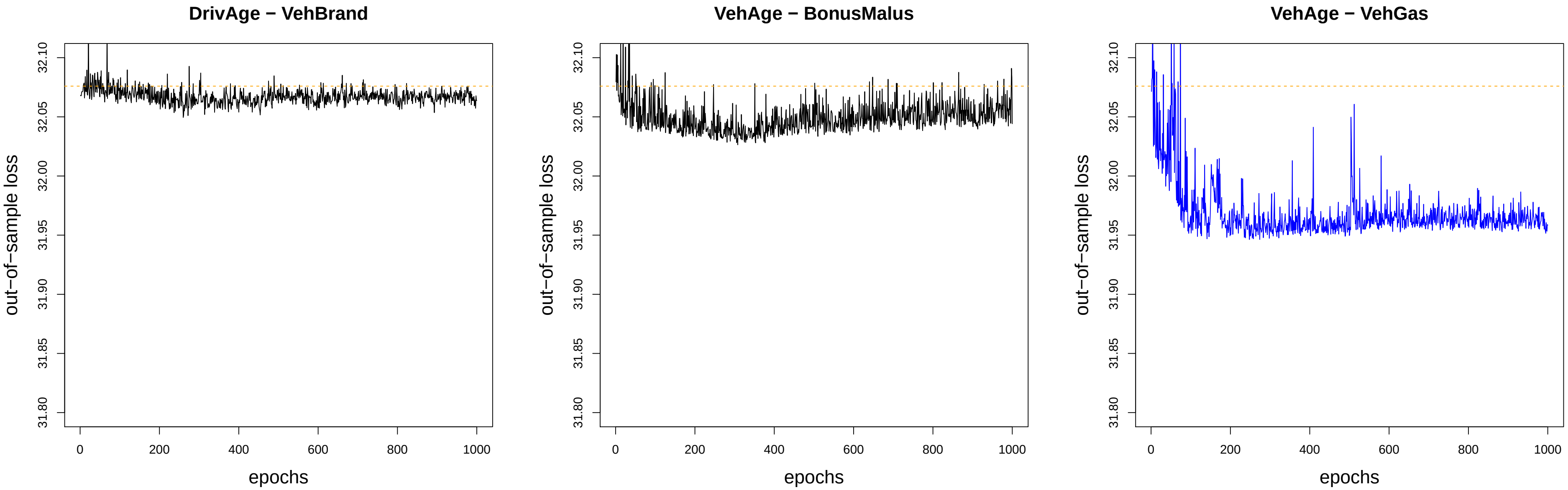
$$\mu^{\mathrm{CANN}}_{(\boldsymbol{\beta},\theta)}(\boldsymbol{x}) = \exp\left\{\langle\boldsymbol{\beta}, \boldsymbol{x}\rangle + \left\langle\theta^{(d+1)}, \boldsymbol{z}^{(d)} \circ \cdots \circ \boldsymbol{z}^{(1)}(\boldsymbol{x})\right\rangle\right\}.$$



- Gradient descent method provides $(\widehat{\boldsymbol{\beta}}, \widehat{\theta})$ w.r.t. deviance loss $(\boldsymbol{\beta}, \theta) \mapsto \mathcal{L}_{\mathcal{D}}(\boldsymbol{\beta}, \theta)$.

- Initialize gradient descent with $\widehat{\boldsymbol{\beta}}^{\mathrm{MLE}}$ and $\theta^{(d+1)} = 0$!

Combined Actuarial Neural Network

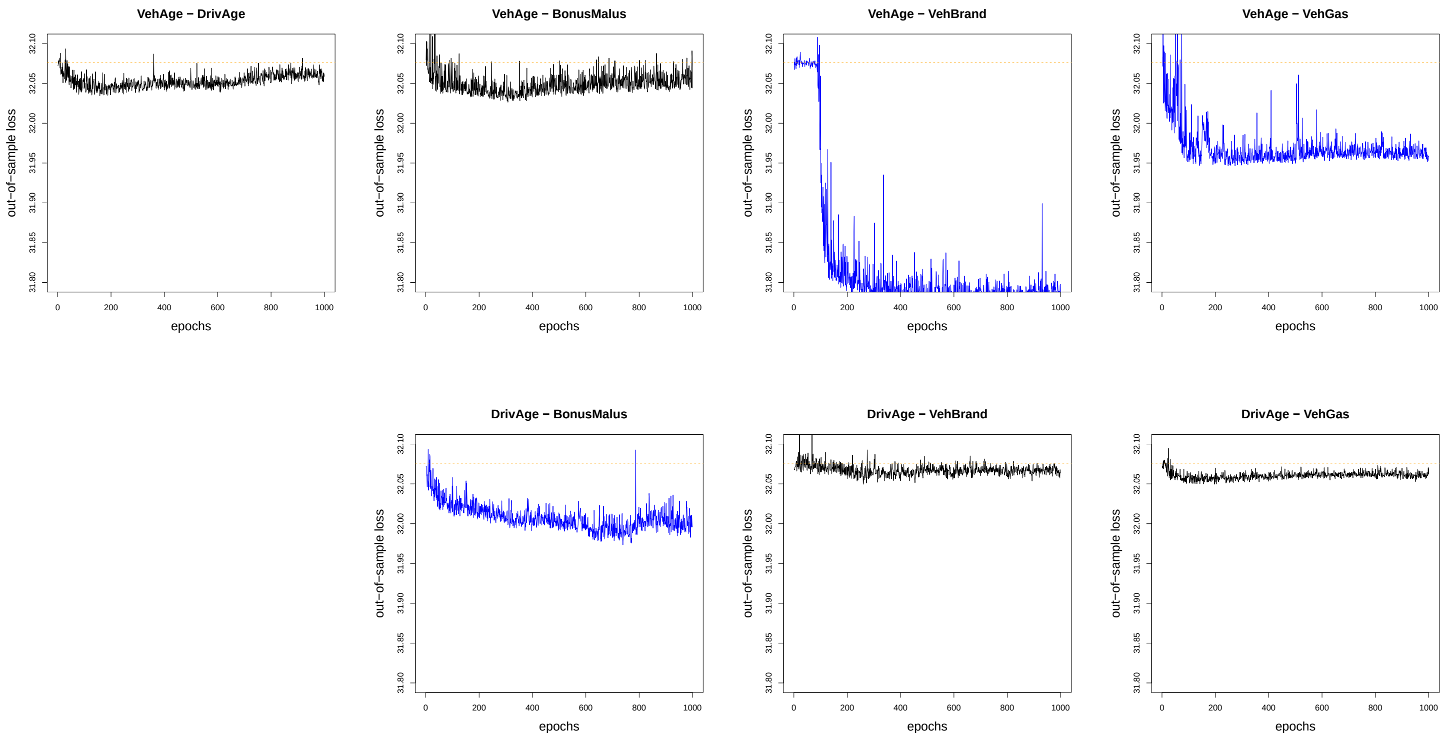


CANN example: car insurance frequencies

	run time	# param.	in-sample loss	out-of-sample loss
homogeneous ($\mu \equiv \text{const.}$)	0.1s	1	32.935	33.861
Model GLM	17s	48	31.257	32.149
CANN (2-dim. embeddings)	117s	792	30.476	31.566
NN (2-dim. embeddings)	365s	792	30.165	31.453

CANN example: binary interactions

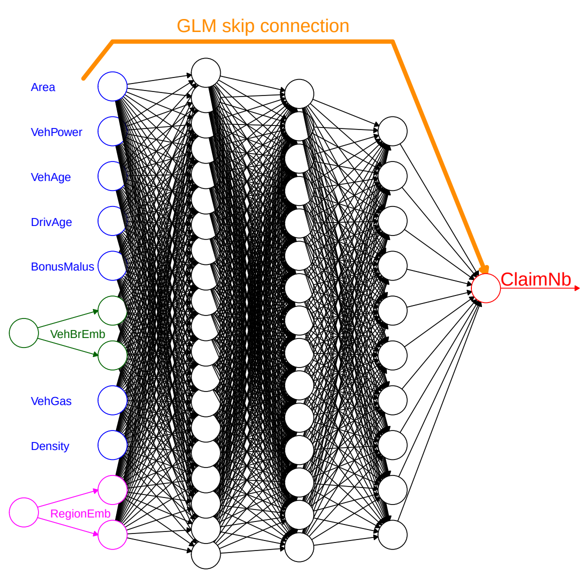
$$\boldsymbol{x} \;\mapsto\; \mu^{\mathrm{CANN}}_{(\boldsymbol{\beta},\theta)}(\boldsymbol{x}) = \exp\left\{\langle \boldsymbol{\beta}, \boldsymbol{x}\rangle + \left\langle \theta^{(d+1)}, \boldsymbol{z}^{(d)} \circ \cdots \circ \boldsymbol{z}^{(1)}(x_k, x_l)\right\rangle\right\}.$$



CANN example: learning across portfolios

Consider different insurance portfolios $\chi \in \{1, \ldots, K\}$.

$$(\boldsymbol{x}, \chi) \;\mapsto\; \mu^{\rm CANN}_{(\boldsymbol{\beta},\theta)}(\boldsymbol{x}, \chi) = \exp\left\{ \sum_{k=1}^{K} \mathbb{1}_{\{\chi=k\}} \langle \boldsymbol{\beta}_k, \boldsymbol{x} \rangle + \left\langle \theta^{(d+1)}, \boldsymbol{z}^{(d)} \circ \cdots \circ \boldsymbol{z}^{(1)}(\boldsymbol{x}) \right\rangle \right\}.$$



Thanks to my co-authors:

Christoph Buser (smile.direct versicherungen)
Andrea Ferrario (Mobiliar Lab for Analytics)
Andrea Gabrielli (RiskLab, ETH Zurich)
Michael Merz (University of Hamburg)
Alexander Noll (PartnerRe)
Ronald Richman (AIG)
Robert Salzmann (Signal Iduna)
Jürg Schelldorfer (Swiss Re)

Yorkshire Actuarial Society

14/03/2019

17:45 – 18:05 Tea & Coffee

Data Science and Machine Learning meets Mortality Modelling

Prof. Dr. Gareth W. Peters (FIOR, YAS-RSE)
Chair of Statistics for Risk and Insurance,
Department of Actuarial Mathematics and Statistics,
Heriot-Watt, Edinburgh, UK

March 12, 2019

Table of contents

Mortality Modelling Context

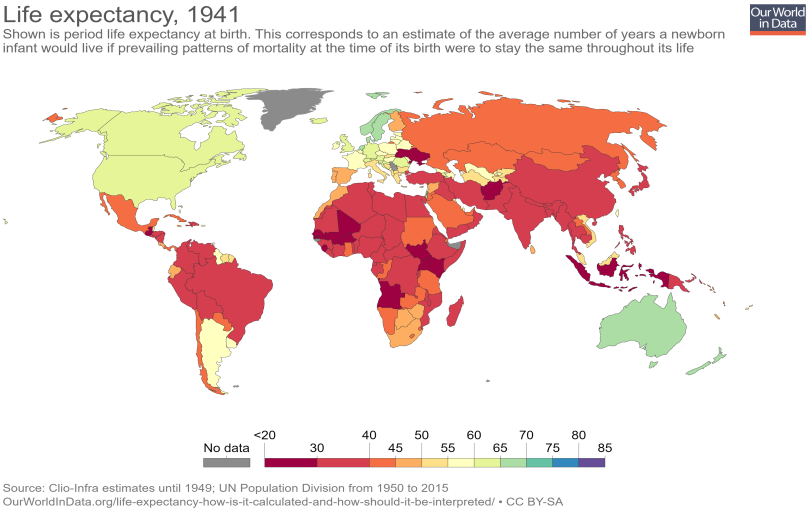


Figure: Global life expectancy by country in 1941

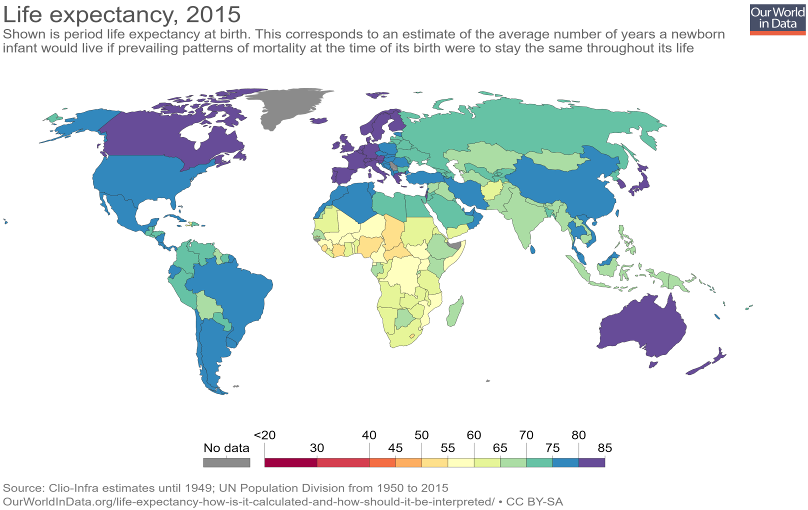
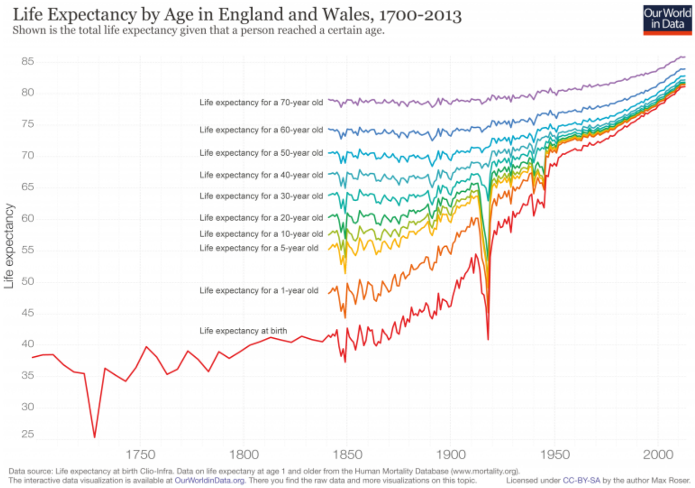


Figure: Global life expectancy by country in 2015


Life Expectancy by Age in England and Wales, 1700-2013
Shown is the total life expectancy given that a person reached a certain age.
Our World in Data
Life expectancy for a 70-year old
Life expectancy for a 60-year old
Life expectancy for a 50-year old
Life expectancy for a 40-year old
Life expectancy for a 30-year old
Life expectancy for a 20-year old
Life expectancy for a 10-year old
Life expectancy for a 5-year old
Life expectancy for a 1-year old
Life expectancy at birth
Life expectancy
85
80
75
70
65
60
55
50
45
40
35
30
25
1750
1800
1850
1900
1950
2000
Data source: Life expectancy at birth Clio-Infra. Data on life expectany at age 1 and older from the Human Mortality Database (www.mortality.org).
The interactive data visualization is available at OurWorldinData.org. There you find the raw data and more visualizations on this topic.
Licensed under CC-BY-SA by the author Max Roser.

Mortality Modelling Context

- Ageing populations are a major challenge for many countries.
  - Fertility rates are declining while life expectancy is increasing.
- longevity risk: the adverse financial outcome of people living longer than expected $\Rightarrow$ possibility of outliving their retirement savings.
  - long term demographic risk has significant implications for societies and manifests as a systematic risk for pension plans and annuity providers.
- Policymakers rely on mortality projection to determine appropriate pension benefits and regulations regarding retirement.

Mortality Modelling Context

Enhancing mortality models requires an understanding of common features of mortality behaviour [Cairns, Blake and Dowd, 2008]

- Mortality rates have fallen dramatically at all ages.
- Rate of decrease in mortality has varied over time and by age group.
- Absolute decreases have varied by age group.
- Aggregate mortality rates have significant volatility year on year.

Unexplored Features

Are there other statistical data features not yet explored in mortality data that will improve modelling and mortality projection?

Table of contents

DATA DRIVEN INNOVATION:

- A series of empirical studies in [Yan et al., 2019]

  "Evidence for Persistence and Long Memory Features in Mortality Data."

  were undertaken to demonstrate that long memory features are present in national level mortality data.

  - systematic evidence provided for 16 countries from the Human Mortality Database (HMD).

- Stratification of data by: gender, age and country all demonstrate prevalence of long memory structure

  $\Rightarrow$ confirmed by non-parametric estimation via Hurst exponent.

Capturing and Modelling New Features

Understanding and incorporating these long memory features may provide more accurate modelling and reliable forecasting of mortality rates!

HERIOT WATT UNIVERSITY

What is a Long Memory Feature?

- Long memory basically refers to the level of statistical dependence between two points in a time series.

- Given a stationary time series process $Y \equiv \{Y_t\}_{t=1:T}$, with $Y \in (\mathbb{N} \cup \{0\})^T$, [Beran, 1994] defined a condition for long memory stationary process via divergence of the autocorrelation function (ACF):

$$\lim_{n\to\infty} \sum_{j=-n}^{n} |\rho(j)| \to \infty \quad \text{where} \quad \rho(j) = \frac{\mathrm{Cov}(Y_t, Y_{t+j})}{\sqrt{\mathrm{Var}(Y_t), \mathrm{Var}(Y_{t+j})}}.$$

- Since the early work of [Hurst, 1951], long memory phenomena has been well recognized in diverse fields of application.

To understand Long Memory:

we need to think about back shift and difference operators in time series to obtain temporal Integration of a time series.

- Consider a time series $\{Y_t\}_{t=1:n}$ then the backshift operator and difference operator give

$$\begin{aligned} BY_t &= Y_{t-1} \\ (1-B)Y_t &= Y_t - Y_{t-1} \end{aligned}$$

- Typically one considers in an Integrated ARI(d)MA model a differencing operator $(1-B)^d$ for integers $d \in \mathbb{N}$ which is used to "differentiate" temporal trends from the time series to make it weakly stationary.

In a long memory model one replaces integer difference operators with fractional differential operators
$\Rightarrow$ by setting d to be fractional ($0 < d < 1/2$)

Extending to fractional differences:

- The fractional difference operator has a generator, expressed in integer powers of back shift operator B:

$$(1-B)^{-d} = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} B^j$$

  which has a lag decay much slower than exponential (with rate depending on d)
  - the ARFIMA$(0, d, 0)$ model describes a long memory stationary process with a hyperbolic decay of the ACF as compared to the geometric decay for an ARIMA model.
- This fractional parameter d captures the persistence for a given long memory model.

Extending to fractional differences:

- Oscillatory autocovariance is obtained using the generalised difference operator:

$$(1 - 2uB + B^2)^{-d} = \sum_{j=0}^{\infty} \Psi_j(u) B^j$$

where $\Psi_j(u)$ is the j-th order Gagenbauer orthogonal polynomial

$$\Psi_j(u) = \sum_{q=0}^{[j/2]} \frac{(-1)^q (2u)^{j-2q} \Gamma(d - q + j)}{q!(j - 2q)!\Gamma(d)}.$$

Long Memory Patterns in Mortality Data

NOTE: It was found that for moratlity data applications it sufficies to consider the non-oscillatory long memory patterns.

How to detect long memory features in mortality time series?

- Statistically testing directly for persistence/long memory properties in mortality time series, can be done via the Hurst exponent (denoted by H).

Developing Estimators for Long Memory Features:

- Three estimators for H that were explored included:
  - rescaled range analysis (R/S),
  - detrended fluctuation analysis (DFA), and
  - periodogram regression (PR) methods.

Connecting Hurst Exponent H to Fractional Difference d

Next we connect the Hurst exponent H in Fractional Brownian Motion to the long memory parameter d in the ARFIMA model

Developing Estimators for Long Memory Features:

[Mandelbrot,1968] defined fractional Brownian motion

$$W_H = \{W_H(t) : 0 \le t < \infty, W_H(t) \in \mathbb{R}\},$$

by its stochastic representation given below.

Definition (fractional Brownian motion)

Let $W_H(t)$ be a random process defined as the solution to following diffusion

$$W_H(t) := \frac{1}{\Gamma(H+\frac{1}{2})}\left(\int_{-\infty}^{0}\left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right)dW(s) + \int_0^t (t-s)^{H-\frac{1}{2}}dW(s)\right),$$

where the Gamma function $\Gamma(\alpha) := \int_0^\infty x^{\alpha-1}\exp(-x)dx$, $W(t)$ denotes standard Brownian motion and $H \in (0,1)$.

Developing Estimators for Long Memory Features:

- Fractional Brownian motion of exponent H is a moving average of $dW(t)$, in which past increments of $W(t)$ are weighted by the kernel $(t-s)^{H-\frac{1}{2}}$.
- The "span of interdependence" between fractional Brownian motion increments can be said to be infinite.
- In addition, $W_H(t)$ will reduce to a standard Brownian motion $W(t)$ when $H = 0.5$.
- Given the initial value $W_H(0) = 0$ at $t = 0$, the expected value and variance are given by

$$\mathbb{E}(W_H(t)) = 0 \quad \text{and} \quad \mathrm{Var}(W_H^2(t)) = t^{2H}, \quad \text{for} \quad t > 0,$$

respectively.

Developing Estimators for Long Memory Features:
The covariance function derived by [Nualart,2006] is given by

$$\mathrm{Cov}_H(t,s) = \mathbb{E}\left[W_H(s)W_H(t)\right] = \frac{1}{2}\left(t^{2H} + s^{2H} - (t-s)^{2H}\right), \text{ for } 0 < s \leq t.$$

An increment of the process in an interval $[s,t]$ has a normal distribution with zero mean and variance

$$\mathbb{E}\left[\left(W_H(t) - W_H(s)\right)^2\right] = |t-s|^{2H}.$$

Developing Estimators for Long Memory Features:

- For $H = 1/2$, the covariance $\mathrm{Cov}_{1/2}(t, s) = \mathsf{min}(s, t)$ and increments of the process in disjoint intervals are independent.
- For $H \neq 1/2$, increments are not independent.

Consider a unit increment discrete time process $\{\Xi_n := W_H(n) - W_H(n-1), n \in \mathbb{N}\}$, which gives a Gaussian stationary sequence with unit variance and ACF

$$\begin{aligned} \rho_H(n) &= \frac{1}{2}\left[(n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\right] \\ &\approx H(2H-1)n^{2H-2} \to 0 \;\; \text{as} \;\; n \to \infty. \end{aligned}$$

Developing Estimators for Long Memory Features:

- For $\frac{1}{2} < \mathrm{H} < 1$ and large enough n, we have $\rho_\mathrm{H}(\mathrm{n}) > 0$ and $\sum_{\mathrm{n}=1}^{\infty} \rho_\mathrm{H}(\mathrm{n}) \to \infty$.
  - $\Rightarrow$ indicates long memory in a time series, meaning that a high value in the series will more likely be followed by another high value and such an effect is likely to continue a long time into the future.

- For $0 < \mathrm{H} < \frac{1}{2}$ and large enough n, we have $\rho_\mathrm{H}(\mathrm{n}) < 0$ and $\sum_{\mathrm{n}=1}^{\infty} |\rho_\mathrm{H}(\mathrm{n})| < \infty$.
  - $\Rightarrow$ indicates a time series which is more likely to switch between high and low values in adjacent pairs, and such anti persistence will last a long time into the future.
  - This process can be used to model sequences with intermittency.

- A value of $\mathrm{H} = \frac{1}{2}$ can indicate a standard Brownian motion which is a short memory process.

HERIOT WATT UNIVERSITY

Relationship between long memory d and Hurst exponent H

Weak convergence of fractional time series to fBM:

Define a fractional difference characterizing process

$$Q_t = (1 - B)^{-d}\varepsilon_t, \quad \varepsilon_t \overset{\text{i.i.d}}{\sim} N(0, 1), \quad |d| < 1/2,$$

then the Q_t is a stationary and weakly dependent process.

Now consider the following scaled partial sum constructed from this process,

$$Z_T(\xi) := \frac{\sum_{t=1}^{[T\xi]} Q_t}{\sigma_T}, \quad 0 \le \xi \le 1,$$

where $\sigma_T^2 = \mathbb{E}\left[\left(\sum_{t=1}^{T} Q_t\right)^2\right]$ and $[\cdot]$ denotes integer part.

[Davidson, 2000] showed the following theorem

Theorem

The following weak convergence holds for all $|d| < 1/2$ and all $0 \leq \xi \leq 1$,

$$Z_T(\xi) \xrightarrow{d} V_d^{-\frac{1}{2}} W_d(\xi), \quad \text{as} \quad T \to \infty,$$

with the scale constant

$$V_d = \frac{1}{\Gamma(d+1)^2} \left(\frac{1}{2d+1} + \int_0^\infty ((1+\tau)^d - \tau^d)^2 d\tau \right),$$

chosen to ensure that $\mathbb{E}\left[W_d(1)^2\right] = 1$. For $|d| < 1/2$, $W_d(\xi)$ is a fractional Brownian motion with the following representation

$$W_d(\xi) = \frac{1}{\Gamma(d+1)} \left(\int_{-\infty}^0 \left((\xi - s)^d - (-s)^d \right) dW(s) + \int_0^\xi (\xi - s)^d dW(s) \right),$$

where W is the standard Brownian motion.

The relationship between d and H is then given by $d = H - 0.5$.

Connecting Hurst Exponent H to Fractional Difference d

Rescaled Range Analysis Estimation of Hurst Exponent

$\Rightarrow$via SIMPLE LINEAR REGRESSION!

Rescaled Range Analysis R/S Estimators:

Given a time series $Y_{t \in (1,2,3,\cdots,T)}$, the sample mean and the standard deviation process are given by

$$\overline{Y}_T = \frac{1}{T} \sum_{j=1}^{T} Y_j \quad \text{and} \quad S_t = \sqrt{\frac{1}{t-1} \sum_{j=1}^{t} (X_j)^2},$$

where the mean adjusted series $X_t = Y_t - \overline{Y}_T$.

Then a cumulative sum series is given by $Z_t = \sum_{j=1}^{t} X_j$ and the cumulative range based on these sums is

$$R_t = \mathrm{Max}\left(0, Z_1, \cdots, Z_t\right) - \mathrm{Min}\left(0, Z_1, \cdots, Z_t\right).$$

Rescaled Range Analysis:

The following proposition proposes an estimator of H as derived in [Mandelbrot, 1975].

Theorem

Consider a time series $Y_t \in \mathbb{R}$ with std. dev. process S_t and cumulative range process R_t, then $\exists$ $C \in \mathbb{R}$ such that following asymptotic property of the rescaled range R/S holds

$$[R/S](T) = \frac{1}{t}\sum_{t=1}^{T} R_t/S_t \sim CT^H, \quad \text{as} \quad T \to \infty.$$

HERIOT WATT UNIVERSITY

Rescaled Range Analysis (SMALL SAMPLES):

[Annis, 1976] showed that for small samples T (as can arise in mortality data), the rescaled range R/S can also be approximated by following equation

$$[\mathrm{R/S}](\mathrm{T}) = \begin{cases} \frac{\mathrm{T}-1/2}{\mathrm{T}} \frac{\Gamma((\mathrm{T}-1)/2)}{\sqrt{\pi}(\mathrm{T}/2)} \sum_{j=1}^{\mathrm{T}-1} \sqrt{\frac{\mathrm{T}-j}{j}}, & \text{for} \quad \mathrm{T} \leq 340 \\ \frac{\mathrm{T}-1/2}{\mathrm{T}} \frac{1}{\sqrt{\mathrm{T}\pi/2}} \sum_{j=1}^{\mathrm{T}-1} \sqrt{\frac{\mathrm{T}-j}{j}}, & \text{for} \quad \mathrm{T} > 340 \end{cases}$$

where the $\frac{\mathrm{T}-1/2}{\mathrm{T}}$ term was added by [Peters, 1994].

Simple Linear Regression Estimator for H

The H estimate can be obtained by a simple linear regression

$$\log \mathrm{R/S}(\mathrm{T}) = \log \mathrm{C} + \mathrm{H} \log \mathrm{T}.$$

Rescaled Range Analysis:

Hence, we have the following definition for the estimator of H.

Definition (Estimator $\hat{H}$ by R/S)

The estimator $\hat{H}$ based on the rescaled range R/S analysis is given by

$$\hat{H}_{R/S} = \frac{T(\sum_{t=1}^{T} \log R/S(t) \log t) - (\sum_{t=1}^{T} \log R/S(t))(\sum_{t=1}^{T} \log t)}{T(\sum_{t=1}^{T} (\log t)^2) - (\sum_{t=1}^{T} \log t)^2}.$$

The empirical confidence interval of $\hat{H}_{R/S}$ with sample length $T = 2^N$ [Weron, 2002] is

$$(0.5 - \exp(-7.33 \log(\log N) + 4.21), \exp(-7.20 \log(\log N) + 4.04) + 0.5),$$

Connecting Hurst Exponent H to Fractional Difference d

Detrended Fluctuation Analysis Estimation of Hurst Exponent

$\Rightarrow$via SIMPLE LINEAR REGRESSION!

Detrended Fluctuation Analysis DFA ([Peters, 1994]):

Conceived as a method for detrending local variability in a time series to insight into long-term variations.

First partition the mortality data time series:

- time series $Y_{t \in (1,2,3,\cdots,T)}$ is divided into K non-overlapped subseries $Y_{l,k}$ of length L for $l = 1, 2, \cdots, L$ and $k = 1, 2, \cdots, K$.

Next construct cumulative series and RMSF:

Construct cumulative time series $Z_{j,k} = \sum_{l=1}^{j} Y_{l,k}$ with $j = 1, \cdots, L$ and the sample average of the root mean square fluctuation for all K sub-series of length L:

$$\bar{F}(L) := \frac{1}{K} \sum_{k=1}^{K} \sqrt{\frac{1}{L} \sum_{j=1}^{L} (Z_{j,k} - a_k j - b_k)^2},$$

where $a_k j + b_k$ is the least-squares line $\tilde{Z}_k(j) = a_k j + b_k$ fitted to points $(j, Z_{j,k}), j = 1, \cdots, L$.

Detrended Fluctuation Analysis:

[Taqqu, 1995] proved the following result for DFA to estimate H.

Theorem

Consider a time series Y_t and $\bar{F}(L)$ defined in Equation (29), $\exists \; C \in \mathbb{R}$ such that the following asymptotic property for $\bar{F}(L)$ holds:

$$\bar{F}(L) \sim CL^{H}, \quad \text{as} \quad L \to \infty.$$

As a consequence of this asymptotic result, the estimated value of H can be obtained by running a linear regression

$$\log(\bar{F}(L)) = \log(C) + H \log(L),$$

and this leads to an alternative statistical estimator for H.

HERIOT WATT UNIVERSITY

Detrended Fluctuation Analysis:

Definition (Estimator $\hat{H}$ by DFA)

The estimator $\hat{H}$ developed by adopting DFA approach is given by

$$\hat{H}_{DFA} = \frac{L(\sum_{l=1}^{L} \log \bar{F}(l) \log l) - (\sum_{l=1}^{L} \log \bar{F}(l))(\sum_{l=1}^{L} \log l)}{L(\sum_{l=1}^{L} (\log l)^2) - (\sum_{l=1}^{L} \log l)^2}.$$

The empirical confidence interval of $\hat{H}$ given in Equation (6) with sample length $T = 2^N$ for $N \in [10, 50]$ [Weron, 2002] is

$$(0.5 - \exp(-2.33 \log N + 3.25), \exp(-2.46 \log N + 3.38) + 0.5),$$

and the empirical confidence interval for $N > 50$ is

$$(0.5 - \exp(-2.93 \log N + 4.45), \exp(-3.10 \log N + 4.77) + 0.5).$$

Connecting Hurst Exponent H to Fractional Difference d

Periodogram Estimation of Hurst Exponent

$\Rightarrow$ via SIMPLE LINEAR REGRESSION!

HERIOT WATT UNIVERSITY

Periodogram Regression:

- OBSERVATION: the spectral density function of a general fractionally integrated model with parameter d is identical to that of a fractional Gaussian noise with Hurst exponent $H = d + 0.5$.

Hence, periodogram regression (PR) was proposed by [Geweke, 1983] to estimate H.

Definition

A periodogram of time series $Y_{t \in (1,2,3,\cdots,T)}$ with frequencies $\omega_j = \frac{j}{T}$ and $j = 1, \cdots, [T/2]$ can be represented as

$$I_T(\omega_j) = \frac{1}{T} \left| \sum_{t=1}^{T} Y_t e^{-2\pi i (t-1)\omega_j} \right|^2 .$$

Periodogram regression:

A simple linear regression at low frequencies $\omega_j, \quad j = 1, \cdots, J \leq [T/2]$ can be applied to estimate H according to the linear model,

$$\log\left(I_T(\omega_j)\right) = C - d \log\left(4 \sin^2\left(\frac{\omega_j}{2}\right)\right) + \varepsilon_j,$$

where C is a constant and $Y_t = \log\left(4 \sin^2\left(\frac{\omega_j}{2}\right)\right)$.

Periodogram regression:

The least squares estimate of the slope yields an estimator for d.

Definition (Estimator $\hat{d}$ by periodogram regression)

Consider a time series Y_t with the periodogram $I_T(\omega_j)$ defined in Equation (7), the estimator $\hat{d}$ developed by adopting periodogram regression is given by

$$\hat{d} = \frac{J\left[\sum_{j=1}^{J} \log\left(I_T(\omega_j)\right) \log\left(4\sin^2\left(\frac{\omega_j}{2}\right)\right)\right] - \left[\sum_{j=1}^{J} \log\left(I_T(\omega_j)\right)\right]\left[\sum_{j=1}^{J} \log\left(4\sin^2\left(\frac{\omega_j}{2}\right)\right)\right]}{J\sum_{j=1}^{J}(\log\left(4\sin^2\left(\frac{\omega_j}{2}\right)\right))^2 - \left[\sum_{j=1}^{J} \log\left(4\sin^2\left(\frac{\omega_j}{2}\right)\right)\right]^2},$$

with asymptotic distribution

$$\hat{d} \sim N\left(d, \frac{\pi^2}{6\sum_{t=1}^{T}(Y_t - \bar{Y}_t)^2}\right).$$

Periodogram regression:

Hence, an estimate of H using periodogram regression can be calculated by $\hat{H} = \hat{d} + 0.5$ and the 95% empirical confidence interval for $\hat{d}$ is calculated with the sample length $L = 2^N$ and $J = [L^{0.5}]$ [Weron, 2002] by

$$(0.5 - \exp(-0.71 \cdot N^{2/3} + 2.04), \exp(-0.68 \cdot N^{2/3} + 1.78) + 0.5).$$

- The periodogram regression approach is the only approach with known asymptotic properties.

Table of contents

The mortality data sets of 16 countries are downloaded from the Human Mortality Database (HMD) which provides detailed mortality and population data to the public.

Table: Data length and abbreviation for country names

Full name	Australia	Belgium	Canada	Denmark
Abbrev.	AU	BE	CA	DK
Data length	94	170	91	180
Full name	Finland	France	Iceland	Italy
Abbrev.	FI	FR	IS	IT
Data length	138	199	176	141
Full name	Netherlands	Norway	Spain	Sweden
Abbrev.	NL	NO	ES	SE
Data length	163	169	107	264
Full name	Switzerland	U.K.	U.S.A.	Japan
Abbrev.	CH	GB	US	JP
Data length	139	92	83	68

Long memory pattern across age groups, gender and countries

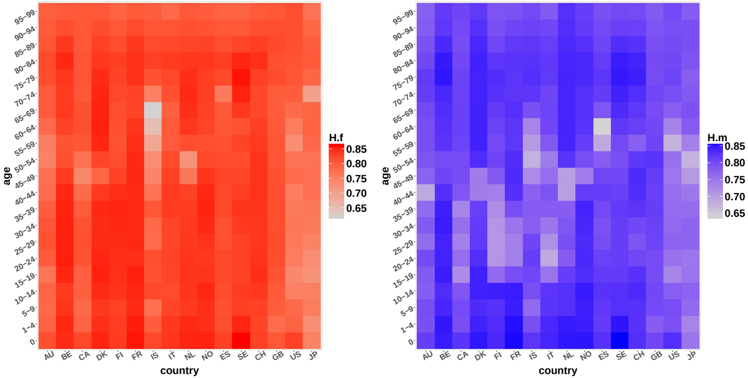


Figure: Heat map of estimated H across countries and age groups for female (left) and male (right).

All countries demonstrate some degree of long memory across all age groups.

Long memory pattern across age group by gender

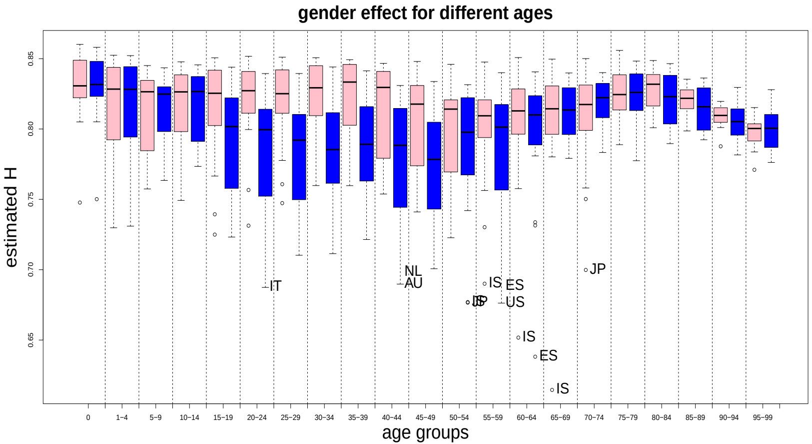


Figure: Boxplot of estimated H across age groups aggregated over countries to show gender effect. For each age group, the first boxplot is female (pink) and the second is male (blue).

Long memory pattern across age group by gender

- Results indicate that the mortality change (improvement) over the years is more persistent for female than male and for younger than middle age groups.

- Improvements vary more across countries for male than female.

- Senior age groups ranging between 80-99, there is a consistent downward trend of estimated H for both genders.
  - Suspect mortality persistence is more difficult to detect in a smaller population of seniors.

Long memory pattern across countries by gender

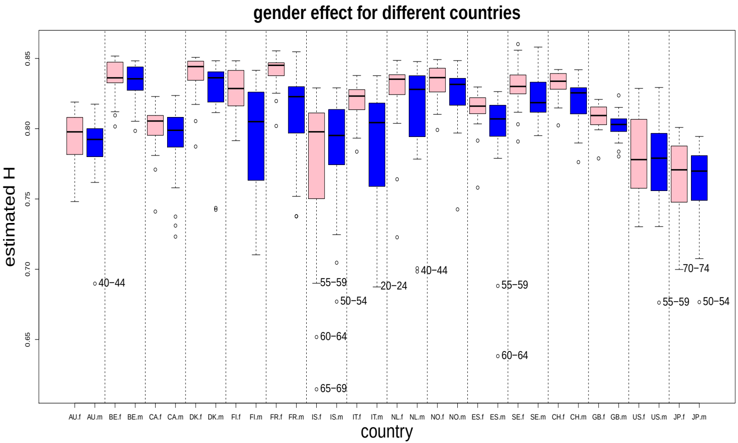


Figure: Boxplot of estimated H across countries aggregated over ages to show gender effect. For each age group, the first boxplot is female (pink) and the second is male (blue).

Long memory pattern across countries by gender

- Compared the estimated H across 16 countries by gender aggregated over 21 age groups, which indicate mortality persistence over the years.
- Again, the estimated H has lower values and wider spreads for male populations.
- Difference in spread is most apparent for Finland, France, Italy and Netherlands.
- Countries like Belgium and Denmark display high level of mortality persistence over the years whereas Iceland, US and Japan typically exhibit lower levels of mortality persistence.
- Spread of H is wider for male than female in Finland, France, Italy, Netherlands and Spain.

The uncertainty in future death rates can be divided into two components:

- Unsystematic mortality risk. Even if the true mortality rate is known, the number of deaths, will be random.
    - larger population $\Rightarrow$ smaller unsystematic mortality risk (due to pooling of offsetting risks - diversification).
- Systematic mortality risk. This is the undiversifiable component of mortality risk that affects all individuals in the same way.
    - Forecasts of mortality rates in future years are uncertain.
- [Lee and Carter, 92] proposed a stochastic mortality model (LC) to forecast the trend of age-specific mortality rates.
- Several extensions to Lee-Carter model have been proposed, overview in [Fung et al. 2017].

Consider random vector
$Y_t = (Y_{x_1,t}, Y_{x_2,t}, \cdots, Y_{x_g,t})$ with $Y_{x,t} \in (\mathbb{N} \cup \{0\})^T$
the set of T dimensional death counts for age group
$x \in \{x_1, \cdots, x_g\}$ and years $t \in \{1, \cdots, T\}$.

Consider random vector of central exposure

$$E_t = (E_{x_1,t}, E_{x_2,t}, \cdots, E_{x_g,t})$$

Define filtrations:

$$\mathcal{F}_{1:t-1} = \sigma(Y_1, Y_2, \ldots, Y_{t-1}),$$
$$\mathcal{F}^E_{1:t-1} = \sigma(E_1, E_2, \ldots, E_{t-1}).$$

Long Memory Mortality Modelling

Two extended Bayesian models for multivariate LCC mortality models are developed:

- MELCC: no long memory with period and cohort effect;
- LMLM: long memory in trend and in cohort effect.

Definition (GLGARMA model)

Given a discrete stationary time series process $\mathrm{Y}_{\mathrm{t}\in\{1,2,3,\cdots,\mathrm{T}\}}$, a GLGARMA model with order $(\mathrm{p},\mathrm{d},\mathrm{q})$ is defined by

$$\mathrm{Y}_{\mathrm{x},\mathrm{t}}|\mathcal{F}_{\mathrm{t}-1} \sim \mathrm{F}(\mathrm{Y}_{\mathrm{x},\mathrm{t}};\mu_{\mathrm{x},\mathrm{t}},\nu_{\mathrm{x}}), \tag{1}$$

$$(1-2\mathrm{uB}+\mathrm{B}^2)^{\mathrm{d}}\,\Phi(\mathrm{B})\,(\log(\mu_{\mathrm{x},\mathrm{t}})-\mathrm{c}) = \Theta(\mathrm{B})\varepsilon_{\mathrm{x},\mathrm{t}}, \tag{2}$$

$$\Phi(\mathrm{B}) = 1-\phi_1\mathrm{B}-\cdots-\phi_{\mathrm{p}}\mathrm{B}^{\mathrm{p}} \quad\text{and}\quad \Theta(\mathrm{B}) = 1+\theta_1\mathrm{B}+\cdots+\theta_{\mathrm{q}}\mathrm{B}^{\mathrm{q}}.$$

and F represents a discrete count distribution with parameters $\mu_{\mathrm{x},\mathrm{t}}$ and ν_{x} which denotes the mean function and dispersion level respectively.

NOTE:

- Mean function is a Gegenbauer long memory time series
  $\Rightarrow$ a slowly damping autocorrelation function with oscillatory pattern.
- We consider a special case with $\mathrm{p}=\mathrm{q}=0$ so that $\Phi(\mathrm{B})=\Theta(\mathrm{B})=1$.

We may rewrite the trend as follows

$$(\log(\mu_{\mathrm{x},\mathrm{t}})-\mathrm{c}) = (1-2\mathrm{uB}+\mathrm{B}^2)^{-\mathrm{d}}\varepsilon_{\mathrm{x},\mathrm{t}} \equiv \sum_{\mathrm{j}=0}^{\infty}\psi_{\mathrm{j}}\varepsilon_{\mathrm{x},\mathrm{t}-\mathrm{j}}.$$

REMARK:

- Gegenbauer polynomial coefficients ψ_{j} are functionally dependent on d and u that control the strength of long memory and the oscillatory pattern respectively.
- Given the constraint $|\mathrm{u}| < 1$, this process is stationary if $\mathrm{d} < 1/2$, invertible if $\mathrm{d} > -1/2$ and has long memory if $0 < \mathrm{d} < 1/2$.
- $\mathrm{ARFIMA}(\mathrm{p}, \mathrm{d}, \mathrm{q})$ model with mean μ is a special case of GARMA when $\mathrm{u} = 1$.

Coefficients ψ_{j} are easily computed using the Rodrigues formula:

$$\psi_{\mathrm{j}} = 2\mathrm{u}\left(\frac{\mathrm{d}-1}{\mathrm{j}} + 1\right)\psi_{\mathrm{j}-1} - \left(2\frac{\mathrm{d}-1}{\mathrm{j}} + 1\right)\psi_{\mathrm{j}-2},$$

where the first three terms are $\psi_0 = 1$, $\psi_1 = 2\mathrm{du}$ and $\psi_2 = -\mathrm{d} + 2\mathrm{d}(1+\mathrm{d})\mathrm{u}^2$.

- Mortality models can be beneficial to incorporate both under- and over-dispersion features in distribution $Y_{x,t}|\mathcal{F}_{t-1} \sim F(Y_{x,t}; \mu_{x,t}, \nu_x)$
- Generalised Poisson (GP) distribution, [Consul, 1989] is adopted and nests Poisson as a special case.

Definition (Generalised Poisson)

Let $Y \sim GP(\mu, \nu)$ be a random variable, taking support on $\mathbb{N} \cup \{0\}$, where the pmf is given by

$$\begin{aligned} f(y; \mu, \nu) &= \mu(1-\nu)[\mu(1-\nu) + \nu y]^{y-1} e^{-\mu(1-\nu)-\nu y}/y!, \\ \mathbb{E}(Y) &= \mu \ \text{ and } \ \mathrm{Var}(Y) = \mu(1-\nu)^{-2}, \end{aligned}$$

where $\mu > 0$ is the mean parameter and $\nu \in [-1, 1)$ is the dispersion parameter. The GP distribution is over-, under- and equi-dispersed when $\nu \in (-1, 1)$ is greater than, less than and equal to 0 respectively.

HERIOT WATT UNIVERSITY

Multivariate Extended Lee-Carter Cohort (MELCC) Model.

Definition (MELCC model)

$$\mathrm{Y_{x,t}}|\mathcal{F}_{1:t-1}, \mathrm{E_{x,t}}, \mu_{x,t}, \kappa_t, \zeta_{x,t} \sim \mathrm{GP}(\mathrm{Y_{x,t}}; \mathrm{E_{x,t}}\mu_{x,t}, \nu_x),$$

$$\ln \mu_{x,t} = \alpha_x + \beta_x \kappa_t + \beta_x^{\zeta} \zeta_{x,t} + \varepsilon_{x,t}, \qquad \varepsilon_{x,t} \overset{\text{i.i.d}}{\sim} \mathrm{N}(0, \sigma_\varepsilon^2),$$

$$\kappa_t = \eta \kappa_{t-1} + \varsigma^{\kappa} + \varepsilon_t^{\kappa}, \qquad \varepsilon_t^{\kappa} \overset{\text{i.i.d}}{\sim} \mathrm{N}(0, \sigma_\kappa^2),$$

$$\zeta_{x_1,t} = \lambda \zeta_{x_1,t-1} + \varsigma + \varepsilon_t^{\zeta}, \qquad \varepsilon_t^{\zeta} \overset{\text{i.i.d}}{\sim} \mathrm{N}(0, \sigma_\zeta^2),$$

$$\zeta_{x_i,t} = \zeta_{x_{i-1},t-1}, \qquad i = 2, 3, \cdots, g$$

- $\boldsymbol{\alpha} = [\alpha_{x_1}, \cdots, \alpha_{x_g}] \in \mathbb{R}^g$ represents the profile of age groups on the log mortality rates
- $\boldsymbol{\beta} = [\beta_{x_1}, \cdots, \beta_{x_g}] \in \mathbb{R}^g$ measures the interaction of age group and time effect on the log mortality rates,
- $\mathrm{E_{x,t}}\mu_{x,t}$ is the mean function and the dispersion parameter $\nu_x \in (-1, 1)$ for GP distribution.

Definition (LMLM model)

Assume $Y_t|\mathcal{F}_{1:t-1}, E_t, \boldsymbol{\mu}_t(\boldsymbol{\zeta}'_t)$ forms a Markov process with long memory period and cohort effect structures:

$$
\begin{aligned}
&Y_{x,t}|\mathcal{F}_{1:t-1}, E_{x,t}, \mu_{x,t}(\zeta'_{x,t}), \zeta'_{x,t} \sim GP(Y_{x,t}; E_{x,t}\,\mu_{x,t}(\zeta'_{x,t}), \nu_x),\\
&\Phi_x(B)\ln \mu_{x,t} = \alpha_x + \zeta'_{x,t} + \Theta_x(B)((1-2uB+B^2)^{-d}\varepsilon_{x,t}),\\
&\zeta'_{x_1,t} = \varsigma' + (1-2u'B+B^2)^{-d'}\varepsilon_t^{\zeta'},\\
&\zeta'_{x_i,t} = \zeta'_{x_{i-1},t-1},\ i \in \{2,3,\cdots,g\},\\
&\qquad \varepsilon_{x,t} \overset{i.i.d}{\sim} N(0,\sigma^2_{x,\varepsilon}),\quad \varepsilon_t^{\zeta'} \overset{i.i.d}{\sim} N(0,\sigma^2_{\zeta'})
\end{aligned}
$$

- fractional difference parameters $d \in (0, 0.5)$ and $d' \in (0, 0.5)$ control the strength of long memory
- Gegenbauer parameters u with $|u| < 1$ and u' with $|u'| < 1$ control the cycle of oscillatory autocorrelation function (ACF) for the mortality rate process and cohort effect process respectively

Measuring forecast performance of MELCC vs LMLM models

$\Rightarrow$ out-sample forecast split death counts $Y_{x,1:T}$ into two parts:

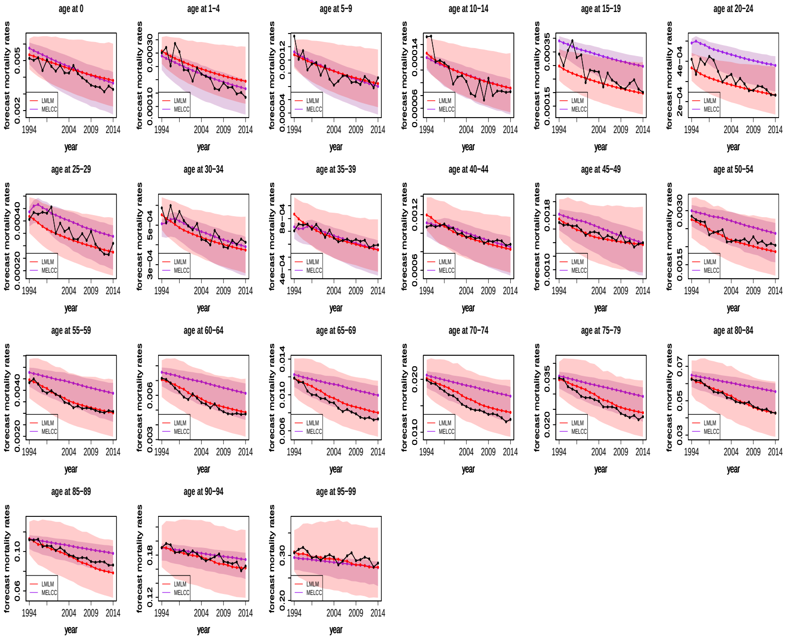
- training $Y_{x,1:(T-20)}$; &
- forecast $Y_{x,(T-19):T}$.

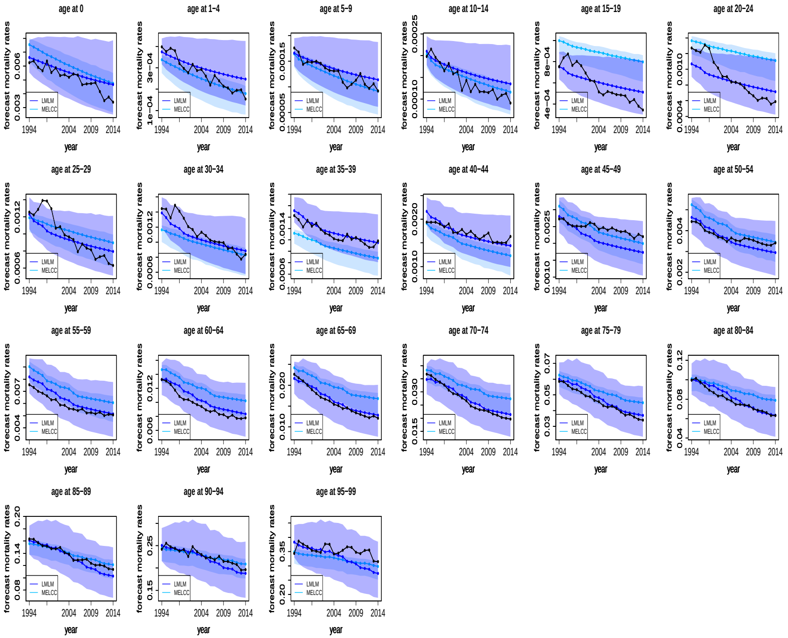
We visualise the performance of LMLM model compared with MELCC model for AU data by plotting the time series of:

- the observed $\boldsymbol{\mu}_{x,(T-19):T}$ (black line)
- the forecasted $\hat{\boldsymbol{\mu}}_{x,(T-19):T}$ by MELCC model (in red female, light blue male); &
- the forecasted $\hat{\boldsymbol{\mu}}_{x,(T-19):T}$ by LMLM model (in purple female, dark blue male)

with 95% credible interval for age groups (0-99).

The forecast performance of LMLM model with long memory component is better than MELCC model for all age groups!


age at 0
age at 1–4
age at 5–9
age at 10–14
age at 15–19
age at 20–24
age at 25–29
age at 30–34
age at 35–39
age at 40–44
age at 45–49
age at 50–54
age at 55–59
age at 60–64
age at 65–69
age at 70–74
age at 75–79
age at 80–84
age at 85–89
age at 90–94
age at 95–99
forecast mortality rates
year
1994
2004
2009
2014
LMLM
MELCC


age at 0
age at 1–4
age at 5–9
age at 10–14
age at 15–19
age at 20–24
age at 25–29
age at 30–34
age at 35–39
age at 40–44
age at 45–49
age at 50–54
age at 55–59
age at 60–64
age at 65–69
age at 70–74
age at 75–79
age at 80–84
age at 85–89
age at 90–94
age at 95–99
forecast mortality rates
year
1994
2004
2009
2014
LMLM
MELCC

The background material in this presentation are based on a selection of papers written for the actuarial audience over the last few years.

- Preprints of the works of Prof. Gareth W. Peters: (https://tinyurl.com/y9nrt4ar)

1. Fung M.C., Peters G.W., Shevchenko P.V. A unified approach to mortality modelling using state-space framework: characterisation, identification, estimation and forecasting. Annals of Actuarial Science. 2017 May:1-47.

   (SSRN: http://dx.doi.org/10.2139/ssrn.2786559)

2. Fung M.C., Peters G.W. and Shevchenko P.V. Cohort Effects in Mortality Modelling: A Bayesian State-Space Approach (March 24, 2017). Annals of Actuarial Science.

   (SSRN: https://ssrn.com/abstract=2907868)

3. Toczydlowska D., Peters G.W., Fung M.C. and Shevchenko P.V. Stochastic Period and Cohort Effect State-Space Mortality Models Incorporating Demographic Factors via Probabilistic Robust Principle Components (May 30, 2017) Risks, 5(3), 1-77. [42].

   (SSRN: https://ssrn.com/abstract=2977306)

HERIOT WATT UNIVERSITY

This talk is based around the following papers:

4 Fung M.C. and Peters G.W. and Shevchenko P.V. A State-Space Estimation of the Lee-Carter Mortality Model and Implications for Annuity Pricing (July 31, 2015). In MODSIM 2015: 21st International Congress on Modelling and Simulation : Proceedings (pp. 952-958). Canberra: Modelling & Simulation Society Australia & New Zealand.

(SSRN: https://ssrn.com/abstract=2699624)

5 Yan, Hongxuan and Peters, Gareth and Chan, Jennifer. Mortality Models Incorporating Long Memory Improves Life Table Estimation: A Comprehensive Analysis (March 26, 2018).

(SSRN: http://dx.doi.org/10.2139/ssrn.3149914.)

6 Yan, Hongxuan and Peters, Gareth and Chan, Jennifer. Multivariate Long Memory Cohort Mortality Models. (April 22, 2018).

(SSRN: http://dx.doi.org/10.2139/ssrn.3166884)

7 Yan, Hongxuan and Peters, Gareth and Chan, Jennifer. Reducing Model Risk and Improving Mortality Forecasts for Life Insurance Product Pricing (January 20, 2019).

(SSRN: https://ssrn.com/abstract=3319355)

8 Yan, Hongxuan and Peters, Gareth and Chan, Jennifer. Evidence for Persistence and Long Memory Features in Mortality Data. (January 25, 2019).

(SSRN: https://ssrn.com/abstract=3322611)

Yorkshire Actuarial Society

14/03/2019

19:00 – 20:00 Food & Drink

Thank you for attending

Please send your feedback
via the event invite