

MODELOS DE TABLAS DE MORTALIDAD EN ESPAÑA Y SITUACION ACTUAL

Amancio Betzuen Zalbidegoitia, Angie Felipe Checa y Monserrat
Guillén Estany.

*Director del Instituto de Estudios Financiero-Actuariales de la
Universidad del País Vasco. Profesoras de Econometría, Estadística y
Economía Española de la Universidad de Barcelona.*

ABSTRACT:

This paper surveys the most recent models of mortality applied to the Spanish population (or any subset); although it also reviews the most current models in Europe. We focus on models which incorporate both the age variation in mortality and the underlying time trends in the mortality rates.

KEYWORDS: Mortality, Forecasting, Spain, Graduation, Mortality Trends.

INTRODUCCION

¿Por qué en estos momentos hay una discusión tan grande sobre las tablas de mortalidad en España? ¿Qué expresiones matemáticas debemos utilizar para modelizar la probabilidad de fallecimiento?

Respecto de la primera pregunta diremos que no sólo en España, sino en toda Europa ha surgido la necesidad metodológica de crear unas tablas de mortalidad que incorporen la evolución que se viene observando en la mortalidad a finales de siglo. En Europa, el Estado de Bienestar ha promovido, por un lado, mejoras sanitarias importantes y, por otro, el "ahorro" para la jubilación bien sea de forma privada o pública. Sin embargo, prácticamente todos los países se han encontrado con la desagradable sorpresa de que los supuestos iniciales bajo los que se estableció el ahorro colectivo no se han cumplido. Todo lo contrario, las mejores perspectivas se han visto sobrepasadas por la realidad: la probabilidad de fallecimiento para las diferentes edades ha sufrido variaciones muy importantes cuando consideramos el medio y el largo plazo; aún más, no todas ellas del mismo signo. Gráficamente diríamos que no se ha producido simplemente un cambio de nivel en la curva de la probabilidad

de fallecimiento (q_x), sino también, modificaciones en la forma de la propia curva.

Es interesante remarcar las palabras con las que Navarro (1991) acaba su libro:

“... , es de esperar que, siguiendo la tónica general de este siglo, la mortalidad actual sea ya algo inferior a la recogida en el presente estudio.”

Respecto a la segunda cuestión planteada, los modelos propuestos más recientemente para el territorio español y por orden cronológico son los de: Navarro (1991, modelo “Logit Gompertz-Makeham”), Rué (1992, modelo “Heligman-Pollard”), Prieto y Fernández (1994, modelo “Makeham”), Betzuen (1997, modelo “Splines”) y Artís y Felipe (1997, modelo “Heligman y Pollard”). Cada uno de ellos partiendo de hipótesis comunes modeliza la supervivencia de una generación (personas nacidas en un mismo año) con diferentes expresiones funcionales utilizando las técnicas del análisis individual para la obtención del modelo.

De todos estos modelos dos son dinámicos: el primero, de Prieto y Fernández (1994), incorpora la dinamicidad en la esperanza de vida; el segundo, Artís y Felipe (1997), no sólo es un modelo dinámico para la probabilidad de fallecimiento sino que además permite la construcción de escenarios para evaluar en el futuro los diferentes impactos epidemiológicos (SIDA, enfermedades coronarias, etc..), políticas sanitarias o diferentes políticas de marketing, en el caso de una cartera de asegurados.

En cuanto a la utilización de las tablas de mortalidad de otros países ha quedado demostrado (Betzuen, 1997; Martínez 1997 y Artís et al., 1998) que sobrevaloran sistemáticamente la mortalidad española para edades mayores, con el consiguiente riesgo financiero que representa para la entidad aseguradora la dotación insuficiente de reservas⁸ matemáticas de vida y de jubilación.

⁸ Aunque el nombre contable actual en las entidades aseguradoras es provisiones, hemos mantenido el de reservas por ser el utilizado por el Banco de España haciendo referencia a la Contabilidad Nacional.

Por tanto, la respuesta a la segunda pregunta no es única, aunque creemos que tratándose de la población general, la graduación bien pudiera ser nacional: sea cual sea el modelo elegido debe aplicarse, en un primer estudio, a la población (o subpoblaciones) española.

A continuación presentamos una revisión sucinta sobre los modelos usuales en España y en los países de nuestro entorno inmediato, de forma que podamos situar donde se encuentran ubicados los modelos españoles en el contexto metodológico internacional.

EXPRESIONES PARA LAS $\{q_x\}$ DE USO MAS COMÚN

El objetivo que se persigue es el de conocer q_x . Para ello modelizamos la probabilidad de fallecimiento q en función de la edad x . Parece lógico plantearse la hipótesis: $q=f(x, \dots)$. Se trata de hallar una expresión o modelo de comportamiento de la probabilidad de fallecimiento para las diferentes edades. Muchas han sido las expresiones matemáticas propuestas y un buen análisis de ellas puede hallarse en Forfar, McCutcheon y Wilkie (1988) (en adelante FMW) y Benjamin y Pollard (1989), entre otros. Nosotros vamos a hacer un breve resumen de éstas dividiéndolas en tres partes: expresiones para un rango de edades, expresiones para todo el rango de edades y expresiones splines; para con posterioridad hacer hincapié en la componente dinámica.

PARA UN RANGO DE EDADES

Son expresiones que se aplican a un rango concreto de edades, aunque con posterioridad puedan concatenarse hasta conseguir diferentes expresiones que abarquen todas las edades. Existen diferentes procesos de modelización en cuanto que unos eligen como información el tanto instantáneo de mortalidad o el tanto anual de mortalidad o el tanto central de mortalidad e incluso, también, un ratio entre tantos anuales. Como decimos, algunas de ellas modelizan el tanto instantáneo de mortalidad μ_x , también llamada fuerza de mortalidad a la edad x (siguiendo la descripción inglesa).

Ésta se define como el ratio de los fallecidos en el intervalo infinitesimal entre la edad x y $x+dt$ y los supervivientes al principio del período. La

relación entre la probabilidad de fallecimiento y el tanto instantáneo de mortalidad para una persona se establece como:

$$q_x = 1 - \exp\left[-\int_0^1 \mu_{x+t} dt\right]$$

Ecuación 1

Las más conocidas son:

- **Gompertz (1825):** utilizada básicamente para las edades adultas-seniles.

$$\mu_x = Bc^x \quad \text{con} \quad q_x = 1 - \exp\left[\frac{-B(c-1)}{\ln c} c^x\right]$$

Ecuación 2

- **Makeham (1867):** constituye una mejora de la ley de Gompertz. También es aplicada preferentemente en edades adultas-seniles. En España, las tablas de mortalidad elaboradas por Prieto fueron construidas aplicando siempre esta ley para la población española, incluso Prieto y Fernández (1994) la utilizaron recientemente.

$$\mu_x = A + Bc^x \quad \text{con} \quad q_x = 1 - \exp\left[-A - \frac{B(c-1)}{\ln c} c^x\right]$$

Ecuación 3

- **Barnett (CMI⁹-74):**

Ésta es la primera expresión en que se ajusta el ratio de probabilidades, aunque no llega a ser un logit puesto que al otro lado de la igualdad encontramos una expresión de tipo polinómica.

$$\frac{q_x}{p_x} = A + Hx + Bc^x$$

Ecuación 4

⁹ Continuous Mortality Investigation.

• **Wilkie (CMI-76):**

$$\frac{q_x}{p_x} = \exp\{pol(x)\}$$

Ecuación 5

donde $pol(x)$ es el polinomio en x apropiado con un grado bajo: lineal o cuadrático.

Todas ellas pueden considerarse casos particulares de una expresión más general. La expresión a la que nos referimos se puede deducir y expresar en la forma que a continuación presentamos.

Dados:

1. r y s números enteros no negativos, no ambos cero y
2. $\alpha = (\alpha_1 \alpha_2 \dots \alpha_{r+s})$ el vector de coeficientes

Se define la fórmula **Gompertz-Makeham** de tipo (r,s) como:

$$GM_{\alpha}^{r,s}(x) = \sum_{i=1}^r \alpha_i x^{i-1} + \exp\left\{ \sum_{i=r+1}^{r+s} \alpha_i x^{i-r-1} \right\}$$

si $r = 0$

$$GM_{\alpha}^{0,s}(x) = \exp\left\{ \sum_{i=1}^s \alpha_i x^{i-1} \right\}$$

si $s = 0$

$$GM_{\alpha}^{r,0}(x) = \sum_{i=1}^r \alpha_i x^{i-1}$$

Ecuación 6

Como casos más notables hemos destacado $r=0$ y $s=0$.

La Ecuación 6 también puede expresarse como:

$$GM_{\alpha}^{r,s}(x) = pol_1(x) + \exp\{pol_2(x)\}$$

Ecuación 7

con $\text{pol}_1(x)$ de orden r y $\text{pol}_2(x)$ de orden s .

Se define la fórmula ‘**Logit Gompertz-Makeham del tipo (r,s)**’¹⁰ como:

$$LGM_{\alpha}^{r,s}(x) = \frac{GM_{\alpha}^{r,s}(x)}{1 + GM_{\alpha}^{r,s}(x)}$$

Ecuación 8

entonces $q_x = GM_{\alpha}^{r,s}(x)$ o bien $q_x = LGM_{\alpha}^{r,s}(x)$; que frecuentemente produce ajustes satisfactorios, a menudo con valores pequeños de r y s .

Según FMW (1988):

“El rango de valores correspondiente a $LGM_{\alpha}^{r,s}(x)$, al ser un Logit, va de 0 a 1, por ello esta fórmula es más adecuada para los valores de q^A (probabilidad de fallecimiento ajustada) en vez de la fórmula Gompertz-Makeham”.

El LGM ha sido utilizado para la población española por Navarro (1991) para el período 81-82 con $r=3$ y $s=3$. Para el tramo inferior se estimaron dos cúbicas enlazadas en la edad β , en el tramo principal el LGM(3,3) entre las edades n y m , en el tramo superior dos cúbicas enlazadas en τ . En total 22 parámetros, más la deducción de las edades de enlace entre las diferentes expresiones. Para las edades superiores a 95 años, Navarro ajusta una parábola.

PARA TODO EL RANGO DE EDADES

Son expresiones matemáticas que tratan de abarcar todo el rango de edades: desde el primer año de vida. Un estudio bastante completo de dichas leyes lo encontramos en Hartmann (1983).

¹⁰ Esta expresión es utilizada por FMW (1988), aunque para un análisis más exhaustivo de las funciones link para q_x ver Renshaw (1991).

- **Thiele (1872):** Fue el primer autor en proponer un modelo de este tipo. La expresión estaba compuesta de tres sumandos que se correspondían a tramos de edades diferentes: infantil, juvenil-adulta y adulta-senil, correspondiéndose esta última con la ley de Gompertz. Sólo precisaba 7 parámetros. Una expresión de estas características es la que se presenta a continuación referida al tanto instantáneo:

$$\mu_x = a_1 e^{-b_1 x} + a_2 e^{\left[-\frac{1}{2} b_2 (x-c)^2\right]} + a_3 e^{b_3 x}$$

Ecuación 9

- **Perks (1932):**

$$\mu_x = \frac{A + Bc^x}{Kc^{-x} + 1 + Dc^x}$$

Ecuación 10

- **Logit de Brass (Zaba, 1979):** Basándose en estudios empíricos, se contrastó que realizando una transformación Logit de las funciones de supervivencia en las tablas de vida, al menos de una forma aproximada, estaban linealmente relacionadas.

$$\text{logit } l_x = \alpha + \beta \text{ logit } l_x^s \quad \text{y} \quad \text{logit } l_x = 0.5 \cdot \ln \frac{l_0 - l_x}{l_x}$$

Ecuación 11

donde l_x^s es la función de supervivencia (ajustada) de una tabla estándar (el mismo Brass propuso dos), o de otra tabla que se desee comparar.

- **Heligman y Pollard (1980):** inspirándose en Thiele ajustan una nueva ley a la mortalidad de Australia de la postguerra. Su expresión genérica es:

$$\frac{q_x}{1 - q_x} = \sum_{i=1}^n A_i e^{-B_i \{f_i(x) - C_i\}^{D_i}}$$

Ecuación 12

donde $i=1, \dots, n$ son los sumandos de la expresión del modelo. Normalmente con $n=3$ se obtienen buenos ajustes. Los parámetros a

estimar son los A_i, B_i, C_i, D_i con $i=1, \dots, n$; en el caso de tres sumandos serían doce parámetros a estimar. Respecto a $f_i(x)$ suele utilizarse x o $\ln x$.

LAS LEYES DE HELIGMAN Y POLLARD

La expresión que realmente se ajustó a la mortalidad australiana fue:

$$\frac{q_x}{p_x} = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^{(x-x_0)}$$

Ecuación 13: Ecuación 1 de H-P.

expresión que consideran no puede distinguirse de:

$$q_x = A^{(x-B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + GH^x}$$

Ecuación 14: Ecuación 1 a de H-P.

Además, para mejorar el ajuste en edades seniles propusieron y ajustaron las siguientes expresiones:

$$q_x = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + KGH^x}$$

Ecuación 15: Ecuación 2 de H-P.

$$q_x = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^{x^k}}{1 + GH^{x^k}}$$

Ecuación 16: Ecuación 3 de H-P.

La curva básica es la de la Ecuación 13 y presenta las siguientes ventajas:

1. Es una función continua.
2. Es aplicable a todo el rango de edades $0 \leq x < \infty$.
3. La probabilidad ajustada está entre los límites apropiados:
 $0 \leq q_x^A \leq 1$.
4. Tiene pocos parámetros si consideramos que es para todo el rango de edades.
5. Los parámetros tienen una interpretación demográfica o biológica.

6. Es suficientemente flexible para ajustar una gran variedad de modelos de mortalidad.

Cada uno de los tres términos de la Ecuación 13 representa un componente distinto de la mortalidad.

- 1) El primer sumando es una exponencial de decrecimiento rápido. Refleja la caída de la mortalidad durante los primeros años infantiles, conforme el niño se adapta a su nuevo ambiente y gana inmunidad contra las enfermedades del mundo exterior.

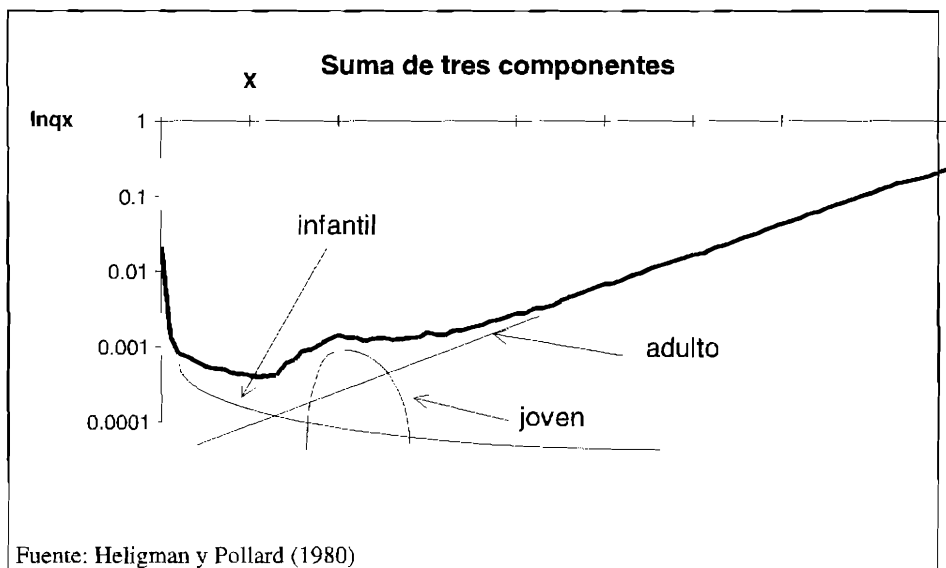


Ilustración 1: Componentes de la Ley Heligman y Pollard.

El parámetro A tiene un valor similar a la q_1 , la probabilidad de fallecimiento durante el primer año de vida. Mide el nivel de mortalidad.

El parámetro B mide la localización de la mortalidad en el primer año de vida entre el rango $(q_1, 0.5)$. Cuando $B=0$ $q_0=0.5$, conforme aumenta B , ceteris paribus, más próximo está q_0 de q_1 . En la práctica B está próximo

a 0 (entre 0.01 y 0.03) y su efecto sobre las probabilidades de fallecimiento en edades distintas de 0 es prácticamente inexistente¹¹.

El parámetro C mide la disminución de la probabilidad de fallecimiento en la infancia para edades mayores de un año: es la rapidez con la cual el niño se adapta a su entorno. A mayor C más rápidamente decrece la mortalidad conforme aumenta la edad.

- 2) El segundo sumando es una función similar a la lognormal. Refleja la mortalidad por accidentes para hombres y accidentes más mortalidad maternal para la población femenina. Se trata de una mortalidad adicional, sobrepuesta a la “curva natural de la mortalidad” que le correspondería según la edad tal y como describen los otros dos términos de la Ecuación 13. La “joroba de los accidentes” se encuentra en todas las poblaciones y aparece o bien como una joroba distinguible en la curva de mortalidad o bien como un mantenimiento (meseta) de las probabilidades de fallecimiento. Generalmente se da entre los 15 y los 30 años, aproximadamente.

El parámetro D establece la severidad de la joroba: a mayor D mayor será el efecto de los accidentes sobre la mortalidad total.

El parámetro E indica el inverso de la dispersión: a mayor E menor dispersión.

El parámetro F indica la localización de la “joroba de los accidentes”, el máximo.

- 3) El tercer sumando es la ley de Gompertz. Refleja el crecimiento casi geométrico de la mortalidad en edades adultas y representa la deterioración del cuerpo con los años, es decir, la mortalidad senil.

El parámetro G indica el nivel base de la mortalidad senil. Técnicamente sería la mortalidad adulta-senil al nacer.

¹¹ Aunque ésta es la interpretación dada por los autores, algunos estudios de simulación posteriores matizan dicha interpretación.

El parámetro H es el tanto de crecimiento de dicha mortalidad en edades adultas.

La edad x_0 es una edad cercana a ω (edad límite). Es aquella edad para la cual $q_x=0.5$ (también llamada 'span of life').

Es importante señalar lo que dice Hartmann (1979) respecto de la interpretación de los parámetros:

“..la interpretación de los parámetros es puramente mecánica. Dicho de otra forma, mientras las interpretaciones de los parámetros al uso indican cómo afectan al modelo de la ley de mortalidad, éstos no pueden utilizarse para explicar la naturaleza de las circunstancias biológicas o ambientales que subyacen en la mortalidad empírica.”

La Ecuación 15 y la Ecuación 16 mejoran el ajuste especialmente para edades mayores: en el caso de que $\ln(q_x/p_x)$ sea convexo (curvatura hacia arriba) la Ecuación 15 es la más adecuada; mientras que si $\ln(q_x/p_x)$ es cóncavo hacia abajo la Ecuación 16 es más apropiada. Cuando $K=1$ ambas ecuaciones se convierten en la Ecuación 14. Conforme mayor es K mayor es la curvatura hacia abajo, cuando K decrece mayor es la curvatura hacia arriba. Es el único parámetro que puede llegar a tomar valores negativos.

Heligman y Pollard (1980) consideran que la interpretación realizada para los parámetros de la Ecuación 13 sigue siendo válida para las demás expresiones. No obstante, los valores de G y H para la Ecuación 16 están muy afectados por la dirección y extensión de la curvatura K; y no queda claro que todavía tengan una interpretación demográfica, aunque x^K puede entenderse como una transformación de la escala de las edades.

Rué (1992) aplicó la Ecuación 14, la Ecuación 15 y la Ecuación 16 a la población española para los datos agrupados de 1979-1982. El mejor ajuste lo obtuvo para la Ecuación 15.

En Artís y Felipe (1997) se ajustó la expresión descrita en la Ecuación 15 para cada uno de los 19 años en estudio: 1975-1993. Previamente se había contrastado la dinamicidad de la mortalidad mediante los test habituales, de forma que no podía considerarse el mismo proceso generador de datos para

dos años consecutivos. Al ajustar la misma forma funcional para cada uno de los años se consigue una continuidad en las tablas obtenidas, que además pueden utilizarse para la obtención de predicciones. De esta forma, las tablas de mortalidad obtenidas para la población española, sin recargos de ningún tipo, pueden ser utilizadas por las entidades aseguradoras como tablas de referencia básica en la elaboración de tablas propias de mortalidad, estableciéndose posteriormente los recargos de seguridad que la entidad estime oportunos, en los niveles que se consideren adecuados para la capacidad financiera o la aversión al riesgo de cada entidad.

SPLINES: COMO FÓRMULA ALTERNATIVA

No sólo los modelos $GM_{\alpha}^{r,s}(x)$, $LGM_{\alpha}^{r,s}(x)$ o Heligman y Pollard pueden aplicarse en tablas de mortalidad, ya que se pueden utilizar muchas otras funciones. Un instrumento alternativo, actualmente muy al uso como primer paso en modelos dinámicos, es la utilización de splines, que puede ser útil en una amplia variedad de situaciones prácticas. Con ellos se han ajustado las dos últimas tablas de mortalidad nacionales inglesas y el modelo austríaco de tablas de mortalidad, el alemán y el suizo se han servido de ellos en su ajuste. En España, Betzuen (1997) los ha utilizado en las tablas de mortalidad de colectivos de trabajadores activos ocupados.

Más que un ajuste, el spline realiza un suavizado¹², en nuestro caso respecto de las edades, de las q_x observadas.

Un spline es una función polinómica definida a intervalos (o a trozos) para la cual existe el máximo número de derivadas posible.

Sean x las edades o datos de edades, t los nudos y k los grados del polinomio. Entonces, dada una secuencia de números reales estrictamente creciente $a=t_0 < t_1 < t_2 \dots < t_{m+1} = b$. Un spline $s(x)$ de grado k , con “nudos” internos t_1, t_2, \dots, t_m es una función definida en el intervalo $[a, b]$ que tiene las siguientes propiedades:

¹² Según describían la característica de suave FMW (1988)

- a) En cada intervalo $[t_i, t_{i+1}]$ ($i=0,1,2,\dots,m$) $s(x)$ viene dado por un polinomio en x de grado k o inferior.
- b) $s(x)$ y sus derivadas de órdenes $1,2,\dots,k-1$ son continuas en cualquier parte del intervalo $[a,b]$.

Es esta última condición la que distingue los splines de otro tipo de funciones polinómicas a trozos.

De todo lo anterior se desprende que:

$$s(x)=p_i(x)$$

Ecuación 17

En particular, el polinomio $p_\theta(x)$ de grado k sobre el intervalo $[x_0,x_1]$ se define por $(k+1)$ parámetros. En la práctica, los splines más utilizados son los cúbicos, esto es $k=3$.

La forma práctica de operar con funciones spline es la siguiente: puesto que el valor de la función $p_\theta(x)$ y sus primeras $(k-1)$ derivadas están definidas en el nudo x_1 éstos deberían ser los valores de $p_j(x)$ y sus primeras $(k-1)$ derivadas en x_1 . El siguiente paso consiste en especificar un nuevo parámetro para el polinomio $p_j(x)$ tal como el valor de su k -ésima derivada en x_j .

Procediendo de esta manera se alcanza:

$$p_i(x)=p_0(x)+\sum_{j=1}^n C_{r+1+j}(x-t_j)^r$$

Ecuación 18

con,

$$p_0(x)=\sum_{j=0}^r C_{j+1} x^j$$

Ecuación 19

sujeto a, en los puntos de unión, las siguientes condiciones:

$$\begin{aligned}
 p_{j-1}(t_j) &= p_j(t_j) \\
 p'_{j-1}(t_j) &= p'_j(t_j) \\
 p''_{j-1}(t_j) &= p''_j(t_j)
 \end{aligned}$$

Ecuación 20

La función a optimizar (minimizar en este caso) se puede escribir de la forma:

$$F = \sum_{x_1}^t w_x [q_x^* - p_0(x)]^2 + \dots + \sum_{t_m}^{x_n} w_x [q_x^* - p_m(x)]^2$$

Ecuación 21

siendo,

q_x^* : los valores brutos, y

w_x : las ponderaciones.

El tratamiento matricial permite simplificar la fórmula anterior, minimizando F para averiguar el vector de coeficientes

$$C = (A^T W A)^{-1} A^T W Q$$

Ecuación 22

siendo,

A: la matriz de componentes

W: la matriz de ponderaciones

Q: el vector de probabilidades brutas

con

$$w_j = \frac{n_j}{\bar{n}}$$

Ecuación 23

n_j : número de observaciones a la edad j

\bar{n} : número medio del total de observaciones,

y

$$(x_i - x_j)_-^r = \begin{cases} (x_i - x_j)^r & \text{si } x_i - x_j > 0 \\ 0 & \text{si } x_i - x_j \leq 0 \end{cases}$$

Ecuación 24

cuando estas componentes constituyen una parte de A.

Recientemente se está llevando a la práctica la técnica de “nudos variables”.

Para unos nudos internos especificados $\{t_1, t_2, \dots, t_m\}$, se necesitan n coeficientes para definir un spline s de grado k , por consiguiente $n=m+k+1$. (El spline es simplemente una combinación lineal de las funciones ‘base’ adecuadas).

Podemos escribir $s(x)=f_\alpha(x)$ con $\alpha=(\alpha_1 \alpha_2 \dots \alpha_n)$ el vector de coeficientes que define el spline. $f_\alpha(x)$ es lineal en cada coeficiente $\{\alpha_1 \alpha_2 \dots \alpha_n\}$. Al considerar la posición de los nudos como más coeficientes, podemos definir un spline con m -nudos de grado k en el intervalo $[a,b]$. En este caso sea $n=m+(m+k+1)=2m+k+1$ y $s(x)=g_\alpha(x)$ con $\alpha=(\alpha_1 \alpha_2 \dots \alpha_n)$ el vector de coeficientes. Ahora los coeficientes $\alpha_1 \alpha_2 \dots \alpha_m$ deben estar en el intervalo (a,b) y especificar la posición de los nudos, mientras que $\alpha_{m+1} \alpha_{m+2} \dots \alpha_n$ determinan el spline en relación a esos nudos (debidamente ordenados) para una base apropiada.

Esta última situación da lugar a la técnica de ‘nudos-variables’ para el ajuste por splines. Es importante darse cuenta que aunque en el caso de $g_\alpha(x)$ no lineal en $\alpha_1 \alpha_2 \dots \alpha_m$ (la posición de los nudos), no hay diferencias en principio respecto a las fórmulas más tradicionales y los splines. En todo caso la fórmula para el ajuste de curvas es definida por un vector de coeficientes α apropiados.

MODELOS QUE INCORPORAN LA DINAMICIDAD DE LAS $\{q_x\}$

Cuando se trata de modelizar directamente la dinamicidad de la mortalidad las propuestas son variadas. Nolfi (1960) considera que la mortalidad decrece de forma exponencial para cada edad. Así para un momento $t > t_0$ la expresión que relaciona las probabilidades sería:

$$q_x(t) = q_x(t_0) \cdot e^{-\alpha(t-t_0)}$$

Ecuación 25

siendo α diferente para cada x , y representa la intensidad de decrecimiento relativo de la mortalidad. Siendo su expresión:

$$\alpha_x = \frac{\ln 2}{T_x}$$

Ecuación 26

Con T_x el tiempo necesario para que el tanto de mortalidad a la edad x disminuya a la mitad del valor considerado a inicio.

Este método ya fue propuesto en España por Lecina (1990) y ha sido utilizado en la elaboración de las tablas de mortalidad suizas GRM/GRF 1995 (Maeder, 1997).

En España Prieto y Fernández (1994) modelizan la dinamicidad de las tablas de mortalidad, aunque en este caso a través de la esperanza de vida para cada edad considerando la siguiente expresión:

$$e_x(t) = \alpha + \beta [1 - e^{-\delta t}]$$

Ecuación 27

Y la relación conocida:

$$q_x = \frac{\frac{1}{2} - e_{x-1}}{e_x - \frac{1}{2}}$$

Ecuación 28

Para las tablas alemanas DAV 1994 (Schütz, R; 1997) se ha considerado que los $\ln q_x$, para cada edad, presentaban tendencia lineal decreciente en el tiempo ajustada a través de mínimo cuadrados ordinarios. El modelo representativo para cada edad:

$$\ln q_x(t) = \alpha_x - \beta_x t$$

Ecuación 29

En definitiva, se modelizan, fijada la edad, las probabilidades de fallecimiento a lo largo del tiempo mediante un proceso no estacionario en varianza y no estacionario en media. La primera se corrige al tomar logaritmo neperiano; la segunda se considera debida a la existencia de una tendencia lineal determinista.

Renshaw, Haberman y Hatzopoulos (1996) modelizan la dinamicidad de la mortalidad a través del tanto instantáneo de mortalidad μ_x . Dichos autores consideran el $\ln \mu_x$ como un polinomio (fórmula Gompertz-Makeham) en la edad más un polinomio en el tiempo. Estos polinomios no están expresados como sucesivas potencias de la edad x , sino que se ha cambiado la base por los polinomios ortogonales de Legendre¹³, los cuales no requieren ningún tipo de ponderación pero sí un cambio de escala en las edades¹⁴ y el tiempo calendario para que se circunscriban al intervalo $[-1,1]$. El modelo finalmente ajustado ha sido:

$$\mu_x(t) = \exp\left\{\beta_0 + \sum_{j=1}^5 \beta_j L_j(x')\right\} \exp\left\{\left[\alpha_1 + \sum_{j=1}^3 \gamma_{1j} L_j(x')\right]t + \left[\alpha_2 + \sum_{j=1}^3 \gamma_{2j} L_j(x')\right]t^2\right\}$$

Ecuación 30

Expresión que modeliza a la vez el tanto instantáneo de mortalidad respecto de la edad y del tiempo, a través de 14 parámetros únicamente. Respecto del tiempo, los autores también están considerando un proceso no estacionario en varianza y media, con la tendencia cuadrática.

McNown y Rogers (1988) modelizan la dinamicidad del comportamiento de la mortalidad para el rango completo de edades. Para ello utilizan el modelo de Heligman-Pollard 1b (1980) ajustado para un período de 84 años empezando en 1990 y consideran que el tiempo ha afectado a la estructura del modelo. Modelizan los ocho parámetros mediante métodos de series

¹³ Ver Polinomios ortogonales en Anexo.

¹⁴ Ver escalonado de edades en Anexo.

temporales, aplicando el análisis univariante y de esta forma predicen la mortalidad para cada una de las edades. En este caso la expresión sería:

$$q(x, t) = A_t^{(x+B_t)^{C_t}} + D_t e^{-E_t(\ln x - \ln F_t)^2} + \frac{G_t H_t^x}{1 + G_t H_t^x}$$

Ecuación 31

con, por ejemplo, el siguiente modelo ARIMA para el parámetro A:

$$\Delta^2 \ln A(t) = \xi(t) - 1.08\xi(t-1) - 0.28\xi(t-2)$$

(0.18) (0.19) $\xi \approx N(0, \sigma)$

Ecuación 32

En Artís y Felipe (1997) se sigue esta línea de investigación pero, por primera vez, aplicada a la Ecuación 15. Ésta incorpora un parámetro adicional: el parámetro K relacionado con la curvatura de la probabilidad de fallecimiento en edades adultas-seniles no recogidas por la ley de Gompertz. Por tanto, el total de parámetros en estudio son nueve. Además se obtienen bandas de confianza sobre los parámetros futuros que serán repercutibles sobre la ley de mortalidad predicha para el futuro.

La expresión finalmente ajustada es:

$$q(x, t) = A_t^{(x+B_t)^{C_t}} + D_t e^{-E_t(\ln x - \ln F_t)^2} + \frac{G_t H_t^x}{1 + K_t G_t H_t^x}$$

Ecuación 33

Donde el período en estudio, t, va de 1975 a 1993 y el período de predicciones, T, va de 1994 al 2010.

La mayoría de los parámetros presentan un comportamiento autorregresivo, por tanto de memoria finita. Ello comporta que las predicciones realizadas con dicho modelo sólo tengan validez, en tanto en cuanto no exista ningún "shock" posterior al último año observado, es decir que haya algún avance médico-sanitario importante o algún proceso epidemiológico nuevo (recordemos la aparición del SIDA en 1981 y sus posteriores efectos difícilmente previsibles). En este punto la capacidad de pronóstico del modelo se hace mucho más relevante, pues se puede realizar diferentes

escenarios con los shocks que el investigador o la institución que lo utilice, puedan prever.

En conclusión este modelo se ajusta bien a la probabilidad de los hombres, aunque no acaba de explicar toda la variabilidad observada en las mujeres; y en su vertiente dinámica funciona: es capaz de realizar predicciones verosímiles y pronósticos.

RESUMEN DE LOS MODELOS ESPAÑOLES

Para finalizar, a la luz de la revisión metodológica realizada, valdría la pena repasar uno a uno los modelos que hoy en día se han implementado en España.

UNESPA tradicionalmente utiliza las tablas de mortalidad elaboradas por el INE como datos de base para ajustar la Ley de Makeham, procedimiento que han seguido utilizando Prieto y Fernández (1994). Aunque con posterioridad, dichos autores obtienen las esperanzas de vida para cada edad y año de calendario de forma que sobre ellas aplican una función de crecimiento exponencial que les sirve para predecir los valores futuros de las esperanzas de vida y a través de ellas las probabilidades de supervivencia, siempre dentro de la Ley de Makeham. Esta utilización de la esperanza de vida para modelizar la dinamicidad de la probabilidad de fallecimiento contiene algunos aspectos aún no resueltos: el cálculo de la esperanza de vida en una determinada edad, puede quedar muy alterado en función del coeficiente de reparto de las defunciones utilizado¹⁵. Cabe estudiar si éstas presentan un componente estacional como algunos autores apuntan (Vinuesa, 1994). Además la esperanza de vida sólo se ha calculado hasta los 100 años, cuando es muy posible que el “span of life”¹⁶ también presente una evolución en el tiempo.

Navarro (1991) fue el primero en aplicar el trabajo de los actuarios británicos a las tablas de mortalidad ajustando un logit de Gompertz-Makeham para los datos del censo de 1981. Los métodos estadísticos

¹⁵ Esto es, del comportamiento de las defunciones dentro de un año para las diferentes edades

¹⁶ Llamamos duración de la vida a la edad $x_0/q_x=0.5$

empleados establecen un antes y un después a su trabajo. En él se aplican, por primera vez en España, intervalos de confianza y una serie de test sobre el ajuste. Las principales críticas a su trabajo se circunscriben a los datos empleados: se utiliza la experiencia ponderada de dos años, sin realizar ningún test previo; las probabilidades de fallecimiento son obtenidas bajo la hipótesis de linealidad en los fallecimientos; los intervalos de confianza no tienen en cuenta la propia función de distribución de la variable y son obtenidos en base a la convergencia con la distribución normal. Por último, su estudio no incorpora ningún tipo de dinamicidad.

Seguidamente, Rué (1992) incorpora los avances de Navarro (1991) pero sobre los modelos de Heligman y Pollard (1980) para la población española y catalana en los años censales. Se trata de un trabajo pionero, que investiga con exhaustividad las diferentes fórmulas de Heligman y Pollard (1980) para ambas poblaciones.

Por otra parte, Betzuen (1997) aplica por primera vez en nuestro país, la técnica de las funciones splines a la “medida de la mortalidad”, si bien y también de forma novedosa, a un colectivo de población activa ocupada.

Finalmente, Artís y Felipe (1997) asumiendo parte de los resultados de Rué (1992) para la población española elaboran un modelo dinámico basado en la segunda Ley de Heligman y Pollard (1980) y los estudios de McNown y Rogers (1988) para los EEUU. En su trabajo además, se incorporan contrastes de hipótesis para establecer la dinamicidad de la probabilidad de fallecimiento, y se obtienen intervalos de confianza para la probabilidad de fallecimiento observada y ajustada. No existe ninguna hipótesis de linealidad en los datos utilizados para los cálculos de probabilidades, sino que estos datos han sido elaborados por el INE especialmente para la investigación llevada a cabo.

No se tiene constancia de que se hayan aplicado las técnicas de Whittaker-Henderson para la graduación de los valores brutos de mortalidad (aunque Artís y Felipe (1997) aplican un caso particular de dicho método de graduación). Esta técnica se utiliza, sin embargo, con cierta frecuencia en otros países.

ANEXO

POLINOMIOS ORTOGONALES.

En su forma general $GM_{\alpha}^{r,s}(x)$ contiene dos polinomios. En la Ecuación 6 anterior, estos polinomios son definidos (con el vector α) en términos de sucesivas potencias de x : es decir $\{1, x, x^2, x^3, \dots\}$. Estas pueden considerarse las bases obvias para los polinomios. No obstante, en ciertas situaciones, puede considerarse más apropiado usar unas bases alternativas. Cualquier secuencia de polinomios $\{p_0(x), p_1(x), p_2(x), \dots\}$ con $p_i(x)$ ($i=0,1,2,3,\dots$) de grado i , forman una base para el conjunto de polinomios.

Consideremos por ejemplo las primeras expresiones de los polinomios de Chebycheff $\{C_i(x)\}$ definidos como:

$$\begin{aligned}C_0(x) &= 1 \\C_1(x) &= x \\C_2(x) &= 2x^2 - 1 \\C_3(x) &= 4x^3 - 3x \\C_4(x) &= 8x^4 - 8x^2 + 1\end{aligned}$$

y la relación de recurrencia:

$$C_{n+1}(x) = 2xC_n(x) - C_{n-1}(x) \text{ para } n \geq 1$$

Ecuación 34

Si $p(x)$ es cualquier polinomio dado de orden n (grado $n-1$), entonces $p(x)$ puede expresarse de forma unívoca como:

$$p(x) = \sum_{i=1}^n \alpha_i C_{i-1}(x)$$

Ecuación 35

es esto lo que la caracteriza como base.

La idea de una base ortogonal surge de forma natural al ajustar una curva a un conjunto de datos dados. Sea m los puntos observados $\{(x_j, y_j) \ j=1 \dots m\}$ y

$\{w_1 \dots w_m\}$ un conjunto de ponderaciones positivas. Podemos intentar encontrar $f_n(x)$ como el polinomio de orden n que mejor ajuste los datos en el sentido de que los valores:

$$\sum_{j=1}^m w_j [y_j - f_n(x_j)]^2$$

Ecuación 36

sean minimizados (mínimos cuadrados generalizados). Si utilizamos la base natural $\{1, x, x^2, x^3, \dots\}$ para definir $f_n(x)$ escribiremos

$$f_n(x) = \sum_{i=1}^n \alpha_i x^{i-1}. \text{ De forma similar también podríamos hallar } f_{n+1}(x) \text{ el}$$

polinomio que mejor ajusta de orden $n+1$, entonces $f_{n+1}(x) = \sum_{i=1}^{n+1} \alpha'_i x^{i-1}$.

Pero los vectores $\alpha' = (\alpha'_1, \alpha'_2, \dots, \alpha'_n, \alpha'_{n+1})$ y $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ tendrán valores diferentes, en general, para los primeros n coeficientes. Así, cuando se utiliza la base natural, la consecuencia de incrementar en uno el grado del polinomio a ajustar no equivale a considerar un coeficiente α'_{n+1} adicional, sino que varían todos los parámetros $\alpha'_i \neq \alpha_i$.

Sin embargo, es posible elegir una base ortogonal (que dependerá de $\{(x_i, w_i) \mid i=1 \dots m\}$) de tal forma que el polinomio que mejor ajusta de orden $n+1$ se obtiene de añadir al polinomio de orden n un múltiplo (digamos α_{n+1}) de la $(n+1)$ -ésima función de la base.

En esta situación el valor del coeficiente adicional puede ser interpretado de forma mucho más simple.

Sea los puntos $\{(x_j) \mid j=1 \dots m\}$ y los pesos positivos $\{w_1 \dots w_m\}$, se escoge una base de polinomios ortogonales de forma que:

$$\sum_{j=1}^m w_j p_r(x_j) p_s(x_j) = \begin{cases} 0 & \text{si } r \neq s \\ e_r & \text{si } r = s \end{cases} \quad e_r \in \mathfrak{R}^+$$

normalmente $e_r = 1$, y

Ecuación 37

$$f_n(x) = \sum_{i=1}^n \alpha_i p_{i-1}(x) \text{ tal que } s = \sum_{j=1}^m w_j \left[y_j - \sum_{i=1}^n \alpha_i p_{i-1}(x_j) \right]^2 \text{ es mínimo}$$

Ecuación 38

entonces,

$$\alpha_k = \frac{\sum_{j=1}^m w_j y_j p_{k-1}(x_j)}{\sum_{j=1}^m w_j [p_{k-1}(x_j)]^2} \text{ para } k = 1, 2, \dots, n$$

Ecuación 39

que no depende de n y α' será $\alpha' = (\alpha_1 \alpha_2 \dots \alpha_n \alpha'_{n+1})$.

La forma continua para polinomios ortogonales en relación a un intervalo $[a, b]$ y una función de peso $w(x)$ se obtiene al construir una secuencia de polinomios $\{p_i(x) \ i=0, 1, \dots\}$ tal que:

$$\int_a^b w(x) p_r(x) p_s(x) dx = \begin{cases} 0 & \text{si } r \neq s \\ e_r & \text{si } r = s \end{cases} \quad e_r \in \mathfrak{R}^+$$

Ecuación 40

Para el intervalo $[-1, 1]$ los polinomios de Chebycheff descritos son ortogonales con función de peso:

$$w(x) = \frac{1}{\sqrt{1-x^2}}$$

Ecuación 41

Para el intervalo $[-1, 1]$ con $w(x)=1$ obtendremos los polinomios de Legendre $\{L_i(x)\}$ como base ortogonal.

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x \\ L_2(x) &= 3x^2 - 1 \\ L_3(x) &= 5x^3 - 3x \end{aligned}$$

y la relación de recurrencia:

$$(n+1) L_{n+1}(x) = (2n+1)x L_n(x) - n L_{n-1}(x) \quad n \geq 1$$

Ecuación 42

ESCALONADO DE EDADES

A veces conviene expresar $GM_{\alpha}^{r,s}(x)$ y $LGM_{\alpha}^{r,s}(x)$ en términos de un conjunto de polinomios ortogonales. Entonces se requiere escalar la edad. Los polinomios se expresan en potencias no de x sino de $t=(x-u)/v$ con u y v apropiadamente elegidos para que $t \in [-1,1]$ o aproximadamente. Por ejemplo si trabajamos con el rango de edades entre 60 y 110 entonces $t=(x-85)/25$ de forma que t estará en el intervalo requerido.

Si lo expresamos todo a la vez tendremos: Dado $\{p_i(x) (i=0,1,2,3,\dots)\}$ la base de polinomios ortogonales escogida y dados u y v convenientemente elegidos; para un vector de coeficientes $\alpha=(\alpha_1 \alpha_2 \dots \alpha_{r+s})$ definimos:

$${}_{u,v}^{(p)}GM_{\alpha}^{r,s}(x) = \sum_{i=1}^r \alpha_i p_{i-1}\left(\frac{x-u}{v}\right) + \exp\left\{\sum_{j=r+1}^{r+s} \alpha_j p_{j-r-1}\left(\frac{x-u}{v}\right)\right\}$$

Ecuación 43

y

$${}_{u,v}^{(p)}LGM_{\alpha}^{r,s}(x) = \frac{{}_{u,v}^{(p)}GM_{\alpha}^{r,s}(x)}{1 + {}_{u,v}^{(p)}GM_{\alpha}^{r,s}(x)}$$

Ecuación 44