Modelos de Mortalidad

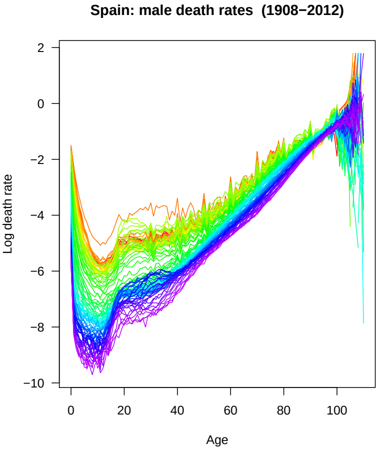
Riskcenter

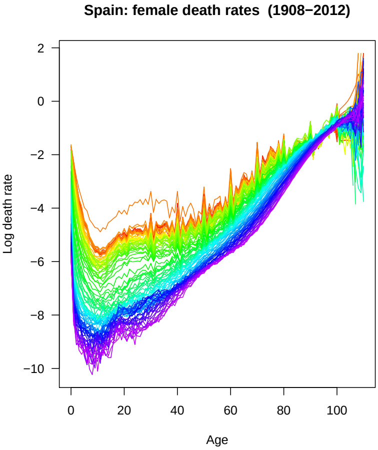
Universidad de Barcelona

http://www.ub.edu/riskcenter

15 de marzo de 2021

Gráficos de tasa brutas





Motivación.

- En el curso del siglo XX, la población del mundo industrializado experimentó un cambio en la mortalidad:
  - Por enfermedades infecciosas, afectando de manera notable a las edades jóvenes.
  - Reducción de mortalidad infantil y en edades adultas.

Motivación.

¿Por qué se necesita modelizar la mortalidad?

- La mortalidad en el futuro es incierta.
- Buena gestión del riesgo.
- Cálculo de Reservas.
- Valores vinculados a la mortalidad.

Notación.

- El año calendario se define como el tiempo que transcurre entre t y $t+1$.
- $m(t,x)$ es la tasa bruta de mortalidad para la edad x en el año calendario t,

$$m_{(x,t)} = \frac{D(x,t)}{E(x,t)}$$

  donde $D(x,t)$ y $E(x,t)$ son las muertes y las personas expuestas al riesgo en la edad x en el año t.
- En general los modelos se hacen respecto a $\log m_{(x,t)}$

Modelo Lee-Carter

- En lugar de resumir los datos en pocos parámetros, se crean parámetros para describir la mortalidad a cada edad y periodo además de un índice para el tiempo, para resumir las tendencias del pasado.
- Este índice se extrapola al futuro, obteniendo tablas de vida proyectadas.
- El modelo de Lee-Carter ha sido adoptado por el *US Bureau of the Census*, por lo que se ha convertido en una técnica de referencia para predecir la mortalidad.

Modelo Lee-Carter (1992)

$$\log m_{(x,t)} = \alpha_x + \beta_x \cdot \kappa_t + \epsilon_{x,t},$$

- $\alpha_x =$ la media del $\log m_{x,t}$ en los años calendario t, por lo tanto $\exp \alpha_x$ es la forma general de la mortalidad.
- κ_t representa la tendencia de la mortalidad en el tiempo. El Forecasting se hace tratando κ_t como una serie temporal.
- β_x describe las variaciones de $\log m_{(x,t)}$ en la edad x cuando κ_t varía en el tiempo.
- El de error $\epsilon_x(t)$ iid $\sim \mathcal{N}(0, \sigma^2)$ y reflejan los efectos que no son capturados por el modelo.

Modelo Lee Carter: Definición.

Los parámetros satisfacen que,

$$\sum_{t=t_{min}}^{t_{max}} \kappa_t = 0 \quad \text{y} \quad \sum_{x=x_{min}}^{x_{max}} \beta_x = 1$$

para asegurar "identifiability", que asegura que diferentes parámetros estimados devuelven diferentes probabilidades

Estimación de los parámetros.

- Los parámetros se obtienen por mínimos cuadrados ordinarios, es decir, minimizando:

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\kappa}}) = arg \min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa}} \sum_{x,t} \big(\ln m_{(x,t)} - \alpha_x - \beta_x \kappa_t \big)^2.$$

- Las restricciones impuestas a los parámetros garantizan la unicidad de la solución.
- Con la descomposición en valores singulares se obtienen las estimaciones de los parámetros.

Estimación de los parámetros.

- Derivando la función objetivo respecto de α_x se tiene,

$$\sum_{t=t_{min}}^{t_{max}} \ln \hat{m}_{(x,t)} = (t_{max} - t_{min} + 1)\alpha_x + \beta_x \underbrace{\sum_{t=t_{min}}^{t_{max}} \kappa_t}_{=0} .$$

- Las α_x se estiman como,

$$\hat{\alpha}_x = \frac{1}{t_{max} - t_{min} + 1} \sum_{t=t_{min}}^{t_{max}} \ln m_{(x,t)}.$$

Estimación de los parámetros.

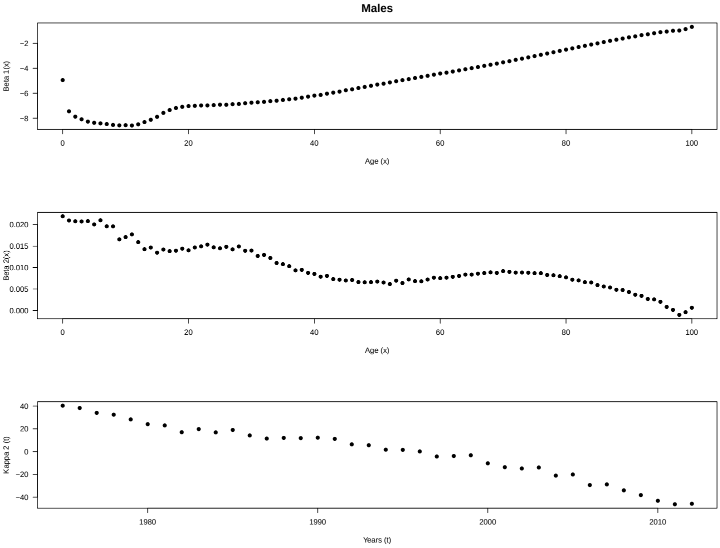
- Sea $\mathbf{Z}$ una matriz de dimensión $(x_{max} - x_{min} + 1) \times (t_{max} - t_{min} + 1)$ con elementos,

$$z_{xt} = \ln m_{(x,t)} - \hat{\alpha}_x.$$

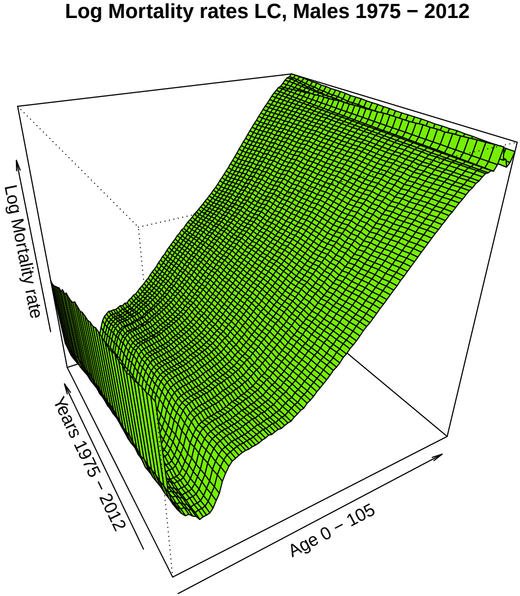
- Sea $\mathbf{u}_1$ y $\mathbf{v}_1$ vectores propio normalizado de $\mathbf{Z}^t\mathbf{Z}$ y $\mathbf{Z}\mathbf{Z}^t$; λ_1 el valor propio más grande .
- Entonces,

$$\hat{\boldsymbol{\beta}} = \frac{\mathbf{v}_1}{\sum_j v_{1j}}, \quad \hat{\boldsymbol{\kappa}} = \lambda_1 \left(\sum_j v_{1j} \right) \mathbf{u}_1.$$

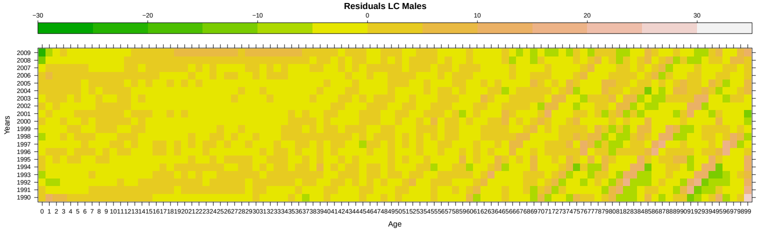
Modelo Lee-Carter - Parámetros estimados



Modelo Lee-Carter - Tasas de Mortalidad estimadas



Modelo Lee-Carter - Residuales



Modelo Renshaw-Haberman

lintenta representar el efecto cohorte, que se define como $c = t - x$.

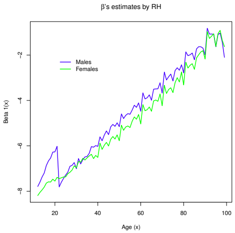
$$\log[m(x,t)] = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_{t-x}^{(3)} + \epsilon_{x,t}. \tag{1}$$

condicionado a

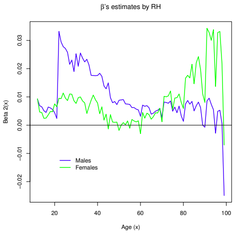
$$\sum_t \kappa_t^{(1)} = \sum_t \gamma_t^{(3)} = 0 \quad \text{y} \quad \sum_x \beta_x^{(2)} = \sum_x \beta_x^{(3)} = 1, \tag{2}$$

Los parámetros se interpretan al igual que en el modelo Lee-Carter. $\beta_x^{(2)}$ es la tasa de cambio de la mortalidad en el tiempo pero en referencia a la cohorte.

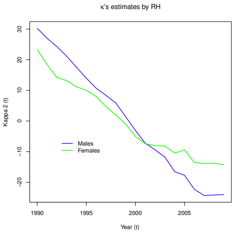
Modelo Renshaw-Haberman - Parámetros estimados



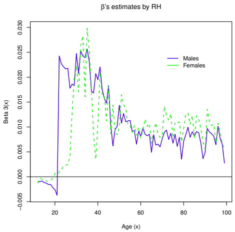
(a) $\hat{\beta}_x^{(1)}$



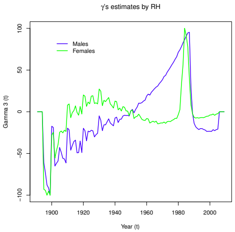
(b) $\hat{\beta}_x^{(2)}$



(c) $\hat{\kappa}_t^{(2)}$

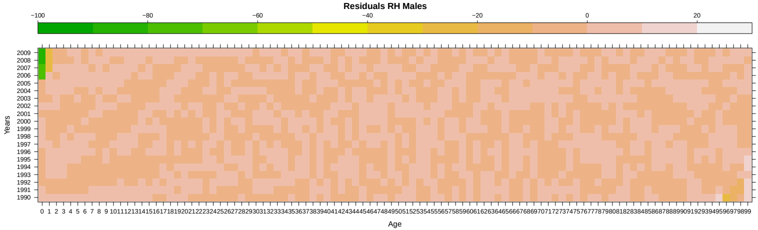


(d) $\hat{\beta}_x^{(3)}$



(e) $\hat{\gamma}_{t-x}^{(3)}$

Modelo Renshaw-Haberman - Residuales



Modelo Age-Period-Cohort (Currie, 2006)

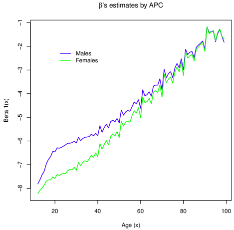
Los efectos de la edad, el período y la cohorte influyen en las tasas de mortalidad de forma independiente.

$$\log[m(x,t)] = \beta_x^{(1)} + \frac{1}{n_a}\kappa_t^{(2)} + \frac{1}{n_a}\gamma_{t-x}^{(3)} + \epsilon_{x,t}, \tag{3}$$

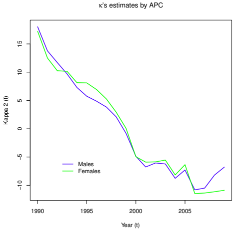
donde n_a es el total de edades. Se imponen las siguientes restricciones:

$$\sum_t \kappa_t^{(2)} = 0 \quad \text{and} \quad \sum_{x,t} \gamma_{t-x}^{(3)} = 0. \tag{4}$$

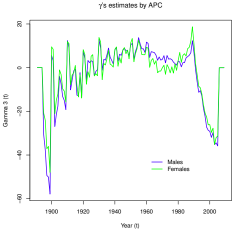
Modelo Age-Period-Cohort - Parámetros estimados



(f) $\hat{\beta}_x^{(1)}$

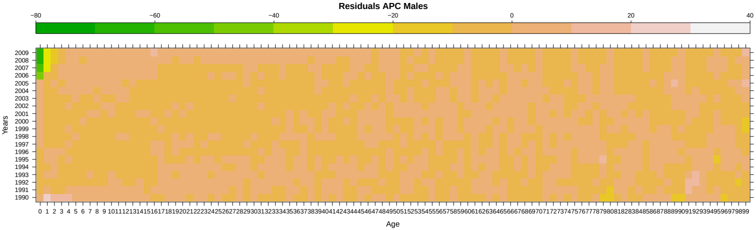


(g) $\hat{\kappa}_t^{(2)}$



(h) $\hat{\gamma}_{t-x}^{(3)}$

Modelo Age-Period-Cohort - Residuales



Ventajas e inconvenientes de las distintas aproximaciones.

- M5, CBD (2006).

$$\text{logit } q(t,x) = \log \frac{q(t,x)}{1-q(t,x)} = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)}.$$

- M6, Primera Generalización del CBD.

$$\text{logit } q(t,x) = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\kappa_{t-x}^{(3)}.$$

- M7, Segunda Generalización del CBD.

$$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \kappa_t^{(3)}\left((x-\bar{x})^2 - \hat{\sigma}_x^2\right) + \gamma_{t-x}^{(4)}.$$

- M8, Tercera Generalización del CBD.

$$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \gamma_{t-x}^{(3)}(x_c - x).$$

Ventajas e inconvenientes de las distintas aproximaciones.

- Todos los modelos tienen en cuenta que los efectos de la edad, periodo y la cohorte son qualitativamente diferentes y por lo tanto es necesario modelizarlos de diferente manera.
- Reconocen que existe aleatoriedad en las tasas de mortalidad en cada edad de un año a otro, causadas por factores externos como pueden ser los factores ambientales.
- Usando un modelo con splines se puede determinar una relación de suavizado entre las tasas de mortalidad de edades adyacentes dentro del mismo año.

Selección de modelos

Bayes Information Criterion (BIC)

- BIC penaliza la sobre parametrización de los modelos,

$$BIC = \hat{l}_x - \frac{1}{2} n \log N$$

donde $\hat{l}_x$ es el valor máximo de la función de verosimilitud, n es el número de parámetros y N el número de observaciones.

Selección de modelos

- Cálculo de los residuos estandarizados,

$$Z(t,x) = \frac{D(t,x) - E(t,x)\hat{m}(x,t;\hat{\Theta})}{\sqrt{E(t,x)\hat{m}(x,t;\hat{\Theta})}},$$

  se asume $Z(x,t)$ independientes e idénticamente distribuidos según una distribución normal standard $N(0,1)$

Selección de modelos

Likelihood ratio test,

- Ho: El modelo anidado es el modelo correcto vs. Ha: El modelo general es el modelo correcto
- Rechazar la hipótesis nula si

$$2(\hat{l}_2 - \hat{l}_1) > \chi^2_{\nu_2 - \nu_1, \alpha}$$

- Por ejemplo el Modelo LC está anidado en el Modelo RH. Las posibles comparaciones son: LC vs. RH, APC vs. RH, M5 vs. M6, M5 vs. M7, M5 vs. M8, M6 vs. M7 y M6 vs. M8,

Tablas de mortalidad para una población asegurada

Protocolo para tratamiento de los datos.

- Identificar inicio y final de la póliza para establecer el número de expuestos correctamente.
- Identificar fallecimientos y evitar duplicidades por existencia de asegurados con más de un contrato.
- Análisis gráfico, identificación de outliers.
- Graduación o suavizado.

Modelización, ajuste y contrastes.

- Tablas estáticas vs. dinámicas.
- Elección del modelo.
- Test de ajuste.
- Modelo elegido frente a modelos alternativos.
- Estabilidad y validez a futuro: (robustez) del modelo.

Variables adicionales de estudio a considerar en la construcción de tablas.

- **Pandemia**. Un incremento aislado de la mortalidad en el año, sin que exista cambio en las expectativas de mortalidad futuras.

  $\Rightarrow$ Se puede proponer como solución un aumento determinado de la tasa de mortalidad durante un año $q_x + \%$, basándose en el promedio de los escenarios optimistas y pesimistas.

Variables adicionales de estudio a considerar en la construcción de tablas.

- **Tendencia**. Un deterioro de las expectativas de mortalidad a largo plazo, el cual puede estar asociado a un cambio en el comportamiento de la población. Por ejemplo, aumento de obesidad que conduce a diabetes, infartos, hipertensión, etc.

  $\Rightarrow$ Se puede proponer como solución un aumento de la tasa de mortalidad en un cierto % por año durante todos los años futuros.

Variables adicionales de estudio a considerar en la construcción de tablas.

- **Incertidumbre** derivada de utilizar fuentes de información con datos escasos o poco fiables.
- **Volatilidad**. La mortalidad real fluctuará alrededor del nivel de mortalidad esperada. Las compañías grandes tenderán a experimentar menos volatilidad que las compañías pequeñas.

$\Rightarrow$ Método de simulaciones de Monte Carlo.

Referencias

Referencias.

- Ayuso, M.; Corrales, H.; Guillén, M. et al. Estadística Actuarial Vida. Barcelona: Universitat de Barcelona, 2001.
- Brouhns, N. y Denuit,M. *Risque de longévité et rentes viagères I. Evolution de la mortalité en Belgique de 1880 à nos jours*. Belgian Actuarial Bulletin, 2002, Vol.2, No. 1, 26-48.
- Brouhns, N. y Denuit,M. *Risque de longévité et rentes viagères II. Tables de mortalité prospectives pour la population belge*. Belgian Actuarial Bulletin, 2002, Vol.2, No. 1, 49-63.
- Brouhns, N. y Denuit,M. *Risque de longévité et rentes viagères III. Elaboration de tables de mortalité prospectives pour la population assurée belge, et évaluation du coût de l'antisélection*. Belgian Actuarial Bulletin, 2002, Vol.2, No. 1, 64-72.

Referencias.

- Cairns, A.; Blake, D.; Dowd, K. et al. *A Quantitative Comparison of Stochastic Mortality Models Using Data from England & Wales and the United States*. North American Actuarial Journal, 2009, Vol.13, No. 1, 1-35.
- Delwarde, A. y Denuit, M. *Importance de la période d'observation et des âges considérés dans la projection de la mortalité selon la méthode de Lee-Carter*. Belgian Actuarial Bulletin, 2003, Vol.3, No. 1, 1-21.
- Haberman, S.; Pitacco, E. Actuarial Models for Disability Insurance. United States: Cahpman & Hall/CRC, 1999.
- Knorr, F.E. *Multidimensional Whittaker-Henderson graduation*.Transactions of Society of Actuaries, 1984, Vol.36, 213-255.

Referencias.

- Pitacco, E.; Denuit, M. et al. Modelling Longevity Dynamics for Pensions and Annuity Business. United States: Oxford University Press, 2009.
- Renshaw, A.E. y Haberman, S. *On the forecasting of mortality reduction factors*. Insurance: Mathematics and Economics, 2003, Vol.32, No. 3, 379-401.