

## **Estudio de la mortalidad en Colombia, ajuste del modelo de Lee-Carter para su análisis y predicción.**

Díaz Rojo, Gisou; gdiazrj@ut.edu.co<sup>1</sup>, Debón Aucejo, Ana; andeau@eio.upv.es<sup>2</sup>,  
<sup>1</sup> *Universidad del Tolima, Colombia.*  
<sup>2</sup> *Centro de Gestión de la Calidad y del Cambio, Universitat Politècnica de València, Spain,*

### **RESUMEN**

En la actualidad resulta de gran importancia el análisis de los fenómenos como el crecimiento poblacional y la reducción de la mortalidad por la repercusión económica y social que dichos procesos tienen en el desarrollo de los países. En este sentido las tablas de mortalidad constituyen una herramienta para comprender, a través de las probabilidades de muerte, la esperanza de vida y otros indicadores, la dinámica poblacional. Lee y Carter (1992), plantearon un modelo, cuyo ajuste permite a los analistas obtener una visión dinámica del comportamiento de la mortalidad durante un periodo de análisis.

En este trabajo se hace uso de este modelo para estudiar la mortalidad en Colombia, utilizando tablas de mortalidad construidas a partir de información suministrada por la base de datos internacional Latin American Human Mortality Database. En los resultados se presentan las comparaciones entre los modelos obtenidos en el periodo 1973-2005. Se observa una reducción de la mortalidad en el periodo analizado y se logra identificar la estructura de la mortalidad para hombres y mujeres, así como una tendencia en la reducción de sus diferencias.

## ABSTRACT

At present, is of great interest the analysis of phenomena such as population growth and reduced mortality by economic and social impact these processes have on developing countries. In this sense, the life tables are a tool for understanding, through the probability of death, life expectancy and other indicators, the population dynamics.

Lee and Carter (1992), made a proposal for modeling, whose setting enables to the analysts obtain a dynamic view of the behavior of mortality during a period of analysis. In this work this model is used to study the mortality in Colombia, taking mortality tables, which are constructed from information provided by the international database Latin American Human Mortality Database. The results, show the comparisons among the models obtained in the period 1973-2005. It is observed a reduction in the mortality rate, identifying the structure of mortality for men and women, as well as a tendency to reduce the differences between them.

***Palabras claves:*** tablas de mortalidad; modelo de Lee-Carter; Colombia.

***Área temática:*** Aspectos Cuantitativos de Problemas Económicos y Empresariales con incertidumbre.

## **1. INTRODUCCIÓN**

El estudio de temas demográficos es de actual interés en muchas áreas del conocimiento tales como la demografía, las ciencias económicas, la biología o las ciencias actuariales y finanzas. La dinámica de la población en ocasiones se refleja en el crecimiento de su esperanza de vida, el envejecimiento de la población o la disminución de la fecundidad. Para comprender y evaluar esta dinámica es necesario analizar tanto el crecimiento de la población, como la incidencia de la mortalidad, para lo cual han sido creados un conjunto de índices correspondientes tales como la tasa de mortalidad por edad, la esperanza de vida global y residual, entre otros, los cuales pueden ser obtenidos desde cálculos a realizar con los insumos que proporciona una tabla de mortalidad.

Para el análisis del comportamiento de la mortalidad en Colombia debemos tener en cuenta diferentes cuestiones de las cuales mencionaremos algunas. El crecimiento de la población ha estado acompañado de un cambio en la distribución por edades y un aumento en el promedio de edad de la población, mostrando síntomas de envejecimiento (Acosta y Romero, 2014). La reducción de las tasas de mortalidad puede atribuirse -como en otros países- a un incremento de los ingresos lo que permite unas mejores condiciones de vida y nutrición para la población, esto acompañado de un mejor acceso a la salud y sus avances científicos.

Según Acosta y Romero (2014), Colombia es un país con características propias de ciclos avanzados y primarios de la transición epidemiológica. Mientras que a inicios del siglo XX el país tenía como principal causa de muertes factores de riesgo enfermedades de tipo infecciosa y parasitaria, en los últimos años se está moviendo hacia enfermedades relacionadas con el sistema circulatorio y los cánceres, propias de las edades más avanzadas. No obstante, las causas externas como los homicidios y accidentes de transporte terrestre aún se encuentran dentro de las principales causas de muerte. De acuerdo con la definición propuesta por Horiuchi (1999), estas características ubican a Colombia en un revés de la transición epidemiológica, en la que coexisten enfermedades propias de las etapas avanzadas

de la transición, como los cánceres, con un número significativo de muertes por homicidios y producto de alienaciones sociales.

Podríamos decir de modo general que Colombia es un país con un aumento progresivo de la población y su expectativa de vida (producto de la caída en de la tasa de mortalidad) y con una reducción de la base de la pirámide poblacional (o envejecimiento de la población), los cuales tienen consecuentes cambios sobre su perfil epidemiológico. Por este motivo consideramos que este tema se debe tener en cuenta para el delineamiento de las políticas de salud y pensión llevadas a cabo en los países que se encuentren en etapas avanzadas de la transición demográfica -como es el caso de Colombia- de allí la importancia de su estudio.

Con el objetivo de evaluar la dinámica de la mortalidad para la población colombiana en el periodo 1973-2005, el presente trabajo presenta el ajuste del modelo Lee-Carter sobre las tablas de mortalidad construidas a partir de información suministrada por la base de datos internacional Latin American Human Mortality Database. Este modelo incorpora un componente temporal para el análisis de la mortalidad a través de un periodo de tiempo, permitiendo generar conclusiones sobre la esperanza de vida y las dinámicas de mortalidad sobre la población. La población total es subdividida por género (hombres, mujeres) analizando las diferencias observadas en su mortalidad en el periodo de observación. El ajuste del modelo se implementa en el software estadístico R 3.2.3.

## **2. TABLAS DE MORTALIDAD**

### **2.1. Conceptos y funciones biométricas**

La tabla de mortalidad o tabla de vida es la herramienta que usualmente se utiliza para la modelización y predicción de la supervivencia humana de un colectivo o población en específico. Su contenido está compuesto por estimaciones basados en observaciones reales; dichas estimaciones pueden ser estudiadas desde la perspectiva de la edad a la que ocurre el fallecimiento del individuo o los años que le restan por vivir después de cierta edad.

Su contenido es una abstracción matemática que representa la evolución y decrecimiento constante de la población. La estructura básica, como describe Palacios (1996), debe estar compuesta por al menos cinco indicadores, representados por los siguientes símbolos:  $x$ ,  $l_x$ ,  $d_x$ ,  $q_x$  y  $p_x$ . A continuación, los definimos brevemente.

$x$ : define la cantidad de años cumplidos del individuo, en el rango  $x \in [0, \omega]$ , donde  $\omega$  representa la edad límite donde no hay más sobrevivientes.

$l_x$ : representa el número de sobrevivientes con edad  $x$ .

$d_x$ , representa el número de defunciones ocurridas de individuos con edad  $x$ .

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

$q_x$ : indica el tanto de fallecimientos ocurridos en la edad  $x$ .

$$q_x = \frac{d_x}{l_x}.$$

$p_x$ : indica el tanto de sobrevivientes en la edad  $x$ .

$$p_x = \frac{l_{x+1}}{l_x}.$$

A partir de estas funciones, se puede estimar otros indicadores de mortalidad de la población que permiten analizar otras características como la esperanza de vida que se refiere a la cantidad de años que le resta vivir a un individuo con edad  $x$ . Su cálculo viene dado por la expresión

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

donde  $T_x$  es el total de años que esperan vivir todos los individuos con edad  $x$ . Su valor se obtiene aplicando  $T_x = \sum_{i \geq x} L_i$ , siendo  $L_i$  el número de personas-años, dado por  $L_x = l_{(x+1)} + d_x/2$ .

La secuencia numérica obtenida de los indicadores de mortalidad son el instrumento habitual para representar el número esperado de supervivientes  $l_x$  con edad  $x$  pertenecientes a una cohorte ficticia  $l_0$ . En ella hay una función de supervivencia  $S(x)$  definida como la probabilidad de que el individuo tenga vida a  $x$  edad. La función de supervivencia también podemos expresarla en función del tiempo,  $s(t)$ , que denota la probabilidad de que el individuo sobreviva durante los próximos  $t$  años. La expresión

$$s(t) = 1 - G(t)$$

relaciona la función de supervivencia con la función que representa la probabilidad de que individuo muera dentro de los  $t$  años siguientes,  $G(t)$ . Por tanto, podemos relacionar las funciones de supervivencia o de mortalidad, con el tanto de fallecimientos o con el tanto de supervivientes en cierta edad  $x$  con la expresión

$${}_tq_x = G(t) = 1 - s(t),$$

o con

$${}_tp_x = s(t) = 1 - G(t).$$

Tanto las tablas de vida como la función de supervivencia son el resultado de cálculos estadísticos que parten de las observaciones reales de mortalidad y que estima, en primer lugar, la tasa central de mortalidad  $m_x$  y, en segundo lugar, una probabilidad anual de muerte  $q_x$ . Los cálculos de ambas funciones se basan en observaciones periódicas, es decir, sobre la frecuencia de muertes observadas durante varios años a lo largo de un periodo dado, que por lo general es de un año. Dado que el año de fallecimiento es un factor que afecta la modelización de la mortalidad debe agregarse el término  $t$  a cada una de los indicadores para representar el año calendario en estudio.

## 2.2. Tablas de mortalidad del HMD y LAHMD

Los datos de mortalidad a nivel mundial los podemos encontrar en *The Human Mortality Database*. Proyecto que inició en el año 2000 con la colaboración del Departamento de Demografía de la Universidad de California, Berkeley (EE.UU.), el Instituto Max Planck de Investigación Demográfica (MPIDR) en Rostock (Alemania), el Instituto Francés de Estudios Demográficos (INED) y el Instituto Nacional sobre el Envejecimiento (EE.UU.). Esta base de acceso abierto proporciona datos detallados de mortalidad y población de 38 países además de promover información proveniente de investigadores, estudiantes, periodistas, analistas políticos y otras personas interesadas en la historia de la longevidad humana.

A nivel Latinoamérica tenemos la web Latin American Human Mortality Database (Urdinola y Queiroz, 2015) que es una plataforma similar a la anterior, 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. La web Latin American Human Mortality Database muestra los datos con los periodos disponibles para cada país razón por la cual estos tienen un formato o información diferentes para cada país.

En el caso de Colombia, las tablas están organizadas en 18 grupos de edades que van desde 0 hasta 84 años, estando el resto de edades superiores a 84 agrupadas en una única categoría (85+). Las defunciones se publican para el periodo 1970–2010. Los datos referentes a la población corresponden a los periodos de 1973, 1985, 1993 y 2005 por ser los únicos años con censos disponibles. Los grupos de edades en estudio para este trabajo serán de 0 hasta 84 años; los datos agrupados en 85+ están altamente influenciados por fluctuaciones aleatorias, razón por la que se le debe aplicar un tratamiento distinto que al resto de edades.

Dado que los datos de población disponibles son pocos, construimos una base de datos para la población interpolada, utilizando la expresión propuesta por Delwarde y Denuit (2003) para calcular la población en los años donde no hay registros de censo (de 1974 a 1984, 1986 a 1992 y de 1994 a 2004). La expresión utilizada es la siguiente:

$$P_{x,t} = P_{x,t_1} + \left( \frac{t-t_1}{t_2-t_1} \right) (P_{x,t_2} - P_{x,t_1}), \quad (2.1)$$

donde  $t_1$  y  $t_2$  son los dos periodos censales consecutivos,  $t$  es el año para el dato a estimar,  $x$  es la edad y  $P$  es la población. Dicha interpolación nos permite estimar las tasas de mortalidad para cualquier periodo  $t$  comprendido entre 1973 y 2005.

### 3. AJUSTE Y PREDICCIÓN DE LA 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 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 son los más utilizados en la actualidad por la propiedad de modelar la mortalidad a través del tiempo.

El modelo Lee-Carter ha tenido una gran aceptación en el ámbito de las ciencias actuariales desde que llegó a explicar el 93% de la variación de los datos de mortalidad de US, 1933-1987. Desde entonces, se ha convertido en el modelo líder en la literatura demográfica. En Lee & Carter (1992) usaron los datos de mortalidad clasificados por edad y año de muerte, luego modelaron la fuerza de mortalidad en términos de estas dos variables; las predicciones fueron obtenidas utilizando el año de muerte como una serie temporal y haciendo predicciones de las estimaciones de estos parámetros.

#### 3.1. 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

$$m_{xt} = \exp(a_x + b_x k_t + \epsilon_{xt}), \quad (3.1)$$

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.

$\epsilon_{xt}$  representa el error aleatorio, con media 0 y varianza  $\sigma_\epsilon^2$ , que no se explica por el modelo.

El modelo se puede expresar de forma equivalente a (3.1) con la expresión

$$\ln(m_{xt}) = a_x + b_x k_t + \epsilon_{xt}, \quad (3.2)$$

Las expresiones (3.1) y (3.2) son versiones reducidas del modelo de Lee-Carter. Una forma más general se expresaría en función de la probabilidad de muerte  $q_{xt}$ :

$$\ln(q_{xt}) = a_x + \sum_{i=1}^r b_x^i k_t^i + \epsilon_{xt} \quad (3.3)$$

donde  $r$  es el rango de la matriz  $\ln(q_{xt}) - a_x$ . En este trabajo se modelizará el  $\text{logit}(q_{xt})$ , como propuesto por Debón et al. (2008)

$$\text{logit}(q_{xt}) = \ln\left(\frac{q_{xt}}{1-q_{xt}}\right) = a_x + \sum_{i=1}^r b_x^i k_t^i + \epsilon_{xt} \quad (3.4)$$

## 3.2. Ajuste del modelo

### 3.2.1. Lee-Carter-SVD

El modelo de Lee-Carter no puede ser ajustado por las técnicas habituales de regresión puesto que los valores del índice  $k_t$  no son observables. La estimación de los parámetros en (3.4) se puede llevar a cabo mediante la descomposición en valores singulares (SVD) de la matriz  $\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) - \hat{a}_x$ .

De esta forma tenemos que

- a) La estimación para  $a_x$  es

$$\hat{a}_x = \frac{\sum_t \ln\left(\frac{q_{xt}}{1-q_{xt}}\right)}{T},$$

con  $T$  número de años.

- b) Los valores de  $b_x$  y  $k_t$  son estimados por el primer término de SVD,

$$\ln\left(\frac{q_{xt}}{1-q_{xt}}\right) - \check{\alpha}_x = \sum_{i=1}^{\min(n,T)} s^i u_x^i v_x^i$$

donde  $n$  es el número de años,  $s^i$  los valores singulares y  $u_x^i$  y  $v_x^i$  los vectores singulares.

Tenemos entonces que  $\hat{k}_t^{SVD} = s^1 v_t^1$  y  $\hat{b}_x^{SVD} = u_x^1$ . Como  $\sum_x \hat{b}_x^{SVD} = c \neq 1$ ,

normalizamos con la transformación  $\hat{b}_x = \frac{\hat{b}_x^{SVD}}{c}$  y  $\hat{k}_t = c \hat{k}_t^{SVD}$ .

- c) Los valores estimados en b) pueden mostrar diferencias entre las muertes observadas y los predichas causadas por el uso del logit en las tasas de mortalidad. Para resolver este problema,  $k_t$  se reestima usando  $\hat{k}_t$  como un valor inicial para las medias de la ecuación

$$D_t = \sum_x \left( E_{xt} \frac{\exp(\check{\alpha}_x + \sum k_t \hat{b}_x)}{1 + \exp(\check{\alpha}_x + \sum k_t \hat{b}_x)} \right) \quad (3.5)$$

donde  $E_{xt}$  es el número inicial de expuestos al riesgo con edad  $x$  en el año  $t$  y  $D_t = \sum_x D_{xt}$  es el total de muertes en el año  $t$ , con  $D_{xt}$  el número de muertes con edad  $x$  en el año  $t$ . La solución de la ecuación anterior (3.5) es  $\tilde{k}_t$ . La solución final requiere la sustitución

$$\hat{k}_t = \tilde{k}_t - \text{means}(\hat{k}_t) \text{ y } \hat{\alpha}_x = \check{\alpha}_x + \hat{b}_x \text{means}(\tilde{k}_t).$$

### 3.2.2. Lee-Carter-Máxima verosimilitud

El método de máxima verosimilitud para estimación de la versión log-bilineal Poisson del modelo Lee-Carter, está desarrollado en Wilmoth (1993) y complementado por Renshaw & Haberman (2009). Como  $D_{xt} \sim \text{Poisson}(L_{xt} m_{xt})$  entonces la función log-verosimilitud basada en los datos  $d_{xt}$  (muertes) y  $E_{xt}$  (expuestos al riesgo) haciendo  $\lambda_{xt} = E_{xt} m_{xt}$  está dada por

$$l(d, \lambda) = \sum_{x,t} d_{xt} \ln(\lambda_{xt}) - \lambda_{xt} - \ln(d_{xt}) \quad (3.6)$$

Se busca encontrar los valores en la matriz  $\lambda = [\lambda_{xt}]$  que maximicen la expresión anterior.

Como se cumple

$$\lambda_{xt} = E_{xt} m_{xt} = E_{xt} e^{a_x + b_x k_t}$$

entonces maximizar (3.6) con respecto  $\lambda_{xt}$  equivale a maximizar

$$l(a, bk) = \sum_{x,t} d_{xt} \ln(E_{xt} e^{a_x + b_x k_t}) - E_{xt} e^{a_x + b_x k_t} \quad (3.7)$$

El estimador de máxima verosimilitud en el contexto de modelos lineales generalizados se obtiene minimizando la función deviance. Si  $d_{xt}$  es el conteo de muertes observadas y  $\hat{d}_{xt} = E_{xt} e^{\hat{a}_x + \hat{b}_x \hat{k}_t}$ , la función deviance se define como

$$D(d_{xt}, \hat{d}_{xt}) = 2 \sum_{x,t} \omega_{xt} \left( d_{xt} \log \left( \frac{d_{xt}}{\hat{d}_{xt}} \right) - (d_{xt} - \hat{d}_{xt}) \right), \quad (3.8)$$

donde  $\omega_{xt} = 0$  si  $d_{xt} = 0$  y  $\omega_{xt} = 1$  si  $d_{xt} > 0$ . La minimización de la función de deviance se hace mediante el método de Newton-Raphson unidimensional.

### 3.3. Predicción con el modelo de Lee-Carter

Luego de ajustar el modelo de Lee-Carter, el siguiente paso es ajustar una serie temporal a los valores de los índices de mortalidad,  $\{\hat{k}_t\}$ , obteniéndose buenos resultados en muchas de las aplicaciones con el modelo

$$\hat{k}_t = p + \hat{k}_{t-1} + u_t$$

con un  $p$  constante y  $u_t$  un ruido blanco.

La predicción para los años  $t_n + s$ ,  $s = 1, 2, \dots$ , posteriores al último  $t_n$ , se realiza sustituyendo en el modelo de Lee-Carter la predicción  $\hat{k}_{t_n+s}$  obtenida a partir de la serie temporal ajustada,

$$\ln \left( \frac{\hat{q}_{x,t_n+s}}{1 - \hat{q}_{x,t_n+s}} \right) = a_x + \hat{b}_x \hat{k}_{t_n+s}, \quad s > 0.$$

La crítica fundamental al modelo de Lee-Carter es que los parámetros  $a_x$  y  $b_x$  dependen solo de la edad y que la predicción de futuros valores de la mortalidad se basa únicamente en  $k_t$ , lo que supone aceptar que no hay interacción entre la edad y el tiempo

(cronológico). Por otro lado, es importante mencionar la ventaja de la fácil interpretación de sus parámetros y su parsimonia, como mencionan Booth, Maindonald y Smith (2002). El modelo de Lee-Carter actualmente presenta gran popularidad debido a sus buenos resultados y a su simplicidad, por lo que podemos encontrar bastante literatura con sus aplicaciones y mejoras.

### **3.4. Estimación del modelo de Lee-Carter en R- librería *gnm***

Los modelos no lineales generalizados (GNM) son una extensión de los GLM. Estos modelos se utilizan en el caso de que  $m_x$  no sea lineal, es decir que algunos de los términos del predictor sean no lineales. En el programa estadístico R, la librería *gnm* creada por Turner & Firth (2015) provee una plataforma flexible para la especificación y estimación de modelos generalizados con términos no lineales. La librería facilita la especificación del modelo incluyendo funciones como Exp, Inv y Mult, utilizadas para especificar expresiones simbólicas en la fórmula del modelo (exponente, inversa y producto de predictores, respectivamente).

La librería *gnm* es la que utilizamos en el presente trabajo para analizar el modelo Lee-Carter (1992) aplicado a datos colombianos. Puede consultarse una descripción detallada de los modelos lineales y no-lineales generalizados, así como su implementación en Currie (2014).

## **4. ANÁLISIS DE LA MORTALIDAD EN COLOMBIA**

### **4.1. Análisis descriptivo**

Los Gráficos 1 y 2 muestran la probabilidad de muerte  $q_{xt}$  en cada grupo de edad  $x$  y año  $t$  para cada uno de los sexos en las edades infantiles: 0-1, 1-4, 5-9 años y para las edades mayores: 65-69, 70-74, 75-79 y 80-84 años, respectivamente. Ambos gráficos exponen que la mortalidad de Colombia tiene un comportamiento similar a la tendencia global: las edades infantiles tienen una tasa de mortalidad alta que decrece rápidamente

después del primer año de vida, luego se estabiliza y después va aumentando paulatinamente hasta envejecer.

En el Gráfico 1 podemos ver que en el primer año de vida la probabilidad de muerte es muy elevada, pero a medida que avanza la edad, esta probabilidad decrece abruptamente, tanto para hombres como para mujeres. Esta probabilidad a través de los años analizados ha ido disminuyendo, lo que resulta una mejora en la esperanza de vida de la población.

El Gráfico 2 muestra que en las edades avanzadas la probabilidad de muerte se aumenta a medida que avanza la edad, tanto para hombres como para mujeres. Estas probabilidades son menores para las mujeres en cualquier edad.

Gráfico 1. Probabilidad de muerte para edades tempranas.

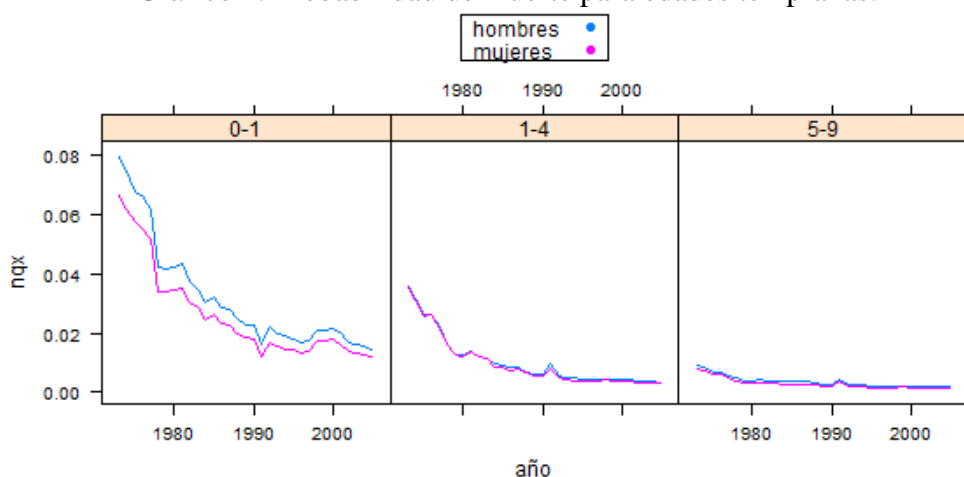
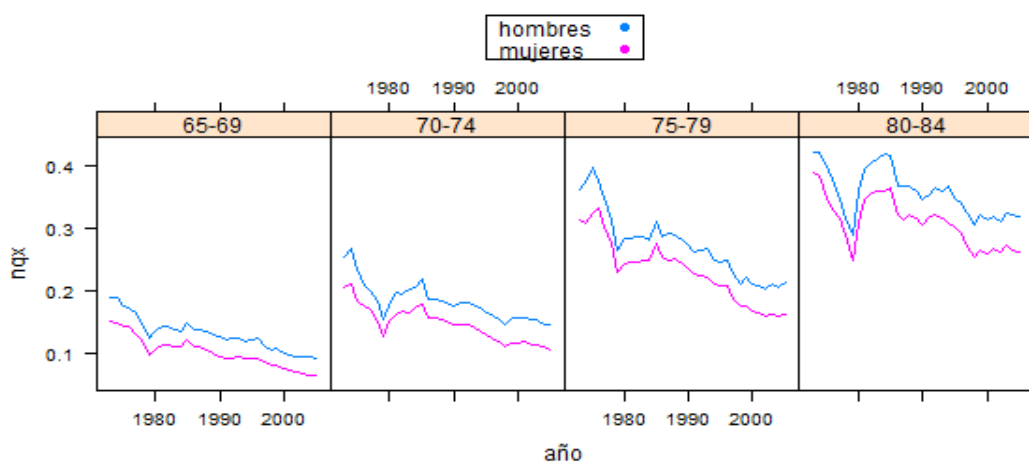


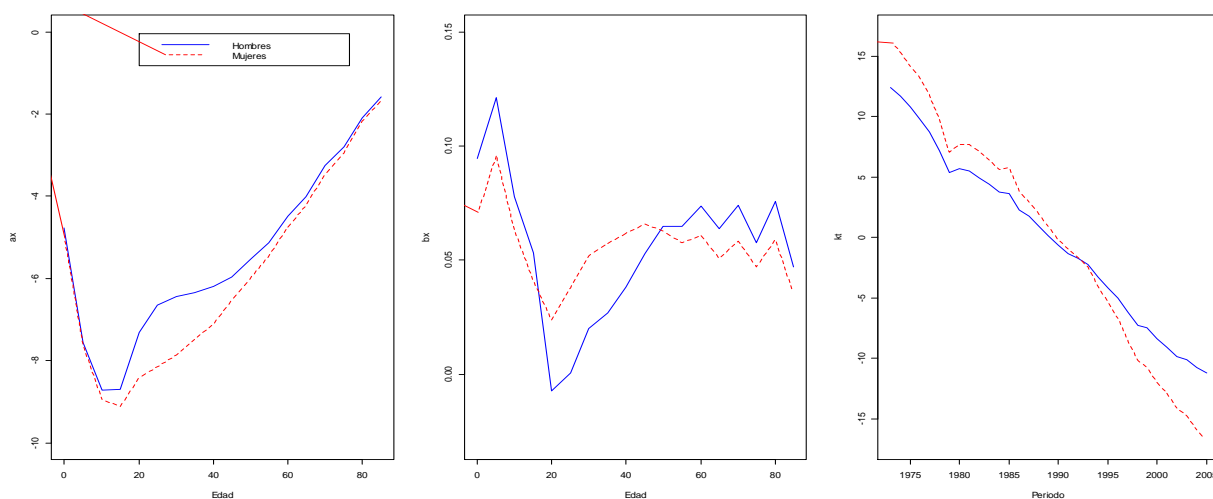
Gráfico 2. Probabilidad de muerte para edades adultas



#### 4.2. Ajuste del modelo de Lee-Carter para Colombia

Sobre las tablas de mortalidad de la población colombiana durante el periodo 1973-2005, se ajustan los parámetros del modelo propuesto por Lee-Carter (1992), con la modificación de la variable respuesta propuesta por Debón et al (2008). Las estimaciones obtenidas para el conjunto de parámetros  $(a_x, b_x, k_t)$  para hombres y mujeres se presentan de manera simultánea en el Gráfico 3.

Gráfico 3. Parámetros estimados para el modelo Lee-Carter.



Desde el comportamiento de los coeficientes  $a_x$  se observan las fases habituales de la mortalidad poblacional. Específicamente se observa que durante el primer quinquenio de edad (0-5 años), el riesgo de mortalidad desciende rápidamente, sin que existan mayores diferencias entre hombres y mujeres. A partir del 5to año de edad, el riesgo de mortalidad de las mujeres empieza a notarse como inferior al de los hombres, siendo amplia esta diferencia entre los 15 - 50 años de edad. A partir de los 50 años de edad el riesgo de muerte de ambas poblaciones tiende a ser nuevamente similar.

Adicionalmente se observa la forma de “joroba” en el riesgo de mortalidad para las edades comprendidas entre los 15- 25 años de edad, la cual se presenta tanto para hombres como para mujeres, sin embargo, es mucho más alta para la población de hombres, lo cual puede ser un reflejo de su condición de exposición al riesgo, común en este rango de edad, asociada a un mayor nivel de riesgo de mortalidad.

El parámetro  $b_x$  es un parámetro de sensibilidad que expresa como responde la tasa de mortalidad a cambios en  $k_t$ . Los parámetros  $b_x$  indican la velocidad con la que varía la mortalidad cuando ha habido cambios en  $k_t$ . Por ende, cuando  $b_x$  toma valores pequeños quiere decir que la mortalidad no varía mucho en ese intervalo de edad  $x$ , y cuando toma valores altos, entonces la mortalidad tiene una alta variabilidad en el intervalo  $x$ , como es el caso de las edades entre 10 y 20 años y luego para edades entre 65y 75 años.

La dinámica de 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 similarmente a las subpoblaciones de hombres y mujeres durante los años 1973 - 2005.

### **4.3. Bondad de ajuste**

Una de las formas para evaluar el nivel de bondad de ajuste del modelo es el análisis de los residuos deviance que mide la desviación del modelo ajustado. Usualmente en los

GLM para medir el ajuste entre el predictor lineal y función enlace se analiza el comportamiento de los residuos deviance, en nuestro caso lo haremos gráficamente. En el análisis de todas las gráficas utilizamos el intervalo  $[-2,2]$  como referencia de un buen ajuste del modelo.

Los Gráficos 4, 5 y 6 presentan el comportamiento de los residuos deviance para los hombres en función de la edad, el año y la cohorte respectivamente, y los Gráficos 7, 8 y 9 presentan el comportamiento de los residuos deviance en función de la edad, el año, pero para las mujeres. Los intervalos donde se distribuyen los residuos deviance frente al periodo, edad y cohorte son más amplios para los hombres ya que en los tres casos se sitúan entre el intervalo  $[-20,20]$ , (ver Gráficos 4, 5 y 6), mientras que, para las mujeres, los residuos deviance frente a la edad, periodo y cohorte se distribuyen en el intervalo más pequeño (Gráficas 7, 8 y 9).

Gráfico 4. Residuos deviance en función de la edad para los hombres.

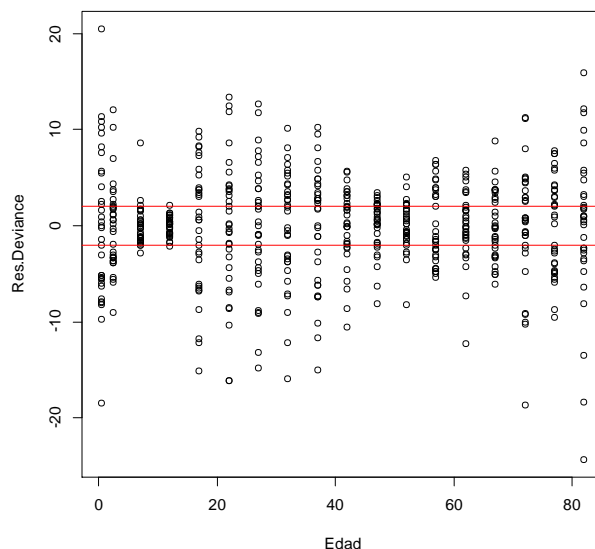




Gráfico 5. Residuos deviance en función del periodo para los hombres.

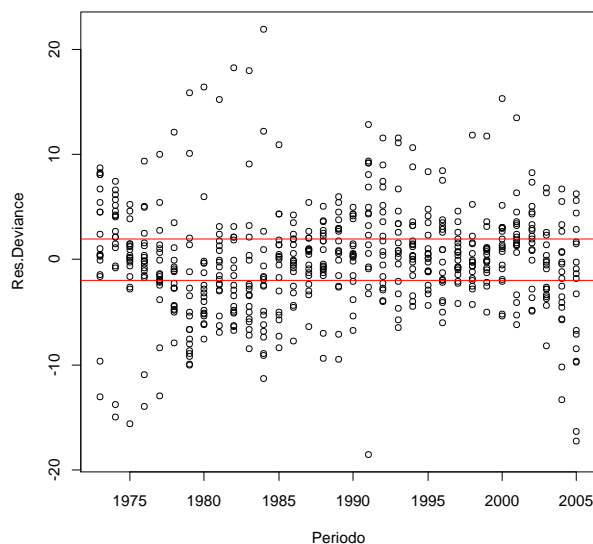
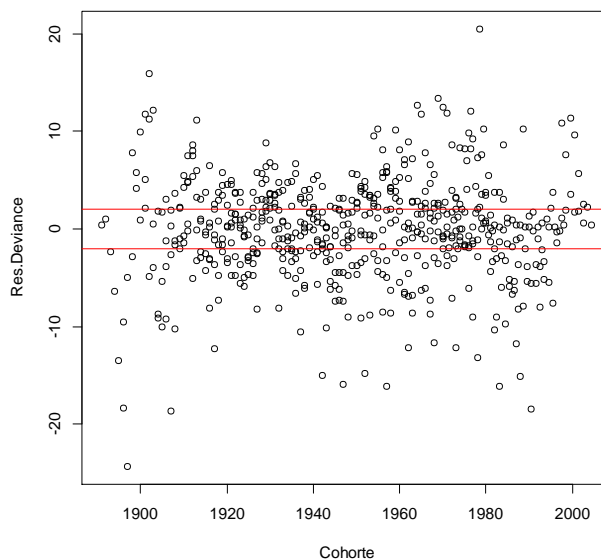


Gráfico 6. Residuos Deviance en función de la cohorte para los hombres.



En general, los residuos del modelo son bastante irregulares puesto que los intervalos donde se estabilizan los residuos deviance de ambos sexos son más amplios que el intervalo de referencia  $[-2,2]$ . Se observa que los residuales presentan franjas con ampliación de variabilidad y residuos muy elevados de manera general, siendo mayores en los hombres con respecto a las mujeres.

Gráfico 7. Residuos deviance en función de la Edad para las mujeres.

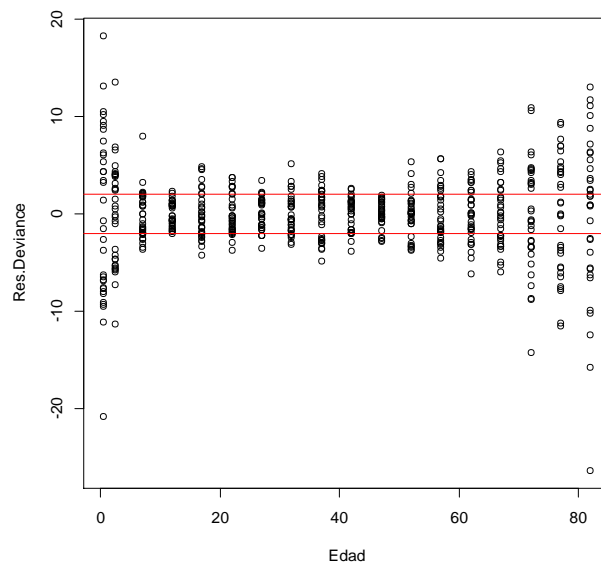


Gráfico 8. Residuos deviance en función del Periodo para las mujeres.

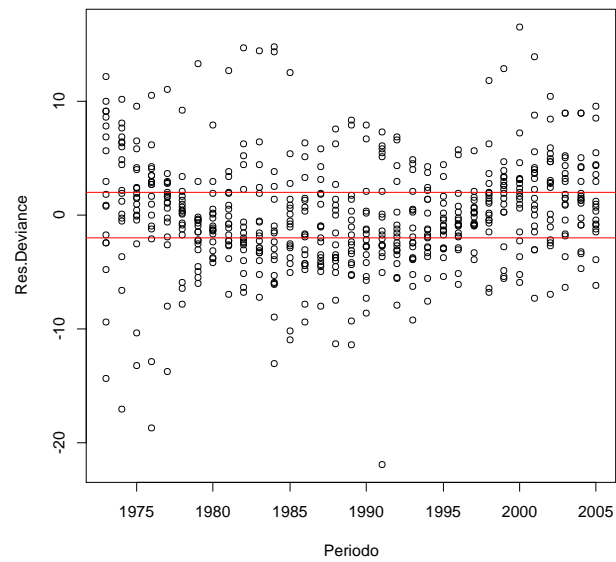
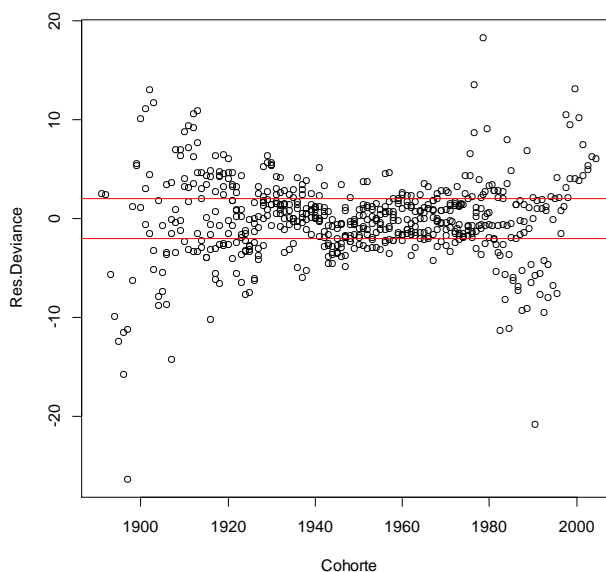


Gráfico 9. Residuos deviance en función de la Cohorte para las mujeres.



Las predicciones a futuro para los índices  $k_t$  del modelo Lee Carter se realizaron mediante el ajuste de la serie temporal observada en sus respectivas estimaciones, ajustando un modelo ARIMA para cada proceso temporal y empleando la ecuación de pronóstico respectiva para cada caso como un proyector de futuros valores de estos índices. En su implementación se utilizaron las funciones `auto.arima` y `forecast`, de la librería `forecast` (Hyndman, 2008) de R, que proporcionan de manera automática el ajuste del correspondiente modelo ARIMA y la predicción de futuros valores.

En los Gráficos 10 y 11 se proyectan las predicciones obtenidas para el índice  $k_t$  para el periodo 2006 – 2025 para los hombres y las mujeres respectivamente, así como los intervalos de confianza asociados. Como puede observarse en ambos casos, hombres y mujeres, obtenemos una tendencia de reducción de la mortalidad.

Gráfico 10. Proyecciones del índice  $k_t$  del modelo de Lee- Carter, hombres.

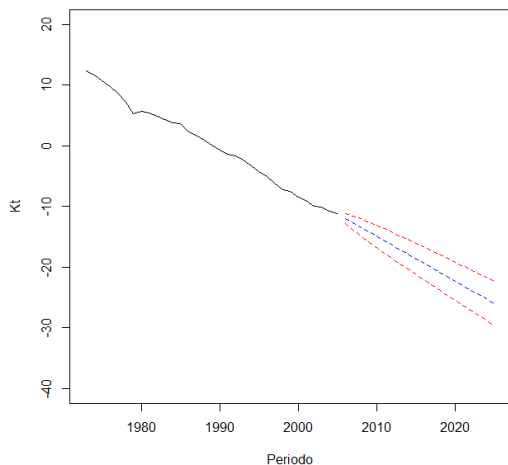
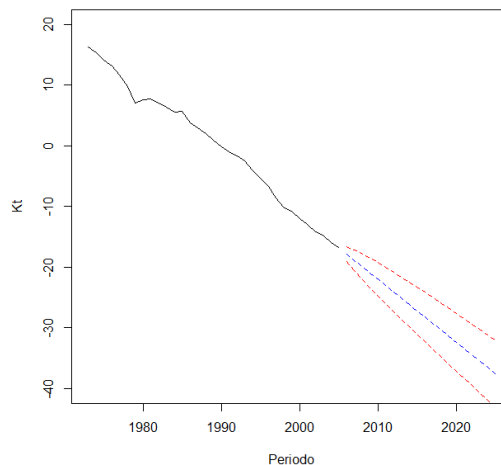


Gráfico 11. Proyecciones del índice  $k_t$  del modelo de Lee- Carter, mujeres.



## 5. CONCLUSIONES

A partir del análisis planteado por este modelo, obtuvimos parámetros tanto para los hombres como para las mujeres permitieron llegar a las siguientes conclusiones:

- El parámetro  $a_x$  nos indica que la tendencia de la mortalidad colombiana es similar a la tendencia de mortalidad global: es elevada en las edades infantiles (0-5 años), a partir del 5to año de edad, el riesgo de mortalidad de las mujeres empieza a notarse como inferior al de los hombres, siendo más amplia esta diferencia entre los 15 - 50 años de edad. A partir de los 50 años de edad el riesgo de muerte de ambas poblaciones tiende a ser nuevamente similar incrementándose conforme envejece el individuo.
- El parámetro  $b_x$  muestra que la mortalidad presenta una alta variabilidad para las edades entre 10 y 20 años y luego para edades entre 65y 75 años, tanto para hombres como para mujeres.
- El parámetro  $k(t)$  señala que la mortalidad tanto de los hombres como de las mujeres ha decrecido a través del tiempo, aunque el decrecimiento ha sido mayor en el caso de las mujeres

Las estimaciones obtenidas para el modelo y el análisis de la dinámica temporal, permiten deducir que la población colombiana se ha visto inmersa en un fenómeno de mejoramiento gradual de sus condiciones de vida, según lo cual su esperanza de vida se ha incrementado en el tiempo. Desde las posibilidades de análisis que generan los parámetros del modelo Lee-Carter, se deduce la potencia de esta herramienta para la modelación y evaluación de la mortalidad, así como realizar análisis predictivos.

## **6. REFERENCIAS BIBLIOGRÁFICAS**

- ACOSTA K. y ROMERO J. (2014) “Cambios recientes en las principales causas de mortalidad en Colombia”. Banco de La República, Colombia.
- BOOTH, H., MAINDONALD, J., y SMITH, L. (2002). “Applying Lee-Carter under conditions of variable mortality decline”. *Population Studies*, 56, 3, pp. 325–336.
- CURRIE, I. (2014). “On fitting generalized linear and nonlinear models of mortality”. *Scandinavian Actuarial Journal*, 14, pp. 1-28.
- DEBÓN, A., MONTES, F., y SALA, R. (2009). “Tablas de mortalidad dinámicas. Una aplicación a la hipoteca inversa en España”. 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.
- HORIUCHI, S. (1999). “Epidemiological transition in human history”. *Health and mortality issues of global concern*, Nueva York. pp.54-71.
- HYNDMAN, R. J. (2008). “*forecast*: Forecasting functions for time series”. *R package version 1.11*.

- LEE, R. (2000). “The Lee-Carter method for forecasting mortality, with various extensions and applications”. *North American Actuarial Journal*, 4, 1. pp.80–91.
- LEE, R. y CARTER, L. R. (1992). “Modelling and Forecasting U.S. Mortality”. *Journal of the American Statistical Association*, 87, pp. 659-671.
- PALACIOS, H. (1996). *Introducción al Cálculo Actuarial*. Mapfre, Madrid.
- RENSHAW, A. E. Y HABERMAN, S. (2009). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance Mathematic and Economics*.
- TURNER, H. y FIRTH, D. (2015). “Generalized nonlinear models in r: An overview of the *gnm* package”. v1.0-8. <http://CRAN.R-project.org/package=gnm>.
- URDINOLA B.P. y QUEIROZ B.L. (2015). *Latin American Human Mortality Database*. <http://www.lamortalidad.org>. [Consultada el 17/10/2015].
- WILMOTH, J. (1993). Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. Technical Report. Department of Demography, University of California, Berkeley.