

Tendencias y comportamiento de la mortalidad en Colombia entre 1973 y 2005*

Gisou Díaz Rojo

Universidad del Tolima. Colombia

Ana Debón Aucejo

Centro de Gestión de la Calidad y del Cambio. Universitat Politècnica de València. España

Resumen

Estudiar los cambios y tendencias de la mortalidad resulta de gran importancia dado que se registran fenómenos como el aumento progresivo de la población, el envejecimiento poblacional y la reducción de la mortalidad, los cuales tienen un impacto en términos de desarrollo económico en el desarrollo general de los países. Con este trabajo se pretende estudiar el comportamiento de la mortalidad en Colombia para el período 1973-2005 utilizando tablas de mortalidad construidas a partir de Latin American Human Mortality Database y modelizar la mortalidad en Colombia con el modelo de Lee-Carter y algunas extensiones. Los resultados obtenidos por los diferentes modelos permitieron identificar que la estructura de la mortalidad de Colombia se asemeja a la de países desarrollados.

Palabras clave: modelos de mortalidad; modelo de Lee-Carter; Colombia; tablas de vida.

Clasificación AMS: 97K80, 62P05

Trend and patterns of mortality in Colombia: 1973-2005

Abstract

Nowadays, it is great importance to study the changes and trends of the mortality. The population aging and the reduction of the mortality have an economic impact in the general development of the countries. The aim of this work is to study the patterns of mortality in Colombia for the period 1973-2005. Data from Latin American Human Mortality Database was used for constructing life tables. Then, the Lee-Carter model and some extensions were fitted for modeling the life tables in Colombia. The results obtained by those different models allowed concluding that the structure of the Colombia's mortality is similar to the developed countries.

* La investigación llevada a cabo en este trabajo ha sido financiada por el Ministerio de Educación, Cultura y Deporte de España, con el proyecto MTM2013-45381-P.

Keywords: Colombian mortality; Lee–Carter models; life tables

AMS Classification AMS: 97K80, 62P05

1. Introducción

Comprender los cambios demográficos de los países es de actual interés en muchas áreas del conocimiento tales como la demografía, las ciencias económicas, la biología o las ciencias actuariales y financieras. Para conocer esta dinámica es necesario analizar tanto el crecimiento de la población, como la incidencia de la mortalidad.

Durante el siglo XX en Colombia se dieron grandes cambios demográficos y socioeconómicos como consecuencia del proceso de urbanización y de industrialización. Eventos como el aumento del nivel educativo de la población, el desarrollo científico y tecnológico, la reducción de la mortalidad y la mejora de la calidad de vida de la población, se relacionan con la transformación demográfica y el envejecimiento poblacional que se percibió en el país (MinSalud, 2013). Estos cambios demográficos estuvieron en concordancia a los datos registrados para América Latina en la segunda mitad del siglo pasado: incremento constante en el crecimiento de las generaciones, descenso acentuado en la tasa de mortalidad infantil, aumento de la esperanza de vida al nacer, disminución de la tasa de fecundidad y disminución de la tasa anual de crecimiento poblacional (Saad *et al.*, 2009).

Según el Centro Latinoamericano y Caribeño de Demografía (CELADE, 1996), aunque se observa heterogeneidad entre los países latinoamericanos en cuanto a su situación demográfica, también se pueden apreciar similitudes en su comportamiento en la fecundidad, la mortalidad y migraciones internacionales. Teniendo en cuenta la tipología propuesta por CELADE (1996) que señala las etapas de la transición demográfica para los países de América Latina, Colombia se ubica en el grupo de países en plena transición demográfica -países con natalidad moderada y mortalidad moderada o baja-, entre los que se encuentran además Brasil, Costa Rica, Ecuador, México, Panamá, Perú, República Dominicana y Venezuela.

Por otra parte, Acosta y Romero (2014a) plantean que Colombia presenta características de ciclos avanzados y primarios de la transición epidemiológica. A inicios del siglo XX, el país tenía entre sus principales causas de muertes enfermedades de tipo infecciosa y parasitaria, mientras que en los últimos años la principal causa de muerte se relacionan con enfermedades del sistema circulatorio y cánceres, enfermedades acorde con edades avanzadas. De igual forma, las causas externas como los homicidios y accidentes de transporte terrestre aún se encuentran dentro de las principales causas de muerte. La Tabla 1 muestra de manera resumida algunos datos relacionados con las causas de muerte para los últimos años en Colombia.

Tabla 1

Distribución porcentual de la mortalidad por causas de 1990 a 2012.

<i>Causas/Año</i>	<i>Mujeres</i>			<i>Hombres</i>		
	<i>1990</i>	<i>2000</i>	<i>2012</i>	<i>1990</i>	<i>2000</i>	<i>2012</i>
Enfermedades del sistema circulatorio	36,5	34,0	33,7	25,1	23,4	28,0
Neoplasias	17,0	19,1	22,1	10,5	12,0	16,2
Enfermedades del sistema respiratorio	9,5	9,4	11,5	7,2	7,3	9,7
Enfermedades infecciosas y parasitarias	4,2	3,8	2,5	3,7	4,2	3,5
Causas externas	7,9	7,8	4,9	35,3	35,5	23,1
Homicidios	2,9	2,8	1,5	23,9	23,5	12,7
Accidentes Terrestres	1,7	2,0	1,3	3,8	5,1	4,5

Fuente: Acosta y Romero (2014a)

Según Horiuchi (1999), las características mencionadas anteriormente ubican a Colombia en un revés de la transición epidemiológica, espacio donde coexisten enfermedades propias de las etapas avanzadas de la transición y también un número importante de muertes por homicidios y producto de alienaciones sociales. Podemos decir además, que la reducción de las tasas de mortalidad puede atribuirse -como en otros países- a un incremento de los ingresos, lo cual permite unas mejores condiciones de vida y nutrición para la población; así como un mejor acceso a la salud y a sus avances científicos y tecnológicos.

De acuerdo con el Ministerio de salud y protección social de Colombia (MinSalud, 2013), en las últimas décadas Colombia ha experimentado un aumento progresivo de la población y de su expectativa de vida (producto de la caída en de la tasa de mortalidad), con una reducción de la base de la pirámide poblacional (o envejecimiento de la población), los cuales tienen consecuentes cambios sobre su perfil demográfico.

La Tabla 2 presenta la evolución de algunos indicadores demográficos para Colombia. Como se puede observar, la población ha crecido en cifras absolutas desde 1973 aunque la velocidad de crecimiento va variando en los diferentes periodos. La tasa de mortalidad infantil decreció de manera importante, la tasa bruta de mortalidad también disminuye notablemente, la esperanza de vida aumenta y la tasa bruta de natalidad decrece así como la tasa total de fecundidad. Estos cambios -propios de países en una transición demográfica avanzada- indican que, la población colombiana presenta características demográficas similares a la de países desarrollados.

Otros investigadores como Reyes (2010) indican que en Colombia la mortalidad tiene dos características fundamentales; la primera es que la mortalidad está subregistrada como en la mayoría de los países subdesarrollados y la segunda, que presenta sobremortalidad masculina por causa del conflicto interno y la violencia juvenil. La autora menciona el hecho de que tanto los registros de defunciones como los de nacimientos en Colombia presentan omisiones, señalando un incremento de las muertes

violentas en la población masculina así como una posible influencia del conflicto armado sobre factores de producción.

Tabla 2

Evolución de los indicadores demográficos en Colombia de 1973 a 2005.

<i>Indicadores</i>	<i>1973</i>	<i>1985</i>	<i>1993</i>	<i>2005</i>	<i>2012</i>
Población censada	20.666.920	27.837.932	33.109.839	41.468.384	46.581.823
Tasa mortalidad infantil (‰)	75,0	53,3	33,9	25,6	17,2
Tasa bruta de mortalidad (‰)	9,4	7,4	7,2	4,4	5,6
Esperanza de vida (años)	60,9	68,0	70,9	72,6	73,8
Tasa bruta de natalidad (‰)	37,6	32,6	27,2	22,1	19,1
Tasa bruta de fecundidad (‰)	4,5	3,2	3,0	2,4	2,3

Fuente: DANE (2007), MinSalud (2013), UNICEF (2016) y Carmona-Fonseca (2005).

Apoyando la idea anterior sobre los problemas de subregistro, Acosta y Romero (2014b) señalan que el subregistro en Colombia en 2011 se estimaba en 20,3%, lo que ubicaba al país entre los de mayor subregistro en América Latina. En este mismo sentido Carmona-Fonseca (2005), resalta las deficiencias que presenta Colombia en sus registros poblacionales o la falta de los mismos durante varios períodos de tiempo, cuestión que impide hacer en varias ocasiones un análisis demográfico adecuado debido a los diferentes valores que presentan las fuentes y los diferentes métodos para ajustar los datos provenientes de los censos o hacer proyecciones.

En este trabajo utilizando tablas de mortalidad construidas a partir de Latin American Human Mortality Database para el periodo 1973-2005 se estudia el comportamiento de la mortalidad en Colombia. Para ello se modeliza la mortalidad haciendo uso de tres modelos conocidos: Lee-Carter, Lee-Carter con dos términos y el modelo Edad-Periodo-Cohorte. El software estadístico utilizado para el ajuste de los modelos fue R Core Team (2015).

La estructura de este trabajo es la siguiente: la Sección 2 muestra las fuentes de datos demográficos para Colombia, la Sección 3 describe la estructura de la tabla de vida y la construcción de dichas tablas para Colombia a partir de interpolación lineal entre censos, la Sección 4 presenta los tres modelos para el análisis de la mortalidad en Colombia: modelo de Lee-Carter (Lee y Carter, 1992), Lee-Carter con dos términos (Booth *et al.*, 2002) y el modelo Edad-Periodo-Cohorte (Cairns *et al.*, 2009); y la Sección 5 expone los resultados de la aplicación de los modelos analizando el comportamiento de los parámetros durante el período 1973-2005. Por último, en la Sección 6 se presenta las conclusiones obtenidas en el trabajo.

2. Fuentes de datos demográficos

Diferentes fuentes de información se pueden usar para obtener datos demográficos provenientes de encuestas y censos nacionales. Según el Departamento Administrativo

Nacional de Estadística (DANE), Colombia cuenta con una larga tradición de censos de población (DANE, 2007) desde el siglo XVIII, sin embargo la historia institucional de estadísticas comienza en 1906 cuando se creó la Dirección General de Estadísticas, vinculada al Ministerio de Hacienda, institución que fue convertida en 1923 en la Dirección General de Estadística, la cual tuvo como misión desde 1935 recopilar, analizar y producir las cifras oficiales del país (Carmona-Fonseca, 2005). Los últimos censos han sido realizados por el DANE, entidad independiente creada en 1953 y adscrita a la presidencia de la República.

Las cifras para Colombia están disponibles en algunas web que mencionamos a continuación.

- Departamento Administrativo Nacional de Estadística (DANE).

Esta Institución colombiana produce y difunde información estadística de calidad para la toma de decisiones y la investigación del país. Es la encargada de desarrollar el Sistema Estadístico Nacional y tiene como objetivos garantizar la producción, disponibilidad y calidad de la información estadística estratégica, dirigir, planear, ejecutar, coordinar, regular y evaluar la producción y difusión de información oficial básica (www.dane.gov.co).

- Latin American Human Mortality Database (LAHMD).

Esta web es una plataforma donde se encuentran disponibles datos para cinco países de Latinoamérica: Argentina, Brasil, Colombia, México y Perú relacionados con el total de defunciones y causas de muerte, tanto a nivel nacional como regional para cada país. En esta página de fácil acceso, se muestran los datos bajo un mismo formato y de manera organizada, tanto por años como por temas, en los periodos disponibles para cada país, lo que ofrece cierta ventaja sobre otras fuentes (www.lamortalidad.org).

- Latin American Mortality Database (LAMBdA).

Esta base de datos contiene información relacionada con censos y estadísticas vitales de la población y el número de muertes para 18 países: Argentina, Brasil, Chile, Colombia, Costa Rica, Cuba, República Dominicana, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, Uruguay y Venezuela. Conjuntamente muestra tablas de vida cada cinco y diez años, tasas de mortalidad por causas de defunción y tablas de vida de cohortes reconstruidas (estimadas) para cohortes de individuos nacidos a principios del siglo pasado (www.ssc.wisc.edu/cdha/latinmortality).

- Banco Mundial.

Esta página de acceso abierto y gratuito, contiene datos sobre desarrollo para muchos países en el mundo. Permite crear gráficos, descargar información en diferentes formatos y consultar los diferentes indicadores económicos, sociales y de salud, así como resultados de investigaciones relacionadas con los temas de estudio (www.bancomundial.org).

- Bases de datos y publicaciones estadísticas de la Comisión Económica para América Latina y el Caribe (CEPALSTAT).

CEPALSTAT permite acceder a toda la información estadística para los países de América Latina y el Caribe recolectada, sistematizada y publicada por la Comisión Económica para América Latina y el Caribe (CEPAL). Además, facilita estadísticas e indicadores periódicos a través de la consulta de tablas y gráficos pre-definidos mediante una consulta en línea en tiempo real (www.cepal.org/es/datos-y-estadísticas).

3. La tabla de vida

La tabla de vida, también conocida como tabla de mortalidad, es la herramienta que usualmente se utiliza para medir las probabilidades de vida o de muerte de una población o un subgrupo de ésta, en función de la edad para un período establecido (Debón *et al.*, 2009). Su estructura está compuesta por estimaciones basadas en observaciones reales datos de registro de mortalidad y censos de población- las cuales pueden ser estudiadas desde la perspectiva de la edad a la que ocurre el fallecimiento del individuo o desde los años que le restan por vivir después de cierta edad. Es habitual que una tabla de vida presente información relacionada con la mortalidad, la sobrevivencia y la esperanza de vida, elementos básicos para el análisis de la situación de un país y de sus regiones en el tema de la mortalidad.

Las tablas de vida pueden clasificarse de varias formas teniendo en cuenta diferentes aspectos (Zarruk *et al.*, 2011). Cuando en la construcción de tablas se observa la mortalidad en un período de tiempo específico y por tanto la probabilidad de muerte depende solo de la edad alcanzada por el individuo, se denominan tablas de mortalidad de *período*. En cambio, si se asume que las tasas de mortalidad además de depender de la edad, dependen del año calendario en que se alcanza dicha edad, se denomina tablas de mortalidad *dinámicas*.

Además podemos clasificar la tabla de vida como *completa*, cuando se construye presentando la información para cada una de las edades, es decir, con las edades año a año, desde el nacimiento hasta la última edad disponible. En caso de que la información se presente de manera agrupada por rangos de edades, entonces la tabla de vida se clasifica como *abreviada*. En las tablas de vida abreviadas, la agrupación que habitualmente se utiliza para la edad es: menores de 1 año, de 1 a 4 años y el resto, en grupos quinquenales de edad hasta el intervalo abierto final. Las tablas abreviadas se utilizan cuando se dispone de los datos de mortalidad en tasas referidas a grupos quinquenales de edad y no las tasas de mortalidad de cada año de edad. Cuando se trabaja con este tipo de tablas se asume que las muertes se distribuyen uniformemente en cada intervalo de edad (Ayuso *et al.*, 2007).

3.1 Estructura de la tabla de vida

En cuanto a la estructura de la tabla de vida, el Instituto Nacional de Estadística (INE, 2009) indica que “la tabla de vida está compuesta por un conjunto de funciones biométricas definidas sobre una cohorte ficticia de individuos”, y destaca la importancia

de la comprensión de cada una de ellas para su interpretación. A continuación se definen los conceptos y las funciones biométricas que generalmente se presentan en una tabla de vida.

- Edad exacta (x): Define la cantidad de años cumplidos del individuo. Para tablas abreviadas se señala la edad inferior de cada uno de los intervalos de edad.
- Amplitud del intervalo de edad (n): Incluye los años en cada intervalo de edad. En las tablas completas $n = 1$.
- Número de sujetos en el intervalo (P_x): Población correspondiente al rango de edad en el año (calendario) analizado.
- Número de defunciones contabilizadas en el intervalo (d_x): Incluye el número de sujetos muertos en ese rango de edad durante el año analizado.
- Promedio de años vividos por los que mueren entre las edades x y $x + n$ (${}_n a_x$): Se refiere al tiempo promedio vivido con edad cumplida x por aquellos individuos de la cohorte ficticia que mueren con dicha edad. En una tabla de vida completa, se utiliza el valor de 0.5 para edades mayores a 5 años, por correspondencia en una tabla de vida abreviada se asume que los que mueren en un intervalo de 5 años, viven un promedio de 2.5 años. En relación con las muertes infantiles, este valor se corrige en los intervalos 0 y 1-4 teniendo en cuenta valores de una tabla de vida modelo como la realizada por Coale *et al.* (1983).
- Tasa central de mortalidad (${}_n m_x$): Es el resultado de calcular el cociente entre las defunciones en el intervalo $x, x + n$ (es decir, d_x) y el número de personas en este grupo de edad (P_x).

$${}_n m_x = \frac{d_x}{P_x} \quad [1]$$

- Probabilidad de muerte entre las edades x y $x+n$ (${}_n q_x$): Las probabilidades de muerte se calculan a partir de las tasas de mortalidad específicas para cada grupo etario. Debe interpretarse como la probabilidad de fallecer entre las edades del intervalo para el sujeto que ha sobrevivido hasta la edad x .

$${}_n q_x = \frac{n \cdot {}_n m_x}{1 + (n - {}_n a_x) \cdot {}_n m_x}$$

- Supervivientes a la edad exacta x (${}_n l_x$): Representa el número de individuos de la cohorte ficticia inicial que llegan con vida a la edad x ; l_0 es el número de nacidos que componen la generación y usualmente se le asigna un valor de 100 000 por convenio.
- Defunciones teóricas entre las edades x y $x + n$ (${}_n d_x$): Número de defunciones de la cohorte ficticia inicial que tienen lugar en individuos entre las edades x y $x + n$. Para calcular esta función y la anterior, se utilizan las expresiones

$${}_n d_x = {}_n l_x \cdot {}_n q_x$$

$${}_{n+1} l_x = {}_n l_x - {}_n d_x$$

- Población estacionaria entre las edades x y $x + n$ (${}_n L_x$): Representa el tiempo total vivido (en años) por los individuos de la generación ficticia con edad cumplida x .

$${}_n L_x = n \cdot {}_{n+1} l_x + n a_x \cdot {}_n d_x$$

$L_w = \frac{d_w}{m_w}$ donde w representa la edad más avanzada en la tabla.

- Total de años vividos a partir de la edad x (${}_n T_x$): Indica el número total de años vividos por los supervivientes l_x desde el aniversario x hasta la extinción de la generación, de esta manera ${}_n T_x = \sum_{i \geq x} L_i$

$${}_n T_x = {}_{n+1} T_x + {}_n L_x$$

$T_w = L_w$, donde w representa la edad más avanzada en la tabla.

- Esperanza de vida a la edad x (${}_n e_x$): Representa el número medio de años que les quedan por vivir a los supervivientes a la edad x . La esperanza de vida al nacer (e_0) es el número medio de años vividos por una generación de nacidos bajo condiciones de mortalidad dadas.

$${}_n e_x = \frac{{}_n T_x}{{}_n l_x}$$

Las tablas de vida de periodo anual generalmente se construyen con el objetivo de describir el comportamiento de la mortalidad de la población residente en un país, ya sea por sexo o de forma conjunta para ambos sexos. Para esto, se somete a una cohorte ficticia de 100000 individuos al patrón de mortalidad por edad definido, es decir, por las tasas específicas de mortalidad observadas sobre la población en estudio en el año de referencia y luego se calculan sobre la misma el resto de funciones biométricas de la tabla (Zarruk *et al.*, 2011).

Por otra parte, es importante recordar que la tabla de vida se construye a partir de los datos de población provenientes de censos y de los registros de mortalidad existentes para cada país o región, por lo cual cualquier dificultad en su calidad afectará la validez de las estimaciones que se realicen a partir de la tabla de vida según la Organización Panamericana de la Salud (OPS, 2003).

3.2 Interpolación lineal de la población

En la construcción de la tabla de vida para Colombia, se utilizó la información proveniente de Latin American Human Mortality Database (Urdinola y Queiroz, 2015), donde las defunciones se publican para el periodo 1970-2012, organizadas para grupos

de edades (desde 0 hasta 84 años), con el resto de edades superiores a 84 agrupadas en una única categoría (85+). En cuanto a los datos referentes a la población, encontramos información sólo para los años censales 1973, 1985, 1993 y 2005.

Dado que se dispone de datos de población para los últimos cuatro censos, se complementó la información haciendo uso de la interpolación lineal a trozos, utilizando la expresión (1) propuesta por Delwarde y Denuit (2003) para calcular la población entre censos (1974 a 1984, 1986 a 1992 y de 1994 a 2004),

$$P_{xt} = P_{xt_1} + \left(\frac{t-t_1}{t_2-t_1} \right) (P_{xt_2} - P_{xt_1}), \quad t \in [t_1, t_2] \quad [2]$$

donde t_1 y t_2 son los dos periodos censales consecutivos.

La interpolación permitió obtener tablas de vida para el periodo 1973-2005 sustituyendo las cifras de defunciones y población en la expresión [1].

4. Modelos de mortalidad

Los modelos para la estimación de las tablas dinámicas de mortalidad se dividen en dos grupos: modelos paramétricos (que pueden ser estructurales o no, dependiendo de si se asume que los parámetros han sido influenciados por el tiempo del calendario o si se incorpora el tiempo cronológico como una variable), o no paramétricos (generalizaciones de técnicas smoothing dependientes tanto de la edad como del tiempo). Ambos tipos de modelos ofrecen diversas herramientas para la construcción y graduación de tablas de mortalidad. Sin embargo, los modelos paramétricos suelen ser más utilizados en la actualidad pues es habitual que los parámetros permitan realizar predicciones sobre la mortalidad futura de una manera más sencilla que los métodos no paramétricos (Debón *et al.*, 2009).

El modelo Lee-Carter ha tenido una gran aceptación en el ámbito de las ciencias actuariales desde que consiguiera explicar en 1992 el 93% de la variación de los datos de mortalidad de EEUU, para el periodo 1933-1987 (Lee y Carter, 1992).

4.1 Modelo de Lee-Carter

El modelo Lee-Carter consiste en expresar la tasa central de mortalidad como una función exponencial que depende de la edad y del tiempo

$$m_{xt} = \exp(a_x + b_x k_t + \epsilon_{xt}) \quad [3]$$

donde, m_{xt} : corresponde al valor central de la mortalidad para una determinada edad y un determinado año, conocida como la tasa de mortalidad específica o fuerza de mortalidad. a_x : es el parámetro de forma que depende de la edad.

b_x : es el parámetro de sensibilidad que expresa como responde la tasa de mortalidad a cambios en k_t a lo largo de la edad.

k_t : es el índice que representa la tendencia de la mortalidad a lo largo del tiempo.

ϵ_{xt} : representa el error aleatorio, con media 0 y varianza σ_ϵ^2 , que no se explica por el modelo.

El modelo se puede expresar de forma equivalente a [3] con la expresión

$$\ln(m_{xt}) = a_x + b_x k_t + \epsilon_{xt} \quad [4]$$

La expresión [3] es una versión reducida del modelo de Lee-Carter, una forma más general se expresaría en función de la probabilidad de muerte q_{xt} como lo proponen Debón *et al.* (2008)

$$\text{logit}(q_{xt}) = \ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = a_x + \sum_{i=1}^r b_x^{(i)} k_t^{(i)} + \epsilon_{xt} \quad [5]$$

La transformación logit es más adecuada para q_{xt} ya que las estimaciones realizadas con [4]. pueden dar como resultado estimaciones donde q_{xt} supera el valor 1 como señala Lee (2000), y como indican Booth *et al.* (2002) y Renshaw y Haberman (2006), la interacción entre la edad y el tiempo se pueden expresar mejor agregando términos a [4].

Este modelo tiene una estructura invariante bajo cualquiera de las siguientes transformaciones de los parámetros $(a_x, \frac{b_x}{c}, ck_t)$ o $(a_x + cb_x, b_x, k_t - c), \forall c..$ Las restricciones propuestas por Lee y Carter (1992) consisten en normalizar los parámetros para obtener una única solución del modelo:

$$\sum_x b_x = 1 \quad y \quad \sum_t b_t = 0$$

En este trabajo, utilizaremos el modelo con un solo término sujeto a las restricciones $\sum_x b_x = 1$ y $k_{t_0} = 0$, para poder comparar con los parámetros k_t resultantes de los otros modelos que estudiaremos y poder implementarlo de forma más sencilla en la librería gnm (Turner y Firth, 2015). Esta librería permite estimar los parámetros del modelo de Lee-Carter por máxima verosimilitud.

La predicción de las probabilidades de muerte para edades específicas se obtienen prediciendo k_t mediante un modelo de serie temporal y fijando los a_x y b_x estimados.

El modelo de Lee-Carter es un método de graduación paramétrica, ajusta bien los datos históricos y es sencillo de utilizar. A pesar de ello podemos señalar como desventaja que los parámetros del modelo a_x y b_x están fijos, siendo k_t el término en el cual se basa el método para realizar las predicciones.

4.2 Modelo de Lee-Carter con dos términos

El modelo Lee-Carter con dos términos es un modelo alternativo que consiste en agregar un término a la expresión con el objetivo de mejorar su ajuste. El modelo se expresa en términos de logit mediante la siguiente expresión:

$$\text{logit}(q_{xt}) = a_x + b_x^1 k_t^1 + b_x^2 k_t^2 + \epsilon_{xt} \quad [6]$$

La expresión [6] se obtiene de la expresión [5] para $r=2$. Este modelo es similar al modelo original de Lee-Carter con un término, la diferencia radica en la inclusión del segundo término en la estructura del modelo, por lo que los parámetros de este nuevo modelo tienen una interpretación semejante a la de los parámetros del modelo de Lee-Carter original. Como en el modelo anterior es necesario establecer algunas restricciones, en este trabajo hemos utilizado

$$\sum_x b_x^i = 1 \quad y \quad k_{t_0}^i = 0 \quad i = 1,2$$

Para ajustar por máxima verosimilitud se utilizó la misma librería gnm de antes.

4.3 Modelo Edad-Periodo-Cohorte (APC)

El uso de la cohorte (año de nacimiento) dentro de los modelos de mortalidad ha sido recomendado por diferentes investigadores. Según Debón *et al.* (2009) los modelos Edad-Periodo-Cohorte (APC por sus siglas en inglés) han mostrado buenos resultados en el campo de la medicina y la demografía, siendo considerados en la literatura actuarial a partir de su propuesta por Currie *et al.* (2006). Estos modelos son el resultado de una evolución natural de los modelos dinámicos al incorporar el efecto de la cohorte ($c = t-x$). El modelo APC se expresa de la siguiente manera:

$$\text{logit}(q_{xt}) = a_x + k_t + \gamma_c \quad [7]$$

Este modelo ha sido ajustado bajo las siguientes restricciones,

$$k_{t_0} = 0, \quad \gamma_{c_0} = 0$$

utilizando una vez más la librería gnm.

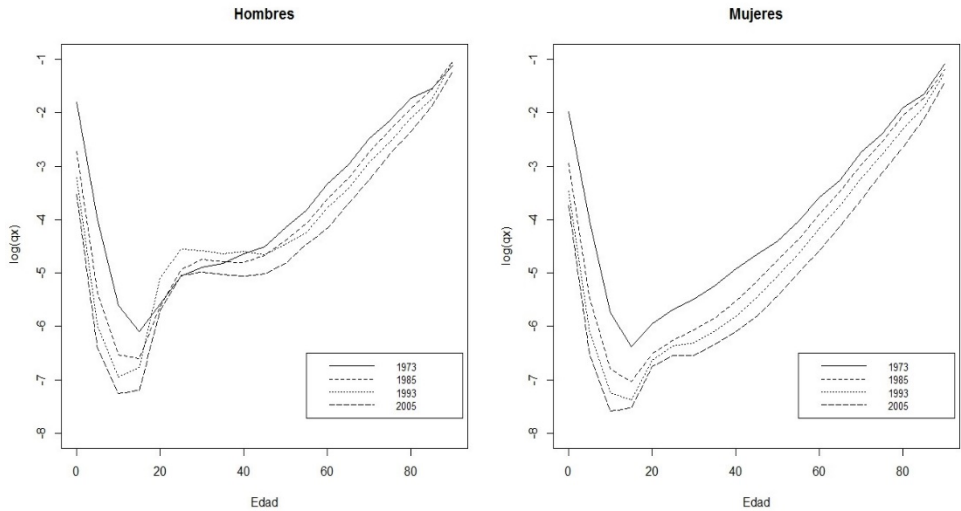
5. Resultados de los ajustes para Colombia

5.1 Análisis descriptivo

A partir del cálculo de la tabla de vida mortalidad para Colombia en los censos disponibles (1973, 1985, 1993 y 2005) es posible observar algunos patrones del comportamiento de la mortalidad que se verán reflejados en los modelos que después ajustaremos. La Figura 1 muestra el comportamiento de las probabilidades de muerte para cada sexo. Como puede observarse (Figura 1) la probabilidad de muerte es muy alta en el grupo de los recién nacidos decreciendo de manera abrupta a partir del primer año cumplido del individuo, se estabiliza para edades intermedias y vuelve a incrementarse a partir de los 60 años para ambos sexos. Además, la Figura 1 muestra un descenso de dicha probabilidad a lo largo de los años.

Figura 1

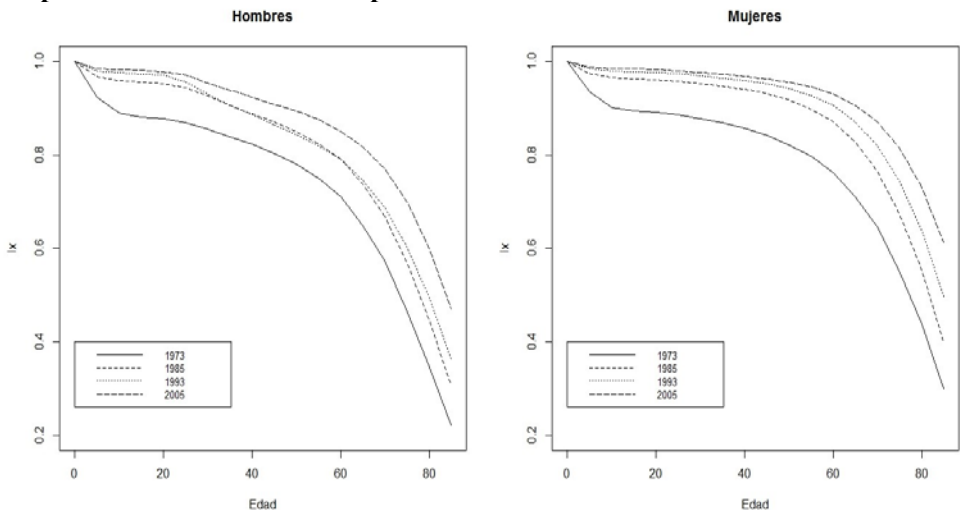
Probabilidades de muerte sq_{xt} en Colombia para los años censales.



Puede también observarse la rectangularización de la curva de supervivientes en la Figura 2. La rectangularización se relaciona con el desplazamiento de la curva de supervivientes hacia la esquina superior derecha (Olivieri, 2001) siendo un fenómeno demográfico que se presenta en los países más desarrollados. En el caso colombiano, este fenómeno se observa de manera más acentuada en las mujeres.

Figura 2

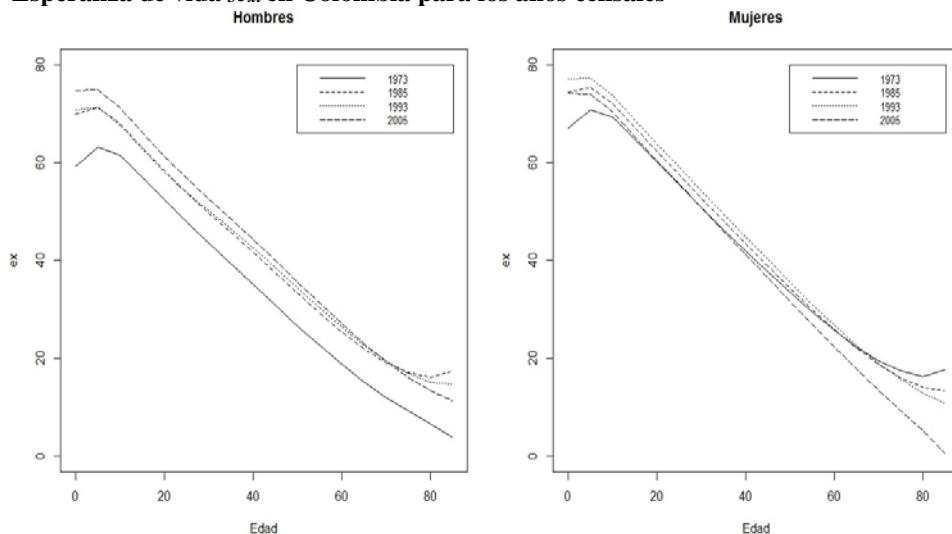
Supervivientes sl_{xt} en Colombia para los años censales



Finalmente, en la Figura 3 se representa la esperanza de vida (en años) la cual se incrementó de manera paulatina a través de los años tanto en hombres como en mujeres, pero presentando mayores valores para las mujeres en todas las edades y en los cuatro años considerados.

Figura 3

Esperanza de vida ${}_{5e_{xt}}$ en Colombia para los años censales



5.2 Estimación de los parámetros

Después de obtener las tablas de mortalidad de la población colombiana, se ajustan los parámetros del modelo de Lee-Carter (LC), el modelo de Lee-Carter con dos términos (LC2) y el modelo Edad-Periodo-Cohorte (APC). Los parámetros a_x , b_x , b^2_x capturan la información de la mortalidad según la edad de los individuos, los parámetros k_t y k_t^2 permiten observar la evolución de la mortalidad a través del tiempo, y el parámetro k_c brinda información de la mortalidad según la cohorte. Las estimaciones obtenidas para el conjunto de parámetros en los diferentes modelos estudiados, tanto para hombres como para mujeres, se representan gráficamente en esta sección.

5.2.1 Estimación del parámetro a_x

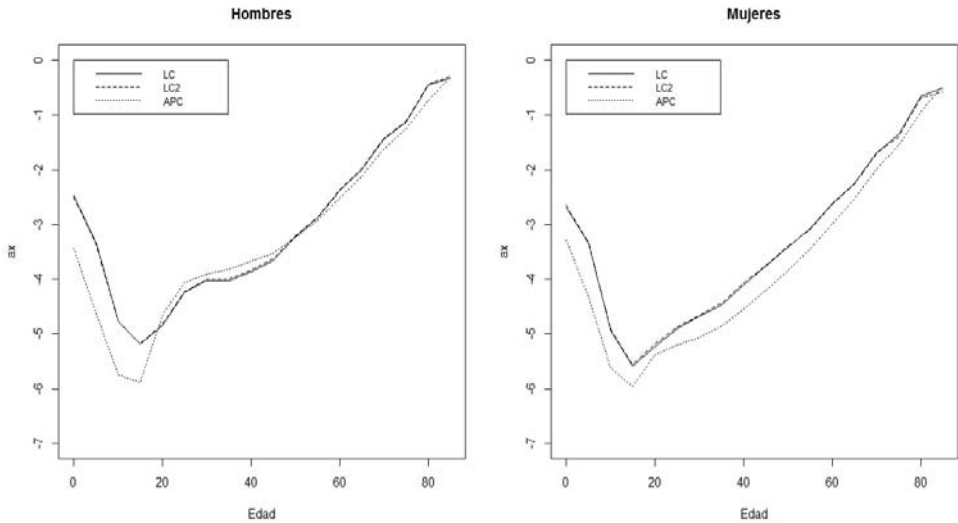
Los valores estimados de los parámetros a_x , tanto para hombres como para mujeres, se muestran en la Figura 4. En esta figura, el comportamiento de los coeficientes a_x evidencia las fases habituales de la mortalidad poblacional. Así, en los primeros años de vida se observa una mortalidad alta que desciende de manera rápida hasta los 15 años de edad, a partir de la cual aumenta lentamente, tanto para hombres como para mujeres. Dicho aumento de la mortalidad es progresivo hasta las últimas edades, no obstante - para el caso de los hombres es más evidente- a este envejecimiento natural se le superpone el efecto del incremento de la mortalidad por accidentes y otras causas

propias del rango de edades 15-35, dicho fenómeno es conocido como “joroba de accidentes” o “joroba social”.

En la Figura 4, para cada sexo, el comportamiento del parámetro a_x es similar en los modelos LC y LC2, y es diferente entre estos y el modelo APC.

Figura 4

Parámetro a_x para para los modelos evaluados

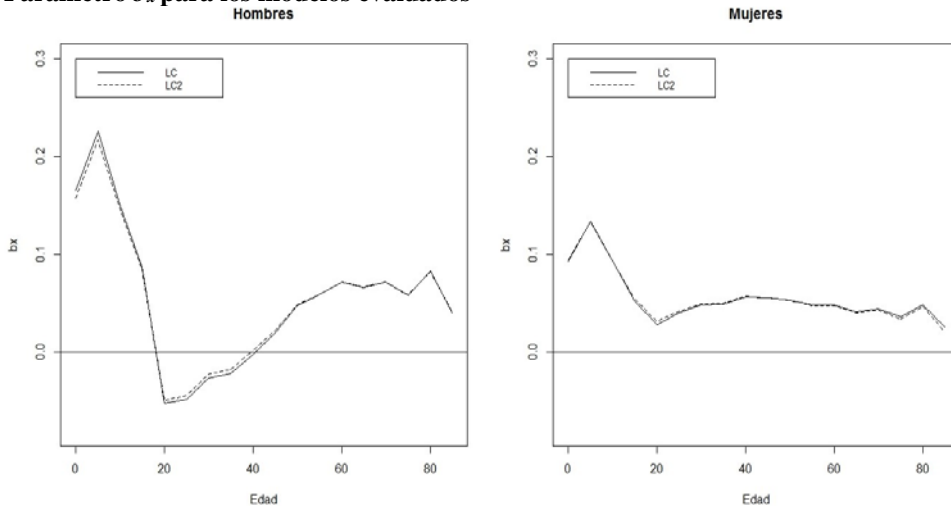


5.2.2 Estimación del parámetro b_x

El parámetro b_x es un parámetro de sensibilidad que indica el comportamiento de la mortalidad a cada edad. Como se puede apreciar en la Figura 5, para los hombres entre los 20 y 40 años, toma valores negativos, indicando que para estas edades la mortalidad se incrementa al variar el tiempo. Por otro lado, en las mujeres se observa que el parámetro es positivo para todas las edades, lo que indica una disminución de la mortalidad a través del tiempo independientemente de la edad. Nuevamente se aprecia que el parámetro b_x tiene un comportamiento similar en los modelos LC y LC2 en ambos sexos.

Figura 5

Parámetro b_x para los modelos evaluados



5.2.3 Estimación del parámetro k_t

El parámetro k_t muestra la tendencia que ha tenido la mortalidad a través del año de defunción. La Figura 6 muestra el comportamiento de las estimaciones del índice temporal de mortalidad para Colombia, que presentan una clara tendencia a disminuir a través del tiempo a través de los modelos LC y LC2.

Figura 6

Parámetro k_t para los modelos evaluados



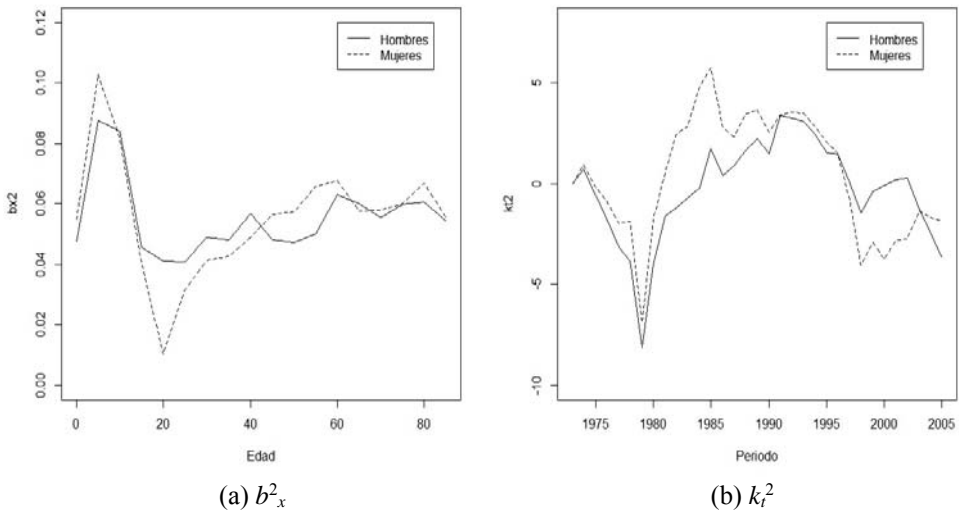
Las estimaciones obtenidas para los parámetros asociados al componente de interacción edad-tiempo, $b_x k_t$, permite deducir que durante el periodo de observación la población colombiana ha presentado un fenómeno sistemático y sostenido de reducción de su mortalidad, que ha afectado de manera conjunta a hombres y mujeres durante los años 1973 - 2005, salvo para los hombres de edades comprendidas entre los 20 y 40 años.

5.2.4 Estimación de los parámetros b_x^2 y k_t^2 del modelo LC2

En la Figura 7 se observan los valores correspondientes a los parámetros b_x^2 y k_t^2 del modelo de Lee-Carter con dos términos. La Figura 7a muestra las estimaciones para el término b_x^2 , el cual alcanza sus valores más altos desde las primeras edades hasta los 15 años, y muestra una ligera tendencia creciente para el resto de edades. Por otro lado, en la Figura 7b se presentan los resultados correspondientes al término k_t^2 . Este término no presenta ninguna tendencia clara, aunque se aprecia que entre 1975 y 1980 los valores decrecen de manera considerable en ambos sexos.

Figura 7

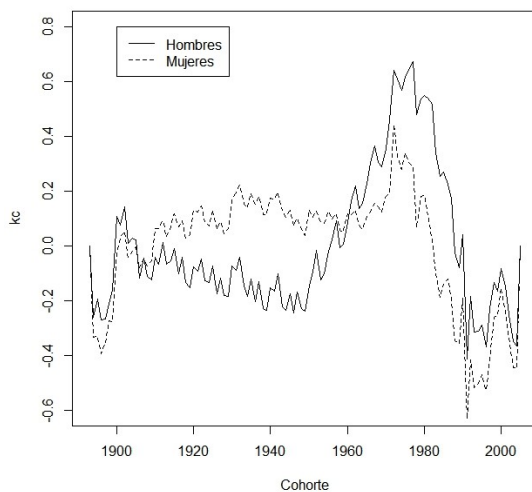
Parámetros b_x^2 y k_t^2 para el modelo LC2



5.2.5 Estimación del parámetro γ_c

La Figura 8 muestra los resultados del término referente a la cohorte del modelo APC, γ_c , los cuales fluctúan sin marcar alguna tendencia. En esta figura lo más importante a destacar es una acusada caída del parámetro entre 1970 y 1990, lo que indica un descenso de la mortalidad para las cohortes en dicho periodo. En la última década de este estudio, de 1995 a 2005, la tendencia de disminución de la mortalidad se ha visto afectada, no obstante, no está claro si dicha disminución se mantiene aunque a un ritmo más lento.

Figura 8

Parámetro γ_c para el modelo APC**5.3 Evaluación de los residuos y comparación de los modelos**

Para analizar la bondad de ajuste de los tres modelos estudiados utilizamos los gráficos de los residuos *deviance*. Las Figuras 9, 10 y 11 muestran el comportamiento de los residuos de los modelos de LC, LC2 y APC, tanto para hombres como mujeres, en función de la edad, el periodo y la cohorte.

La Figura 9 muestra el comportamiento de los residuos para el modelo de LC. En general, los intervalos donde se distribuyen los residuos no son óptimos dado que se distribuyen en el intervalo $[-30, 20]$ para ambos sexos. Además, se puede observar que estos valores son mucho más amplios en los hombres que en las mujeres. Las edades donde existe mayor variabilidad de los residuos son los grupos de edades infantiles y de edades avanzadas (mayores a 70 años), aunque en los hombres, entre los 15 y 40 años, se perciben valores elevados. Por otra parte, es válido señalar que los residuos presentan mayor variabilidad de 1978 a 1985, y para las cohortes alrededor de 1900, tanto en hombres como en mujeres.

En la Figura 10 podemos observar los residuos para el modelo de LC2, los cuales toman valores en el intervalo $[-20, 20]$ para los hombres y $[-15, 15]$ para las mujeres; lo que indica un comportamiento bastante disperso, aunque relativamente mejor que el modelo LC. Nuevamente se aprecia que los residuos en función de la edad, tanto para hombres como para mujeres presentan valores muy elevados en las edades tempranas y en las edades avanzadas. Además, los residuos en función del periodo y la cohorte, presentan un comportamiento similar a lo observado con el modelo LC. Nuevamente en las mujeres el ajuste es un poco mejor dado que los residuos son menores en términos absolutos.

Figura 9

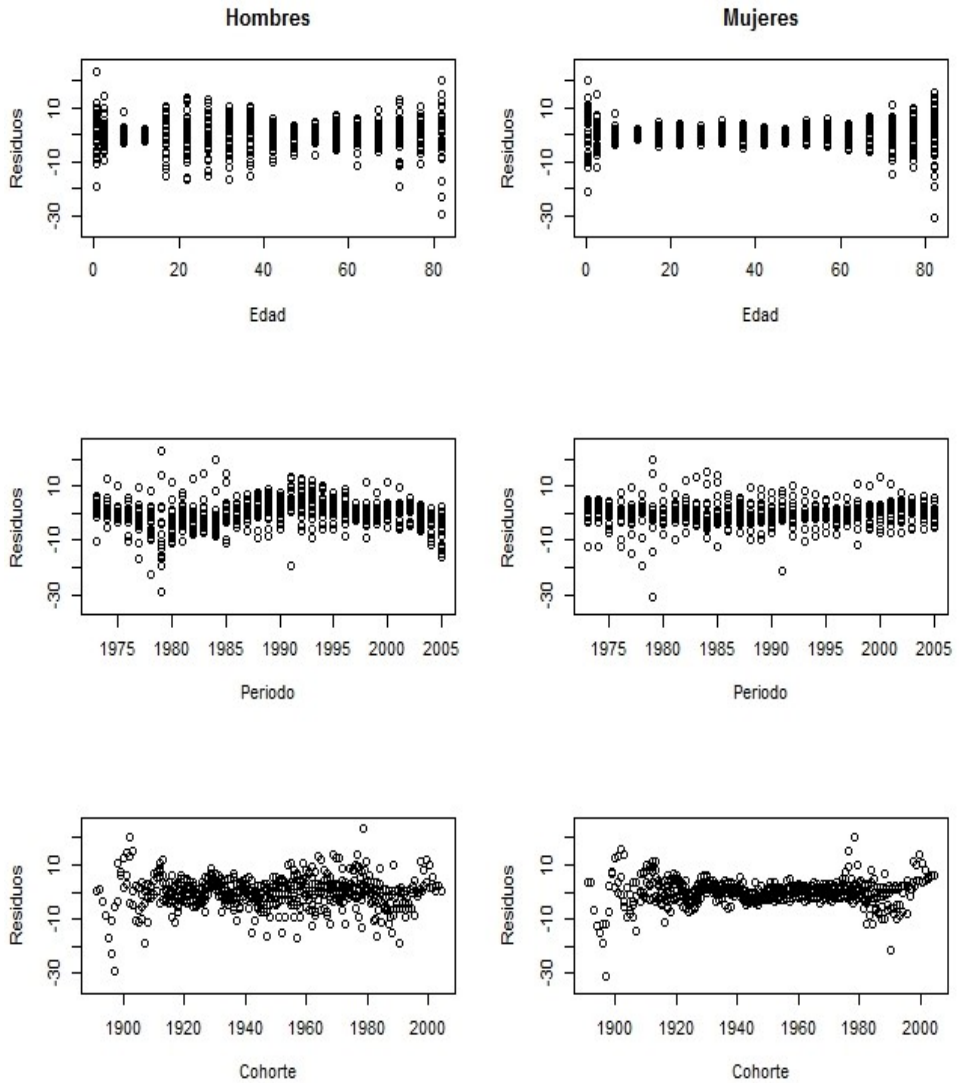
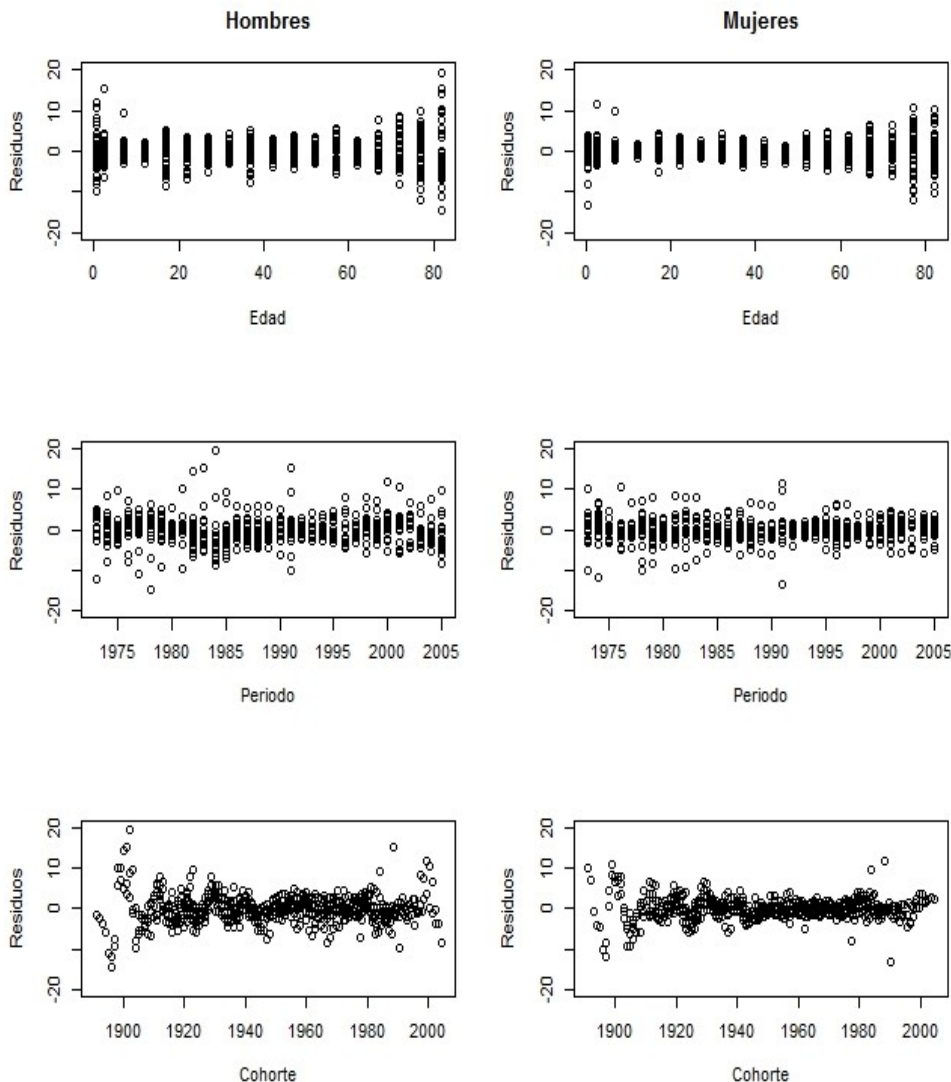
Residuos *deviance* para el modelo Lee-Carter

Figura 10

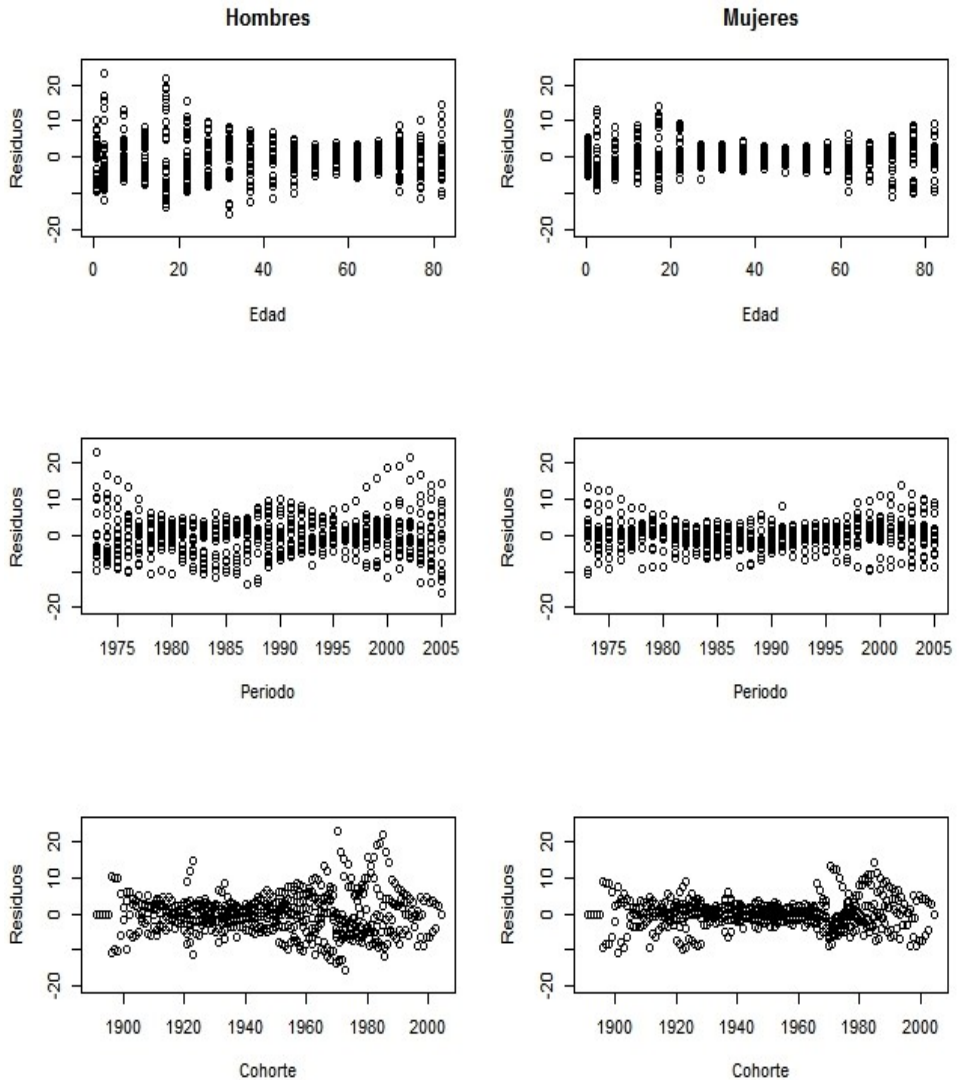
Residuos *deviance* para el modelo Lee-Carter con dos términos



El análisis de los residuos para el modelo APC se presenta en la Figura 11. Así, en el caso de los hombres, los residuos se distribuyen en el intervalo $[-20,25]$ presentando una alta variabilidad. En el caso de las mujeres los residuos se distribuyen en el intervalo $[15,15]$, y aunque son más homogéneos, siguen presentando alta variabilidad. Los residuos en función de la edad tienen un comportamiento similar a los observados con

los modelos anteriores aunque presentan mayor variabilidad en los hombres para las edades entre 15 y 45 años. En cuanto al comportamiento de los residuos en función de la cohorte, los mayores valores se presentan en tres momentos: los primeros años del siglo XX, en torno a 1920 y a partir de 1970. Pese a que el comportamiento es similar para hombres y para mujeres, se reconoce un mejor ajuste en las mujeres.

Figura 11

Residuos *deviance* para el modelo APC

Los modelos se compararon utilizando dos medidas de bondad ajuste, el Error Cuadrático Medio (ECM) y el Error Porcentual Absoluto Medio (EPAM). Estas medidas que han sido utilizadas con anterioridad para modelos de mortalidad en Felipe *et al.* (2002) y Debón *et al.* (2008), permiten determinar cuál de los modelos aplicados presenta un mejor ajuste al comparar sus estimaciones con los datos observados. Los valores correspondientes a estas medidas se muestran en la Tabla 3.

$$ECM(\hat{q}_{xt}) = \sqrt{\frac{\sum_x^{n_x} \sum_t^T (q_{xt} - \hat{q}_{xt})^2}{n_x \cdot T}}$$

$$EPAM(\hat{q}_{xt}) = \left(\frac{\sum_x^{n_x} \sum_t^T \frac{|q_{xt} - \hat{q}_{xt}|}{q_{xt}}}{n_x \cdot T} \right) \cdot 100 \%$$

donde n_x es el número de grupos de edades y T el número total de años.

En la Tabla 3 podemos observar que el modelo de Lee-Carter con dos términos es el que ofrece mejor ajuste, puesto que presenta valores menores para ambas medidas tanto para hombres como para mujeres. De acuerdo con con lo anterior, sería aconsejable utilizar el modelo de Lee-Carter con dos términos para futuras predicciones.

Tabla 3

Estadísticos de bondad de ajuste de los modelos estimados

Modelos	ECM		EPAM	
	Hombres	Mujeres	Hombres	Mujeres
LC	0,1186	0,1062	6,81	13,58
LC2	0,0770	0,0685	4,91	7,47
APC	0,1671	0,1874	14,73	27,43

6. Conclusiones

En este trabajo se aplica el modelo de Lee-Carter, el modelo de Lee-Carter con dos términos y el modelo Edad-Periodo-Cohorte a tablas de mortalidad construidas a partir de datos de población interpolados para el periodo 1973-2005 en Colombia. El ajuste de estos tres modelos permiten analizar el comportamiento de la mortalidad.

La mortalidad colombiana muestra similitudes con el comportamiento de la mortalidad para países desarrollados:

- La mortalidad es mayor en las edades infantiles, luego se estabiliza a partir de los 15 años y posteriormente se incrementa conforme el individuo envejece.
- La mortalidad presenta una disminución con el tiempo independientemente de la edad.
- La mortalidad de las mujeres es inferior a la de los hombres.

- La mortalidad de los hombres colombianos es mayor entre los 15 y 35 años, lo que usualmente se conoce como “joroba social” asociada a una mayor exposición al riesgo.

Los gráficos de los residuos en función del tiempo, la edad y la cohorte muestran, en todos los casos, que los datos de mortalidad de las mujeres son los que presentan un mejor ajuste de los modelos, siendo el modelo Lee-Carter con dos términos el que mejor explica los efectos de la mortalidad para ambos sexos en Colombia. Los residuos para los hombres presentaron mayor variabilidad en general.

A partir de la comparación de los tres modelos con las medidas de bondad de ajuste se encuentra que el modelo de Lee-Carter con dos términos es el que presenta un mejor ajuste tanto en hombres como en mujeres.

Referencias

- ACOSTA, K. Y ROMERO, J. (2014a). «Cambios recientes en las principales causas de mortalidad en Colombia». Número 209. *Banco de La República*, Colombia.
- ACOSTA, K. Y ROMERO, J. (2014b). «Estimación indirecta de la tasa de mortalidad infantil en Colombia, 1964-2008». Número 199. *Banco de La República*, Colombia.
- AYUSO, M., CORRALES, H., GUILLÉN, M., A.M. PÉREZ-MARÍN, M., Y ROJO, J. (2007). «Estadística actuarial vida». *Edicions de la Universitat de Barcelona*.
- BOOTH, H., MAINDONALD, J., Y SMITH, L. (2002). «Applying Lee-Carter under conditions of variable mortality decline». *Population Studies*, 56(3):325–336.
- CAIRNS, A., BLAKE, D., DOWD, K., COUGHLAN, G., EPSTEIN, D., ONG, A., Y BALEVICH, I. (2009). «A quantitative comparison of stochastic mortality models using data from England and Wales and the United States». *North American Actuarial Journal*, 13(1):1–35.
- CARMONA-FONSECA, J. (2005). «Cambios demográficos y epidemiológicos en Colombia durante el siglo xx». *Biomédica*, 25:464–480.
- CELADE (1996). «Impacto de las tendencias demográficas sobre los sectores sociales en América Latina: contribución al diseño de políticas y programas». Technical Report 45, CEPAL/CELADE/BID, Santiago de Chile.
- COALE, A., DEMENY, P., Y VAUGHAN, B. (1983). «Regional Model Life Tables and Stable Populations». *Academic press, inc*, segunda edición.
- CURRIE, I., DURBAN, M., Y EILERS, P. (2006). «Generalized linear array models with applications to multidimensional smoothing». *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(2):259–280.
- DANE (2007). «Estimación de la mortalidad 1985-2005». «Conciliación censal 1985-2005», Colombia. Technical report, DANE.

- DEBÓN, A., MARTÍNEZ-RUIZ, F., MONTES, F., Y MOSHUK, M. (2015). «Proyección de los indicadores de mortalidad para España». *Estadística Española*, 57(188):291–321.
- DEBÓN, A., MONTES, F., Y PUIG, F. (2008). «Modelling and forecasting mortality in Spain». *European Journal of Operation Research*, 189(3):624–637.
- DEBÓN, A., MONTES, F., Y SALA, R. (2009). «Tablas de mortalidad dinámicas. Una aplicación a la hipoteca inversa en España». *Fundación ICO. Publicaciones de la Universitat de València, València*.
- DELWARDE, A. Y DENUIT, M. (2003). «Importance de la période d'observation et des âges considérés dans la projection de la mortalité selon la méthode de Lee-Carter». *Belgian Actuarial Bulletin*, 3(1):1–21.
- FELIPE, A., GUILLÉN, M. Y PEREZ-MARIN, A.M. (2002). «Recent mortality trends in the Spanish population». *British Actuarial Journal*, 8(4):7
- HORIUCHI, S. (1999). «Epidemiological transition in human history». En *Health and mortality issues of global concern*, Nueva York.
- INE (2009). «Tablas de mortalidad. Metodología». *Instituto Nacional de Estadística (INE), Madrid*.
- LEE, R. (2000). «The Lee-Carter method for forecasting mortality, with various extensions and applications». *North American Actuarial Journal*, 4(1):80–91.
- LEE, R. Y CARTER, L. (1992). «Modelling and forecasting U.S. mortality». *Journal of the American Statistical Association*, 87:659–671.
- MINSALUD (2013). «Envejecimiento Demográfico. Colombia 1951-2020. Dinámica demografica y estructuras poblacionales». *Ministerio de Salud y Protección Social, Colombia*.
- OLIVIERI, A. (2001). «Uncertainty in mortality projections: an actuarial perspective». *Insurance: Mathematics and Economics*, 29(4):231–245.
- OPS (2003). «Sobre la estimación de tasas de mortalidad para países de la región de las Américas». *Boletín Epidemiológico. Organización Panamericana de la Salud*, 24(4):1–16.
- ORNELAS, A. (2015). «La Mortalidad y la Longevidad en la Cuantificación del Riesgo Actuarial para la Población de México». Tesis doctoral, Universitat de Barcelona.
- R CORE TEAM (2015). «R: A Language and Environment for Statistical Computing». *R Foundation for Statistical Computing, Vienna, Austria*.
- RENSHAW, A. Y HABERMAN, S. (2006). «A cohort-based extension to the Lee-Carter model for mortality reduction factors». *Insurance:Mathematic and Economics*, 38(3):556–570.

- REYES, A. (2010). «Una aproximación al costo fiscal en pensiones como consecuencia del envejecimiento de la población en Colombia y el efecto de la sobre-mortalidad masculina». *Tesis de máster, Universidad Nacional de Colombia*.
- SAAD, P., MILLER, T., Y MARTÍNEZ, C. (2009). «Impacto de los cambios demográficos en las demandas sectoriales en América Latina». *Revista Brasileira de Estudos de População*, 26(2):239–261.
- TURNER, H. Y FIRTH, D. (2015). «Generalized nonlinear models in R: An overview of the gnm package». R package version 1.0-8.
- UNICEF (2016). www.unicef.org, 3 de Diciembre de 2016.
- URDINOLA, B. Y QUEIROZ, B. (2015). www.lamortalidad.org. Latin American Human Mortality Database, 21 de Noviembre de 2015.
- VILLEGAS, A. M., MILLOSSOVICH, P., Y KAISHEV, V. K. (2016). *StMoMo: An R Package for Stochastic Mortality Modelling*. R package version 0.3.1.
- ZARRUK, A., VILLEGAS, A., Y ORTIZ., F. (2011). «Tablas de mortalidad. Evolución en el sector asegurador colombiano». Technical report, Fasescolda.