TÉCNICAS ESTADÍSTICAS ORIENTADAS A LA ESTIMACIÓN DE MORTALIDAD EN EL SEGURO DE VIDA

José María Sánchez López
Departamento de Economía Financiera, Contabilidad y Comercialización
Universidad Rey Juan Carlos
e-mail: jmsl@fcjs.urjc.es

Ana Isabel Cid Cid
Departamento de Economía Financiera, Contabilidad y Comercialización
Universidad Rey Juan Carlos
e-mail: anacid@fcjs.urjc.es

Resumen

El sector asegurador, en el ramo de vida, presenta: gran interés económico, riesgos específicos, inversión del proceso productivo, repercusiones sociales y largo plazo. Es imprescindible un completo análisis del fenómeno de la mortalidad mediante métodos cuantitativos estadísticos para su correcto tratamiento (fijar primas y reservas, evitar insolvencias).

En el entorno actuarial, el análisis se abordará supeditado al fin que se persigue y a los datos disponibles en las aseguradoras. Se busca una metodología para construcción de tablas de mortalidad (estimaciones de mortalidad aplicables) que se adapte al problema real: plantear el uso de datos de pólizas (evitando sesgos derivados del uso de tablas de población general), analizar alternativas para las estimaciones iniciales (según propiedades respecto de los modelos de supervivencia), rectificar con distintos métodos las estimaciones iniciales (graduación paramétrica y no paramétrica en base a información adicional), estudiar grupos diferenciados según la información complementaria (análisis multivariante y credibilidad), y proyectar las estimaciones a futuro (dinamicidad del fenómeno).

Se presentará una visión actual del problema y de las soluciones: valorando las alternativas, señalando las mejores técnicas y ofreciendo nuevas posibilidades.

Palabras clave: Modelos de supervivencia, estimación actuarial, graduación, análisis de mortalidad, credibilidad, tabla de mortalidad, proyección de mortalidad.

1. Introducción.

Desde un “enfoque actuarial”, la estimación de la probabilidad de muerte para un perfil de individuos se basa en los datos brutos observados. Además, los datos se debe tomar a partir de colectivos análogos a los que se aseguran (informaciones recogidas en pólizas de la cartera). Los elementos de interés para una correcta recopilación de los datos brutos son: la precisión, el intervalo de registro, el periodo elegido para la recogida de datos, la independencia entre los individuos, y la recogida y análisis de varios tipos de variables. Estos elementos están relacionados. Se propugna, frente al habitual empleo de datos demográficos generales para la estimación de las tablas de mortalidad, el uso de los datos recogidos en las pólizas. Las razones para esta decisión se basan en que el perfil riesgo de estos individuos es más análogo al de los posibles próximos asegurados y en que la calidad de los datos recogidos (precisión y características reflejadas) es mayor. En base a ellos las técnicas estadísticas deben ofrecer:

- plantear la estimación delimitando caso de modelo tabular y dato incompleto,

- seleccionar grupos homogéneo y emplear credibilidad en valoraciones,

- obtener estimaciones rectificadas según información adicional, y

- permitir proyectar la mortalidad.

2. Estimación de modelos de supervivencia.

“Discusión sobre situación real de dato incompleto”. Los datos completos se dan en situaciones muy concretas. Los datos normalmente disponibles para los cálculos actuariales en las empresas de seguros vida, son incompletos. En primer lugar, existen otros fenómenos (distintos de la muerte) de naturaleza aleatoria, que pueden concretarse en sucesos que conlleven la "retirada" de los individuos de la muestra observada antes del momento de la muerte. Aunque no se considere significativo el número de retiradas, en la mayoría de los estudios actuariales se encuentra el problema de observaciones de las que no se realiza un seguimiento hasta la total extinción del grupo seleccionado. Los estudios longitudinales, a edades no avanzadas, sólo se consideran viables si se analizan muestras pequeñas con corto tiempo de vida (análisis de supervivencia en ensayos clínicos o investigaciones médicas). Los estudios actuariales son mayoritariamente de corte transversal. Al finalizar el periodo de observación son muchos los individuos de la

muestra que permanecen vivos, son los llamados "finalistas", que tendrán (en el momento de cierre de observación) una "edad de finalización predeterminada".

"Discusión sobre situación real de modelos tabulares". En los modelos de supervivencia, si las probabilidades de supervivencia $S(x)$ se dan según una función matemática se dice que la función está en **forma paramétrica** y genera un modelo paramétrico que dependerá de uno o más parámetros. Sin embargo, la forma más tratada en la literatura y más empleada en el trabajo actuarial ha sido la denominada **forma tabular**. El modelo tabular clásico utiliza valores enteros de x (variable edad) para los que se calcula su correspondiente valor según $S(x)$. La falta de valores intermedios para valores no enteros de x se resuelve estableciendo unas hipótesis de mortalidad o métodos de interpolación entre valores enteros consecutivos que permiten considerar el modelo completamente especificado (especificado para todos los valores positivos de x). El modelo tabular se da a conocer como "tabla de mortalidad" , "tabla de supervivencia" o "tabla de vida".

"Discusión sobre opciones respecto de la exposición al riesgo". De todos los datos recogidos para la estimación tabular, interesa la precisión en las fechas relevantes en el cálculo de la exposición individual al riesgo de muerte, ya que serán decisivas en la técnica de estimación. Partiendo de la contribución individual a la exposición al riesgo de muerte se estiman probabilidades por periodos recogiendo datos de todos los individuos. Según el tipo de exposición considerada se pueden utilizar distintas técnicas de estimación.

Se debe contemplar la necesidad de recoger tres edades importantes para los análisis posteriores: la edad de entrada al estudio y_i, la edad de salida programada del estudio z_i y la edad real de salida del periodo de estudio. Esta última situación puede darse por muerte a una edad q_i o puede darse por retirada a una edad f_i, en ambos casos la situación debe producirse antes de fin del periodo de estudio u observación. Las variables q_i y f_i toman valores cero si no se da la situación de muerte o retirada que las define. Así, a cada persona objeto de estudio se le puede asignar un vector que la caracterice de la forma $v_i' = [y_i, z_i, q_i, f_i]$ que recoge la información necesaria sobre la edad (ya sea real o virtual) para los posteriores procesos de estimación.

Un primer tratamiento de los datos llevaría a determinar la contribución de cada persona al intervalo de estimación que, en principio, se considera unitario y de la forma $(x, x+1]$. Se eliminan los vectores de los individuos que no contribuyen o permanecen en el

intervalo, esto es, si $y_i \geq x+1$ o si $z_i \leq x$ o si $\theta_i \leq x$ o si $\phi_i \leq x$. Por último, se transforma cada vector v_i' en un vector de duración en $(x,x+1]$, de la forma $u_{i,x}' = [r_i, s_i, l_i, k_i]$:

$$r_i \begin{cases} 0 & si \quad y_i \leq x \\ y_i - x & si \quad x < y_i < x+1 \end{cases}$$

$$s_i \begin{cases} z_i - x & si \quad x < z_i < x+1 \\ 1 & si \quad z_i \geq x+1 \end{cases}$$

$$l_i \begin{cases} 0 & si \quad \theta_i = 0 \\ \theta_i - x & si \quad x < \theta_i \leq x+1 \\ 0 & si \quad \theta_i > x+1 \end{cases}$$

$$k_i \begin{cases} 0 & si \quad \phi_i = 0 \\ \phi_i - x & si \quad x < \phi_i \leq x+1 \\ 0 & si \quad \phi_i > x+1 \end{cases}$$

Se puede ya, a partir de los datos así presentados, obtener la exposición al riesgo de muerte para un individuo en el intervalo de estudio u observación, esto es, la amplitud del periodo en el cual un individuo está bajo observación, condicionando la posibilidad de muerte.

Se considera la **exposición exacta** como $\begin{bmatrix} s_i \\ l_i \\ k_i \end{bmatrix} - r_i$, eligiendo el menor valor distinto de cero en el vector columna.

Se considera la **exposición programada** como $\begin{bmatrix} s_i \\ k_i \end{bmatrix} - r_i$, eligiendo el menor valor distinto de cero en el vector columna. Existe la posibilidad, en los casos de muerte, de utilizar una exposición distinta a la anterior: en vez de considerar la exposición hasta la salida programada del estudio se toma la exposición hasta una retirada programada fijada (la fecha de finalización de la póliza por ejemplo) o hasta una retirada aleatoria (edad máxima aleatoria en función del comportamiento de toda la cartera).

Se considera la **exposición actuarial** como $\begin{bmatrix} s_i \\ k_i \\ 1 \end{bmatrix} - r_i$, eligiendo el menor valor distinto de cero en el vector columna, salvo en casos de muerte que se toma el valor uno.

Para hallar la exposición total se han de considerar todos los individuos sumando sus exposiciones al riesgo de muerte. Los distintos tipos de exposiciones al riesgo de muerte se utilizarán en los procesos de estimación en las situaciones de datos incompletos.

"Discusión sobre alternativas de estimación". Con el método de los momentos se identifica el número de muertes esperado (calculado a través de la variable aleatoria número de muertes) con el número de muertes observado (obtenido con la información muestral) en un determinado intervalo de tiempo conocido. Se propone alguna hipótesis que lleve a una expresión que refleje el número de muertes esperado y se resuelve la ecuación para hallar el estimador deseado. Este procedimiento de estimación se justifica, siguiendo el teorema de Kintchine, por la convergencia en probabilidad de los momentos muestrales a los poblacionales, supuesta la i.i.d. de las variantes que se tienen en los elementos de la muestra, cuya esperanza finita se supone igual a la de la población. El método no utiliza de forma directa y completa la información sobre la distribución poblacional.

Para obtener probabilidades estimadas para años enteros hay que utilizar alguna **hipótesis de distribución** de los valores de la probabilidad en el interior del intervalo. Se presentan tres situaciones: fuerza de mortalidad decreciente (según comportamiento hiperbólico de la función de supervivencia), fuerza de mortalidad constante (comportamiento exponencial de la función de supervivencia) y fuerza de mortalidad creciente (comportamiento lineal de la función de supervivencia).

Por otra parte, el número de muertes observado en la muestra será representado por d_x.

Al igualar esperanza de la variante y número de muertes observado, la estimación de la probabilidad de muerte durante el año x se obtendrá bajo los tres supuestos considerados:

$$E(D_x) = d_x \quad \Rightarrow \quad \hat{q}_x = \frac{d_x}{\sum_{i=1}^{n_x} (s_i - r_i)},$$

$$E(D_x) = d_x \quad \Rightarrow \quad n_x - \sum_{i=1}^{n_x} (1 - \hat{q}_x)^{s_i - r_i} = d_x,$$

$$E(D_x) = d_x \quad \Rightarrow \quad \hat{q}_x \sum_{i=1}^{n_x} \frac{s_i - r_i}{1 - (r_i \bullet \hat{q}_x)} = d_x.$$

El método de máxima verosimilitud se basa en el conocimiento que se tiene (o se supone válido) de la distribución de probabilidad poblacional, binomial en este caso, se construye la función de verosimilitud para expresar mediante ésta la posibilidad u orden de preferencia en cuanto al valor concreto que pueda tomar el parámetro desconocido. Para ello, se toma la información contenida en la muestra: los valores muestrales aportados por la empresa aseguradora se toman como constantes y el parámetro desconocido varía en el espacio paramétrico [0,1] al tratarse de una probabilidad. Con todo, se llega a una función de verosimilitud definida como proporcional a la probabilidad de aparición de la muestra condicionada al parámetro. El valor del parámetro es el que hace máxima la función.

Con grandes muestras, tal como se tienen en el caso que se estudia, los estimadores de máxima verosimilitud tienen interesantes propiedades. Son asintóticamente insesgados y consistentes, si existe un estimador eficiente es el obtenido por máxima verosimilitud, siempre se llega a la normalidad y eficiencia asintóticas, si un parámetro tiene un estadístico suficiente el estimador máximo-verosímil es función del mencionado estadístico, y la estimación máximo-verosímil es invariante ante una transformación del parámetro.

Se puede incorporar, si está disponible, la información sobre la edad precisa de muerte. Se distingue entre la situación de "dato total", cuando se utiliza la edad precisa de muerte; y, la situación de "dato parcial", cuando no se dispone de la edad precisa de muerte. Para una situación, más realista, de dato total se construye la función de verosimilitud para un individuo: $L_i = {}_{t_i - r_i}p_{x+r_i} \left(m_{x+r_i}\right)^{d_i}$.

Apareciendo los siguientes elementos:

Li es la verosimilitud (aunque definida proporcional a la probabilidad del elemento muestral se prescinde del coeficiente de proporcionalidad),

r_i expresa el tiempo de entrada, desde el momento x, que no se considera variable aleatoria,

t_i se supone la realización de la variante T_i o de tiempo de cese de la observación (topado a uno que es la amplitud predeterminada del intervalo considerado y sujeta a la posibilidad de muerte o retirada previa),

d_i variable indicadora de muerte en el intervalo $(x,x+1]$ tratado (igual a uno si el individuo i muere en el intervalo, cero en otro caso).

Si la persona i-ésima no muere en el intervalo, la verosimilitud refleja la probabilidad de supervivencia entre $x+r_i$ y $x+t_i$. Si muere, la verosimilitud es la probabilidad de muerte en $x+t_i$ estando vivo en $x+r_i$. Todo ello a falta de fijar el parámetro correspondiente. Para la verosimilitud de toda la muestra, de tamaño n_x, se supone independencia entre los distintos elementos muestrales (individuos asegurados), lo que permite obtener la expresión: $L = \prod_{i=1}^{n_x} {}_{t_i-r_i}p_{x+r_i}\left(m_{x+r_i}\right)^{d_i}$.

Para llegar a resultados explícitos desde datos totales empíricos es necesario efectuar suposiciones sobre la distribución en el intervalo analizado. Se presentan dos situaciones: fuerza de mortalidad constante y fuerza de mortalidad creciente (con incremento lineal de la función de supervivencia).

Si se admite la hipótesis de un comportamiento exponencial en la mortalidad, lo que supone una fuerza de mortalidad constante, se obtiene:

$$L = m^{d_x}\prod_{i=1}^{n_x} e^{-(t_i-r_i)m} \quad \Rightarrow \quad \ell = \ln L = d_x \ln m - m\sum_{i=1}^{n_x}(t_i - r_i)$$

$$\frac{d\ell}{dm} = 0 \quad \Rightarrow \quad \hat{m} = \frac{d_x}{\sum_{i=1}^{n_x}(t_i - r_i)} \quad \Rightarrow \quad \hat{q}_x = 1 - e^{-\hat{m}} \qquad .$$

En este caso se estima la fuerza de mortalidad y desde ella se calcula la probabilidad de muerte para el intervalo considerado.

Si se presupone un comportamiento lineal de la función de supervivencia, tal como expresa la ley de Moivre, se llega a una fuerza de mortalidad creciente dentro del intervalo de estimación (situación más realista y equivalente a una interpolación lineal entre valores consecutivos de la función de supervivencia), queda:

$$L = \left(q_x\right)^{d_x}\prod_{i=1}^{n_x}\left(1 - r_i q_x\right)^{-1}\prod_{j-e}\left(1 - t_i q_x\right)$$

$$\ell = \ln L = d_x \ln\left(q_x\right) + \sum_{i=1}^{n_x}\ln\left(1 - r_i q_x\right)^{-1} + \sum_{j-e}\ln\left(1 - t_i q_x\right)$$

$$\frac{d\ell}{dq_x} = 0 \quad \Rightarrow \quad \frac{d_x}{q_x} + \sum_{i=1}^{n_x}\frac{r_i}{\left(1 - r_i q_x\right)} - \sum_{j-e}\frac{t_i}{\left(1 - t_i q_x\right)} = 0$$

donde (j-e) indica operación para todos los elementos muestrales excepto los correspondientes a individuos fallecidos observados. Se obtiene q_x estimado por iteración.

En la construcción del **estimador Kaplan-Meier** no se distingue entre salidas programadas (finalistas) y retiradas aleatorias, se utilizan "terminaciones" para nombrar cualquiera de las dos situaciones. Requiere información sobre la edad precisa de muerte, esto es, necesita del dato total, para poder ordenar la secuencia de muertes que se produzca.

Se divide el intervalo inicial de estimación (amplitud anual normalmente) en subintervalos, coincidiendo con los momentos-puntos temporales donde se produzca una entrada o una terminación.

En un subintervalo cualquiera i, la probabilidad de muerte, condicionada a estar vivo al inicio del subintervalo, se estima por máxima verosimilitud como:

$$\hat{q}_i = \frac{d_i}{n_i},$$

siendo d_i el número de muertes observadas en el subintervalo y n_i el tamaño de la muestra para ese subintervalo. Es preciso adoptar algún criterio en el caso de coincidencia entre momento de entradas o terminaciones y momento de muerte (por ejemplo asignar la muerte al intervalo que se tenga en el instante anterior).

Para todo el intervalo inicial se define un estimador llamado estimador límite producto o de Kaplan-Meier, de la forma:

$$\hat{q}_x = 1 - \prod_{i=1}^{m}\left(1 - \hat{q}_i\right) = 1 - \prod_{i=1}^{m}\left(\frac{n_i - d_i}{n_i}\right),$$

considerando m subintervalos.

Para muestras grandes que presenten un gran número de entradas y terminaciones se formarán numerosos subintervalos, lo que supone un problema práctico. Esto ocurre en los datos que aportan las empresas aseguradoras. Los procesos automáticos de cálculo y un registro preciso de los momentos de muerte disminuyen este problema.

Otro planteamiento que permite llegar a un estimador del mismo tipo límite-producto se obtiene si se parte el intervalo en cada punto de muerte, estableciéndose k subintervalos. Para cada subintervalo existe un grupo de riesgo integrado por el número de personas que permanecen vivas justo en el momento anterior a la muerte que marca el fin del subintervalo. Las terminaciones simultáneas al momento de la muerte forman parte del grupo de riesgo, mientras que las entradas simultáneas al momento de la muerte que delimita el extremo superior del intervalo no se consideran en el grupo de riesgo del subintervalo mencionado. Con todo, la probabilidad de muerte en el subintervalo j-ésimo será:

$$\hat{q}_j = \frac{1}{r_j} \text{ ó } \hat{q}_j = \frac{d_j}{r_j},$$

según se produzcan una o más de una muertes (simultáneas) en el subintervalo entre las r_j personas expuestas que forman el grupo de riesgo. Si en el último subintervalo no se producen muertes (muertes en $x+1$) se estima que la probabilidad de muerte en dicho subintervalo es cero. Si existe algún subintervalo sin grupo de riesgo no se podrá efectuar la estimación (por ejemplo cuando todos los individuos fallecen antes de $x+1$).

En todo el intervalo inicial se considera el estimador límite-producto con la expresión:

$$\hat{q}_x = 1 - \prod_{j=1}^{k} \left(1 - \hat{q}_j\right) = 1 - \prod_{j=1}^{k} \left(\frac{r_j - 1}{r_j}\right).$$

En general, el estimador límite-producto es insesgado y consistente. Un trabajo técnico que analiza en detalle las condiciones de convergencia débil y fuerte para completar aspectos asintóticos de la teoría de supervivencia puede encontrarse en Chen, Kani and Lo, Shaw-Hwa (1.997).

Se recomienda por sus adecuadas propiedades asintóticas y por su independencia respecto de las hipótesis de mortalidad intraintervalo el estimador Kaplan-Meier. En caso de no ser viable, el estimador por máxima verosimilitud suministra buenos resultados en muestras suficientemente grandes.

3. Revisión de las estimaciones.

"Planteamiento de la graduación". Una vez realizadas las estimaciones iniciales se plantean los métodos de rectificación o graduación de las estimaciones según las ideas

previas que se tengan sobre la secuencia en todo el rango de edades del individuo representativo del grupo de riesgo. Tras observar la naturaleza estadística de la graduación se presentan graduaciones no paramétricas (el método de las medias móviles, la fórmula de Whittaker, el método de Kimeldorf-Jones, la graduación bayesiana isotónica) y distintas graduaciones paramétricas.

A veces se ha identificado graduación con mecanismo simple de suavización o alisamiento de una secuencia de datos tratados. Sin embargo, se debe aclarar siguiendo a London, D. (1.985) que "se reconoce la validez del objetivo de alisamiento, pero se expande la idea hacia el concepto más general que refleja todos los elementos de opinión previa en un proceso de graduación".

Así, estas técnicas que proporcionan una mejora en la estimación de tasas anuales de mortalidad, tienen una naturaleza bayesiana. Ahora bien, no siempre se basarán en un proceso bayesiano formal. Esto requiere, además de una opinión previa, la valoración de esta creencia en forma de una distribución de probabilidad.

Los resultados graduados no se separarán mucho de las estimaciones iniciales ya que éstas, si están bien construidas, son indicadores razonables del verdadero valor. Las rectificaciones en base al conocimiento previo sobre el fenómeno serán menores cuanto mayor sea la muestra y menor sea la varianza de la estimación inicial (de forma parecida a las rectificaciones por credibilidad).

El planteamiento de la graduación se puede establecer con los siguiente elementos:

x índice definiendo la secuencia (según la edad en el caso de seguro de vida),

t_x secuencia de valores verdaderos desconocidos que se deben estimar,

n_x tamaño de la muestra,

w_x ponderaciones,

U_x variable aleatoria estimador de t_x ,

E_x variable aleatoria error de estimación,

u_x realizaciones de la variable aleatoria U_x ,

e_x realizaciones de la variable aleatoria E_x ,

v_x valores graduados.

La relación entre el fenómeno verdadero y las estimaciones iniciales se pueden establecer de la forma $U_x = t_x + E_x \Rightarrow u_x = t_x + e_x$.

Si se emplea un operador de graduación:

$$G(u_x) = v_x = G(t_x + e_x) = t_x + e'_x \Rightarrow |e'_x| < |e_x| \quad \forall x .$$

Esto es, se supone que disminuyen los errores cometido.

"Método de las medias móviles". Fue uno de los primeros métodos desarrollados. Aunque se inició de forma intuitiva, sólo en busca de un objetivo práctico y de fácil aplicación para la suavización de la secuencia de estimaciones, se llega a una expresión más formal en estudios posteriores. Para dar solución al tratamiento de los extremos de la secuencia aparecen ciertos estructuras matriciales en Hoem, J.M. and Linnemann, P. (1.988). Por su mayor alcance destacan las extensiones desarrolladas utilizando el proceso de estimación de kernel tal como se encuentran en Gavin, John; Haberman, Steven and Verral, Richard (1.993), y Gavin, John B.; Haberman, Steven and Verral, Richard J. (1.994).

A pesar de la sencillez del método de las medias ponderadas móviles, se considera una herramienta flexible que proporciona resultados más robustos que métodos más valorados en la literatura actuarial, tal como la graduación por fórmula matemática. En definitiva, suponen una ventaja cuando se desea aceptar un menor número de supuestos de partida ya que debido a su robustez no se alteran de forma significativa los resultados que proporciona.

Se comienza con el planteamiento y la solución para el problema de la determinación de **coeficientes en los valores centrales de la secuencia**. El valor graduado se halla mediante una media ponderada de un cierto número de valores no graduados consecutivos. Siendo v_x los valores ajustados, u_x las estimaciones iniciales, t_x el verdadero valor subyacente, e_x el error cometido en la estimación inicial y a_r las ponderaciones simétricas en torno al año x considerado, se representa:

$$v_x = \sum_{r=-n}^{n} a_r \cdot u_{x+r} = \sum_{r=-n}^{n} a_r \cdot t_{x+r} + \sum_{r=-n}^{n} a_r \cdot e_{x+r} = \sum_{r=-n}^{n} a_r \cdot t_{x+r} + e_x'$$

$$con \quad u_x = t_x + e_x \quad , \quad e_x' = \sum_{r=-n}^{n} a_r \cdot e_{x+r} \quad y \quad u_x = t_x + e_x$$

tomando en consideración un rango de $2n+1$ estimaciones iniciales para formar cada valor graduado. Se intenta minimizar el error estadístico representado por e_x', tras reproducir t_x. La suavización en los saltos entre las estimaciones consecutivas debe ser el inicio y la consecuencia del proceso.

En primer lugar, con una hipótesis funcional (guiada por la opinión previa de suavidad) de la secuencia de los t_x en el rango $[x-n,x+n]$ según un polinomio de grado cúbico se obtiene la deseada reproducción en los verdaderos valores t_x (propiedad reproductiva) con sólo dos restricciones:

$$t_x = \sum_{r=-n}^{n} a_r \cdot t_{x+r} \quad si \quad \begin{cases} \sum_{r=-n}^{n} a_r = 1 \\ \sum_{r=-n}^{n} r^2 \bullet a_r = 0 \end{cases}.$$

Además, considerando el carácter simétrico de los coeficientes, $a_r = a_{-r}$ para los valores de r en el rango, se tienen:

$$\begin{cases} \sum_{r=-n}^{n} r \bullet a_r = 0 \\ \sum_{r=-n}^{n} r^3 \bullet a_r = 0 \\ \sum_{r=1}^{n} r^2 \bullet a_r = 1 \\ a_0 + 2\sum_{r=1}^{n} a_r = 1 \end{cases}$$

En segundo lugar, para minimizar e_x' se considera que es una concrección de la variable aleatoria E_x' . El valor probable será:

$$E[E_x'] = E[t_x - V_x] = t_x - E[V_x] = t_x - E\left[\sum_{r=-n}^{n} a_r \cdot U_{x+r}\right] =$$

$$= t_x - \sum_{r=-n}^{n} a_r \cdot E(U_{x+r}) = t_x - \sum_{r=-n}^{n} a_r \cdot t_{x+r} = t_x - t_x = 0$$

Para esto se admite U_x como proporción binomial con $E[U_x]=t_x$ y que t_x cumple la propiedad reproductiva. Como la variable aleatoria E_x' es insesgada, se centra el problema en obtener la menor varianza posible:

$$Var[E_x'] = Var[t_x - V_x] = Var[V_x] = Var\left[\sum_{r=-n}^{n} a_r \cdot U_{x+r}\right] =$$

$$= Var\left[\sum_{r=-n}^{n} a_r \cdot E_{x+r}\right] = \sigma^2 \cdot \sum_{r=-n}^{n} (a_r)^2 \quad .$$

$$con \quad Var(E_{x+r}) = Var(U_{x+r}) = \sigma^2 \quad \forall r$$

Como se observa se incorpora la hipótesis de homocedasticidad en la variante de estimación. Dado que la varianza va a depender del tamaño muestral n_x podrían incorporarse correcciones de la forma:

$$Var[E_x + r] = \sigma^2 \cdot \frac{n_x}{n_{x+r}} \quad siendo \quad Var[E_x] = \sigma^2 .$$

Con este desarrollo surgen la combinación de valores que se han de minimizar con las restricciones expuestas. Esta expresión final se conoce como ratio de varianzas debido a que se puede obtener como cociente entre las varianzas de la variante graduada y sin graduar:

$$R_0^2 = \sum_{r=-n}^{n} (a_r)^2 = \frac{Var[V_x]}{Var[U_x]}.$$

Su significado es claro: debe ser menor que uno para que el proceso de graduación mejore la estimación estadística.

En la práctica, para aumentar la suavidad en el ajuste, se define una expresión a minimizar de la forma:

$$R_z^2 = \frac{Var(\Delta^z V_x)}{Var(\Delta^z U_x)},$$

considerándola una generalización de la expresión inicial y recomendándose z=2,3,4.

El numerador y denominador de la expresión anterior se obtiene (con supuestos de incorrelación y equivarianza) como $Var(\Delta^z V_x) = \sigma^2 \cdot \sum_{-n-z}^{n} (\Delta^z a_r)^2$

y como $Var(\Delta^2 U_x) = \binom{2z}{z} \cdot \sigma^2$.

Se llega a:

$$R_z^2 = \frac{1}{\binom{2z}{z}} \sum_{r=-n-z}^{n} (\Delta^z a_r)^2$$

Otra alternativa sería dar énfasis a la suavización de los datos y minimizar una expresión del tipo

$$\sum_x (\Delta^z v_x{}^2) \quad con \quad \Delta^z v_x = (-1)^z \sum_{-n-z}^{n} \Delta^z a_r \cdot u_{x+r+z}$$

donde se usan concreciones en lugar de variables aleatorias.

Se han desarrollado otros métodos para minimizar el ratio de varianzas teniendo en cuenta las restricciones.

"Fórmula de Whittaker". Método iniciado por Whittaker, E.T. y Henderson, R. Se observa como **subyace una racionalidad bayesiana formal**. Posteriores trabajos obtienen el desarrollo bayesiano formal junto con la obtención del parámetro de suavidad según Carlin, Bradley P. and Klugman, Stuart A. (1.993); y, además, se generaliza e interpreta con Taylor, Greg (1.992) y Verrall, R.J. (1.993). Por último, Broffit, James D. (1.996) generaliza el proceso en un entorno multidimensional para las medidas de nivel ajuste y nivel de suavidad.

Tal como expresan sus autores iniciales, la opinión previa sobre la secuencia de los verdaderos valores de t_x, realizaciones de la secuencia de variables aleatorias T_x, se concreta expresando la densidad de probabilidad de una secuencia dada de t_x, sería:

$$f_T(t_x) = c_1 \cdot e^{-lS} \quad con \quad S = \sum_{x=1}^{n-z} (\Delta^z t_x)^2 .$$

Siendo l una constante que puede determinar la medida de la confianza que se tiene en la opinión previa y c_1 otras constante que toma el valor que hace de la función una verdadera función de densidad. En principio, una secuencia de valores de t_x es más probable si S toma valores menores (más suave). Ahora bien, Whittaker considera que

la secuencia más probable previa a la información suministrada por el dato observado, está indefinida.

Por otra parte, la variable aleatoria E_x, que informa de la desviación entre la variable aleatoria estimación inicial de t_x y el verdadero valor subyacente desconocido, se considera con distribución normal de media cero y con desviación típica finita. Por ello, la función de densidad para un valor determinado y es:

$$f_{E/T}(e_y/t_y) = c_y \cdot e^{-\frac{1}{2}\left(\frac{e_y}{s_y}\right)^2}$$

$$f_{U/T}(u_y/t_y) = c_y \cdot e^{-\frac{1}{2}w_y(t_y-u_y)^2}$$

siendo w_y la recíproca de la varianza de U_y.

Para toda la secuencia, asumiendo la independencia para todos los valores de x (todas las edades):

$$f_{U/T}(u_x/t_x) = c_2 \cdot e^{-\frac{1}{2}\sum_{x=1}^{n} w_x(t_x-u_x)^2}$$

interpretado como la probabilidad de obtener la secuencia entera de estimaciones iniciales supuesto que la secuencia de los valores verdaderos es la secuencia t_x.

A continuación se puede obtener la función de densidad de t_x condicionada a las estimaciones iniciales:

$$f_{T/U}(t_x/u_x) = \frac{f_{U/T}(u_x/t_x) \bullet f_T(t_x)}{f_U(u_x)} = \frac{c_3 \cdot e^{-lS-\frac{1}{2}\sum_{x=1}^{n} w_x(t_x-u_x)^2}}{f_U(u_x)}$$

con c_3=$c_1 \cdot c_2$.

Se considera, entonces, la secuencia más probable de verdaderos valores condicionado por la información que suministran las estimaciones iniciales. Para ello se maximiza la expresión anterior minimizando:

$$lS + \frac{1}{2}\sum_{x=1}^{n} w_x\left(t_x - u_x\right)^2 = lS + \frac{1}{2}F \quad .$$

Finalmente se opta por minimizar una expresión de la forma: $2lS + F = hS + F$.

Se llega a una fórmula que recoge los niveles de ajuste y de suavidad en la graduación. Se obtienen los valores rectificados de la estimación (valores de v_x) minimizando, entonces, la expresión:

$$M = F + hS = \sum_{x=1}^{n} w_x (v_x - u_x)^2 + h \sum_{x=1}^{n-z} (\Delta^z v_x)^2 .$$

El primer sumando valora el nivel de ajuste entre las estimaciones iniciales y las estimaciones corregidas. Pondera las desviaciones cuadráticas asignando pesos distintos a cada desviación. El segundo sumatorio es una medida de la suavidad de la secuencia de las estimaciones rectificadas (considerada como cualidad necesaria en la secuencia de estimaciones).

Una forma sencilla de establecer las ponderaciones es identificarlas con el tamaño de la muestra disponible (a mayor tamaño más peso) en cada estimación u_x o establecer una relación de la forma: $w_x=(n_x)/(n)$, siendo n la media aritmética de todos los tamaños de muestras utilizados para la estimación de toda la serie de valores considerada. Con datos de mortalidad, cuando se identifica ponderación y tamaño de la muestra, resulta que

$$\sum_{x=1}^{n} w_x (v_x - u_x) = \sum_{x=1}^{n} x \bullet w_x (v_x - u_x) = 0 .$$

Esto es: el número total de muertes, las edades totales de muerte y la edad media de muerte coinciden para los datos observados y para los datos graduados (considerando n_x; v_x como las muertes esperadas).

Otra manera de proceder para fijar los valores de w_x es seguir las consideraciones ya manifestadas sobre el uso del recíproco de la varianza estimada de u_x, para lo que se emplean los valores de v_x en lugar de los desconocidos verdaderos valores t_x (una simplificación aun mayor podría llevar a utilizar las estimaciones iniciales u_x):

$$w_x = \frac{n_x}{v_x(1-v_x)} \approx \frac{n_x}{u_x(1-u_x)} .$$

El parámetro h (número real positivo) es un elemento de control de la importancia relativa que se atribuye a los dos sumandos de la expresión, es decir, sirve para dar mayor o menor relevancia a la suavidad sobre el ajuste o viceversa. Si se hiciera h=0 no se corregirían las estimaciones iniciales. Si h tomara un valor suficientemente grande se prescindiría del proceso de ajuste.

“Procesos bayesianos formales”. La estadística bayesiana permite formalizar la opinión previa que se tenga mediante la utilización de distribuciones de probabilidad "a priori" y "a posteriori" de la muestra de datos obtenida mediante observación. La aproximación bayesiana es natural en el problema de las tasas de mortalidad debido a los numerosos y extensos estudios existentes. Sería un error no usar esta voluminosa cantidad de información. La información suministrada por una muestra elegida puede cambiar la idea del comportamiento probabilístico que se tuviera sobre el parámetro (debido a investigaciones anteriores y/o a otras muestras utilizadas en el pasado) y, en este sentido, es posible aceptar que exista una distribución "a posteriori" del parámetro, donde se recoge la modificación de la distribución de probabilidad de dicho parámetro cuando se dispone de la información muestral.

Por ello, el proceso bayesiano aplicado a la estimación en el análisis de mortalidad requiere asignar distribuciones de probabilidad al parámetro desconocido t_x (según la opinión previa a la obtención de la muestra) ya que se considera es una variable aleatoria: será la distribución "a priori". La selección de la función de densidad previa de t_x sería una manera de caracterizar o explicitar el grado de conocimiento sobre el parámetro t_x. Se podría plantear también la idea de un parámetro t_x fijo aunque desconocido (no aleatorio en el sentido conceptual estricto) al que se tratase matemáticamente como una variable aleatoria. Se considerará T_x como variable aleatoria del parámetro t_x, T como el vector aleatorio para todos los valores de x, y t como el vector de realizaciones o concreciones de T. Dado el carácter multidimensional del problema se obtendrá una función multivariante de densidad previa para los parámetros que expresan el comportamiento aleatorio de la variante tasa de mortalidad para los distintos x (normalmente edades enteras y para un periodo de un año). Esta función debería contener en su estructura, teóricamente, todos los conocimientos y opiniones previas sobre T. En la práctica se consigue una aproximación a esa densidad ideal utilizando un miembro de una familia adecuada de distribuciones de probabilidad que presente propiedades matemáticas convenientes.

Tras formular la distribución de probabilidad previa se precisa seleccionar una expresión para la distribución condicionada de los datos observados, supuesta la secuencia t_x. Se trata del modelo experimental o modelo estocástico que siguen los datos en función o teniendo en cuenta la presentación de t. Los datos de este modelo muestral serán las estimaciones iniciales obtenidas con la información almacenada por las

compañías de seguros (recogidos en sus polizas). En estos datos se puede considerar su distribución (sin apoyarse en un modelo probabilístico según cierto *t*).

Con todo ello, se llega a la obtención de la distribución "a poteriori" de *T* condicionada a las estimaciones iniciales:

$$f_{T/U}(t/u) = \frac{f_{U/T}(u/t)}{f_U(u)} f_T(t).$$

Para el análisis y tratamiento posterior de la distribución "a posteriori" es conveniente una forma adecuada que se puede conseguir con el empleo de distribuciones conjugadas. El vector de valores graduados se obtiene a partir de la información contenida en esta distribución a "posteriori". Si interesa el valor esperado se hallará la media, si interesa el valor más probable se hallará la moda, etc.

El método de Kimeldorf-Jones aparece en los trabajos de Kimeldorf, G.S. and Jones, D.A. (1.967) y se considera el primer método formal de graduación bayesiana. Aplica las líneas generales expuestas, se concreta la forma de las distintas funciones de densidad que intervienen, se analiza el significado de los parámetros y se dan directrices para su estimación.

Otro sistema de graduación con base en metodología bayesiana es la graduación bayesiana isotónica que aplican Broffit, James D. (1.988). Se corrige una primera función de verosimilitud obtenida a partir de los datos con la distribución previa de los parámetros (obligados a un cierto orden creciente) para obtener la distribución corregida que conduce a un estimador bayesiano. La influencia del modelo (que representa la información previa) evita que los resultados se apoyen exclusivamente en los datos, que ante una muestra poco representativa lleva a una estimación no adecuada.

“Métodos paramétricos”. La secuencia de valores revisados o graduados en los métodos no paramétricos se explicitaba en forma de tabla o forma numérica. En los métodos paramétricos se expresará como una función de argumento *x*, siendo v_x una función que contiene los parámetros que deben fijarse desde los datos que configuran las estimaciones iniciales. La justificación de este tipo de graduación se basa en la necesidad de corregir las imprecisiones de la estimación (debidas a datos mal ubicados en el tiempo y a la aleatoriedad del muestreo) mediante el ajuste de una función matemática que se supone representa las verdaderas tasas de mortalidad.

La opinión previa se manifiesta en la determinación de la forma funcional elegida (en la elección de dicha función también participa la información que suministran los datos). Las técnicas que se utilizan en el ajuste de las curvas para obtener los valores graduados (estimaciones corregidas) son las características de ajustes de curvas y de regresión en casos generales. Para iniciar una graduación por fórmula matemática, se debe elegir el tipo de funciones. Se distinguen distintas posibilidades según el rango de edades en el cual se desea utilizar el método.

Si se aplica para un "rango de edades no extenso" existen múltiples funciones básicas que se pueden utilizar. Son las mismas funciones que se detallaron en el análisis de mortalidad y que fueron empleadas, en ciertos casos, como distribución de muertes dentro del intervalo anual, con el fin de realizar las estimaciones iniciales de mortalidad en el año considerado. Las funciones mencionadas eran: ley de Moivre; primera, segunda y tercera ley de Dormoy; ley de Sang; ley de Gompertz; primera y segunda ley de Makeham; ley de Lazarus; ley de Risser; ley de Barnett; ley de Wilkie; y, ley de Weibull.

Ciertas expresiones ofrecen la posibilidad de extenderse a un "amplio rango de edades", son expresiones conocidas como "fórmulas Gompertz- Makeham de tipo (r,s)",

$$GM_a^{r,s}(x) = \sum_{i=1}^{r} \boldsymbol{a}_i x^{i-1} + e^{\sum_{i=r+1}^{r+s} a_i x^{i-r-1}},$$

o como "fórmulas Logit Gompertz- Makeham de tipo (r,s)",

$$LGM_a^{r,s}(x) = \frac{GM_a^{r,s}(x)}{1+GM_a^{r,s}(x)},$$

que ofrece la ventaja de ser un logit con rango de valores posible entre cero y uno. Estas funciones son estudiadas dentro del marco de los modelos lineales y no lineales generalizados de utilidad en el ámbito actuarial por Renshaw, A.E. (1.991).

Extenderse a "todo el rango de edades", desde el primer año de vida al último posible, supone considerar funciones más flexibles. No existen muchos intentos en esta dirección ya que la solución más habitual lleva a dividir el rango de edades utilizando para cada rango una adecuada función.

Una descripción de las expresiones matemáticas que tratan de abarcar todo el rango de edades, se encuentra en Betzuen Zalbidegoitia, Amancio; Felipe checa, Angie y Guillén Estany, Monserrat (1.997), se destaca el modelo de Heligman y Pollard. Otro estudio de interés se encuentra en Yuen, Kam C. (1.997).

La expresión debida a Heligman, L and Pollard, L.H. (1.980) es:

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

Para utilizar esta fórmula hay que determinar el número de sumandos, determinar la función y ajustar los parámetros resultantes. En la población de Australia se utilizó la expresión:

$$\frac{q_x}{1-q_x} = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^{(x-x_0)} .$$

Se trata de una función continua y que toma valores entre cero y uno, se aplica a todo el rango de edades (aunque no incluya muchos parámetros), y tiene una deseable interpretación demográfica: el primer sumando es una exponencial de crecimiento rápido que indica el descenso de mortalidad en los primeros años; el segundo sumando recoge mediante una función similar a la lognormal la mortalidad por accidentes; el tercer sumando manifiesta la mortalidad en edades avanzadas con una ley de Gompertz.

El significado de los parámetros se detalla a continuación. El parámetro *A* tiene un valor semejante al que toma la probabilidad de fallecimiento durante el primer año de vida. El parámetro *B* mide la localización de la mortalidad en el primer año de vida, teniendo efecto sólo a la edad de cero años. El parámetro *C* valora la disminución de la probabilidad de muerte en la infancia (edades superiores al año). El parámetro *D* mide la magnitud de la típica "joroba de accidentes" para individuos jóvenes. El parámetro *E* es inverso a la dispersión. El parámetro *F* localiza el máximo de la mencionada "joroba". El parámetro *G* refleja el nivel base en la mortalidad senil, esto es, la mortalidad de tipo adulta-senil que se considera existe al nacer. El parámetro *H* indica el tanto de aumento de la mortalidad anterior para edades adultas. El parámetro x_0 especifica una edad para la cual $q_x = 0'5$, muy próxima a la máxima considerada (la edad que se suele representar como *w*).

Se recomiendan, a nivel teórico, los métodos bayesianos formales del tipo Kimeldorf-Jones cuando se tenga información que permita estimar las covarianzas de las distribuciones a priori y condicionadas. En caso contrario, se propone un tipo particular de graduación paramétrica basada en la ley de Heligman y Pollard (con parámetros con significación real). A nivel práctico, se recomienda que todo proceso de graduación debe seguirse de pruebas que garanticen su bondad, permitiendo diferenciar a posteriori, según resultados, eligiendo así las técnicas que proporcionan mejores valores. Las pruebas de bondad más adecuadas serán: test de intervalos de confianza, test de las desviaciones acumuladas, test signos, test de cambios de signo, test de la chi-cuadrado y test de Kolmogorov-Smirnov.

4. Grupos homogéneos y credibilidad.

La separación en grupos de riesgos incluirá todas las variables disponibles para efectuar los análisis multivariantes. Se prefiere la separación en grupos de riesgo frente al tratamiento conjunto mediante modelos multivariantes debido a los objetivos perseguidos y al tipo de variables, muchas veces cualitativas, que se recogen como información de partida relevante en las pólizas.

Se condiciona el análisis al nivel de información disponible. Un primer análisis lleva a distinguir entre las posibles causas de muerte como forma de asignar riesgos y valorar los individuos más expuestos. No es que en principio interese la causa de muerte (las indemnizaciones pactadas en las pólizas se basan casi siempre en la muerte sin distinguir causas) sino que se desea buscar las causas de muerte para obtener a continuación los factores que incidan en su aparición.

Es evidente que siempre existirán factores de riesgo no observables por lo que las clases no serán realmente homogéneas. Todo esfuerzo en este campo será siempre una aproximación. Además, la separación en grupos supone una discriminación (al menos económica) cuya justificación científica será siempre discutible: por los datos utilizados, por los factores causales o aproximados considerados, por el método elegido y por la imposibilidad de conocer el fenómeno real subyacente para verificar el método. Las respuestas que se pueden dar ante esta situación se basan en la necesidad real por motivos económicos (evitar antiselección, insolvencia, injusticia en precios...) y en la posibilidad técnica de establecer métodos objetivables (aunque no indiscutibles).

La discriminación de los distintos grupos de riesgo para la posterior estimación de la mortalidad en cada uno de ellos es uno de los activos inmateriales más importantes de la empresa aseguradora de vida, supone una garantía de éxito en un mercado competitivo. Un ejemplo que recoge los fundamentos de la clasificación de riesgos en seguro de vida, las prácticas actuales en las aseguradoras y formas objetivas de valoración se encuentra en el trabajo "Risk classification in life insurance" de Cummins, J.D.; Smith, B.D.; Vance, R.N. and VanDerhei, J.L. (1.983). Destaca el uso de la regresión con un modelo logístico múltiple como método cuantitativo para asignar a la póliza anual una determinada probabilidad de muerte según las características del sujeto asegurado en el momento de la suscripción. Una alternativa parecida se obtiene a través de los modelos Probit. También se persigue diferenciar grupos y asignar puntuaciones que midan el incremento de riesgo debido a combinaciones de factores.

La detección de anomalías se basará en el estadístico propuesto en el trabajo de Serrano, Gregorio R. (1.995). Se aplica en modelos de elección binaria según la influencia de cada elemento muestral en la estimación de máxima verosimilitud. Su utilidad fundamental radica en ofrecer un método para detectar datos heterogéneos en un grupo o, al menos, indica la influencia relativa de cada dato en el proceso de estimación. En la selección de grupos homogéneos para análisis de mortalidad en una cartera de pólizas es útil en cuanto detecta los individuos que distorsionen la estimación dentro del grupo.

Según el tamaño de los grupos de riesgo se puede necesitar en los procesos de estimación recurrir a la Teoría de Credibilidad para encontrar estimaciones de compromiso. Se adapta bien a la situación estudiada la denominada teoría de la credibilidad clásica o de escuela americana.

La **solución de compromiso** se puede reflejar de la forma: $\boldsymbol{q}_j^a = z\hat{\boldsymbol{q}}_j + (1-z)\hat{\boldsymbol{q}}$

siendo $\hat{\boldsymbol{q}}_j$ la estimación con la información existente para la subclase especificada, $\hat{\boldsymbol{q}}$ la estimación para todo el grupo (con su heterogeneidad) y $\boldsymbol{q}_j^a$ la solución de compromiso obtenida con las anteriores estimaciones combinadas según el factor de credibilidad.

Siguiendo los argumentos de la **teoría de la credibilidad clásica** o de escuela americana se puede buscar un margen de fluctuación (o de error en la estimación) limitado de la forma $P\left[\left|\hat{\boldsymbol{q}}_j - \boldsymbol{q}_j\right| \le k\boldsymbol{q}_j\right] \ge 1-\boldsymbol{e}$

según el número *n* de observaciones utilizadas para estimar el parámetro θ_j.

Para ello se recurre a una aproximación a la distribución normal del estadístico estimador

$$\hat{\theta}_j \xrightarrow{d} N\left(\theta_j, \sqrt{\frac{\theta_j(1-\theta_j)}{n}}\right)$$

que permita obtener un umbral n_0 que garantice un límite en el valor $\left|\hat{\theta}_j - \theta_j\right|$ medido según una probabilidad prefijada. Operando con los dos resultados anteriores se llega a

$$P\left[\left|N(0,1)\right| \le \frac{k\theta_j}{\sqrt{\frac{\theta_j(1-\theta_j)}{n}}}\right] \ge 1-\varepsilon \Rightarrow \frac{k\theta_j}{\sqrt{\frac{\theta_j(1-\theta_j)}{n}}} \ge \lambda_{\varepsilon/2}$$

siendo $\lambda_{\varepsilon/2} / P\left[\left|N(0,1)\right| \le \lambda_{\varepsilon/2}\right] = 1-\varepsilon$.

Desde este resultado se obtiene la condición para una credibilidad total de la estimación y el umbral de credibilidad total:

$$n \ge \frac{(\lambda_{\varepsilon/2})^2(1-\theta_j)}{k^2 \cdot \theta_j} \quad , \quad n_0 = \frac{(\lambda_{\varepsilon/2})^2(1-\theta_j)}{k^2 \cdot \theta_j}.$$

Aparecen dos problemas prácticos: el desconocimiento del valor real del parámetro (se sustituye en las expresiones el valor del parámetro por su estimación) y los altos valores que alcanza el umbral para valores muy pequeños del parámetro (caso del seguro de vida que lleva a la necesidad de grandes carteras de pólizas registradas).

Si no se cumple la condición para la credibilidad total, el problema consistirá en hallar un valor del factor de credibilidad *z* con el fin de poder utilizar la mencionada solución de compromiso: $\theta_j^a = z\hat{\theta}_j + (1-z)\hat{\theta}$.

Para ello se admite que el error cometido al utilizar la combinación de estimaciones propuesta por la teoría de la credibilidad será: $\theta_j^a - \theta_j = z\left(\hat{\theta}_j - \theta_j\right) + (1-z)\left(\hat{\theta} - \theta_j\right)$.

En esta expresión se limita el error debido a la estimación según la experiencia individual (de la subclase)

$$P\left[z\left|\hat{q}_j - q_j\right| \le k q_j\right] \ge 1-e$$

y se aplica la convergencia en distribución a la ley normal:

$$P\left[\left|N(0,1)\right| \le \frac{k q_j}{z\cdot\sqrt{\frac{q_j(1-q_j)}{n}}}\right] \ge 1-e \Rightarrow \frac{k q_j}{z\cdot\sqrt{\frac{q_j(1-q_j)}{n}}} \ge l_{e/2},$$

siendo $l_{e/2} / P\left[\left|N(0,1)\right| \le l_{e/2}\right] = 1-e$.

Utilizando la condición de credibilidad total, el umbral de credibilidad total y sustituyendo en la expresión anterior se llega al valor recomendado para el factor de credibilidad:

$$z = min\left\{\sqrt{\frac{n}{n_0}}, 1\right\}.$$

5. Proyección de estimaciones.

Para finalizar el proceso de estimación se puede intentar efectuar proyecciones a medio plazo, siempre que se disponga de información pasada, para apoyar los cálculos de mortalidad futura efectuados en los seguros de supervivencia a largo plazo.

Parece evidente la evolución en el tiempo cronológico del fenómeno de la mortalidad. En el trabajo de MacDonald, A.S. et al. (1.998) se aportan y analizan estudios empíricos para comparar la evolución y las últimas tendencias apreciadas en la mortalidad en diferentes países. Esto indica que un estudio de mortalidad en base a datos presentes debe completarse, cuando se busque cierta validez a lo largo del tiempo, con una proyección hacia futuro. Los métodos basados en proyecciones para cada valor de la tabla de mortalidad utilizan modelos paramétricos para cada edad, con la variable tiempo calendario de forma explícita. Los métodos basados en proyecciones del modelo estructural utilizado en la tabla de mortalidad obtienen las probabilidades de fallecimiento proyectadas mediante cambios en los parámetros del modelo.

Se destacan y se sigue a continuación, por su aplicación en el campo actuarial, las investigaciones de Benjamin, B. and Soliman, A.S. (1.995) y de Felipe Checa, María de los Ángeles y Guillén Estany, Montserrat (1.999).

Una vez realizadas las estimaciones definitivas, para proceder a modelizar la mortalidad considerando el tiempo de calendario se puede proceder de tres formas: según proyecciones para cada valor de la tabla de mortalidad, según proyecciones del modelo estructural tomado en la tabla de mortalidad y según una proyección atendiendo a las causas del fallecimiento.

Los métodos basados en **proyecciones para cada valor de la tabla de mortalidad** utilizan modelos paramétricos para cada edad en los que aparece la variable tiempo calendario de forma explícita. Son modelos del tipo: $q_x(t) = f(x, \boldsymbol{a}, t)$,

donde x es la edad que se estudia, t la variable tiempo calendario y $\boldsymbol{a}$ un vector de parámetros. Exigen un seguimiento en el tiempo de los estudios de mortalidad.

En las tablas de mortalidad alemanas DAV1.994R se emplea el modelo: $q_x(t) = a_x b_x^t$.

Se linealiza tomando logaritmos y se aplica el método de mínimos cuadrados ordinarios. Los parámetros estimados determinan el modelo que proyecta la probabilidad de muerte según se introducen valores a la variable t.

En las tablas españolas, realizadas para U.N.E.S.P.A. por Fernández Plasencia, M.J. y Prieto Pérez, Eugenio (1.994) para proyectar la mortalidad se utiliza un modelo para la esperanza de vida: $e_x(t) = a_x + b_x\left[1 - e^{d_x t}\right]$,

que permite obtener la probabilidad de muerte mediante la relación

$$q_x = \frac{0'5 - e_{x+1}}{e_x - 0'5}$$

En las tablas de mortalidad suizas GRM/GRF1.995 se emplea el modelo:

$q_x(t) = q_x(t_0) \cdot e^{-d_x(t - t_0)}$, siendo t_0 el tiempo calendario inicial.

De forma conjunta pero con un cambio de escala, se consideran la edad y el tiempo en el trabajo de Haberman, S.; Hatzopoulos P. and Renshaw, A.E. (1.996), mediante una expresión referida al tanto instantáneo de mortalidad y que contiene catorce parámetros:

$$m_x(t) = \exp\left[b_0 + \sum_{j=1}^{5} b_j L_j(x')\right] \exp\left[\left(a_1 + \sum_{j=1}^{3} c_{1j} L_j(x')\right) t' + \left(a_2 + \sum_{j=1}^{3} c_{2j} L_j(x')\right) t'^2\right].$$

Todos estos modelos suponen la ausencia de cambios estructurales en el tiempo en el que se toman las muestras y en el tiempo futuro en que se proyecta. En estos casos, sólo la simulación de escenarios ayudaría a analizar el impacto futuro de cambios estructurales.

Los métodos de estimación que se utilizan en la estimación de los parámetros suponen perturbaciones aleatorias con comportamiento normal de media cero. Esta hipótesis es discutible debido a las transformaciones logarítmicas que se realizan en los modelos. Además, se debería probar la no existencia de autocorrelación.

Los métodos basados en proyecciones del modelo estructural utilizado en la tabla de mortalidad obtienen las probabilidades de fallecimiento proyectadas mediante cambios en los parámetros del modelo. El modelo más empleado en este tipo de proyecciones es el ya mencionado modelo de Heligman-Pollard. Un ejemplo se encuentra en Felipe Checa, María de los Ángeles y Guillén Estany, Montserrat (1.999): se estima la probabilidad de muerte por mínimos cuadrados no lineales ponderados (según la inversa de la probabilidad de muerte observada) para todo el rango de edades y cada uno de los años calendario, después se aplica el análisis univariante de series temporales (ARIMA) para cada uno de los parámetros de la ley Heligman-Pollard con el fin de obtener predicciones de todos ellos, éstas predicciones determinan las predicciones de las probabilidades de muerte (para cada edad en el calendario futuro).

A pesar del interés de las proyecciones del modelo estructural, persiste el problema de la invalidez ante cambios posteriores al periodo de observación. Esta situación se debe al uso de series temporales (los parámetros tienen comportamiento autorregresivo). En la práctica, avances médicos relevantes o epidemias significativas impiden la validez de las proyecciones efectuadas.

Los métodos basados en proyecciones atendiendo a las causas del fallecimiento argumentan que podrían obtenerse mejores predicciones de mortalidad futura proyectando por separado la mortalidad debida a ciertos grupos de enfermedades y la mortalidad un residuo de fallecimiento por causas varias. Finalmente, se sumarían las distintas proyecciones. Benjamin, B. and Soliman, A.S. (1.995) utilizan la clasificación de enfermedades de la O.M.S., las causas se consideran excluyentes y se deja una causa definida como “resto de causas”. La relación entre la probabilidad general de muerte y la probabilidad según las diferentes causas de muerte (*I*,*II*,*VII*,*VIII*,*IX*,*XII*,*R*) se supone de la forma:

$$\ln q_x^t = \ln q_x^t(I)\cdot\ln q_x^t(II)\cdot\ln q_x^t(VII)\cdot$$
$$\cdot\ln q_x^t(VIII)\cdot\ln q_x^t(IX)\cdot\ln q_x^t(XII)\cdot\ln q_x^t(R).$$

La evolución en el tiempo calendario se toma como:

$$\ln q_x(i) = \ln a_x(i) + t\ln b_x(i) \quad con \quad i = I, II, VII, VIII, IX, XII, R.$$

Evidentemente, se necesita información detallada de la causa de muerte para realizar este análisis. Presenta los inconvenientes mencionados en el método denominado de proyección para cada valor de la tabla (sería una variante realmente).

La mejor proyección dependerá de la información disponible y de si se emplea o no modelo estructural en la tabla.

Bibliografía.

Benjamin, B. and Soliman, A.S. (1.995). *Mortality on the move,* Ed. City University Print. Londres.

Betzuen Zalbidegoitia, Amancio (1.999). "La medida de la mortalidad en un colectivo de activos ocupados", *Actuarios,* Núm.17, Dossier, pp.:1-8.

Betzuen Zalbidegoitia, Amancio; Felipe Checa, Angie y Guillén Estany, Monserrat (1.997). "Modelos de tablas de mortalidad en España y situación actual", *Anales del Instituto de Actuarios Españoles*, Tercera época, Núm.3, pp.:79-104.

Breslow, N. and Crowley, J. (1.974). "A large sample study of the life table and product limit estimates under random censorship", *The Annals of Statistics*. Vol.2, Núm.3, pp.:437-453.

Broffit, James D. (1.984). "Maximum likelihood alternatives to actuarial estimators of mortality rates", *Transactions of the Society of Actuaries*. Vol.36, pp.:77-122.

Broffit, James D. (1.988). "Increasing and increasing convex bayesian graduation (with discussion)", *Transactions of the Society of Actuaries,* Vol.XL, pp.:115-148.

Broffit, James D. (1.996). "On smoothness terms in multimensional Whittaker graduation", *Insurance: Mathematics and Economics,* Vol.18, Núm.1, pp.:13-27. North-Holland.

Carlin, Bradley P. and Klugman, Stuart A. (1.993). "Hierachical Bayesian Whittaker graduation", *Scandinavian Actuarial Journal,* Núm.2, pp.:183-196.

Chen, Kani and Lo, Shaw-Hwa (1.997). "On the rate of uniform convergence of the product-limit estimator: strong and weak laws", *The Annals of Statistics,* Vol.25, Núm.3, pp.:1050-1087.

Cummins, J.D.; Smith, B.D.; Vance, R.N. and VanDerhei, J.L. (1.983). *Risk classification in life insurance,* Kluwer Academic Publishers. Boston.

Elandt-Johnson, Regina C. and Johnson Norman, L. (1.980). *Survival models an data analysis.* Wiley. New York.

Felipe Checa, María de los Ángeles y Guillén Estany, Montserrat (1.999). *Evolución y predicción de tablas de mortalidad dinámicas para la población española,* Cuadernos de la Fundación. Editorial Mapfre Estudios.

Fernández Plasencia, M.J. y Prieto Pérez, Eugenio (1.994). *Tablas de mortalidad de la población española de 1.950 a 1.990. Tabla proyectada del año 2.000. Tablas con y sin margen de seguridad,* UNESPA. Madrid.

Gavin, John B.; Haberman, Steven and Verral, Richard J. (1.993). **"**Moving weighted average graduation using kernel estimation", *Insurance, mathematics and economics,* Núm.12, pp.:113-126. North-Holland.

Gavin, John B.; Haberman, Steven and Verral, Richard J. (1.994). "On the choice of bandwidth for Kernel graduation", *Journal of the Institute of Actuaries,* Vol.121, Núm.478, pp.:119-134.

Heligman, L and Pollard, L.H. (1.980). "The age pattern of mortality", *Journal of the Institute of Actuaries,* Núm.107, pp.:49-80.

Hoem, Jan M. and Linnemann, P. (1.988). "**The tails in moving average graduations",** *Scandinavian Actuarial Journal,* Núm.4, pp.:193-229.

London, D. (1.985). *Graduation: the revision of estimates***,** Actex Publications. Connecticut.

London, D. (1.988). *Survival models and their estimation***,** Actex Publications. Connecticut.

MacDonald, A.S. et al. (1.998). "An international comparison of recent trends in poputation mortality", *British Actuarial Journal*, Vol.4, Núm.1, pp.:3-143.

Renshaw, A.E. (1.991). "Actuarial graduation practice and generalised linear and non-linear models", *Journal of the Institute of Actuaries*, Núm.118, Vol.2, pp.:295-312.

Sánchez López, José María y Albarrán Lozano, Irene (1.999). "Graduación actuarial no paramétrica. Una revisión crítica"**,** *Anales del Instituto de Actuarios*, Tercera Época, Núm. 5, pp.:11-50.

Sánchez López, José María (2.001). *Cuantificación de riesgos y análisis global de la empresa aseguradora de vida*, Dykinson, S.L., Madrid.

Sánchez López, José María (2.001). "Construcción de tablas de mortalidad", *Revista Española de Seguros*. Núm.106, pp.:297-306. Madrid.

Serrano, Gregorio R. (1.995). "Estadísticos para la detección de observaciones anómalas en modelos de elección binaria: una aplicación con datos reales", *Revista Estadística Española*, Vol.37, Núm.140, pp.:321-348.

Taylor, Greg (1.992). "A bayesian interpretation of Whittaker-Henderson graduation", *Insurance, Mathematics and Economics*, Núm.11, pp.:7-16. North-Holland.

Verrall, R.J. (1.993). "A state space formulation of Whittaker-Henderson graduation, with extensions", *Insurance, Mathematics and Economics*, Núm.13, pp.:7-14. North-Holland.

Yuen, Kam C. (1.997). "Comments on some parametric models for mortality tables", *Journal of Actuarial Practice*, Vol.5, Núm.2, pp.:253-266.