

# Proyección de mortalidad en España mediante mixturas de modelos y análisis del impacto económico del riesgo de longevidad \*

ANDRÉS BENCHIMOL

*Universidad Carlos III de Madrid, Facultad de Ciencias Sociales y Jurídicas, Calle Madrid, 126, 28903, Getafe, Madrid, España. E-mail: abenchim@est-econ.uc3m.es*

## RESUMEN

En este artículo se proyectan las tasas de mortalidad de España por medio de una mixtura de modelos basada en el criterio de información de Akaike. Se ha trabajado con cuatro de los modelos más difundidos en la literatura actuarial y las proyecciones arrojadas por la mixtura han mostrado mejores propiedades estadísticas que las de cualquier modelo considerado por separado. Finalmente, se exponen las consecuencias económicas del riesgo de longevidad.

*Palabras clave:* Mixturas de Modelos, Proyección de Mortalidad, Riesgo de Longevidad, Tablas de Mortalidad Projectadas.

## Mortality Projection in Spain through Mixtures of Models and Analysis of the Economic Impact of Longevity Risk

## ABSTRACT

In this paper, the mortality rates of Spain are projected through a mixture of models based on the Akaike info criterion. We have worked with four of the most widely used models in the actuarial literature and the projections from the mix have shown better statistical properties than any model considered separately. Finally, the economic consequences of longevity risk are presented.

*Keywords:* Mixtures of Predictive Models, Mortality Projection, Longevity Risk, Projected Life Tables.

Clasificación JEL: C15, C53, J11, J17

---

\* El presente trabajo ha sido patrocinado por el Ministerio de Economía y Competitividad a través del proyecto MTM2014-56535-R. El autor agradece los comentarios y sugerencias de los dos revisores anónimos, que han contribuido a mejorar el manuscrito original.

## 1. INTRODUCCIÓN

El modelo de mortalidad con el que trabajan los actuarios en las compañías de seguros se denomina *tabla de mortalidad*. Estas tablas son instrumentos para estimar probabilidades de supervivencia y fallecimiento y son de utilidad en diversos ámbitos y especialmente en el actuarial. Se utilizan para calcular las primas de los seguros de vida, los valores de rentas vitalicias, los beneficios de planes de pensiones, así como también para determinar los valores de las reservas matemáticas<sup>1</sup> que deben constituir las compañías de seguros de vida como prueba de solvencia. Cualquier vicio o irregularidad que pudiera existir en las tablas de mortalidad utilizadas está destinado a trasladarse en el futuro a los balances y estados de resultados de las compañías.

Un hecho al que se han estado enfrentando los actuarios en las compañías de seguros de vida es a que las personas viven más de lo esperado de acuerdo con las tablas de mortalidad utilizadas. Así, la mayoría de las proyecciones de la mortalidad para la segunda mitad del siglo XX han subestimado el aumento en la longevidad y las mejoras en la mortalidad observadas en la práctica han superado a las mejoras previstas.

Si bien el aumento del tiempo de vida constituye uno de los mayores logros de la humanidad en el último siglo, a la vez, tiene como consecuencia un gran desafío económico y financiero para las próximas décadas. Esto se debe al hecho de que a mayor tiempo de vida mayor es el número de años que las personas que se jubilan pasarán en etapa de retiro, lo cual implica una mayor cuantía de fondos necesarios para solventar las necesidades vitales de estos colectivos.

Desde el punto de vista técnico-actuarial, para garantizar la solvencia de cualquier plan de seguros y su viabilidad es necesario que exista un equilibrio entre el valor actual actuarial<sup>2</sup> de las contribuciones que hacen los asegurados durante su etapa de vida activa y el valor actual actuarial de los beneficios que recibirán en su etapa de vida pasiva. Puesto que se han subestimado las probabilidades de supervivencia, el resultado ha sido que el valor actual actuarial de los fondos provistos por los asegurados sea inferior al valor actual actuarial de los compromisos que tiene la compañía de seguros con ellos y, por tanto, que exista un déficit en la cuantía de fondos necesarios para financiar los beneficios de la etapa pasiva. Así, las compañías de seguros se han visto forzadas a asignar más capital para sostener sus negocios vigentes de rentas

---

<sup>1</sup> Las reservas matemáticas son pasivos de naturaleza contingente que deben constituir las compañías de seguros de vida por cada póliza que emiten.

<sup>2</sup> El valor actual actuarial es la actualización financiera y biométrica de las obligaciones futuras que están sujetas a la ocurrencia de determinados eventos (la obligación de los asegurados es efectuar las contribuciones a la compañía de seguros durante la etapa activa y la obligación de la compañía es el pago de los beneficios a sus asegurados durante la etapa pasiva).

vitalicias con efectos adversos en reservas y rentabilidad y, fundamentalmente, en la solvencia de las compañías. Por tanto, este problema requiere la toma de medidas urgentes.

Esta desviación sistemática y unidireccional de la mortalidad observada respecto a la esperada recibe actualmente el nombre propio de *riesgo de longevidad*. Desde el enfoque actuarial de la gestión de riesgos, el riesgo de longevidad es un riesgo no diversificable, es decir, que no puede ser reducido mediante la emisión de un gran número de pólizas. Por otro lado, el impacto económico del riesgo de longevidad es aún mayor en la medida en que exista concurrencia de pensiones de jubilación y viudedad, tal como ocurre en el Sistema de la Seguridad Social español (ver Alaminos y Ayuso (2015)).

Las reacciones ante el riesgo de longevidad son principalmente dos. Por un lado, los actuarios están intentando proponer modelos más ajustados teniendo en cuenta la mejora en la mortalidad y prestando mayor atención a los niveles de incertidumbre de las predicciones (ver Lee y Carter (1992), Renshaw y Haberman (2006), Cairns *et al.* (2006)). Por otro lado, se está imitando a los mercados de capitales en su forma de compartir el riesgo a través de la aparición de instrumentos financieros ligados a la mortalidad (ver Cox *et al.* (2010), Blake *et al.* (2006) y Lin y Cox (2005)).

El objeto de este artículo está relacionado con la primera de estas dos cuestiones, es decir, con la modelización de la mejora en la mortalidad y el desarrollo de tablas de mortalidad que contemplen las mejoras esperadas. Estas tablas en la actualidad ya reciben el nombre de *tablas de mortalidad proyectadas* o *tablas de mortalidad dinámicas*. En estas tablas las tasas proyectadas de mortalidad ya no son sólo función de la edad y el sexo (como se había estado trabajando) sino también de la cohorte o generación a la que pertenecen los individuos. Así, se propone una metodología para elaborar tablas de mortalidad dinámicas y proporcionar una herramienta para poder cuantificar el impacto económico del riesgo de longevidad.

El riesgo de longevidad es un tema que hasta principios del siglo XXI no se había comenzado a analizar en profundidad, por lo que las aportaciones en este campo están en continua transferencia entre los académicos y la industria aseguradora. Tanto es así, que este riesgo está contemplado en la Directiva de *Solvencia II* (Directiva 2009/138/EC)<sup>3</sup>. Esta Directiva establece un marco regulatorio con injerencia sobre las compañías de seguros y reaseguros de los países miembros de la Unión Europea. La Directiva fue adoptada en noviembre de 2009 por el Consejo de la Unión Europea y el Parlamento Europeo pero no

---

<sup>3</sup> Directiva 2009/138/CE del Parlamento Europeo y del Consejo, de 25 de noviembre de 2009, sobre el seguro de vida, el acceso a la actividad de seguro y de reaseguro y su ejercicio.

entró plenamente en vigor hasta el 1 de enero de 2016 por la Directiva complementaria Omnibus II, del 11 de marzo de 2014.

La Directiva de Solvencia II ha cambiado las normas europeas del Seguro. Su principal objetivo es mejorar el control y la medición de los riesgos (de mercado, operacionales, de crédito y de liquidez) a los que están expuestos las compañías de seguros y reaseguros. Entre otros aspectos, la Directiva de Solvencia II hace referencia al cálculo del *Margen de Solvencia*<sup>4</sup>. De este modo, contempla el establecimiento de una fórmula estándar para su cálculo pero también prevé la posibilidad de que las compañías de seguros desarrollen modelos internos.

Por tanto, la metodología presentada en este artículo se propone como herramienta de aplicación directa por parte de los departamentos técnico-actuariales de las compañías de seguros así como también por parte de los organismos supervisores.

Por último, el riesgo de longevidad traspasa los límites del mercado asegurador y de la Seguridad Social. Por ejemplo, se traslada al mercado de *hipotecas inversas* o *hipotecas revertidas* y también al de *life settlements* (también denominado *mercado secundario de pólizas de seguros de vida*).

Una hipoteca inversa es un crédito con garantía inmobiliaria para personas mayores de 65 años que permite convertir en dinero el valor patrimonial que representa la propiedad de su vivienda. Así, en una operación de este tipo, el propietario de un inmueble recibe mensualmente una renta vitalicia que le permite disfrutar en vida del valor de la vivienda, además de su uso normal. Una vez que ocurre fallecimiento del suscriptor de la hipoteca inversa se pueden producir diferentes escenarios y posibilidades, dado que los herederos recibirán tanto la propiedad de la vivienda como la deuda anexa a la misma. La hipoteca inversa está regulada en España por la Ley 41/2007<sup>5</sup>.

En cuanto al *life settlement*, se trata de una operación financiera que consiste en la venta que el titular de una póliza de seguro de vida realiza a un tercero por una cantidad superior al valor de rescate que obtendría de la compañía, pero inferior a la suma asegurada. El comprador se hace cargo del pago de las primas desde que se concreta la operación y espera tener una ganancia al momento del fallecimiento del asegurado y, por tanto, de cobrar la suma asegurada. En su mayoría, los productos involucrados en estas transacciones son los de *Vida*

---

<sup>4</sup> El Margen de Solvencia es como se denomina al capital mínimo necesario que deben disponer las compañías de seguros para hacer frente a las posibles desviaciones de sus negocios.

<sup>5</sup> Ley 41/2007, de 7 de diciembre, por la que se modifica la Ley 2/1981, de 25 de marzo, de Regulación del Mercado Hipotecario y otras normas del sistema hipotecario y financiero, de regulación de las hipotecas inversas y el seguro de dependencia y por la que se establece determinada norma tributaria, BOE núm. 294, de 8 de diciembre de 2007.

*Universal (Universal Life)* y los asegurados suelen ser personas mayores de 65 años. El mercado de *life settlements* está muy desarrollado en los Estados Unidos.

Tanto en un caso como en el otro, para el inversor es fundamental contar con asesoramiento fiable respecto a la expectativa de vida del propietario del inmueble (en el primer caso) o del titular de la póliza de seguro de vida (en el segundo) de modo que su inversión ofrezca una rentabilidad esperada positiva.

El resto del artículo está estructurado como se detalla a continuación. En la Sección 2 se presentan los principales modelos de proyección de mortalidad en la literatura actuarial. En la Sección 3 se presenta una metodología para obtener una mezcla de modelos basada en el criterio de Akaike y se expone la aplicación a los modelos de mortalidad. En la Sección 4 se aplica la metodología a datos de España, se valida su capacidad proyectiva y se realizan proyecciones a largo plazo. Finalmente, en la Sección 5 se presentan las consideraciones finales.

## 2. MODELOS DE PROYECCIÓN DE MORTALIDAD

Bajo el enfoque estocástico, se considera que la mortalidad observada ha sido sólo uno de los posibles resultados de un fenómeno aleatorio. De la misma manera, el patrón de mortalidad en el futuro es tratado como un proceso estocástico.

Un punto importante a destacar de este enfoque es que, mientras que con los modelos tradicionales se obtienen estimaciones puntuales de las tasas de mortalidad, los métodos estocásticos contemplan fluctuaciones aleatorias en la mortalidad permitiendo, además, la estimación por intervalos.

A continuación, se definen una serie de conceptos y funciones biométricas necesarias para el estudio de la mortalidad.

Considérese un colectivo cerrado de vidas homogéneas e independientes. Sean  $l_x(t)$  el número de personas con vida con edad  $x$  en el año  $t$ ,  $d_x(t)$  el número de fallecimientos entre las edades  $x$  y  $x + 1$  en el año  $t$  y  $L_x(t)$  el número total de años vividos entre las edades  $x$  y  $x + 1$  por los sobrevivientes a la edad  $x$  en el año  $t$  (tiempo total de exposición).

Se define  $q_x(t)$  a la probabilidad de fallecimiento entre las edades  $x$  y  $x + 1$  en el año  $t$  como

$$q_x(t) = \frac{d_x(t)}{l_x(t)}.$$

Y se define  $m_x(t)$  a la tasa central de mortalidad entre las edades  $x$  y  $x + 1$  en el año  $t$  como

$$m_x(t) = \frac{d_x(t)}{L_x(t)}.$$

Teniendo presentes estas definiciones, se presentan a continuación cuatro de los principales modelos de proyección de mortalidad.

## 2.1. El modelo de Lee y Carter (LC)

Lee y Carter (1992) propusieron un modelo para describir el cambio en la mortalidad en función de un índice temporal. La metodología aplicada por estos autores constituye una de las propuestas más influyentes en la literatura reciente sobre proyección de la mortalidad. La especificación del modelo es la siguiente:

$$\log(m_x(t)) = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t},$$

donde los parámetros  $\alpha_x$  describen el patrón de la mortalidad promedio de cada edad que es independiente del tiempo, en tanto que los parámetros  $\beta_x$  describen las desviaciones respecto de tal patrón promedio cuando  $\kappa_t$  varía. La variable independiente  $\kappa_t$  constituye un índice de mortalidad y describe el cambio en el nivel de mortalidad a través del tiempo. Por último,  $\varepsilon_{x,t}$ , el término de error denota las desviaciones aleatorias no explicadas por el modelo y se asume que tiene media 0 y varianza constante.

El índice temporal  $\kappa_t$  es una variable latente y recoge lo que se conoce como *efecto periodo* y añade un nuevo factor al análisis de la mortalidad, además del *efecto edad* que era el principal que se había estado considerando hasta entonces. Este efecto periodo explica los cambios en la mortalidad que se deben a eventos que afectan a todos los individuos de todas las edades a lo largo del tiempo. Los más importantes posiblemente estén vinculados a avances de la medicina tales como antibióticos, vacunas y demás desarrollos que han repercutido favorablemente a lo largo del tiempo en la mortalidad de todas las edades. Sin embargo, también puede haber eventos negativos tales como las catástrofes naturales, las epidemias o las guerras.

Como el sistema resulta indeterminado, se debe imponer un conjunto de restricciones. La indeterminación del modelo se debe a que la verosimilitud asociada al modelo tiene un número infinito de máximos equivalentes, cada uno de los cuales genera las mismas proyecciones. Por lo tanto, la elección de las restricciones es arbitraria y no tiene impacto en la calidad del ajuste ni en las proyecciones de la mortalidad. Lee y Carter (1992) propusieron las restricciones  $\sum_x \beta_x = 1$  y  $\sum_t \kappa_t = 0$ , lo cual implica la selección de uno de esos máximos equivalentes.

Una vez obtenidas las estimaciones de los parámetros del modelo, la predicción de la mortalidad se puede hacer considerando los valores de  $\kappa_t$  como un *random walk* (paseo aleatorio) con *drift* (deriva):

$$\kappa_t = \kappa_{t-1} + \theta + \omega_t.$$

De esta manera, la tasa central de mortalidad  $m_x(t)$  depende de una serie temporal, lo cual permite utilizar la metodología de Box-Jenkins de series temporales para proyectar sus valores futuros.

## 2.2. El modelo de Renshaw y Haberman (RH)

Renshaw y Haberman (2003) argumentaron que el modelo edad-período de Lee-Carter (1992) no siempre ajusta con precisión a los datos empíricos. Años más tarde, Renshaw y Haberman (2006) propusieron una extensión del modelo de Lee-Carter mediante la incorporación de un conjunto de un parámetros que recogen lo que llamaron *efecto cohorte*. El efecto cohorte explica una variación en la mortalidad que experimenta una generación debido a cambios sociales o de contexto respecto a las generaciones previas y/o posteriores y que se prolonga a lo largo del tiempo en dicha generación. Puede surgir por diversos motivos. Por ejemplo, diferentes hábitos de consumo de tabaco entre generaciones, una peor nutrición en los primeros años de vida de las personas durante un período de guerra respecto a los años previos y posteriores (como puede ser el caso de la Segunda Guerra Mundial), generaciones que se vieron beneficiadas por haber nacido mientras se desarrollaba en su país un servicio de salud público o que en un momento en que la investigación médica avanzó de manera considerable en la disminución de las tasas de mortalidad infantil.

Este modelo constituye hoy en día el modelo de referencia en la modelización de la mortalidad con efecto cohorte, habiendo sido el primero de una serie de modelos conocidos como *edad-periodo-cohorte*.

La especificación del modelo RH es la siguiente:

$$\log(m_x(t)) = \alpha_x + \beta_x^{(1)}\kappa_t + \beta_x^{(2)}\gamma_{t-x} + \varepsilon_{x,t}$$

donde los parámetros  $\alpha_x$ ,  $\beta_x$  y  $\kappa_t$  tienen el mismo significado que en el modelo de Lee y Carter, y la novedad es el factor  $\gamma_{t-x}$ , que es una serie temporal latente que recoge el efecto cohorte.

Al igual que en el modelo de Lee-Carter, también es necesario imponer un conjunto de restricciones para que el sistema sea determinado (Ver Renshaw y Haberman (2006)).

## 2.3. El modelo de Cairns, Blake y Dowd (CBD)

El modelo alternativo al de Lee y Carter más difundido en la literatura actuarial es la especificación planteada por Cairns *et al.* (2006). Estos autores, apoyándose en el hecho empírico de que  $\text{logit}(q_x(t))$  es aproximadamente lineal en  $x$  para valores de  $t$  fijos (excepto en las edades jóvenes) asumen que

$$\text{logit}(q_x(t)) = \kappa_t^{(1)} + \kappa_t^{(2)}x$$

donde el vector  $\mathbf{k}_t = (\kappa_t^{(1)}, \kappa_t^{(2)})^T$  se modeliza como una serie temporal bivalente. Así, en este modelo, la edad entra como una variable explicativa continua, en tanto que la pendiente y el intercepto son dinámicos en el tiempo.

Esta especificación no presenta ningún problema de identificabilidad por lo que no es necesario imponer ninguna restricción como en los modelos anteriores.

#### 2.4. El modelo de Cairns, Blake y Dowd con efecto cohorte (M6)

El modelo planteado por Cairns *et al.* (2009) fue la primera generalización del modelo CBD para contemplar el efecto cohorte mediante la inclusión de un factor latente:

$$\text{logit}(q_x(t)) = \kappa_t^{(1)} + \kappa_t^{(2)}x + \gamma_{t-x}.$$

La inclusión del término  $\gamma_{t-x}$  en el modelo original es análoga a la que Renshaw y Haberman (2006) habían hecho sobre al modelo planteado por Lee y Carter (1992).

Este modelo también presenta problemas de identificabilidad que se resuelven mediante un cambio de variables y un conjunto de restricciones (ver Cairns *et al.* (2009)).

### 3. MIXTURAS DE MODELOS DE PROYECCIÓN DE MORTALIDAD

En general, en la práctica estadística, el proceso de selección de modelos consiste en la identificación de cuál es el *mejor* de ellos de acuerdo con algún criterio de selección. Como resultado, los analistas se quedan con dicho *mejor* modelo, que se suele tratar como si fuera el *verdadero* y se procede como si el modelo seleccionado hubiera generado los datos y se utiliza para hacer inferencias. Mientras tanto, los modelos alternativos quedan descartados.

Como una crítica a esta forma de proceder, algunos autores plantean que “la práctica estadística estándar ignora la incertidumbre de modelos” (ver Hoeting (1999)).

El hecho de que este tipo de incertidumbre sea ignorada en el proceso de selección de modelos, se traduce en una subestimación de la incertidumbre de las inferencias (ver Draper (1995)), lo que conduce a un exceso de confianza en las inferencias que se hacen y a decisiones que conllevan más riesgo del que se cree. Por tanto, es necesario que además de la incertidumbre asociada a los parámetros del modelo, también se tenga en cuenta la incertidumbre asociada a la selección del modelo.

Una manera de hacer frente a las dificultades anteriores es trabajar con varios modelos, construir una mixtura y obtener un modelo final que tenga en cuenta a

todos. De esta manera, los modelos individuales actúan como componentes que se combinan linealmente para obtener un nuevo modelo final. Así, las técnicas de mixturas de modelos resuelven el problema de tener que elegir un modelo concreto.

Las técnicas de selección de modelos pueden ser vistas como casos particulares de mixturas cuando el peso del modelo elegido es igual a 1, y el de los demás igual a 0. Es decir, las técnicas de selección de modelos son casos extremos de casos más generales como son las mixturas.

Estas técnicas presentan ciertas ventajas sobre las de selección de modelos. Los promedios tienen en cuenta la incertidumbre y producen inferencias más precisas y realistas (ver Hoeting (1999)).

Existen diferentes propuestas para conseguir mixturas de modelos que se basan en criterios de información. Por ejemplo, se puede usar un método basado en el criterio de información de Akaike (AIC) (ver Buckland (1997)). Este enfoque ha sido ampliamente utilizado en diversas áreas, como ecología (ver Cade (2015)), medicina (ver Schorning *et al.* (2016)) o finanzas (ver Liu y Kuo (2016)).

### 3.1. El método basado en el criterio de información de Akaike (AIC)

El criterio de información de Akaike (AIC) (Akaike, 1974) es una herramienta utilizada para comparar y seleccionar modelos. Este criterio le asigna una medida a cada modelo considerando tanto la verosimilitud como la cantidad de parámetros. En cierto modo, se trata de encontrar un equilibrio entre la bondad de ajuste y la complejidad del modelo, lo que se conoce como *principio de parsimonia*. Así, se proporciona una herramienta útil para selección de modelos (ver Burnham y Anderson (2003)).

Sea  $L$  el máximo valor de la función de verosimilitud de un modelo y sea  $k$  el número de parámetros estimados en el modelo. El valor del  $AIC_C$  del modelo se define como:

$$AIC_C = -2\log(L) + 2k + \frac{2k(k+1)}{n-k-1}.$$

Los dos primeros términos conforman el AIC, y al añadirle el tercer término, el factor de corrección, se obtiene el AIC Corregido ( $AIC_C$ ). Cuando el tamaño muestral  $n$  se incrementa más rápidamente que el número de parámetros  $k$ , el factor de corrección disminuye y se torna despreciable para tamaños muestrales grandes.

La mixtura basada en el criterio de información de Akaike consiste en el cálculo de una media ponderada según los valores de  $AIC_C$  de los resultados obtenidos por un conjunto de modelos. Para determinar los pesos correspondientes de cada

modelo, primero es necesario identificar cuál es el *mejor modelo*, entendiendo por tal aquél que presenta el menor valor de  $AIC_C$ .

Siguiendo la propuesta de Burnham y Anderson (2003), dados los datos y los modelos propuestos, estos autores definen como  $w_i$  al peso relativo del modelo  $i$ ,

$$w_i = \frac{e^{-\frac{1}{2}\delta_i}}{\sum_{j=1}^M e^{-\frac{1}{2}\delta_j}},$$

donde  $M$  es el número de modelos propuestos, y

$$\delta_i = \frac{AIC_i - AIC_{\text{mejor modelo}}}{AIC_{\text{mejor modelo}}},$$

el incremento relativo del  $AIC_C$ .

### 3.2. Especificación de la mixtura modelos de mortalidad

El modelo propuesto en esta sección es una mixtura por medio del método basado en los pesos de Akaike de los cuatro modelos presentados en la Sección 2.

Para pasar de las probabilidades de fallecimiento  $q_x(t)$  a las tasas centrales de mortalidad  $m_x(t)$  se ha trabajado con la hipótesis de tipo instantáneo de mortalidad constante entre edades  $x$  y  $x + 1$ , que implica que

$$m_x(t) = -\log(1 - q_x(t))$$

Las tasas centrales de mortalidad proyectadas  $\hat{m}_x(t)$  se obtienen mediante la combinación lineal

$$\hat{m}_x(t) = \sum_{i=1}^M w_i \hat{m}_x^i(t),$$

Donde  $\hat{m}_x^i(t)$  representa la tasa central de mortalidad para la edad  $x$  en el año  $t$  y según el modelo  $i$ , en tanto que  $w_i$  es el peso que la metodología le asigna al modelo  $i$ .

## 4. APLICACIÓN A DATOS DE ESPAÑA

En esta sección se aplica la mixtura de modelos basada en el  $AIC_C$  al conjunto de modelos de proyección de la mortalidad seleccionado, esto es, a los modelos LC, RH, CBD y M6. Hasta donde se ha investigado, ésta es la primera vez en la que esta metodología es aplicada a este tipo de modelos.

La fuente de datos utilizada es *Human Mortality Database* (disponible en <http://www.mortality.org/>). Esta base de datos fue creada en el año 2000 como una iniciativa conjunta del Departamento de Demografía de la Universidad de California, Berkeley (Estados Unidos) y el Instituto Max Planck de Investigación Demográfica en Rostock (Alemania) y fue lanzada en mayo de 2002. La base contiene datos de mortalidad por sexo, edad y cohorte, y la edad límite está fijada en 110. Actualmente contiene tablas de mortalidad de 38 países. Para más información ver Shkolnikov *et al.* (2005).

Es importante señalar que pueden existir diferencias en el patrón de mortalidad entre la población general y la población asegurada. Estas diferencias se deben a lo que en seguros se conoce como *antiselección* o *selección adversa*. Este concepto tiene que ver con el hecho de que quien contrata un seguro voluntariamente lo hace porque tiene una necesidad particular que está relacionada con su mortalidad o supervivencia esperada. Si bien la antiselección está más presente en los seguros de vida, en el caso de las coberturas por supervivencia también puede existir. Así, el perfil de mortalidad de quien decide comprar una renta vitalicia de manera voluntaria puede no ser el mismo que el de alguien que ha sido incluido de manera obligatoria en un plan de pensiones (por ejemplo, como consecuencia de trabajar en una empresa que lo otorga como beneficio) o el de alguien que se jubila en un sistema público de pensiones.

Generalmente, las rentas vitalicias voluntarias suelen ser adquiridas por personas que gozan de buena salud y consideran que tienen una esperanza de vida alta, mientras que es poco probable que individuos con baja esperanza de vida adquieran una cobertura de este tipo. Esta es la razón por la que quienes adquieren productos voluntariamente en compañías de seguros (es decir, los que conforman la población asegurada) suelen tener un perfil de mortalidad diferente del de la población general.

No obstante, existe una gran dificultad para disponer de datos de población asegurada. Por un lado, las compañías de seguros suelen ser reacias a compartírselos escudándose en las leyes de protección de datos. En particular, en España, está la Ley Orgánica 15/1999 de 13 de diciembre de Protección de Datos de Carácter Personal (LOPD) (fundamentada en el artículo 18 de la Constitución Española de 1978, sobre el derecho a la intimidad familiar y personal y el secreto de las comunicaciones), cuyo principal objetivo es regular el tratamiento de los datos y archivos de carácter personal, los derechos de los ciudadanos sobre ellos y las obligaciones de aquéllos que los crean o tratan. Por otro lado, es necesario que el tamaño de la cartera de las compañías de seguros (o agregación de carteras) sea lo suficientemente grande. Esto está relacionado con la convergencia en probabilidad de las tasas de mortalidad observadas a las verdaderas tasas de mortalidad de asegurados. Y en la medida en que el tamaño

de la cartera no sea lo suficientemente grande, las inferencias carecerán de validez.

Por otro lado, existen ciertos cuestionamientos acerca de calidad de los datos utilizados para proyectar mortalidad. Las tasas presentadas en la Sección 2 se calculan como el cociente de dos cantidades: número de fallecimientos y número de expuestos o tiempo de exposición. Estos datos suelen ser relevados de diferentes fuentes, aunque principalmente de los censos de población, que comúnmente se hacen con una frecuencia decenal. En los censos la principal fuente de error de estimación se puede deber a la falta de respuesta. Asimismo, los valores entre censos de Human Mortality Database se obtienen mediante interpolaciones. Por otro lado, los flujos migratorios suelen ser difíciles de calcular con precisión, y así la aproximación del tiempo de exposición puede acarrear errores de estimación. Algunos autores han comenzado a investigar estos asuntos y a elaborar propuestas correctivas (ver Cairns *et al.* (2016)) y el tema está tomando tal importancia que el Instituto de Actuarios de Reino Unido ha publicado un documento de trabajo considerando estos problemas, así como también cuestionando la utilización de la población con edad  $x$  al último cumpleaños a mitad de año  $t$  como aproximación del valor de  $L_x(t)$ , lo cual podría no ser preciso puesto que depende de la distribución de fallecimiento a lo largo del año  $t$ , o bien proponiendo métodos para detectar anomalías en los datos de mortalidad (ver CMI Mortality Projections Committee (2016)). El problema central radica en que todos estos inconvenientes con los datos pueden condicionar los resultados del modelo.

Teniendo presente lo anteriormente expuesto, se ha trabajado con datos de población general. No obstante, el objetivo es presentar una metodología que también pueda ser utilizada por compañías de seguros trabajando con datos propios de población asegurada.

En lo que a intervalo de tiempo de refiere, se ha decidido trabajar con datos de los años calendario que van desde 1960 hasta 2009, ambos límites incluidos. La determinación del límite inferior del intervalo ha tenido que ver, por un lado, con que la mejora en la mortalidad se tornó más significativa durante la segunda mitad del siglo XX y, por otro, con la decisión de fijarlo en un año no muy próximo a la finalización de la Segunda Guerra Mundial lo cual podría introducir cierto sesgo en las proyecciones.

En cuanto a las edades, se ha trabajado con aquéllas comprendidas entre los 60 y 100 años, ambos límites incluidos. Una vez más, el motivo ha tenido que ver el objetivo primario de ajustar modelos que capten la mejora en la mortalidad en las edades avanzadas, que son las que están presentando los mayores retos a los sistemas de Seguridad Social públicos y planes de pensiones privados. Si bien la *Human Mortality Database* cuenta con datos hasta los 110

años, teniendo en consideración los pequeños tamaños muestrales existentes para edades superiores, el límite superior ha sido determinado en 100.

**4.1. Validación del modelo**

En primer lugar, se han estimado los modelos LC, RH, CBD y M6 con una muestra de entrenamiento consistente en datos recogidos desde el año 1960 hasta 1999, y luego se han obtenido proyecciones para el periodo que va desde los años 2000 hasta 2009. Los modelos han sido ajustados utilizando los paquetes *demography* (Hyndman *et al.* (2014)) y *StMoMo* (Villegas *et al.* (2015)) de R (R Core Team (2015)).

En la Tabla 1 se muestran los valores del  $AIC_c$ , los valores reescalados y los pesos de cada modelo.

**Tabla 1**  
Valores  $AIC_c$  y pesos relativos de los modelos en la mezcla.

	$AIC_c$	$\delta_i$	$w_i$
<b>LC</b>	1318,552	0,026122	0,256712
<b>RH</b>	1497,384	0,165292	0,239456
<b>CBD</b>	1284,986	0,000000	0,260087
<b>M6</b>	1451,768	0,129793	0,243744

Fuente: Elaboración propia.

Con el propósito de validar la capacidad proyectiva del modelo ponderado se han realizado proyecciones de las tasas centrales de mortalidad para el periodo 2000 a 2009 y se han comparado con los valores observados. Las tasas centrales de mortalidad de la mezcla se han calculado como

$$\hat{m}_x(t) = \sum_{i=1}^M w_i \hat{m}_x^i(t),$$

donde  $\hat{m}_x^i(t)$  representa la tasa central de mortalidad para la edad  $x$  en el año  $t$  y bajo el modelo  $i$ .

Por otro lado, las varianzas de las predicciones son

$$Var(\hat{m}_x(t)) = \sum_{i=1}^M w_i^2 Var(\hat{m}_x^i(t)) + \sum_{\substack{i=1 \\ i \neq j}}^M w_i w_j Cov(\hat{m}_x^i(t), \hat{m}_x^j(t)),$$

las cuales no son fáciles de determinar, puesto que incluyen covarianzas de las tasas centrales de mortalidad proyectadas obtenidas según diferentes modelos. Si bien se puede establecer una cota superior (ver Buckland (1997)):

$$Var(\hat{m}_x(t)) \leq \left( \sum_{i=1}^M w_i \sqrt{Var(\hat{m}_x^i(t))} \right)^2,$$

la expresión exacta de la varianza no tiene una forma cerrada.

Una alternativa para obtener valores estimados de la variabilidad de los estimadores es recurrir a técnicas de remuestreo. El remuestreo *bootstrap* fue presentado por Efron (1979) como una alternativa a la técnica *jackknife* (ver Quenouille (1956)), y consiste en el remuestreo con reemplazamiento de una muestra original de observaciones que se usa para obtener aproximaciones de errores estándar e intervalos de confianza de estimadores.

Hay una importante cantidad de artículos en el campo actuarial y en demografía en los que se utiliza el *bootstrap*, tales como: Koissi *et al.* (2006), que aplicaron métodos *bootstrap* para predicción de esperanzas de vida bajo modelos de Lee-Carter, Hoedemakers *et al.* (2003) para estimar intervalos de confianza en reservas de siniestros, o Caswell (2001) que presentó una descripción de métodos de remuestreo para construir intervalos de confianza relacionados con estimaciones demográficas.

Existen varios métodos para determinar intervalos de confianza *bootstrap*. En este caso se han utilizado intervalos *bootstrap* basados en percentiles, puesto que éstos muestran buenas propiedades a la vez de que su cálculo es sencillo sin que sea necesario asumir ningún modelo paramétrico.

A continuación se presentan las ideas básicas acerca del *bootstrap de residuos* para estimar intervalos de confianza de las tasas centrales de mortalidad proyectadas  $\hat{m}_x(t)$  obtenidas mediante la mixtura de modelo. El procedimiento *bootstrap* aplicado se resume en los siguientes pasos:

- Paso 1: Se ha dividido la base de datos original en dos muestras: una muestra de entrenamiento con los datos de los años calendario  $t = 1960, \dots, 1999$  y una muestra de validación con los datos de los años calendario  $t = 2000, \dots, 2009$ .
- Paso 2: Con la muestra de entrenamiento, se han obtenido estimaciones de los parámetros de cada modelo. Luego, para cada modelo, se ha obtenido una matriz de residuos  $R$
- Paso 3: Se generaron  $N$  réplicas (en particular se ha tomado  $N = 1000$ ),  $R^{(n)}$ ,  $n = 1, \dots, N$ , tomando muestras con reemplazamiento de los elementos de la matriz  $R$ . Luego, se ha replicado  $N$  veces la matriz de datos original, y a cada réplica se ha añadido una matriz de residuos  $R^{(n)}$ ,  $n = 1, \dots, N$ , obteniendo así  $N$  muestras *bootstrap*.
- Paso 4: Se determinó

$$\hat{m}_x(t) = \sum_{i=1}^M w_i \hat{m}_x^i(t),$$

para cada edad  $x = 60, \dots, 100$ , cada año calendario  $t = 2000, \dots, 2009$  y cada réplica  $n = 1, \dots, N$ .

- Paso 5: Se estimaron los intervalos de predicción para las tasas centrales de mortalidad  $m_x(t)$  mediante la determinación de los percentiles correspondientes a un determinado nivel de confianza. En particular, se obtuvieron intervalos de predicción al 95% de confianza tomándose los percentiles 2,5 y 97,5 de la distribución muestral de  $\hat{m}_x(t)$  para cada edad  $x = 60, \dots, 100$  y cada año calendario  $t = 2000, \dots, 2009$ .

**Tabla 2**  
Intervalos de predicción de la mixtura de modelos

	x = 60			x = 70		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0108	0,0117	0,0107	0,0277	0,0303	0,0266
2001	0,0105	0,0118	0,0105	0,0266	0,0312	0,0269
2002	0,0102	0,0120	0,0105	0,0252	0,0321	0,0253
2003	0,0099	0,0123	0,0107	0,0241	0,0334	0,0255
2004	0,0097	0,0126	0,0100	0,0228	0,0345	0,0240
2005	0,0094	0,0128	0,0102	0,0215	0,036	0,0235
2006	0,0091	0,0131	0,0102	0,0205	0,0375	0,0223
2007	0,0089	0,0134	0,0098	0,0195	0,0392	0,0228
2008	0,0087	0,0136	0,0099	0,0186	0,0409	0,0217
2009	0,0085	0,0139	0,0090	0,0177	0,0427	0,0209

	x = 80			x = 90		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0767	0,0847	0,0775	0,2018	0,2285	0,2033
2001	0,0724	0,0883	0,0716	0,1900	0,2435	0,2032
2002	0,0686	0,0928	0,0722	0,1782	0,2601	0,1974
2003	0,0649	0,0974	0,0729	0,1677	0,2788	0,2187
2004	0,0614	0,1025	0,0701	0,1573	0,2994	0,1982
2005	0,0578	0,1085	0,0734	0,1475	0,3223	0,2028
2006	0,0544	0,1139	0,0643	0,1365	0,3458	0,1896
2007	0,0514	0,1212	0,0661	0,1276	0,3764	0,1935
2008	0,0482	0,1280	0,0629	0,1194	0,4056	0,1924
2009	0,0452	0,1372	0,0617	0,1107	0,4423	0,1837

	x = 100		
	Límite inferior	Límite superior	Valor real
2000	0,4626	0,5388	0,4597
2001	0,4305	0,5865	0,4576
2002	0,4008	0,6377	0,4551
2003	0,3746	0,7065	0,4818
2004	0,3477	0,7778	0,4567
2005	0,3233	0,8603	0,4662
2006	0,2987	0,9553	0,4473
2007	0,2780	1,0620	0,4503
2008	0,2562	1,1803	0,4516
2009	0,2364	1,3187	0,4414

Fuente: Elaboración propia.

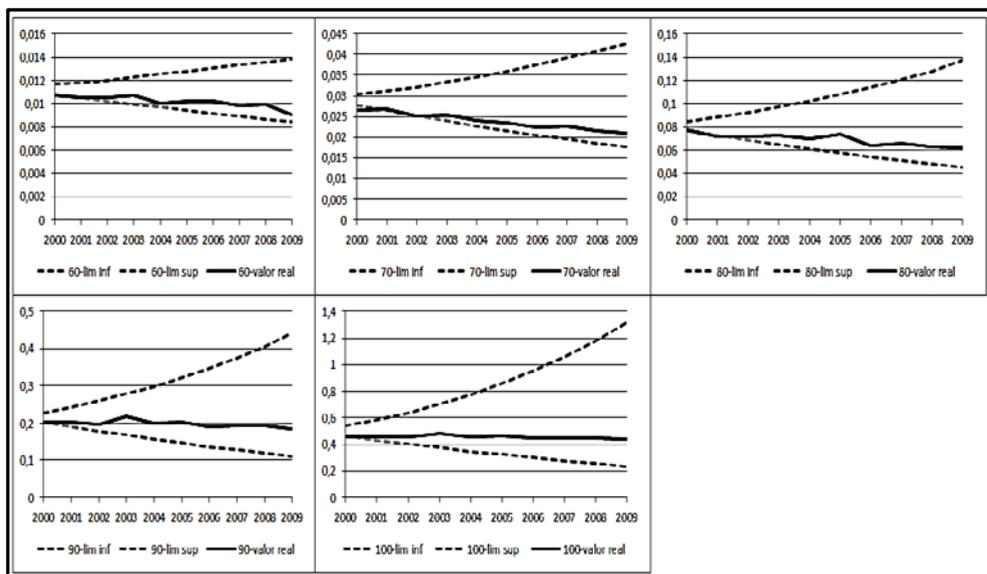
En los Anexos 1 a 4 se muestran los intervalos de predicción bootstrap al 95% de confianza para las tasas centrales de mortalidad según los modelos LC, RH, CBD y M6, respectivamente, y los correspondientes valores observados para las edades  $x = 60, 70, 80, 90$  y  $100$  en los 10 años correspondientes a la muestra de validación.

Puede observarse que las proyecciones del modelo LC resultan sesgadas. En concreto, las proyecciones sobreestiman la mortalidad, excepto en algunos casos de la edad  $x = 100$ . Así, para las edades seleccionadas, sólo el 16% de los intervalos de predicción contienen a los valores realmente observados. En cuanto al modelo RH, el 100% los intervalos de predicción del modelo RH contienen a los valores observados. No obstante, la amplitud de los intervalos crece notablemente conforme se avanza en las edades, ofreciendo una precisión muy escasa. Por su parte, en el modelo CBD la tasa de acierto es del 22% y, contrariamente a lo que ocurre con el modelo LC, la mayor parte de los intervalos que contienen a los verdaderos valores observados están en general en las edades jóvenes. Finalmente, el modelo M6 presenta una tasa de acierto del 32% no encontrándose un comportamiento tan claro en la distribución de los intervalos que contienen a los verdaderos valores observados.

En la Tabla 2 se muestran los intervalos de predicción bootstrap al 95% de confianza para las tasas centrales de mortalidad de la mixtura de modelos y los correspondientes valores observados para las edades  $x = 60, 70, 80, 90$  y  $100$  en los 10 años de la muestra de validación. Se observa que la mixtura de modelos presenta intervalos angostos, y los verdaderos valores están contenidos dentro de los intervalos en el 90% de los casos, lo cual da cuenta de una proyección con menor sesgo y mayor precisión que la que se obtiene con los modelos considerados individualmente.

**Figura 1**

Bandas de predicción y valores observados de las tasas centrales de mortalidad



Fuente: Elaboración propia a partir de los datos históricos disponibles en *Human Mortality Database*.

Por su parte, para una visualización gráfica, en la Figura 1 están representadas las bandas de predicción al 95% de confianza y las trayectorias de mortalidad observadas para las edades  $x = 60, 70, 80, 90$  y  $100$ . Como se puede apreciar, las bandas de predicción generadas a partir de la muestra de entrenamiento contienen a las trayectorias de mortalidad observadas en la mayoría de los casos.

#### 4.2. Proyección de la mortalidad a largo plazo

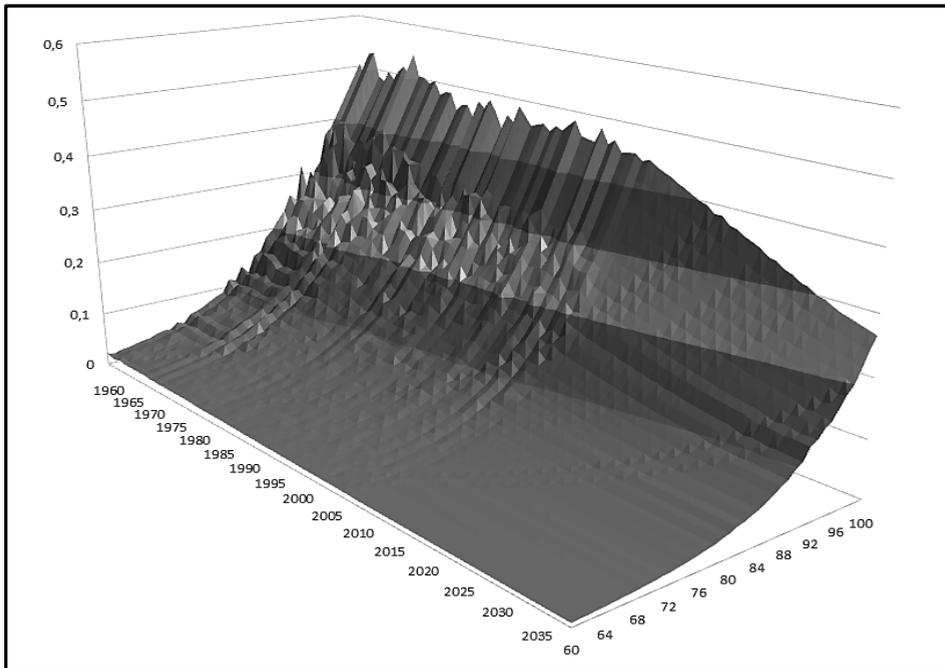
Una vez evaluada la capacidad proyectiva de la mixtura de modelos con el método basado en el AIC, se completa el análisis mediante la proyección de la mejora en la mortalidad a largo plazo. Esto es habitual en los negocios de rentas vitalicias y, en este caso, se considera un periodo de 30 años.

El lector interesado puede solicitar al autor la hoja de cálculo en la que se exponen las tasas de mortalidad proyectadas para población de hombres españoles mayores a 60 años utilizando la metodología propuesta. Esta tabla de mortalidad proyectada puede ser leída de acuerdo a tres arreglos:

- Cada columna de la tabla constituye una *tabla de mortalidad por periodo.*, cuyos elementos son las tasas de mortalidad para cada edad en un año dado.
- Cada diagonal de la tabla constituye una *tabla de mortalidad por cohorte.*, cuyos elementos son las tasas de mortalidad para cada edad en una cohorte dada.
- Cada fila constituye un *perfil de mortalidad.*, cuyos elementos son las tasas de mortalidad en cada año para una edad dada.

Para una visualización gráfica de los valores de la tabla de mortalidad proyectada, en la Figura 2 se representa la superficie de mortalidad observada para la población masculina para los años  $t = 2000, \dots, 2009$  seguida de la superficie de mortalidad proyectada para los años  $t = 2010, \dots, 2039$ . A la izquierda del eje de los años calendario se observan las tasas centrales de mortalidad observadas  $m_x(t)$  representadas por las altura de la superficie para los años  $t = 2000, \dots, 2009$ . Hacia la derecha del mismo eje, las tasas centrales de mortalidad proyectadas para los años  $t = 2010, \dots, 2039$  obtenidas mediante la mixtura de modelos. Cada cambio de color representa una curva de nivel de la superficie de mortalidad. Es decir, el lugar geométrico donde el valor de la función es constante. Desplazándose en el plano de forma paralela al eje de las edades y proyectando hacia la superficie se puede observar que, para cualquier año calendario dado, la mortalidad es creciente con la edad. Y desplazándose de forma paralela al eje del tiempo, que para cualquier edad dada se aprecia el decaimiento de los perfiles de mortalidad de cada edad con el correr del tiempo.

**Figura 2**  
Superficie de mortalidad histórica y proyectada.



Fuente: Elaboración propia a partir de los datos históricos disponibles en *Human Mortality Database*.

## 5. CONSIDERACIONES FINALES

Las coberturas de los seguros sobre la vida de las personas constituyen operaciones financieras aleatorias de larga duración, y asociado a ello surge la necesidad de hacer una valoración de las operaciones en cualquier momento del tiempo.

Las compañías de seguros deben atender a esta cuestión en sus cierres de balances, aunque también puede ser frecuente en otras tareas, por ejemplo, en aquellas relacionadas con el planeamiento y confección de presupuestos. Esto da lugar al concepto de *reservas matemáticas*.

Estas reservas surgen de la necesidad de cuantificar las obligaciones que tiene la compañía de seguros frente a sus asegurados en cualquier momento posterior a la contratación. Así, estas reservas constituyen una partida fundamental del balance general de una compañía de seguros de vida.

Hay dos cuestiones esenciales sobre el valor actuarial de la vida: una vinculada con el componente biométrico, y la otra con el componente financiero. A estas dos cuestiones se las conoce como *bases técnicas* y constituyen las dos hipótesis fundamentales en toda valoración actuarial sobre la vida. El riesgo de

longevidad está relacionado con el desvío sistemático de la primera de estas dos hipótesis.

Así, en este artículo se ha presentado una metodología para contemplar el problema de la incertidumbre respecto al componente biométrico. Si bien dicha técnica se había aplicado en diferentes áreas del conocimiento, es la primera vez que es aplicada en el campo de la proyección de mortalidad.

Se ha trabajado con datos de mortalidad de España y con 4 de los modelos de proyección de mortalidad más difundidos en la literatura actuarial. Con el fin de validar la capacidad proyectual de la mixtura de modelos, se ha dividido la muestra en una de entrenamiento y otra de validación. Con la primera muestra se han estimado los parámetros de los modelos y se ha proyectado la mortalidad, en tanto que la segunda se ha utilizado para contrastar las proyecciones obtenidas con los verdaderos valores observados. Se ha podido concluir, mediante intervalos de predicción al 95%, que la mixtura de modelos ha sido capaz de proyectar correctamente las tasas centrales de mortalidad en la gran mayoría de los casos, con menor sesgo y/o mayor precisión que los modelos considerados individualmente. sobre el mismo conjunto de datos.

Se propone la utilización de esta metodología por parte de las compañías de seguros como así también por organismos de la Seguridad Social, mitigando el riesgo de longevidad y minimizando, en la medida de lo posible, las futuras pérdidas económicas resultantes de una incorrecta calibración de las tablas dinámicas de mortalidad con las que trabajan.

Las futuras líneas de investigación están relacionadas con la proyección de la mortalidad mediante mixturas de modelos bajo enfoque Bayesiano, así como la implementación de otras técnicas estadísticas, como el data cloning o el análisis funcional de datos. En cualquier caso, el objetivo es obtener tablas de mortalidad proyectadas correctamente calibradas con el fin de mitigar el riesgo de longevidad.

## REFERENCIAS BIBLIOGRÁFICAS

---

- AKAIKE, H. (1974). "A new look at the statistical model identification". *IEEE transactions on automatic control*, 19(6), pp. 716-723.
- ALAMINOS, E. y AYUSO, M. (2015). "Una estimación actuarial del coste individual de las pensiones de jubilación y viudedad: Concurrencia de pensiones del Sistema de la Seguridad Social español". *Estudios de Economía Aplicada*, 33(3), pp. 817-838.
- BLAKE, D.; CAIRNS A.; Dowd, K y MACMINN, R. (2006). "Longevity bonds: financial engineering, valuation, and hedging". *Journal of Risk and Insurance*, 73(4), pp. 647-672.

- BUCKLAND, S. T.; BURNHAM, K. P. y AUGUSTIN, N. H. (1997). "Model selection: an integral part of inference". *Biometrics*, 53(2), pp. 603-618.
- BURNHAM, K. P. y ANDERSON, D. R. (2003). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media.
- CADE, B. S. (2015). "Model averaging and muddled multimodel inferences". *Ecology*, 96(9), pp. 2370-2382.
- CAIRNS, A. J. G.; BLAKE, D. y DOWD, K. (2006). "A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration". *Journal of Risk and Insurance*, 2006, 73(4), pp. 687-718.
- CAIRNS, A. J. G.; BLAKE, D.; DOWD, K.; COUGHLAN, G. D.; EPSTEIN, D.; ONG, A. y BALEVICH, I. (2009). "A quantitative comparison of stochastic mortality models using data from England and Wales and the United States". *North American Actuarial Journal*, 13(1), pp. 1-35.
- CAIRNS, A. J. G.; BLAKE, D.; DOWD, K. y KESSLER, A. R. (2016). "Phantoms never die: living with unreliable population data". *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(4), pp. 975-1005.
- CASWELL, H. (2001). *Matrix population models*. Wiley Online Library.
- COX, S. H.; LIN, Y. y PEDERSEN, H. (2010). "Mortality risk modeling: Applications to insurance securitization". *Insurance: Mathematics and Economics*, 46(1), pp. 242-253.
- CMI Mortality Projections Committee (2016). "CMI Mortality Projections Model consultation – technical paper". Institute and Faculty of Actuaries. Working paper.
- DRAPER, D. (1995). "Assessment and propagation of model uncertainty". *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 45-97.
- EFRON, B. (1979). "Bootstrap methods: another look at the jackknife". *The Annals of Statistics*, 7, pp. 1-26.
- HOEDEMAEKERS, T., BEIRLANT, J.; GOOVAERTS, M. J. y DHAENE, J. (2003). "Confidence bounds for discounted loss reserves". *Insurance: Mathematics and Economics*, 33(2), pp. 297-316.
- HOETING, J. A.; MADIGAN, D.; RAFTERY, A. E. y VOLINSKY, C. T. (1999). "Bayesian model averaging: a tutorial". *Statistical science*, 14(4), pp. 382-417.
- Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at [www.mortality.org](http://www.mortality.org)
- HYNDMAN, R. J.; BOOTH, H.; TICKLE, L. y MAINDONALD, J. (2014). "Demography: Forecasting mortality, fertility, migration and population data". R package version 1.18.
- KOISSI, M. C.; SHAPIRO, A. F. y HÖGNÄS (2006). "Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval". *Insurance: Mathematics and Economics*, 38(1), pp. 1-20.
- LEE, R. D. y CARTER, L. R. (1992). "Modeling and forecasting US mortality". *Journal of the American statistical association*, 87(419), pp. 659-671.
- LIN, Y. y COX, S. H. (2005). "Securitization of mortality risks in life annuities". *Journal of risk and Insurance*, 72(2), pp. 227-252.
- LIU, C. A. y KUO, B. S. (2016). "Model averaging in predictive regressions". *The Econometrics Journal*, 19(2), pp. 203-231.
- QUENOUILLE, M. H. (1956). "Notes on bias in estimation". *Biometrika*, 43(3/4), pp. 353-360.

- R CORE TEAM (2015). "R: A Language and Environment for Statistical Computing". Vienna, Austria: R Foundation for Statistical Computing.
- RENSHAW, A. E. y HABERMAN, S. (2003). "Lee-Carter mortality forecasting: A parallel generalized linear modelling approach for England and Wales mortality projections". *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(1), pp. 119-137.
- RENSHAW, A. E. y HABERMAN, S. (2006). "A cohort-based extension to the Lee-Carter model for mortality reduction factors". *Insurance: Mathematics and Economics*, 38(3), pp. 556-570.
- SCHORNING, K.; BORNKAMP, B.; BRETZ, F. y DETTE, H. (2016). "Model selection versus model averaging in dose finding studies". *Statistics in medicine*, 35(22), pp. 4021-4040.
- SHKOLNIKOV, V. M., WILMOTH, J. R. y GLEI, D. A. (2005). "Introduction to the Special Collection 'Human Mortality over Age, Time, Sex, and Place: The 1st HMD Symposium' ". *Demographic Research*, 13, pp. 223-230.
- VILLEGAS, A. M.; KAISHEV, V. K. y MILLOSSOVICH, P. (2015). "StMoMo: An R package for stochastic mortality modelling". In 7th Australasian Actuarial Education and Research Symposium.

## Anexo 1

Intervalos de predicción del modelo LC de las tasas centrales de mortalidad

	x = 60			x = 70		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0119	0,0126	0,0107	0,0290	0,0306	0,0266
2001	0,0118	0,0125	0,0105	0,0287	0,0303	0,0269
2002	0,0117	0,0124	0,0105	0,0283	0,0299	0,0253
2003	0,0116	0,0123	0,0107	0,0279	0,0296	0,0255
2004	0,0114	0,0122	0,0100	0,0275	0,0292	0,0240
2005	0,0113	0,0120	0,0102	0,0272	0,0289	0,0235
2006	0,0112	0,0119	0,0102	0,0268	0,0286	0,0223
2007	0,0111	0,0118	0,0098	0,0265	0,0282	0,0228
2008	0,0109	0,0117	0,0099	0,0261	0,0279	0,0217
2009	0,0108	0,0116	0,0090	0,0258	0,0276	0,0209

	x = 80			x = 90		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0796	0,0838	0,0775	0,2122	0,2213	0,2033
2001	0,0786	0,0829	0,0716	0,2106	0,2199	0,2032
2002	0,0777	0,0820	0,0722	0,2090	0,2185	0,1974
2003	0,0768	0,0812	0,0729	0,2073	0,2170	0,2187
2004	0,0759	0,0804	0,0701	0,2057	0,2156	0,1982
2005	0,0750	0,0796	0,0734	0,2041	0,2142	0,2028
2006	0,0741	0,0788	0,0643	0,2025	0,2127	0,1896
2007	0,0732	0,0780	0,0661	0,2009	0,2113	0,1935
2008	0,0724	0,0771	0,0629	0,1993	0,2099	0,1924
2009	0,0716	0,0763	0,0617	0,1977	0,2085	0,1837

	x = 100		
	Límite inferior	Límite superior	Valor real
2000	0,4506	0,4676	0,4597
2001	0,4492	0,4667	0,4576
2002	0,4478	0,4659	0,4551
2003	0,4464	0,4649	0,4818
2004	0,4450	0,4641	0,4567
2005	0,4437	0,4633	0,4662
2006	0,4422	0,4625	0,4473
2007	0,4407	0,4616	0,4503
2008	0,4392	0,4608	0,4516
2009	0,4378	0,4599	0,4414

Fuente: Elaboración propia a partir de los datos históricos disponibles en *Human Mortality Database*.

## Anexo 2

Intervalos de predicción del modelo RH de las tasas centrales de mortalidad

	x = 60			x = 70		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0107	0,0142	0,0107	0,0245	0,0340	0,0266
2001	0,0103	0,0160	0,0105	0,0210	0,0407	0,0269
2002	0,0096	0,0183	0,0105	0,0176	0,0480	0,0253
2003	0,0089	0,0209	0,0107	0,0152	0,0586	0,0255
2004	0,0083	0,0241	0,0100	0,0124	0,0705	0,0240
2005	0,0077	0,0274	0,0102	0,0105	0,0863	0,0235
2006	0,0072	0,0314	0,0102	0,0089	0,1081	0,0223
2007	0,0067	0,0359	0,0098	0,0077	0,1373	0,0228
2008	0,0062	0,0407	0,0099	0,0066	0,1707	0,0217
2009	0,0058	0,0461	0,0090	0,0056	0,2160	0,0209

	x = 80			x = 90		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0107	0,0142	0,0107	0,1761	0,2892	0,2033
2001	0,0103	0,0160	0,0105	0,1389	0,3818	0,2032
2002	0,0096	0,0183	0,0105	0,1081	0,5095	0,1974
2003	0,0089	0,0209	0,0107	0,0855	0,6953	0,2187
2004	0,0083	0,0241	0,0100	0,0664	0,9551	0,1982
2005	0,0077	0,0274	0,0102	0,0516	1,3305	0,2028
2006	0,0072	0,0314	0,0102	0,0391	1,8393	0,1896
2007	0,0067	0,0359	0,0098	0,0301	2,6116	0,1935
2008	0,0062	0,0407	0,0099	0,0229	3,7491	0,1924
2009	0,0058	0,0461	0,0090	0,0171	5,3171	0,1837

	x = 100		
	Límite inferior	Límite superior	Valor real
2000	0,3910	0,7267	0,4597
2001	0,2950	1,0414	0,4576
2002	0,2202	1,5079	0,4551
2003	0,1692	2,3355	0,4818
2004	0,1260	3,5296	0,4567
2005	0,0926	5,4170	0,4662
2006	0,0681	8,3459	0,4473
2007	0,0495	12,9189	0,4503
2008	0,0353	20,6407	0,4516
2009	0,0254	33,0270	0,4414

Fuente: Elaboración propia a partir de los datos históricos disponibles en *Human Mortality Database*.

### Anexo 3

Intervalos de predicción del modelo CBD y de las tasas centrales de mortalidad

	x = 60			x = 70		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0096	0,0109	0,0107	0,0276	0,0298	0,0266
2001	0,0094	0,0108	0,0105	0,0271	0,0295	0,0269
2002	0,0093	0,0106	0,0105	0,0267	0,0291	0,0253
2003	0,0091	0,0104	0,0107	0,0263	0,0287	0,0255
2004	0,0089	0,0103	0,0100	0,0259	0,0284	0,0240
2005	0,0087	0,0101	0,0102	0,0255	0,0280	0,0235
2006	0,0086	0,0099	0,0102	0,0252	0,0276	0,0223
2007	0,0084	0,0098	0,0098	0,0248	0,0273	0,0228
2008	0,0082	0,0096	0,0099	0,0244	0,0270	0,0217
2009	0,0081	0,0095	0,0090	0,0241	0,0266	0,0209

	x = 80			x = 90		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0772	0,0809	0,0107	0,2038	0,2144	0,2033
2001	0,0763	0,0802	0,0105	0,2026	0,2132	0,2032
2002	0,0755	0,0794	0,0105	0,2011	0,2121	0,1974
2003	0,0747	0,0787	0,0107	0,1998	0,2109	0,2187
2004	0,0739	0,0779	0,0100	0,1984	0,2098	0,1982
2005	0,0731	0,0772	0,0102	0,1971	0,2086	0,2028
2006	0,0724	0,0765	0,0102	0,1957	0,2075	0,1896
2007	0,0716	0,0757	0,0098	0,1945	0,2065	0,1935
2008	0,0708	0,0750	0,0099	0,1932	0,2054	0,1924
2009	0,0700	0,0743	0,0090	0,1918	0,2044	0,1837

	x = 100		
	Límite inferior	Límite superior	Valor real
2000	0,4871	0,5243	0,4597
2001	0,4853	0,5234	0,4576
2002	0,4836	0,5227	0,4551
2003	0,4817	0,5219	0,4818
2004	0,4798	0,5212	0,4567
2005	0,4778	0,5206	0,4662
2006	0,4761	0,5200	0,4473
2007	0,4740	0,5193	0,4503
2008	0,4723	0,5185	0,4516
2009	0,4706	0,5180	0,4414

Fuente: Elaboración propia a partir de los datos históricos disponibles en *Human Mortality Database*.

## Anexo 4

Intervalos de predicción del modelo M6 de las tasas centrales de mortalidad

	x = 60			x = 70		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0094	0,0110	0,0107	0,0270	0,0305	0,0266
2001	0,0093	0,0108	0,0105	0,0267	0,0300	0,0269
2002	0,0091	0,0107	0,0105	0,0262	0,0297	0,0253
2003	0,0089	0,0105	0,0107	0,0259	0,0293	0,0255
2004	0,0087	0,0103	0,0100	0,0254	0,0291	0,0240
2005	0,0086	0,0102	0,0102	0,0248	0,0286	0,0235
2006	0,0084	0,0100	0,0102	0,0243	0,0284	0,0223
2007	0,0082	0,0098	0,0098	0,0244	0,0275	0,0228
2008	0,0081	0,0097	0,0099	0,0239	0,0273	0,0217
2009	0,0079	0,0095	0,0090	0,0236	0,0269	0,0209

	x = 80			x = 90		
	Límite inferior	Límite superior	Valor real	Límite inferior	Límite superior	Valor real
2000	0,0762	0,0820	0,0107	0,2016	0,2160	0,2033
2001	0,0751	0,0813	0,0105	0,2006	0,2154	0,2032
2002	0,0742	0,0808	0,0105	0,1993	0,2135	0,1974
2003	0,0732	0,0799	0,0107	0,1977	0,2122	0,2187
2004	0,0726	0,0791	0,0100	0,1964	0,2113	0,1982
2005	0,0718	0,0787	0,0102	0,1951	0,2105	0,2028
2006	0,0711	0,0778	0,0102	0,1937	0,2097	0,1896
2007	0,0700	0,0772	0,0098	0,1920	0,2085	0,1935
2008	0,0694	0,0763	0,0099	0,1901	0,2075	0,1924
2009	0,0686	0,0754	0,0090	0,1896	0,2066	0,1837

	x = 100		
	Límite inferior	Límite superior	Valor real
2000	0,4853	0,5257	0,4597
2001	0,4813	0,5237	0,4576
2002	0,4811	0,5243	0,4551
2003	0,4790	0,5240	0,4818
2004	0,4772	0,5227	0,4567
2005	0,4751	0,5215	0,4662
2006	0,4738	0,5208	0,4473
2007	0,4712	0,5215	0,4503
2008	0,4701	0,5202	0,4516
2009	0,4682	0,5205	0,4414

Fuente: Elaboración propia a partir de los datos históricos disponibles en *Human Mortality Database*.

