Datos Curriculares:

Doctor en Estudios de Población (con especialización en mortalidad, temas de salud, mortalidad por causas y ajuste matemático de los fenómenos demográficos) por parte de El Colegio de México promoción 2008-2012; Maestro en Demografía (con especialización en mortalidad y mortalidad por causas) por parte de El Colegio de México; Licenciado en Actuaría, por parte del Instituto Tecnológico Autónomo de México.

Fue investigador en el Centro de Información para Decisiones en Salud Pública (CENIDSP) en Instituto Nacional de Salud Pública (INSP), México de septiembre 2010 a octubre 2011, y co-coordinador del seminario de tesis: Población y Salud I, II, III y IV de la maestría en población y desarrollo en La Facultad Latinoamericana de Ciencias Sociales (FLACSO-México) de Mayo 2011 a Julio de 2012.

Entre sus publicaciones recientes se encuentran: el libro "Análisis de la mortalidad por causas en México: Ganancias en esperanza de vida 2000 y 2005 y proyecciones al 2015" publicado en la Editorial Académica Española en 2011. Autor de los artículos "Diabetes en México y Colombia: Análisis de la tendencia de años de vida perdidos, 1998-2007" en la Revista de Salud Pública de la Universidad Nacional de Colombia en 2011; "Diferencias socioeconómicas en los años de vida perdidos por cáncer de mama y cérvico-uterino en Colombia, 1997 y 2007" en la Revista en Gerencia y Políticas en Salud de la Universidad Javeriana en Colombia en 2010; y "Cambios en los años de vida perdidos por cáncer de mama y cérvico–uterino en México según grado de marginación estatal, 1997 y 2007" en la Revista de Salud Pública de Chile en 2010.

Actualmente es profesor investigador dentro del programa de Maestría en Población y Desarrollo de la Facultad Latinoamericana de Ciencias Sociales (FLACSO cede México) ubicada en: Carretera al Ajusco 377, col. Héroes de Padierna, Del. Tlalpan, México D.F. C.P.14200. Tel. (52-55) 30 000 200. Correo electrónico: claudio.davila@flacso.edu.mx

Ajuste matemático de la mortalidad general en México, 2000, 2005 y 2010

Claudio Alberto Dávila Cervantes[1]

Resumen

Se presenta una propuesta del método de splines cúbicos para ajustar la curva de mortalidad general en México por sexo en todo el rango de edad para el 2000, 2005 y 2010. El método consiste en ajustar un polinomio de tercer grado en las primeras 4 edades estimándolas de manera perfecta; entre los 4 y 80 años se ajusta un spline cúbico empleando 5 nodos; se ajusta otro polinomio de tercer grado a partir de 80 años. Se ocupan 24 parámetros para estimar la curva de mortalidad en todo el rango de edad. Presenta una excelente bondad de ajuste para hombres, que se incrementa con el paso de los años; para mujeres, la bondad de ajuste es, en general, superior a hombres. Estimación relativamente parsimoniosa de la mortalidad; respeta los cambios de concavidad de la curva de mortalidad; posibilita realizar proyecciones de la mortalidad en diferentes contextos y causas de muerte.

Palabras clave: Mortalidad, Splines cúbicos, Ajuste matemático, México, Rango de edad

Abstract

I present a methodology based on Cubic Splines to fit the general mortality curve in Mexico by sex at all ages for the years of 2000, 2005 and 2010. This method consists in fitting a polynomial of degree three in the first four ages (fitting them perfectly); between ages 4 and 80 we fit a cubic spline using five nodes; we fit another third-degree polynomial beginning at age 80. We employ 24 parameters to estimate the mortality curve in the full age range. This model presents an excellent goodness of fit in men, which increases over the years; for women, the goodness of fit is, in general, better than men's. It's a relatively parsimonious estimation of mortality; it maintains the mortality curve's changes in concavity; allows the realization of mortality projections in different contexts and causes of death.

Key words: Mortality, Cubic Splines, Mathematical modeling, Mexico, Full age range

[1] Profesor investigador de la Flacso México – claudio.davila@flacso.edu.mx

Introducción

Durante el siglo XX, se presentaron incrementos sustanciales en la esperanza de vida de la mayoría de los países del mundo. Éstos fueron provocados principalmente por avances científicos, médicos y de salud pública y, en general, por una mejora en los estándares de vida. A lo largo de ese periodo, la esperanza de vida se incrementó a más de 75 años en los países de altos ingresos y a más de 60 en muchos países de ingresos medios (Smith et al., 2001: 49). Estos cambios han tenido impactos dramáticos en el tamaño y composición de las poblaciones.

En México, durante ese mismo periodo, se dio también una profunda transformación de la mortalidad. Después de la Revolución Mexicana, se presentó una paulatina disminución de la mortalidad, gracias a muchos factores, entre los que destacan: modificaciones en el nivel de vida de la población y los primeros programas de salud pública (Camposortega, 1992; 1997: 11). Esta tendencia continuó durante todo el siglo y se vio reflejada en el aumento de esperanza de vida durante el periodo mencionado, en el cual este indicador alcanzó un nivel alrededor de 74 años, a final del siglo, cuando su nivel inicial en 1920 era de 30 años aproximadamente. Este aumento en la esperanza de vida se dio de forma acelerada entre los años de 1940 a 1960, debido a que las ganancias en esperanza de vida en esos 20 años, fueron de casi 19 años. Sin embargo, en los siguientes 50 años, si bien se siguieron dando ganancias, el ritmo ha sido considerablemente menor, ya que se ha incrementado 18 años aproximadamente.

Esta situación se dio principalmente porque en los años cuarenta y cincuenta se controlaron con éxito las causas de muerte más susceptibles de controlar: las enfermedades infecciosas y parasitarias. Posteriormente, las enfermedades crónico-degenerativas (como enfermedades cardiovasculares, diabetes, neoplasmas, etc.), así como las muertes por causas externas (accidentes y muertes violentas), tuvieron una cada vez mayor presencia relativa dentro del perfil epidemiológico del país (Camposortega, 1997: 11). Estas causas de muerte son más difíciles de controlar, debido a que requieren de una mayor infraestructura médica y de salud pública, mayores avances técnicos, un mejor nivel de vida de la población y una cultura de prevención; esto se cree que es parte de la explicación del porque el ritmo de aumento en la esperanza de vida se ha vuelto paulatinamente más lento. Cabe mencionar que la relación entre la estructura de causas de muerte de una población, su patrón por edad de la mortalidad y el efecto que se tiene sobre la esperanza de vida de la población, se formalizó en la Teoría de la Transición Epidemiológica (Omran, 1971: 509).

Entonces, la cuestión frente al menor crecimiento de la esperanza de vida es analizar cómo ésta se puede continuar incrementando; parte de la respuesta está en entender el comportamiento del patrón de la mortalidad en una población y ajustarlo de tal forma que se pueda realizar un manejo matemático del fenómeno, lo que puede posibilitar realizar proyecciones (Halli y Rao, 1992) y desarrollar investigaciones sobre efectos, modificaciones y cambios posibles en edades específicas. Esto provoca un creciente interés en la descripción y ajuste matemático de los patrones de mortalidad. Esto es, a través de la estimación de patrones modelo parametrizados se describe matemáticamente la estructura por edad y sexo de la mortalidad y la evolución de ésta a través del tiempo (McNown y Rogers, 1989: 645-646).

En cuanto a la distribución por edad y sexo del total de defunciones en el país, en general, existe una sobremortalidad masculina, particularmente en las edades comprendidas entre los 15 y los 60 años de edad. Asimismo, la mortalidad del país sigue el patrón regular en forma de *j* observado en muchas poblaciones. Este patrón está provocado por una mayor mortalidad en el primer año de vida de las personas, que disminuye hasta alcanzar su mínimo entre los 5 y 10 años de edad; posteriormente se presenta un constante aumento de las probabilidades de muerte a partir de los 10 años de vida de las personas, con un incremento considerable para los hombres en la adolescencia y edades adultas jóvenes, debido principalmente a la mayor mortalidad por accidentes y muertes violentas que se presenta en esas edades; y finalmente un crecimiento exponencial reflejado en un crecimiento casi geométrico de las tasas de mortalidad de en las edades adultas mayores (Heligman y Pollard, 1980: 73).

Existen varios intentos previos para modelar la curva de mortalidad en todo el rango de edad. Entre ellos se destacan el modelo propuesto por Gompertz en 1825 y sus variantes Gompertz-Makeham y Lazarus. Sin embargo, hoy en día es conocido que la ley de Gompertz no ajusta de manera correcta las tasas de mortalidad de ciertas edades, en particular en las primeras edades de la tabla de vida y las edades muy avanzadas; ya que las tasas observadas suelen ser menores que aquellas predichas por el modelo y, con ello, el número de supervivientes en edades extremas es mayor a aquel que la ley de Gompertz ajusta (Gavrilov y Gavrilova, 2002:10).

Asimismo, se han propuesto otras tres de esas leyes de mortalidad que ajustan todo el rango de edad de la mortalidad, las cuales son la de Thiele en 1872, Wittstein en 1883 y la propuesta por Heligman y Pollard en 1980 (Hartmann, 1987: 19). La valía de estas leyes es que intentan modelar todo el rango a diferencia de otras leyes de mortalidad que son parciales puesto que ajustan bien secciones de las curvas de mortalidad, pero tienen problemas representando ciertas características de éstas.

Bajo este marco, el objetivo de este trabajo es presentar una propuesta del método de splines cúbicos como alternativa para ajustar la curva de mortalidad general por sexo en todo el rango de edad para los años 2000, 2005 y 2010. Para ello se calcula la bondad de ajuste entre la serie de supervivientes a edad x (l_x) observada a partir de datos de mortalidad de las estadísticas vitales y aquella estimada a partir de los splines cúbicos propuestos. Se realiza también una comparación entre la esperanza de vida observada y la estimada por medio de los splines cúbicos, con la finalidad de analizar la calidad del ajuste propuesto.

Material y Métodos

Se propone una metodología en la cual se ajusta un polinomio de tercer grado en las primeras 4 edades de la serie de supervivientes ajustándolas de manera perfecta, debido al cambio de concavidad en la curva provocado por la relativamente alta mortalidad infantil (en el primer año de vida) todavía existente en el país, el cual no es bien ajustado por las funciones de supervivencia como la función Gompertz o Gompertz-Makeham. Este primer polinomio presenta una unión suave con el primer polinomio del spline. Entre las edades 4 y 80 se ajusta un spline cúbico empleando 5 nodos, los cuales varían en su ubicación dependiendo el año y el sexo. Se ajusta un polinomio de tercer grado en las últimas edades de la curva de mortalidad, a partir de los 80 años, que ajusta el cambio de concavidad que también existe en esas edades mayores relacionado con la desaceleración y el nivelamiento de la mortalidad en las edades mayores de la tabla de vida (Gavrilov y Gavrilova, 2006:23).

Fuentes de información

La información de este estudio se obtuvo de las Estadísticas Vitales de Mortalidad del Instituto Nacional de Estadística y Geografía (INEGI) y de las proyecciones de población realizadas por el Consejo Nacional de Población de México (CONAPO). Los datos de defunciones se obtuvieron para los años 2000, 2005 y 2010, a nivel nacional, por sexo y por edad desplegada. La población a mitad de año, para calcular las tablas de mortalidad de las cuales se deriva la serie de supervivientes a edad exacta x, se obtuvo de las proyecciones 1990 a 2050 del CONAPO.

Para el análisis, se tuvieron en cuenta las defunciones ocurridas al interior del país, quedando excluidos los siguientes dos casos: registros no especificados por edad y sexo; y muertes ocurridas en el exterior. A partir de esto, se descartaron del total de registros, 2 946 en 2000, 2 402 en 2005 y 3 006 en 2010, que corresponden a un 1.21%, 0.88% y 0.91% del total de las defunciones para hombres; en cambio para mujeres no se tomaron en cuenta 1 417 registros en 2000, 950 en 2005, y 780 en 2010 que corresponden a un 0.73%, 0.43% y 0.30% del total. Cabe resaltar que este estudio se realiza a nivel nacional.

Desarrollo de los Splines

Los splines son funciones definidas por una familia de polinomios *sociables*, es decir, los polinomios que constituyen una función spline están estrechamente vinculados (Barrera et al., 1996). El concepto de los splines se debe a un instrumento que utilizaban los ingenieros navales para dibujar curvas suaves, forzadas a pasar por un conjunto de puntos prefijados. La mayor dificultad radica en construir un interpolante que preserve el comportamiento de los datos, es decir, que mantenga las propiedades de monotonía y/o convexidad de los datos (Barrera et al., 1996).

Un spline es entonces una función ***f*** formada por secciones de polinomios cúbicos, por lo que tiene la forma $P(x) = ax^3 + bx^2 + cx + d$, los cuales se unen con la mayor suavidad posible (sin que ***f*** sea necesariamente un polinomio único) (Barrera et al., 1996). En este caso se debe exigir que cada polinomio $P_i(x)$ del spline cúbico satisfaga:

$$P_i(x_i) = g_i \qquad P_i(x_{i+1}) = g_{i+1}$$
$$P_i'(x_i) = s_i \qquad P_i'(x_{i+1}) = s_{i+1}$$

Donde
g_i es el valor de la función ***g*** a ajustar
s_i es el valor de la primera derivada de ***g*** en un conjunto de puntos { x_i }, con $i = 1,...,n$.

Si se cumple esto, no importa el valor de s_i, la función ***f*** definida por $P_i(x)$ en cada intervalo [x_i , x_{i+1}] es continuamente diferenciable. Por tanto, para que ***f*** sea un spline basta tomar los s_i de modo que la segunda derivada ***f''*** sea continua. Como ***f*** está formada por secciones de polinomios cúbicos, los únicos puntos posibles de discontinuidad de ***f''*** son los puntos x_i donde se empatan dos polinomios cúbicos. La continuidad de ***f''*** es equivalente a exigir que:

$$f''(x_i^-) = f''(x_i^+) \text{ con } i = 2,...,n-1$$

Entonces, después del desarrollo que se presenta en Barrera et al. (2006), las s_i deben de satisfacer las siguientes ecuaciones:

$$a_i s_{i-1} + d_i s_i + (1-a_i)s_{i+1} = b_i \; i = 2,...,n-1$$

Donde

$$d_i = 2$$

$$a_i = \frac{\Delta x_i}{\Delta x_{i-1} + \Delta x_i} \text{ con } \Delta x_i = x_{i+1} - x_i$$

$$b_i = 3\left[\frac{\Delta x_i}{\Delta x_{i-1}} \frac{g_i - g_{i-1}}{\Delta x_{i-1} + \Delta x_i} + \frac{\Delta x_{i-1}}{\Delta x_i} \frac{g_{i+1} - g_i}{\Delta x_{i-1} + \Delta x_i} \right]$$

Suponiendo que s_1 y s_n se eligen de alguna manera, tenemos un sistema de *n-2* ecuaciones lineales para calcular las *n-2* incógnitas $s_2, ..., s_{n-1}$. Para obtener el valor de las incógnitas es necesario resolver el sistema de ecuaciones, el cual tiene una sola solución y puede resolverse por el método de eliminación de Gauss sin pivoteo[2]. Los parámetros s_1 y s_n dependen de las condiciones que se impongan en la frontera. En general existen tres tipos de condiciones.

La primera condición, da pie al *spline cúbico completo*, se deriva de conocer de antemano el valor de ***g'*** en x_1 y x_n, por lo se toma $s_1 = g'(x_1)$ y $s_n = g'(x_n)$. En ese caso el spline no solo interpola a ***g*** en los puntos $x_1, ..., x_n$, sino que también interpola a ***g'*** en x_1 y x_n.

La segunda condición consiste en elegir que la segunda derivada del spline en los extremos sea cero. Esto es, que $f''(x_1) = f''(x_n) = 0$. Las ecuaciones a las que da lugar esta condición se pueden ver en Barrera et al., (1996). Este tipo de spline se conoce como *spline cúbico natural*. Barrera et al., (1996) mencionan que desde cierto punto de vista, éste tipo no es muy recomendable ya que la imposición arbitraria de las segundas derivadas puede provocar que cerca de los extremos x_1 y x_n el error aumenta.

La tercera condición se da cuando no se tiene conocimiento sobre las derivadas de ***g*** en los extremos, entonces se eligen s_1 y s_n de tal manera que P_1 coincida idénticamente con P_2 y que P_{n-1} coincida con P_{n-2}. En otras palabras se trata de escoger s_1 y s_n de modo que los

[2] La explicación y solución de este tipo de sistemas se puede ver en Barrera et al., (1996)

puntos x_2 y x_{n-1} no sean puntos de ruptura de la función ***f***. Como ***f*** es un spline cúbico se sabe que:

$$P_1^{(j)}(x_2) = P_2^{(j)}(x_2);\ P_{n-1}^{(j)}(x_{n-1}) = P_n^{(j)}(x_{n-1}) \text{ para } j = 0,1,2$$

Además, como P_1 y P_2 son polinomios cúbicos se pueden escribir como (Barrera et al., 2006):

$$P_1(x) = P_1(x_2) + P_1'(x_2)(x - x_2) + P_1''(x_2)(x - x_2)^2 / 2 + P_1'''(x_2)(x - x_2)^3 / 6$$
$$P_2(x) = P_2(x_2) + P_2'(x_2)(x - x_2) + P_2''(x_2)(x - x_2)^2 / 2 + P_2'''(x_2)(x - x_2)^3 / 6$$

Por tanto la condición en la frontera es equivalente a exigir que $P_1'''(x_2) = P_2'''(x_2)$, es decir, que ***f’’’*** sea continua en x_2 y de manera análoga $P_{n-2}'''(x_{n-1}) = P_{n-1}'''(x_{n-1})$ (Barrera et al., 2006). En este estudio se van a elegir la tercera condición para la elección de los parámetros s_1 y s_n**,** dado que no se tiene conocimiento sobre la función de supervivientes debido a que los datos se han obtenido a partir de los registros administrativos, y por tanto no se dispone de información sobre la primera derivada de la función.

En el caso particular de este estudio se va a ajustar un spline cúbico con 5 nodos, que es entonces una función ***f*** formada por 4 secciones de polinomios cúbicos, de la forma

$P(x) = ax^3 + bx^2 + cx + d$.

El spline se calcula a partir de la edad 4 (primer nodo) hasta la edad 80 (quinto nodo) y con tres edades intermedias, las cuales indicarían el resto de los nodos (gráfica 1). Esto implica que se van a ajustar 4 polinomios cúbicos (números 1 a 4 en la gráfica 1). Como se puede observar en el punto $[x_2, y_2]$ se intersectan el primer y segundo polinomio, en el punto $[x_3, y_3]$ se cruzan el segundo y el tercero y en $[x_4, y_4]$ se unen el tercero y cuarto polinomios. En cambio en los puntos $[x_1, y_1]$ y $[x_5, y_5]$ se tiene el inicio del primer y cuarto polinomios respectivamente. Esto en términos de ecuaciones se traduce en el siguiente sistema de ocho ecuaciones:

$$y_1 = a_1x_1^3 + b_1x_1^2 + c_1x_1 + d_1 \text{ ... } (1)$$
$$y_2 = a_1x_2^3 + b_1x_2^2 + c_1x_2 + d_1 \text{ ... } (2)$$
$$y_2 = a_2x_2^3 + b_2x_2^2 + c_2x_2 + d_2 \text{ ... } (3)$$
$$y_3 = a_2x_3^3 + b_2x_3^2 + c_2x_3 + d_2 \text{ ... } (4)$$
$$y_3 = a_3x_3^3 + b_3x_3^2 + c_3x_3 + d_3 \text{ ... } (5)$$
$$y_4 = a_3x_4^3 + b_3x_4^2 + c_3x_4 + d_3 \text{ ... } (6)$$
$$y_4 = a_4x_4^3 + b_4x_4^2 + c_4x_4 + d_4 \text{ ... } (7)$$
$$y_5 = a_4x_5^3 + b_4x_5^2 + c_4x_5 + d_4 \text{ ... } (8)$$

Donde los parámetros a_i, b_i, c_i y d_i, con $i = 1,2,3,4$ son los correspondientes a cada uno de los 4 polinomios.

Aquí va la Gráfica 1

Como las discontinuidades se presentan en los puntos $[x_2, y_2]$, $[x_3, y_3]$ y $[x_4, y_4]$, se tiene que pedir la condición de que las primeras derivadas deben coincidir en estos puntos. Para tal caso se obtienen las derivadas con respecto a *x* de las ecuaciones (2), (3), (4), (5), (6) y (7) y se igualan las correspondientes por pares ((2) y (3), (4) y (5), (6) y (7)). Entonces el resultado proporciona un sistema de 3 ecuaciones más, de la forma:

$$3a_1x_2^2 + 2b_1x_2 + c_1 - 3a_2x_2^2 - 2b_2x_2 - c_2 = 0 \text{... } (9)$$
$$3a_2x_3^2 + 2b_2x_3 + c_2 - 3a_3x_3^2 - 2b_3x_3 - c_3 = 0 \text{... } (10)$$
$$3a_3x_4^2 + 2b_3x_4 + c_3 - 3a_4x_4^2 - 2b_4x_4 - c_4 = 0 \text{... } (11)$$

Estas ecuaciones resultan de igualar las derivadas y despejar los términos del lado derecho de las igualdades y por ello quedan igualadas a 0. Para la siguiente condición se recuerda que, como el spline está formado por secciones de polinomios cúbicos, los únicos puntos posibles de discontinuidad de la segunda derivada son los puntos x_i donde se empatan dos polinomios cúbicos y por tanto se pide también que las segundas derivadas sean iguales: $f''(x_i^-) = f''(x_i^+)$ con $i = 2,...,n-1$. Esto se calcula a partir de las ecuaciones (9), (10) y (11), derivándolas una vez más, y de esa forma se obtiene un sistema de 3 ecuaciones de la forma

$6a_1x_2 + 2b_1 - 6a_2x_2 - 2b_2 = 0...$ (12) que se rescribe como $3a_1x_2 + b_1 - 3a_2x_2 - b_2 = 0...$ (12)

$6a_2x_3 + 2b_2 - 6a_3x_3 - 2b_3 = 0...$ (13) que se rescribe como $3a_2x_3 + b_2 - 3a_3x_3 - b_3 = 0...$ (13)

$6a_3x_4 + 2b_3 - 6a_4x_4 - 2b_4 = 0...$ (14) que se rescribe como $3a_3x_4 + b_3 - 3a_4x_4 - b_4 = 0...$ (14)

Hasta el momento se tiene un sistema de 14 ecuaciones para las 16 incógnitas iniciales. Debido a que no se conoce acerca de las derivadas de la función en los puntos extremos, se va utilizar la tercera condición y por tanto se tienen que igualar las terceras derivadas en los puntos $[x_2, y_2]$ y $[x_4, y_4]$ por tanto se tienen que calcular las terceras derivadas de (12) y (14). Entonces el cálculo de estas terceras derivadas es igual a:

$3a_1 - 3a_2 = 0...$ (15)

$3a_3 - 3a_4 = 0...$ (16)

Que se rescriben como:

$$a_1 - a_2 = 0$$

$$a_3 - a_4 = 0$$

Debido a esto, se cuenta ya con un sistema de 16 ecuaciones y 16 incógnitas. En términos matriciales se tiene:

$$AB = C$$

Donde A es una matriz de 16x16, B es una matriz de 16x1 que contiene los parámetros y C también es una matriz de 16x1 (ver Anexo 1). La forma de las matrices se encuentra en el anexo 1. Entonces este problema se resuelve calculando la inversa de la matriz A[3] y por tanto

$$B = A^{-1}C$$

Una vez que se tienen los valores de los parámetros, se procede a estimar los supervivientes a edad exacta x correspondientes a cada uno de los segmentos del spline, por medio de las ecuaciones $P_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i$ donde $i=1,2,3$ *y* 4. En estas ecuaciones cada $P_i(x)$ equivale a la l_x y las x son las edades entre los nodos correspondientes al segmento i del spline. Este resultado se compara con la l_x calculada a partir de los datos del CONAPO y se calcula la bondad de ajuste entre ambas series mediante el estadístico R^2. Cabe recordar que entre más cercano a 1 esté el valor, el ajuste se considera mejor. Finalmente, se calculan ambas esperanzas de vida para comparar el alcance y la calidad del método propuesto.

Resultados

Se ajustaron las series de supervivientes a edad x (l_x) para el total de defunciones, tanto para hombres y mujeres en 2000, 2005 y 2010 utilizando el método propuesto. El ajuste por medio de esta metodología es bueno para todos los años y en ambos sexos, ya que existe una alta similitud entre ambas curvas (gráficas 2 y 3). Esto se corrobora numéricamente al analizar los valores de R^2 calculados entre la serie observada y la estimada (tabla 1).Los resultados que se obtienen a partir de este método de estimación, muestran una bondad de ajuste cercana a un valor máximo posible de uno en todos los casos ajustados, lo que

---

[3] La Matriz A es no singular, excepto en el caso en el cual alguno de los x_1, x_2, x_3, x_4 y x_5 son iguales, por lo tanto los sistemas de ecuaciones tienen solución única.

implica un ajuste satisfactorio de los datos de mortalidad en ambos sexos y para todos los años seleccionados (véase anexo 2).

Aquí va la Gráfica 2 **(**Comparación de la serie de supervivientes a edad exacta *x*, observada y ajustada hombres. México 2000, 2005 y 2010)

Aquí va la Gráfica 3 (Comparación de la serie de supervivientes a edad exacta *x*, observada y ajustada mujeres. México 2000, 2005 y 2010)

Si estos resultados se analizan con más detalle, se tiene que el valor de R^2 para hombres en todos los años se encuentra por encima de 0.9999, y se incrementa ligeramente de un año a otro, proporcionando un mejor ajuste en 2005; y uno todavía mejor en 2010. En el caso de mujeres, la bondad de ajuste es superior a la de hombres en 2000 y 2005, no así en 2010. Otra característica del ajuste en mujeres es que el mejor ajuste se presenta en el 2000, y su valor tiende a descender con el paso de los años (aunque el ajuste continúa siendo realmente bueno).

Aquí va la Tabla 1

Del mismo modo, se analizó la bondad de ajuste de la estimación propuesta utilizando la serie de probabilidades de defunción (q_x) con la finalidad de corroborar si el ajuste mantiene la concavidad de los datos observados a lo largo de todo el rango de edad y si ajusta de manera adecuada otra serie de la tabla de vida. Las gráficas del ajuste de la serie de probabilidades de defunción, tanto para hombres como para mujeres en todos los años que abarca este estudio, se pueden observar en el Anexo 3[4]. La estimación de las probabilidades resulta muy buena especialmente a partir de la edad del segundo nodo, para todos los casos. En las primeras edades, el cambio de concavidad que se presenta en la niñez y la adolescencia es ajustado de manera adecuada (sin dejar de ser un buen ajuste en términos generales). Destaca que se tiene una menor bondad de ajuste que con la serie l_x, pero la estimación sigue siendo buena (tabla 1). En el caso de hombres, se tienen valores de

[4] Las probabilidades de defunción se presentan en escala logarítmica

R^2 superiores a 0.99 y consistentemente el ajuste es progresivamente mejor cada año. En el caso de mujeres se tiene también una alta bondad de ajuste, siendo mejor ésta para el 2000 que en los otros dos años. Lo que resulta interesante ver es que en ambas series de mortalidad, si bien el valor de ambos periodos de R^2 para mujeres es superior al de hombres en 2000, en 2005 y 2010 el ajuste de hombres resulta comparativamente mejor al de mujeres.

En cuanto a la esperanza de vida calculada por medio del ajuste por splines se tienen valores de 71.37 y 76.47 en 2000 para hombres y mujeres respectivamente, los cuales son similares a las cifras oficiales publicadas por CONAPO (tabla 2); para 2005, se tiene un monto de 72.25 y 77.05 para hombres y mujeres respectivamente, valores también cercanos a los oficiales; en 2010, éstos valores son 73.08 para hombres y 77.81 en mujeres. Esto indica que el método propuesto se puede utilizar de manera confiable para realizar estimaciones de este importante indicador en éste y posiblemente en otros contextos o años.

Aquí va la Tabla 2

Asimismo, se presentan las edades en las cuales se situaron los nodos para hombres y mujeres (tabla 3). La ubicación de los nodos es importante en este tipo de estudios debido a que dan cuenta de los cambios que se presentan en la mortalidad. Esto es, un aumento en la edad de un nodo implica una disminución de la mortalidad en el grupo de edad que ajusta. En este estudio, se observa que: el segundo nodo está ubicado en edades adultas similares para hombres (33) y mujeres (37), lo que implica que el primer polinomio del spline ajusta la mortalidad en la niñez, la adolescencia y los adultos jóvenes en ambos sexos; el tercer nodo está ubicado en una edad adulta a los 38 años de edad para hombres y 43 para mujeres lo cual es indicativo de una menor mortalidad femenina en este grupo de edad; el cuarto nodo se ubica en edades distintas para ambos sexos, puesto que en hombres se localiza en la edad 50, mientras que para mujeres en los 61 años, esto implica que al ser menor la mortalidad de las mujeres en esos rangos de edad, no se presenta una caída significativa de la curva de mortalidad, como en el caso de hombres por lo que para alcanzar aproximadamente el mismo nivel de supervivientes en mujeres se hace en una edad mayor y por tanto la edad de los nodos es mayor para las mujeres. En este estudio, se utilizaron los mismos nodos para cada sexo, en todos los años que abarca el estudio debido a que el tiempo de análisis se observa que no es suficiente para indicar un cambio significativo en la mortalidad por cada sexo.

Aquí va la Tabla 3

Discusión

La mortalidad en México ha disminuido de manera constante desde los años 40 hasta la actualidad, y la esperanza de vida ha continuado aumentando durante este periodo. Esto se ha presentado a ritmos diferentes y se refleja en que a finales del siglo XX y principios del siglo XXI ha existido una disminución considerable del ritmo de ganancias en esperanza de vida de la población en general (Camposortega, 1992; 1997: 11). Esto es, si bien continúan registrándose ganancias en este indicador, el ritmo al que actualmente se presentan éstas es considerablemente menor al que se presentaba en los años cincuenta y sesenta[5].

Es por ello que el estudio de la mortalidad, y específicamente de la estimación de los patrones modelo parametrizados, que describan matemáticamente la estructura por edad y sexo de la mortalidad y la evolución de ésta a través del tiempo, puede ayudar a profundizar en el conocimiento sobre el comportamiento de este indicador demográfico. Por tanto, este tipo de investigaciones se considera que son importantes ya que permiten realizar la manipulación matemática de los fenómenos demográficos, específicamente la mortalidad en este caso, lo que puede posibilitar además la realización de proyecciones (Halli y Rao, 1992). Esto es, encontrar las leyes que rigen los fenómenos demográficos para describirlos matemáticamente, permite plantear de forma fundamentada escenarios futuros que sean posibles, probables o deseables si ocurriesen una serie de condiciones específicas y a partir de ello proponer acciones adecuadas para intentar disminuir la mortalidad en esas edades concretas. Los modelos, como el aquí propuesto, son útiles también para desarrollar

[5] Las cifras sobre el incremento de la esperanza de vida a lo largo del siglo XX y principios del XXI se presentaron previamente.

investigaciones sobre efectos posibles en edades específicas, cambios en los determinantes demográficos, y el impacto que tendrán estos sobre la tendencia y el impacto de la mortalidad en la población (Halli y Rao, 1992), lo cual puede a su vez servir como insumo para el establecimiento de políticas públicas o campañas de salud pública dirigidas a mitigar o retrasar la mortalidad.

Se propuso una metodología que involucra un polinomio de tercer grado en las primeras 4 edades y en las últimas de la serie de supervivientes los cambios de concavidad observados en la curva. Estos polinomios presentan una unión suave con los polinomios del spline cúbico de 5 nodos ajustado entre las edades 4 y 80. Entre los principales alcances que se tienen del uso del método propuesto es que éste presenta una buena bondad de ajuste y por tanto una buena descripción de la mortalidad.

Esto es, el ajuste propuesto se cree que es adecuado puesto que toma en cuenta las propiedades intrínsecas de las curvas de mortalidad como son: el cambio de concavidad en la curva en las primeras edades de la tabla de vida al utilizar un polinomio de tercer grado que ajuste perfectamente las primeras cuatro edades de la tabla de vida, especialmente en los países en vías de desarrollo como México, donde la mortalidad infantil y de la niñez es todavía alta, si se compara con otros países del mundo; el crecimiento exponencial de las tasas de mortalidad en edades adultas, lo cual se cumple a partir del uso del spline y la flexibilidad de estimación que conlleva esta técnica; y la desaceleración de la mortalidad que se presenta en las edades extremas de la tabla de vida, también a través del uso de un polinomio de tercer grado para respetar este cambio de concavidad en la curva de mortalidad.

Entonces, el método de splines se considera que estima correctamente patrón general de la mortalidad describiendo adecuadamente sus cinco características principales: una alta tasa de mortalidad en el primer año de vida; una tasa descendente a partir de esa edad durante la niñez; un aumento significativo en las tasas de mortalidad, especialmente a partir de los 20 años aproximadamente; una probabilidad creciente de muerte a medida que la edad avanza

(Congdon, 1993: 242); y finalmente, una desaceleración de la mortalidad que se presenta en las edades extremas de la tabla de vida (Gavrilov y Gavrilova, 2002:10).

Este modelo además de tener la ventaja de ser una estimación con una alta bondad de ajuste de los datos de mortalidad solamente requiere el conocimiento de álgebra de matrices y del cálculo de las tablas de vida para calcularlo; asimismo, su mayor ventaja, gracias a su flexibilidad debida al uso de polinomios en la estimación matemática de la mortalidad, es que presenta la posibilidad de utilizarse para realizar proyecciones del fenómeno demográfico en diferentes contextos, en distintos momentos y puede ser también ocupado en el ajuste de causas de muerte.

Limitaciones

Esta metodología tiene la desventaja principal que se requiere estimar una alta cantidad de parámetros, debido a que por cada polinomio del spline hay 4 parámetros a estimar lo cual da un total de 24 parámetros. Esta situación provoca que este ajuste sea, hasta cierto punto, poco parsimonioso. Otra desventaja, radica en la elección de los nodos, la cual se tiene que realizar de forma manual y no otorga la posibilidad de generalizar, por sexo, para distintas causas de muerte o para diferentes años si se tomara una serie de datos mayor a la analizada aquí, lo que provoca que los ajustes no se puedan realizar de forma automática.

Una limitación adicional se relaciona con la calidad de los datos de mortalidad provenientes de estadísticas vitales, especialmente en las edades mayores. Estos datos tienen ciertas deficiencias, como problemas de cobertura, particularmente en las zonas rurales debido a una cobertura inadecuada del Registro Civil en regiones del país poco accesibles y para defunciones infantiles; el subregistro de defunciones, especialmente infantiles y maternas ha conservado niveles importantes (Camposortega, 1992) y un largo proceso burocrático que provoca el retraso de publicación de la información (Hernández y Narro, 2010: 244). Esto puede provocar que los cálculos no se ajusten de la manera más adecuada a la realidad y exista cierta incertidumbre sobre el nivel real de la mortalidad en las edades extremas especialmente.

Anexos

Anexo 1 - Matrices ocupadas en el cálculo del spline cúbico

$$A=\begin{pmatrix}
x_1^3 & x_1^2 & x_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
x_2^3 & x_2^2 & x_2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & x_2^3 & x_2^2 & x_2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & x_3^3 & x_3^2 & x_3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_3^3 & x_3^2 & x_3 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_4^3 & x_4^2 & x_4 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_4^3 & x_4^2 & x_4 & 1\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_5^3 & x_5^2 & x_5 & 1\\
3x_2^2 & 2x_2 & 1 & 0 & -3x_2^2 & -2x_2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 3x_3^2 & 2x_3 & 1 & 0 & -3x_3^2 & -2x_3 & -1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3x_4^2 & 2x_4 & 1 & 0 & -3x_4^2 & -2x_4 & -1 & 0\\
3x_2 & 1 & 0 & 0 & -3x_2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 3x_3 & 1 & 0 & 0 & -3x_3 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3x_4 & 1 & 0 & 0 & -3x_4 & -1 & 0 & 0\\
1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0
\end{pmatrix}
\qquad
B=\begin{pmatrix} a_1\\ b_1\\ c_1\\ d_1\\ a_2\\ b_2\\ c_2\\ d_2\\ a_3\\ b_3\\ c_3\\ d_3\\ a_4\\ b_4\\ c_4\\ d_4 \end{pmatrix}
\qquad
C=\begin{pmatrix} y_1\\ y_2\\ y_2\\ y_3\\ y_3\\ y_4\\ y_4\\ y_5\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0\\ 0 \end{pmatrix}$$

Anexo 2 – Gráficas por año de la comparación de la serie de supervivientes a edad exacta *x*, observada y ajustada, hombres y mujeres. México 2000, 2005 y 2010

Aquí va la Gráfica Anexo 2 Hombres 2000

Aquí va la Gráfica Anexo 2 Hombres 2005

Aquí va la Gráfica Anexo 2 Hombres 2010

Aquí va la Gráfica Anexo 2 Mujeres 2000

Aquí va la Gráfica Anexo 2 Mujeres 2005

Aquí va la Gráfica Anexo 2 Mujeres 2010

Anexo 3 – Gráficas por año de la comparación de la serie probabilidades de muerte, observada y ajustada, hombres y mujeres. México 2000, 2005 y 2010

Aquí va la Gráfica Anexo 3 Hombres 2000

Aquí va la Gráfica Anexo 3 Hombres 2005

Aquí va la Gráfica Anexo 3 Hombres 2010

Aquí va la Gráfica Anexo 3 Mujeres 2000

Aquí va la Gráfica Anexo 3 Mujeres 2005

Aquí va la Gráfica Anexo 3 Mujeres 2010

Bibliografía

- Barrera, P. et al., 1996. El ABC de los Splines. *Sociedad Matemática Mexicana.* Serie Aportaciones Matemáticas, No.9, p. 245.

- Camposortega Cruz, S., 1992. *Análisis Demográfico de la Mortalidad en México 1940-1980*. Distrito Federal: El Colegio de México.

- Camposortega Cruz, S. 1997. Cambios en la mortalidad / cien años de mortalidad en México. *DemoS*, [En línea], disponible en <http://www.ejournal.unam.mx/dms/no10/DMS01005.pdf>, [Acceso el 12 de marzo de 2012].

- Consejo Nacional de Población (CONAPO) [En línea], disponible en <http://conapo.gob.mx/>, [Acceso el 8 de agosto de 2012]

- Congdon, P., 1993. Statistical Graduation in Local Demographic Analysis and Projection. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 156 (2), pp. 237-270.

- Gavrilov, L. A. y Gavrilova, N. S., 2002. The quest for the theory of human longevity.. *The Actuary*, Living to 100, pp. 10-13.

- Gavrilov, L. A. y Gavrilova, N. S, 2006. Reliability Theory of Aging and Longevity. En: Masoro E.J. & Austad S.N.. (eds.): *Handbook of the Biology of Aging*, 6th ed. San Diego: Academic Press. pp. 3-42.

- Hartmann, M., 1987. Past and recent attempts to model mortality at all ages. *Journal of Official Statistics*, 3 (1), pp. 19-39.

- Halli, S. y Rao K. V., 1992. *Advanced Techniques of Population Analysis*. N.Y: Plenum Press, The Plenum series on demographic methods and population analysis.

- Heligman, L. y Pollard, L.H., 1980. The age pattern of mortality. *Journal of the Institute of Actuaries*, No.107, pp. 49-80.

- Hernández-Bringas, H. y Narro-Robles, J., 2010. El homicidio en México, 2000-2008. *Papeles de Población*, 16 (63), pp. 243-271.

- Instituto Nacional de Estadística y Geografía (INEGI) [En línea], disponible en <http://www.inegi.org.mx/est/contenidos/espanol/proyectos/continuas/vitales/bd/mortalidad/MortalidadGeneral.asp?s=est&c=11144>, [Acceso el 8 de agosto de 2012]

- Omran A., 1971. The Epidemiologic Transition. *Milbank Memorial Fund*, 49 (4), pp. 509-537.

- Smith, S., Tayman, J. and Swanson, D., 2001. *State and Local Population Projections: Methodology and Analysis*. Durham: The Springer Series on Demographic Methods and Population Analysis.