2

Tablas de mortalidad

2.1 Tablas de Mortalidad

Es un hecho bien conocido que la probabilidad de que un individuo concreto fallezca en un determinado periodo depende de muchos factores, como por ejemplo su edad (el único factor que hemos estudiado en el capítulo anterior), sexo, estado de salud, factores genéticos y ambientales, etc. En efecto, es evidente que la mortalidad aumenta con la edad (salvo en el caso de las edades infantiles). También se sabe que la mortalidad femenina, a igualdad de los restantes factores, es inferior a la masculina. Asimismo es evidente que la mortalidad está relacionada con la aparición de enfermedades graves, aparición que puede acelerarse o retardarse por factores genéticos y ambientales.

Por otro lado, las estadísticas y censos relativos a una población suelen registrar las edades y el sexo de sus componentes, pero no su estado de salud ni su posible exposición a factores de riesgo genéticos o ambientales. Además, si la población es suficientemente grande entonces el principal factor determinante de la mortalidad resulta ser la edad de los individuos. Por esta razón, en el capítulo anterior hemos considerado únicamente la edad como factor determinante de la mortalidad. Llamaremos **población homogénea** a una población en la que se verifique la propiedad anterior, es decir, en la que la edad sea el principal factor determinante de la mortalidad de los individuos. A partir de ahora nos referiremos exclusivamente a una población homogénea, que intuitivamente podemos identificar con la población masculina o con la población femenina de un determinado país o región.

Una **Tabla de Mortalidad (o de Supervivencia)** contiene los elementos básicos que permiten calcular las probabilidades de muerte y supervivencia en una población homogénea, a partir de las cuales se llevan a cabo los cálculos actuariales.

Una tabla de mortalidad típica puede tener la siguiente estructura:

x	q_x	d_x	l_x	e_x
0	0.012964	12964.000	1000000.000	72.80
1	0.001011	997.893	987036.000	72.75
2	0.000704	694.171	986038.107	71.82
...	....	....	....	....
107	1.000000	4396.000	4396.000	0.50
108	1.000000	000.000	000.000	0.00

La primera columna representa las edades de los individuos, que únicamente toman valores enteros. La segunda columna representa las probabilidades de que los individuos de edades $x = 0, 1, 2, ...$ mueran antes de un año. En el capítulo anterior hemos denominado q_x a dichas probabilidades. La última columna representa, como sabemos, la esperanza de vida (abreviada) a las distintas edades*.

Conocidos los valores de las q_x se pueden calcular fácilmente las demás probabilidades básicas ${}_tp_x$, ${}_tq_x$, ${}_{s/t}q_x$ (aunque solamente para valores enteros de x, s y t). En efecto, sabemos que $p_x = 1 - q_x$, y en el apartado 1.3 demostramos que ${}_tp_x = p_x\ p_{x+1}...p_{x+t-1}$. Además, es claro que ${}_tq_x = 1 - {}_tp_x$. Finalmente, de nuevo por un resultado de 1.3 tenemos que ${}_{s/t}q_x = {}_sp_x\ {}_tq_{x+s}$.

Por tanto, la tabla de mortalidad nos proporciona la distribución de probabilidad de la variable aleatoria K_x (número completos de años de vida hasta la muerte de (x)).

Incluso la última columna se puede obtener a partir de los q_x, ya que, como sabemos por el apartado 1.5.2, se verifica que $e_x = \sum_{k=0}^{\infty} {}_{k+1}p_x$. En consecuencia, las tablas de mortalidad podrían constar tan solo de una única columna de valores de q_x. Pero tradicionalmente las tablas de mortalidad incluyen dos columnas más, que informan de los valores de dos nuevas variables denotadas como d_x y l_x, y que definiremos a continuación.

Consideremos un grupo de l_0 recién nacidos, por ejemplo $l_0 = 1000000$ (la elección de su valor es arbitraria). La supervivencia hasta la edad x de uno cualquiera de los recién nacidos es un suceso aleatorio que se puede representar como una variable aleatoria de Bernouilli, que toma el valor 1 con probabilidad $s(x)$ y el valor 0 con probabilidad $1 - s(x)$. Definamos una nueva variable aleatoria $\mathcal{L}(x)$ como el número total de recién nacidos que sobreviven hasta la edad x. Es claro que la variable $\mathcal{L}(x)$ resulta ser suma de l_0 variables de Bernouilli, y por tanto, suponiendo independencia entre ellas, $\mathcal{L}(x)$ debe ser una variable aleatoria binomial. Llamaremos l_x (l de *living*, vivos) a la esperanza de la variable binomial

*En algunas tablas, admitiendo la hipótesis de distribución uniforme de la mortalidad (véase ejercicio 4, se suma $\frac{1}{2}$ a la esperanza de vida abreviada para cada edad de forma que se obiene la esperanza de vida completa

$\mathcal{L}(x)$ que, como es bien conocido, es igual a $l_0\ s(x)$. Tenemos entonces:

$$l_x = E(\mathcal{L}(x)) = l_0\ s(x) \tag{2.1}$$

Los valores de las l_x (número esperado de recién nacidos que sobreviven a las distintas edades $x = 0, 1, 2, 3...$) están tabulados en la cuarta columna de nuestro ejemplo de tabla de mortalidad. Vemos, así, que el número esperado de recién nacidos que llegan a cumplir un año de edad es 987036, el de los que llegan a cumplir dos años de edad es 986038.107, etc. Obviamente, los valores de l_x decrecen según aumenta la edad x, hasta llegar a la edad límite ω (108 años en nuestro ejemplo) en donde evidentemente se tiene que $l_\omega = 0$.

Los valores de las l_x parecen algo arbitrarios, ya que dependen de la elección de l_0. Sin embargo, como veremos a continuación, las probabilidades básicas de muerte y supervivencia ${}_tp_x$, ${}_tq_x$, ${}_{s/t}q_x$ pueden calcularse fácilmente a partir de dichas l_x. En efecto,

$$l_x = l_0\ s(x) \Rightarrow s(x) = \frac{l_x}{l_0} \Rightarrow s(x+t) = \frac{l_{x+t}}{l_0}$$

y por tanto:

$${}_tp_x = \frac{s(x+t)}{s(x)} = \frac{l_{x+t}}{l_x} \tag{2.2}$$

$${}_tq_x = 1 -{}_t p_x = \frac{l_x - l_{x+t}}{l_x} \tag{2.3}$$

$${}_{s/t}q_x = \frac{s(x+s) - s(x+s+t)}{s(x)} = \frac{l_{x+s} - l_{x+s+t}}{l_x} \tag{2.4}$$

Por otro lado, sea $\mathcal{D}(x,t)$ una nueva variable aleatoria que representa el número de individuos del grupo inicial de l_0 recién nacidos que mueren entre las edades x y $x+t$. Llamaremos a su esperanza ${}_td_x$ (d de *dead*, muertos):

$${}_td_x = E(\mathcal{D}(x,t))$$

Es evidente que $\mathcal{D}(x,t) = \mathcal{L}(x) - \mathcal{L}(x+t)$. Por tanto, tomando esperanzas,

$${}_td_x = l_x - l_{x+t}$$

En particular, tomando t=1 obtenemos el número esperado de individuos que mueren entre las edades x y $x+1$, y que notaremos como d_x:

$$d_x = l_x - l_{x+1}$$

Los valores de d_x están tabulados en la tercera columna de nuestro ejemplo de tabla de mortalidad, donde podemos comprobar fácilmente que se verifica la relación anterior: si a l_0=1000000 le restamos l_1=987036 obtenemos d_0=12964, si a l_1 le restamos l_2=986038.107 obtenemos d_1=997.893, etc.

En consecuencia, si conocemos l_0 y la columna de d_x podremos calcular la columna de l_x (ya que $l_{x+1} = l_x - d_x$, y por tanto $l_x = l_0 - \sum_{i=0}^{x-1} d_i$) y de ahí obtener todas las probabilidades básicas.

Son interesantes las relaciones:

$$ {}_tq_x = \frac{l_x - l_{x+t}}{l_x} = \frac{{}_td_x}{l_x} \tag{2.5} $$

de donde se deduce que

$$ q_x = \frac{d_x}{l_x} \tag{2.6} $$

Finalmente,

$$ {}_{s/t}q_x = \frac{l_{x+s} - l_{x+s+t}}{l_x} = \frac{{}_td_{x+s}}{l_x} \tag{2.7} $$

Concluímos este apartado mencionando una vez más que las probabilidades básicas ${}_tp_x$, ${}_tq_x$ y ${}_{s/t}q_x$ pueden calcularse en una tabla de mortalidad, o bien a partir de la columna de q_x, o bien a partir de la columna de l_x, o bien a partir de l_0 y la columna de d_x. Es claro que la segunda opción resulta la más recomendable, pues las fórmulas que aparecen son relativamente sencillas y fáciles de recordar. Por esta razón, aunque la información proporcionada puede variar de una tabla a otra, todas las tablas contienen con seguridad la columna de l_x.

2.2 La Función de Supervivencia

Aunque en las tablas solamente aparecen los valores de l_x para x entero, la función l_x puede perfectamente definirse para cualquier x real y positivo. En tal caso l_x se denomina **Función de Supervivencia**. Se trata de una función real de variable real, decreciente y que interseca al eje de ordenadas en l_0 y al de abscisas en ω.

Pero debemos hacer una pequeña precisión. En el apartado 1.2.3 definimos la función de supervivencia como aquella que para cada edad x proporciona la probabilidad de que un recién nacido alcance con vida dicha edad, y la denotamos como $s(x)$. Ahora denominamos función de supervivencia a l_x, número esperado de individuos del colectivo inicial de l_0 recién nacidos que sobreviven a la edad x. Es claro que la función l_x depende de la elección de l_0, y por tanto deberíamos hablar realmente de una colección de funciones, todas ellas proporcionales entre sí, siendo precisamente $s(x)$ una de ellas, la correspondiente a la elección de $l_0 = 1$

(en efecto, puesto que $l_x = l_0.s(x)$, si $l_0 = 1$ entonces $l_x = s(x)$). Puesto que tanto $s(x)$ como las l_x asociadas a cualquier $l_0 \neq 1$ pertenecen a la misma familia y son proporcionales entre sí, a partir de ahora las denotaremos a todas ellas con la misma denominación de Función de Supervivencia.

Al igual que sucedía con el tanto instantáneo o fuerza de mortalidad, muchos investigadores han propuesto expresiones analíticas de la función de supervivencia. De hecho, se trata del mismo problema, ya que obviamente la fuerza de mortalidad y la función de supervivencia se encuentran relacionadas. En efecto, en el apartado 1.4 comprobamos que

$$\mu_x = -\frac{s'(x)}{s(x)}$$

Pero

$$s(x) = \frac{l_x}{l_0} \Rightarrow s'(x) = \frac{l'_x}{l_0}$$

y, por tanto,

$$\mu_x = -\frac{l'_x}{l_x} = -\frac{d}{dx}Ln(l_x) \tag{2.8}$$

luego conocida la función de supervivencia se puede calcular la fuerza de mortalidad, que resulta ser la variación relativa de aquella. Inversamente, el lector puede comprobar que

$$l_x = l_0.e^{-\int_0^x \mu_s ds} \tag{2.9}$$

Así, por ejemplo, cuando la fuerza de mortalidad es constante (ley exponencial) se comprueba fácilmente a partir de la fórmula anterior que

$$l_x = l_0.e^{-\mu x} \tag{2.10}$$

En el apartado 1.6 hemos discutido la hipótesis de una fuerza de mortalidad constante, llegando a la conclusión de que no resulta realista. Sin embargo, esta hipótesis tiene gran importancia histórica. A continuación comentaremos brevemente algunas de las primeras tablas de mortalidad que aparecieron en la segunda mitad del siglo XVII, así como las funciones de supervivencia y leyes de mortalidad subyacentes.

Es bien conocido que el primer texto impreso sobre Teoría de la Probabilidad fue *De Ratiociniis in Ludo Aede*, escrito por **Christian Huygens** y publicado en 1657. En 1662, tan sólo cinco años después, **John Graunt** publicó sus *Observations upon the Bills of Mortality*, trabajo que ha sido posteriormente reconocido como el precursor de la Estadística Demográfica. En él Graunt incluyó la primera tabla de mortalidad de la historia, relativa a la población de Londres. Reproducimos a continuación su famosa tabla:

x	l_x
0	100
6	64
16	40
26	25
36	16
46	10
56	6
66	3
76	1

Los registros de mortalidad a los que tuvo acceso Graunt indicaban la causa de la muerte y el sexo los difuntos, pero no su edad. Graunt registró la proporción de personas que morían de enfermedades infantiles (las cuales presumiblemente han de ser niños), añadiendo la mitad de las que morían de enfermedades como sarampión o viruela (que afectan igualmente a niños y adultos), y concluyendo que 36 de cada 100 personas morían antes de los 6 años. Esto proporciona la segunda fila de su tabla de mortalidad. La hipótesis de que casi nadie sobrevivía a los 76 años de edad proporciona su última fila.

Graunt no explica cómo obtuvo las filas intermedias. Un gran número de investigadores se han planteado este problema, y algunos llegan a la conclusión de que inventó los números. Otros (véase Hacking (1995)) aventuran la hipótesis de que Graunt llevó a cabo una interpolación entre los 6 y los 76 años siguiendo una ley exponencial: en efecto, tomando $\mu = 0.047$ y redondeando a alguno de los enteros más cercanos, podemos reproducir aproximadamente la tabla de Graunt.

-Si $l_6 = 64$ entonces $l_{16} = 64\ e^{-0.047\ (16-6)} = 40$

-Si $l_{16} = 40$ entonces $l_{26} = 40\ e^{-0.047\ (26-16)} = 25$

-Si $l_{26} = 25$ entonces $l_{36} = 25\ e^{-0.047\ (36-26)} = 15.62$ (redondeamos a 16)

-Si $l_{36} = 16$ entonces $l_{46} = 16\ e^{-0.047\ (46-36)} = 10$

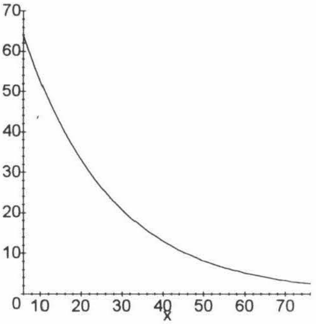
-Si $l_{46} = 10$ entonces $l_{56} = 10\ e^{-0.047\ (56-46)} = 6.25$ (redondeamos a 6)

-Si $l_{56} = 6$ entonces $l_{66} = 6\ e^{-0.047\ (66-56)} = 3.75$ (redondeamos a 3)

-Si $l_{66} = 3$ entonces $l_{76} = 3\ e^{-0.047\ (76-66)} = 1.87$ (redondeamos a 1)

A continuación reproducimos la función de supervivencia de Graunt para las edades comprendidas entre los 6 y 76 años:

$$l_x = 64\ e^{-0.047\ (x-6)}$$



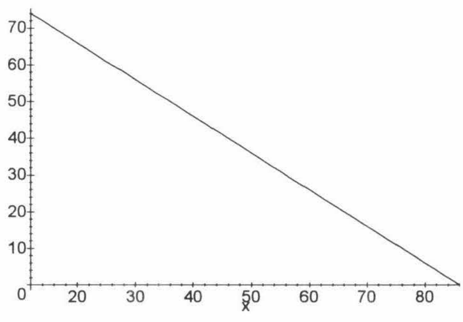
El supuesto de una fuerza de mortalidad constante (a tramos) fue también asumido por otros famosos científicos, como **Jan De Witt** y **Jan Hudde**, quienes construyeron tablas de mortalidad poco después que Graunt. Este supuesto tuvo, pues, gran importancia en los comienzos de la Matemática Actuarial.

En 1693 **Edmond Halley** publicó una tabla de mortalidad construída a partir de registros de nacimientos y muertes de la ciudad alemana de Breslau, en los que se incluía la edad de los difuntos (*An Estimate of the Degrees of the Mortality of Mankind, Drawn from Various Tables of Births and Funerals at the City of Breslau*). Halley no tuvo necesidad de ninguna hipótesis de mortalidad uniforme para obtener su tabla, que por otra parte fue considerada la mejor construída hasta ese momento y se mantuvo en uso durante unos ochenta años.

En 1725 **Abraham De Moivre** ajustó por primera vez una fórmula matemática a una tabla empírica. La tabla fue la de Halley, y la fórmula ajustada fue

$$l_x = 86 - x$$

para las edades x comprendidas entre los 12 y los 86 años.



El lector puede comprobar fácilmente que la función de supervivencia asociada a una fuerza de mortalidad que en el apartado 1.6.2 denominamos de De Moivre,

$$\mu_x = \frac{1}{\omega - x}$$

resulta ser

$$l_x = l_0(1 - \frac{x}{\omega})$$

para x entre 0 y la edad límite ω. Se trata de una recta decreciente que une los puntos $(0, l_0)$ y $(\omega, 0)$. Obviamente la fórmula que realmente ajustó De Moivre a la tabla de Halley es un caso particular de la anterior, con la pequeña modificación de que parte de l_{12} en lugar de l_0.

Al igual que la fuerza de mortalidad de De Moivre, en la práctica la función de supervivencia de De Moivre sólo resulta satisfactoria para pequeños intervalos de tiempo. Posteriormente se han propuesto numerosos modelos alternativos (evidentemente tantos como para la fuerza de mortalidad), de los que únicamente destacaremos la función de supervivencia de **Makeham**, publicada en 1860, con la que a menudo se consiguen buenos ajustes para ciertos tramos de edad.

Recordemos del apartado 1.6.4 que la expresión de la fuerza de mortalidad de Makeham es

$$\mu_x = A + B\ C^x$$

siendo $B > 0, C > 1, A > -B$. La función s(x) resulta ser

$$s(x) = s^x\ g^{C^x - 1}$$

siendo

$$s = e^{-A},\ \ g = e^{-\frac{B}{LnC}}$$

Por tanto, la función de supervivencia tiene por expresión

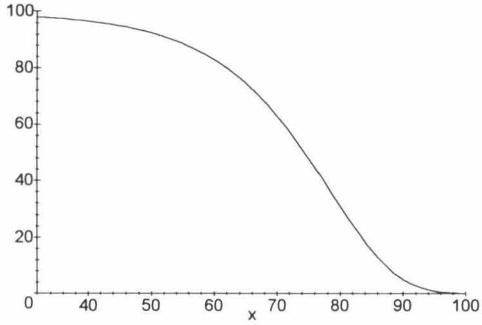
$$l_x = l_0\ s^x\ g^{C^x - 1}$$

o, equivalentemente,

$$l_x = k\ s^x\ g^{C^x}$$

siendo $k = \frac{l_0}{g}$.

Como ya hemos comentado, con esta función de supervivencia se consiguen buenos ajustes para ciertos tramos de edades (que, recordemos, no incluyen las edades infantiles). Así, por ejemplo, en la tabla española PEM70 (Población Española Masculina en el año 1970) se consiguió ajustar una función de Makeham para $x \geq 36$, siendo A=0.0002702165, B=0.000054595, C=1.0996287. A continuación representamos la función de supervivencia resultante:



El lector puede comparar las tres últimas gráficas y llegar a sus propias conclusiones sobre la forma aproximada que debe tener una función de mortalidad realista.

2.3 La Interpretación Determinista

En los apartados anteriores hemos expuesto con detalle la interpretación correcta de los valores que aparecen en una tabla de mortalidad, según la cual, recordemos, q_x resulta ser la probabilidad de que un individuo de edad x muera en el transcurso de un año, y l_x y d_x son, respectivamente, el número medio de individuos vivos a la edad x y el número medio de individuos que fallecen entre las edades x y $x+1$, de un colectivo inicial de l_0 recién nacidos. Se trata, por tanto, de probabilidades y de esperanzas matemáticas asociadas con ciertas variables aleatorias. Por esta razón, a la interpretación anterior (que es la correcta) se le suele denominar **interpretación estocástica** de la tabla de mortalidad.

Existe asimismo una forma alternativa de interpretar una tabla de mortalidad, incorrecta pero más sencilla, según la cual los valores de las l_x coinciden exactamente con el número de individuos del colectivo inicial de l_0 recién nacidos que alcanzan con vida las distintas edades x. Según esta interpretación, d_x sería el número exacto de individuos del colectivo inicial de l_0 recién nacidos que fallecen entre las edades x y $x+1$. Asimismo q_x se interpreta como la proporción de individuos del colectivo inicial de l_0 recién nacidos que habiendo alcanzado con vida la edad x mueren antes de cumplir un año más. Observemos que las probabilidades y esperanzas matemáticas de la interpretación estocástica se han convertido en proporciones y valores exactos. Por esta razón se habla de la **interpretación determinista** (o **clásica**) de la tabla de mortalidad.

Pese a ser incorrecta, la interpretación determinista de la tabla permite obtener las fórmulas de las probabilidades básicas sin necesidad de recurrir a las herramientas del Cálculo de Probabilidades. En efecto, puesto que dichas probabilidades son interpretadas como proporciones relativas a una población cuya evolución a lo largo del tiempo se supone perfectamente conocida, entonces es posible calcularlas como cocientes entre los "casos favorables" y los "casos posibles".

Así, por ejemplo, la proporción de individuos que habiendo alcanzado con vida la edad x continúan vivos a la edad $x+t$, ${}_tp_x$, se puede calcular como el cociente entre el número de "casos favorables" (los individuos vivos a la edad $x+t$, l_{x+t}) y de "casos posibles" (los individuos vivos a la edad x, l_x). Obtenemos, pues,

$$ {}_tp_x = \frac{l_{x+t}}{l_x} $$

Asimismo, la proporción de individuos que habiendo llegado con vida a la edad x fallecen antes de alcanzar la edad $x+t$, ${}_tq_x$, se puede calcular como el cociente entre el número de individuos que fallecen entre las edades x y $x+t$ (que a su vez es igual a ${}_td_x$, que es la diferencia entre los que están vivos a la edad x, l_x, y los que están vivos a la edad $x+t$, l_{x+t}) y el total de individuos que llegan vivos a la edad x, l_x. Obtenemos, pues,

$$ {}_tq_x = 1 - {}_tp_x = \frac{l_x - l_{x+t}}{l_x} $$

Finalmente, la proporción de individuos que habiendo llegado con vida a la edad x fallecen entre las edades $x+s$ y $x+s+t$, ${}_{s/t}q_x$, se calcula como el cociente entre los que mueren entre las edades $x+s$ y $x+s+t$ ($l_{x+s} - l_{x+s+t}$) y el total de individuos vivos a la edad x, l_x. Es decir,

$$ {}_{s/t}q_x = \frac{l_{x+s} - l_{x+s+t}}{l_x} $$

Se obtienen, pues, exactamente las mismas fórmulas para las probabilidades básicas a partir de la interpretación determinista que a partir de la estocástica. Y, por otra parte, es claro que la obtención de dichas fórmulas resulta más sencilla a partir de la interpretación determinista. Sin embargo, a pesar de tales ventajas, la interpretación determinista presenta inconvenientes, que comentamos a continuación.

En efecto, la interpretación determinista sugiere que la columna de las variables l_x es la columna básica de una tabla de mortalidad, a partir de la cual se deducen todas las demás, y que la construcción de dicha tabla debería llevarse a cabo observando la mortalidad de un colectivo real de l_0 recién nacidos desde su nacimiento hasta la muerte del último de ellos. Pero obviamente esta metodología sería imposible de aplicar en la realidad. En efecto, para poder construir una tabla completa serían necesarios más de cien años. Además, dicha tabla no tendría ninguna validez, al haberse construído a partir de datos tan alejados en el tiempo. Finalmente, los resultados dependerían del colectivo elegido, ya que al cambiar de colectivo inicial también cambiaría presumiblemente la mortalidad. Por estas razones se considera incorrecta la interpretación determinista, aunque sin embargo a veces resulte útil recurrir a ella a causa de la sencillez con que se obtienen los resultados.

En el siguiente apartado comentaremos brevemente el proceso que realmente se sigue para construir una tabla de mortalidad.

2.4 Construcción de Tablas de Mortalidad

En la práctica, casi todas las tablas de mortalidad se construyen estimando en primer lugar la columna de los tantos de mortalidad q_x, y calculando posteriormente las demás columnas a partir de aquélla siguiendo los procedimientos que hemos establecido anteriormente. Como es natural, las probabilidades q_x de que los individuos de edad x mueran en el transcurso de un año se estiman a partir de las frecuencias relativas de los mismos sucesos. Es necesario, por tanto, considerar para cada edad x un gran número de individuos de esa edad, y contabilizar cuántos de ellos mueren en el transcurso de un año. Los datos necesarios se pueden obtener del Censo y de las estadísticas de defunciones del Registro Civil.

Las estimaciones de los tantos q_x obtenidas de esta forma no suelen ser muy fiables, ya que los datos de partida pueden estar afectados por diversos tipos de errores. Por ejemplo, los censos o las estadísticas de defunciones pueden estar mal elaborados. También pueden presentar picos o altibajos atribuíbles exclusivamente al azar. Asimismo, tanto los censos como los registros de defunciones suelen sufrir sesgos en la declaración de la edad, ya que a menudo se observa que abundan más las edades terminadas en 0 ó 5 que las restantes edades. Por estas razones se procede, en una segunda fase, a suavizar los datos originales, es decir, a sustituir los valores brutos de los q_x por nuevos valores que presenten un desarrollo más regular, sin altibajos y saltos injustificables.

Existen técnicas estadísticas específicas para suavizar series de datos, siendo quizás la más conocida la técnica de las **medias móviles**. Dicha metodología procede a sustituir cada dato por una media ponderada formada a partir del propio dato y los que le rodean, consiguiendo de esta manera una nueva serie de datos más suave que la original.

Pero los actuarios suelen preferir suavizar los valores brutos mediante el ajuste de alguna función matemática que resulte adecuada para representar la mortalidad. Lo más habitual es recurrir a la expresión de alguna fuerza de mortalidad teórica, y ajustar sus parámetros por el **método de los mínimos cuadrados**, es decir, de forma que se minimice la suma de los cuadrados de las desviaciones entre los valores brutos de los q_x y los correspondientes valores ajustados. Como ya hemos apuntado anteriormente, para edades adultas es habitual considerar la ley de mortalidad de **Makeham** o alguna otra parecida pero con mayor número de parámetros.

En la práctica, los ajustes suelen llevarse a cabo por tramos de edad. Como mínimo es habitual distinguir tres tramos, uno para las edades infantiles, otro para las edades adultas y un tercero para las edades seniles. Precisamente los datos de partida en este último tramo de edades seniles suelen ser los menos fiables.

Finalmente y como ya hemos comentado anteriormente, una vez que se dispone de los valores ajustados de los tantos de mortalidad q_x se procede a calcular los demás elementos de la tabla por procedimientos ya conocidos.

2.5 Cálculo de Probabilidades Básicas para Edades Fraccionarias

Sabemos que una tabla de mortalidad permite calcular las probabilidades básicas de muerte y supervivencia para valores enteros de las variables x y t. Nos planteamos a continuación el problema del cálculo de esas mismas probabilidades para valores no enteros (en la práctica, fraccionarios) de las variables.

Es posible plantear este mismo problema en términos más técnicos: si conocemos las probabilidades asociadas con la variable aletoria discreta K_x, ¿cómo se pueden calcular las probabilidades asociadas con la variable aleatoria continua T_x?.

Si conocemos el tanto instantáneo o fuerza de mortalidad μ_x a partir del cual se han generado los valores de la tabla, el problema no tiene mayor complicación, pues sabemos que la fórmula

$$_tq_x = 1 - e^{-\int_x^{x+t} \mu_s ds}$$

es válida para cualesquiera valores de x y t, tanto si son enteros como si no lo son.

Así, por ejemplo, recordemos que para $x \geq 36$ el tanto instantáneo de mortalidad de la tabla PEM70 era

$$\mu_x = 0.0002702165 + 0.000054595\ 1.0996287^x$$

En consecuencia, para cualesquiera valores de x y t tendremos

$$_tq_x = 1 - e^{-\int_x^{x+t}(0.0002702165+0.000054595\ 1.0996287^s)\ ds}$$

A partir de la expresión anterior podemos calcular los valores que aparecen en la tabla, como por ejemplo

$$q_{60} = 1 - e^{-\int_{60}^{61}(0.0002702165+0.000054595\ 1.0996287^s)ds} = 1.7209\ 10^{-2}$$
$$q_{61} = 1 - e^{-\int_{61}^{62}(0.0002702165+0.000054595\ 1.0996287^s)ds} = 1.8881\ 10^{-2}$$

Pero también podemos calcular valores que no aparecen en la tabla, como por ejemplo la probabilidad de que un individuo de 60 años de edad fallezca antes de seis meses, probabilidad que resulta ser igual a

$$_{0.5}q_{60} = 1 - e^{-\int_{60}^{60.5}(0.0002702165+0.000054595\ 1.0996287^s)ds} = 8.4407\ 10^{-3}$$

En el caso en que desconozcamos la fuerza de mortalidad, deberemos recurrir a técnicas de interpolación para calcular probabilidades correspondientes a edades fraccionarias y que por tanto no figuran directamente en la tabla. Existen en la literatura tres técnicas de interpolación para resolver este problema, que comentaremos a continuación. Las tres técnicas, que se corresponden con tres hipótesis alternativas sobre el comportamiento de la mortalidad en edades fraccionarias, consiguen empalmar de forma continua tanto las probabilidades básicas como la función de supervivencia, pero no así la fuerza de mortalidad, que resulta discontinua en las edades enteras.

Hipótesis I: $_tq_x$ es una función lineal de t entre $t = 1$ y $t = 0$.

Es decir, para calcular $_tq_x$ con $t \in (0,1)$ interpolamos linealmente entre $_1q_x = q_x$ y $_0q_x = 0$, de forma que

$$_tq_x = t\ _1q_x + (1-t)\ _0q_x = t\ q_x \tag{2.11}$$

En el ejemplo anterior, obtendríamos un valor un poco más elevado que el verdadero que ya hemos calculado:

$$_{0.5}q_{60} = 0.5 \times q_{60} = 0.5 \times 0.017209 = 8.6045\ 10^{-3}$$

Como vimos en el apartado 1.6.2, la hipótesis de que $_tq_x$ es una función lineal de t se corresponde con la ley de De Moivre y con la distribución uniforme de las muertes. Por esta razón se la conoce como **Hipótesis de De Moivre**. Este supuesto equivale a asumir una distribución uniforme de las muertes entre un año y el siguiente, supuesto que es falso pero que proporciona buenas aproximaciones a los valores correctos, ya que la mortalidad al final de un año es, para la mayoría de las edades, muy similar a la del principio del mismo año. Así, por ejemplo, el valor de $_{0.5}q_{60}$ que acabamos de obtener bajo la hipótesis de De Moivre (0.0086045) coincide hasta el tercer decimal con el verdadero valor obtenido más arriba (0.0084407).

El lector puede comprobar que se verifica (basta expresar las probabilidades $_tq_x$ y $t\ q_x$ mediante la función de supervivencia)

$$s(x+t) = t\ s(x+1) + (1-t)\ s(x) \tag{2.12}$$

y por tanto, multiplicando por l_0,

$$l_{x+t} = t\ l_{x+1} + (1-t)\ l_x \tag{2.13}$$

luego bajo la hipótesis de De Moivre los valores de l_{x+t} con $t \in (0,1)$ también se interpolan linealmente entre l_{x+1} y l_x. La función de supervivencia l_x resulta ser una curva decreciente, continua y lineal a trozos.

Tal y como hemos indicado, la tabla de mortalidad nos permite obtener la distribución de probabilidad de la variable aleatoria número completo de años de vida hasta la muerte de (x), K_x; analicemos a continuación cómo bajo esta hipótesis es posible obtener la distribución de probabilidad de T_x (vida residual de (x)).

Definamos para ello una nueva variable aleatoria S que representa la fracción de año vivida por (x) en el año de su fallecimiento, se verifica

$$T_x = K_x + S_x$$

Pues bien es posible demostrar (véase ejercicio 2.3) que la hipótesis

$${}_tq_x = t\ q_x$$

es equivalente a que las variables aleatorias K_x y S_x sean independientes y que S_x siga una distribución uniforme y continua en $(0,1)$.

Hipótesis II: La fuerza de mortalidad es constante entre dos años consecutivos.

Como sabemos, en la realidad la fuerza de mortalidad crece (para edades adultas o seniles) o decrece (para las edades infantiles) constantemente con la edad, por lo que la hipótesis que comentamos es falsa. Sin embargo, para la mayoría de las edades esta hipótesis proporciona buenas aproximaciones a la mortalidad real, ya que la fuerza de mortalidad no varía excesivamente entre el principio y el final de un año.

Para cualesquiera x y t, sabemos que se verifica

$${}_tp_x = e^{-\int_x^{x+t}\mu_s ds} = e^{-\int_0^t \mu_{x+s}ds}$$

Pero, según la hipótesis que comentamos, la fuerza de mortalidad es constante entre x y $x+1$, es decir, $\mu_{x+s} = \mu_x$ para $s \in [0,1)$. Por tanto,

$${}_tp_x = e^{-\int_0^t \mu_x ds} = e^{-\mu_x . t} = (e^{-\mu_x})^t$$

$${}_tq_x = 1 - (e^{-\mu_x})^t$$

para $t \in [0,1)$. Si además definimos a μ_x como

$$\mu_x = -Ln(p_x)$$

entonces se tiene

$$p_x = e^{-\mu_x}$$

y conseguimos que las probabilidades se empalmen de forma continua, es decir, que

$$\lim_{t\to 1}(\ {}_tp_x) = p_x$$

En consecuencia, para cada edad entera x y para cada $t \in [0,1]$ se puede calcular ${}_tp_x$, de acuerdo con la hipótesis de fuerza de mortalidad constante, como

$${}_tp_x = (p_x)^t$$

y por tanto

$${}_tq_x = 1 - (p_x)^t$$

Así, por ejemplo, la probabilidad ${}_{0.5}q_{60}$ se calcularía como

$${}_{0.5}q_{60} = 1 - (p_{60})^{\frac{1}{2}} = 1 - (1 - 0.017209)^{\frac{1}{2}} = 8.6418\ 10^{-3}$$

De nuevo observamos una coincidencia de tres cifras decimales entre el valor verdadero y el que hemos obtenido siguiendo la hipótesis de fuerza de mortalidad constante.

Finalmente, bajo esta hipótesis es fácil demostrar que la función $s(x)$ verifica

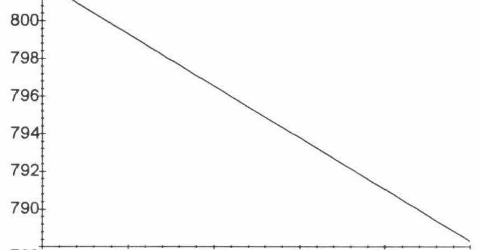
$$s(x+t) = s(x)\ (p_x)^t$$

de donde, multiplicando por l_0, obtenemos que

$$l_{x+t} = l_x\ (p_x)^t$$

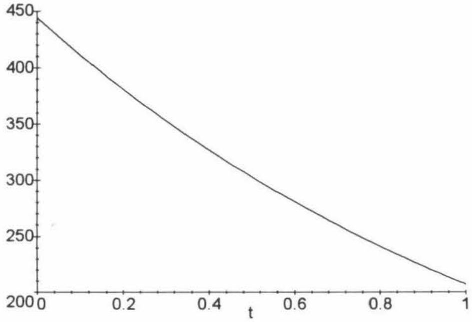
La función de supervivencia resulta ser continua, pero los empalmes entre edades enteras ya no son líneas rectas, como en el caso anterior, sino curvas exponenciales. Sin embargo, en la práctica la curvatura es bastante pequeña. Por ejemplo, la interpolación entre $l_{60} = 802.088$ y $l_{61} = 788.285$ sería, teniendo en cuenta que $p_{60} = 0.98279$,

$$l_{60+t} = l_{60}(p_{60})^t = 802.088\ 0.98279^t$$



Es necesario considerar edades muy avanzadas para poder apreciar algo de curvatura. Así, por ejemplo, la interpolación entre $l_{100} = 444$ y $l_{101} = 206$ sería, teniendo en cuenta que $p_{100} = 0.46612$,

$$l_{100+t} = 444\ 0.46612^t$$



Hipótesis III: Esta hipótesis, conocida como **Hipótesis de Balducci**, también supone una relación lineal, pero no entre ${}_tq_x$ y t como en la hipótesis I, sino entre ${}_{1-t}q_{x+t}$ y t. En consecuencia, se asume que

$${}_{1-t}q_{x+t} = (1-t)\ q_x$$

Es posible demostrar que se verifica

$${}_tq_x = \frac{t.q_x}{1-(1-t)\ q_x}$$

$${}_tp_x = \frac{p_x}{1-(1-t)\ q_x}$$

Así, por ejemplo,

$${}_{0.5}q_{60} = \frac{0.5\ q_{60}}{1-0.5\ q_{60}} = \frac{0.5\ 0.017209}{1-0.5\ 0.017209} = 8.\,6792\ 10^{-3}$$

Obtenemos así un valor aproximado que también coincide en sus tres primeras cifras decimales con el verdadero valor. Debemos observar que la hipótesis de Balducci proporciona valores de ${}_tq_x$ más elevados que los que se obtienen mediante la aproximación lineal de la hipótesis I (e inversamente para los ${}_tp_x$).

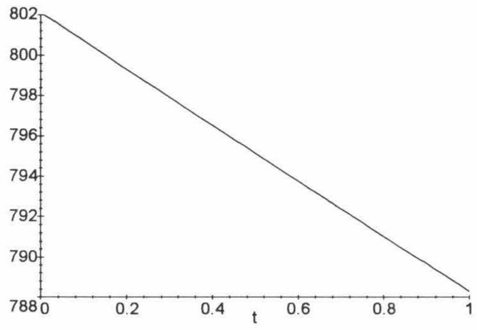
También es posible demostrar que se verifica

$$\frac{1}{s(x+t)} = \frac{t}{s(x+1)} + \frac{1-t}{s(x)}$$

$$\frac{1}{l_{x+t}} = \frac{t}{l_{x+1}} + \frac{1-t}{l_x}$$

En consecuencia, la hipótesis de Balducci puede interpretarse como una interpolación lineal entre los inversos de los valores consecutivos de l_x con x entero. Así, la interpolación entre $l_{60} = 802.088$ y $l_{61} = 788.285$ sería

$$l_{60+t} = \frac{1}{\frac{t}{l_{61}} + \frac{1-t}{l_{60}}} = \frac{1}{\frac{t}{788.285} + \frac{1-t}{802.088}}$$



Se trata, como vemos, de una interpolación no lineal pero con muy poca curvatura.

2.6 Tablas de Seleccionados

Las tablas que hemos estudiado hasta este momento proporcionan las probabilidades de muerte y supervivencia de los individuos en función únicamente de su edad, y se suelen denominar **tablas agregadas**. En ciertos tipos de contratos de seguro, la compañía aseguradora obliga a los posibles clientes a someterse a una revisión médica previa a la firma del contrato, aceptando exclusivamente a aquellos que demuestren tener buena salud. Estos individuos se denominan **seleccionados**, y sus probabilidades de supervivencia son mayores que las de la población en general. De forma que si un individuo de edad x ha resultado seleccionado, su probabilidad de muerte antes de un año (que denominaremos $q_{[x]}$) es menor que el valor q_x de una tabla agregada, ya que en el cálculo de este último valor se toman en cuenta tanto individuos sanos como enfermos.

Las llamadas **tablas de seleccionados** (*Select Tables*) se utilizan para calcular las probabilidades de muerte y supervivencia en este nuevo contexto, y dependen de dos variables, la edad $[x]$ del individuo cuando fue seleccionado y el tiempo t transcurrido desde entonces. Así pues, $q_{[x]+t}$ es la probabilidad de muerte antes de un año de un individuo de edad $[x]+t$ que fue seleccionado a la edad $[x]$. Puesto que el efecto de la selección se difumina con el tiempo, se verifica la siguiente cadena de desigualdades (relativas todas ellas a individuos de la misma edad x):

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \dots$$

Es decir, la probabilidad de muerte antes de un año de un individuo de edad x que acaba de ser seleccionado es estrictamente menor que la de un individuo de la misma edad seleccionado el año anterior, etc. Asimismo es claro que el efecto de la selección suele desaparecer totalmente al cabo de un cierto número de años, r, al que se denomina **periodo de selección**. Por tanto, las desigualdades anteriores acaban convirtiéndose en igualdades:

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} < \dots < q_{[x-r]+r} = q_{[x-r-1]+r+1} = \dots = q_x$$

El último valor q_x debe ser interpretado como la probabilidad de muerte antes de un año de un individuo de edad x que fue seleccionado hace tanto tiempo que se ha anulado el efecto de la selección. Estos valores se muestran en la denominada **tabla definitiva** (*Ultimate Table*), y son bastante parecidos a los que figuran en una tabla agregada relativa a la totalidad de la población.TABLAS DE MORTALIDAD

2.7 Ejercicios

1.- Supongamos que la función de supervivencia de una cierta población es

$$l_x = 20000 - 2x^2$$

para $0 \leq x \leq 100$. **Encuentre los valores de** ${}_{20}d_{40}$, ${}_{20}q_{40}$, ${}_{20/10}q_{40}$.

Solución:

$${}_{20}d_{40} = l_{40} - l_{60} = 20000 - 2\ 40^2 - (20000 - 2\ 60^2) = 4000.$$

$${}_{20}q_{40} = \frac{l_{40} - l_{60}}{l_{40}} = \frac{20000 - 2\ 40^2 - (20000 - 2\ 60^2)}{20000 - 2\ 40^2} = \frac{5}{21} = 0.2381$$

$${}_{20/10}q_{40} = \frac{l_{60} - l_{70}}{l_{40}} = \frac{20000 - 2\ 60^2 - (20000 - 2\ 70^2)}{20000 - 2\ 40^2} = \frac{13}{84} = 0.15476$$

2.- En una cierta tabla de mortalidad se tiene que $q_{65} = 0.022$. Calcúlese ${}_{\frac{1}{3}}q_{65}$ según las hipótesis de De Moivre, de la fuerza de mortalidad constante y de Balducci.

Solución:

Según la hipótesis de De Moivre, se tiene

$${}_{\frac{1}{3}}q_{65} = (1/3)\ q_{65} = \frac{0.022}{3} = 7.3333\ 10^{-3}$$

Según la hipótesis de la fuerza de mortalidad constante, se tiene

$${}_{\frac{1}{3}}q_{65} = 1 - (p_{65})^{0.3} = 1 - (1 - q_{65})^{0.3} = 1 - (1 - 0.022)^{0.3} = 6.6515\ 10^{-3}$$

Según la hipótesis de Balducci, se tiene

$${}_{\frac{1}{3}}q_{65} = \frac{(1/3)\ q_{65}}{1 - ((2/3)\ q_{65})} = \frac{(1/3)\ 0.022}{1 - ((2/3)\ 0.022)} = 7.4425\ 10^{-3}$$

3.- Consideremos la descomposición

$$T_x = K_x + S_x$$

donde la variable aleatoria continua S_x representa la fracción de año vivida por (x) en el año de su fallecimiento, tomando por tanto valores en el intervalo (0,1).

Pruebe que la hipótesis de De Moivre

$${}_tq_x = t\ q_x$$

es equivalente a que las variables aleatorias K_x y S_x sean independientes entre sí y que S_x sea independiente de x y se distribuya uniformemente en $(0, 1)$.

Solución:

Supongamos cierta la hipótesis ${}_tq_x = t.q_x$, y sea $0 < s < 1$; entonces

$$P((K_x = k) \cap (S_x \leq s)) = P(k < T_x \leq k + s) = {}_{k/s}q_x = {}_kp_x\ {}_sq_{x+k} =$$

$$= {}_kp_x\ s\ q_{x+k} = {}_{k/}q_x\ s = P(K_x = k)\ s$$

Puesto que para cualesquiera sucesos A y B se tiene que

$$P(A \cap B) = P(A)\ P(B/A)$$

entonces deducimos que

$$P((S_x \leq s)/(K_x = k)) = s$$

de donde deducimos que S_x es independiente de x (por lo que la llamaremos S) y de K_x, y que además está uniformemente distribuída (ya que si S sigue una distribución uniforme, entonces precisamente se tiene que $P(S \leq s) = s$).

Supongamos ahora que K_x y S sean independientes y S uniforme en $(0,1)$; entonces

$$_{k/s}q_x = P(k < T_x \leq k+s) = P(K_x = k, S \leq s) = P(K_x = k)\ P(S \leq s) = \ _{k/}q_x\ s$$

y tomando $k = 0$ obtenemos

$$_sq_x = s\ q_x$$

4.- Pruebe que bajo la hipótesis I se verifican las igualdades

$$\overset{o}{e}_x = e_x + \frac{1}{2}$$

y

$$Var(T_x) = Var(K_x) + \frac{1}{12}$$

Solución:
Basta tener en cuenta el resultado del ejercicio anterior

$$T_x = K_x + S$$

por lo que

$$E(T_x) = E(K_x) + E(S)$$

y

$$Var(T_x) = Var(K_x) + Var(S)$$

y recordar que si S se distribuye según una uniforme en el intervalo (0,1), entonces

$$E(S) = \frac{1}{2} \quad y \quad Var(S) = \frac{1}{12}$$

5. Demuestre la equivalencia de las tres formulaciones alternativas de la hipótesis de Balducci:

(a)

$$_{1-t}q_{x+t} = (1-t)\ q_x$$

(b)

$$\frac{1}{s(x+t)} = \frac{t}{s(x+1)} + \frac{1-t}{s(x)}$$

(c)

$$_tp_x = \frac{p_x}{1-(1-t)q_x}$$

Solución:

(a)$\Longrightarrow$(b)

En efecto, la expresión

$$_{1-t}q_{x+t} = (1-t)\ q_x$$

equivale a

$$\frac{s(x+t)-s(x+1)}{s(x+t)} = (1-t)\frac{s(x)-s(x+1)}{s(x)}$$

es decir,

$$1-\frac{s(x+1)}{s(x+t)} = 1-\frac{s(x+1)}{s(x)} - t\left(\frac{s(x)-s(x+1)}{s(x)}\right)$$

$$-\frac{s(x+1)}{s(x+t)} = -(1-t)\frac{s(x+1)}{s(x)} - t = -(1-t)\frac{s(x+1)}{s(x)} - t\frac{s(x+1)}{s(x+1)}$$

Multiplicando por -1 y dividiendo por $s(x+1)$, obtenemos la expresión buscada:

$$\frac{1}{s(x+t)} = \frac{t}{s(x+1)} + \frac{1-t}{s(x)}$$

(b)$\Longrightarrow$(c)

Multiplicando la expresión anterior por $s(x)$, obtenemos

$$\frac{s(x)}{s(x+t)} = t\frac{s(x)}{s(x+1)} + (1-t)\frac{s(x)}{s(x)}$$

es decir,

$$\frac{1}{_tp_x} = t\frac{1}{p_x} + (1-t) = t\frac{1}{p_x} + (1-t)\frac{p_x}{p_x} = \frac{t+(1-t)p_x}{p_x}$$

Invirtiendo obtenemos

$$_tp_x = \frac{p_x}{t+(1-t)p_x} = \frac{p_x}{t+(1-t)(1-q_x)} = \frac{p_x}{1-(1-t)q_x}$$

que es la expresión buscada.

(c)$\Longrightarrow$(a)

Sabemos que

$$p_x = {}_t p_x \; {}_{1-t}p_{x+t} = {}_t p_x \left(1 - {}_{1-t}\, q_{x+t}\right)$$

Sustituyendo ${}_t p_x$ por su equivalente obtenemos

$$p_x = \frac{p_x}{1-(1-t)q_x}\left(1 - {}_{1-t}q_{x+t}\right)$$

es decir,

$$1 - {}_{1-t}q_{x+t} = 1-(1-t)q_x$$

y por tanto obtenemos la expresión buscada:

$${}_{1-t}q_{x+t} = (1-t)\; q_x$$