



UNIVERSIDAD AUTÓNOMA METROPOLITANA – IZTAPALAPA

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA

**MODELACIÓN MATEMÁTICA DE LA FUERZA DE
MORTALIDAD Y SU APLICACIÓN A FINANZAS**

Tesis que presenta:

Michelle Carreón Miranda

Para obtener el grado de:

Maestra en Ciencias

(Matemáticas Aplicadas e Industriales)

Asesora: Dra. Patricia Saavedra Barrera

Jurado calificador:

Presidente: Dra. María Asunción Begoña Fernández Fernández

Secretaria: Dra. Patricia Saavedra Barrera

Vocal: Dr. Gabriel Núñez Antonio

México, Ciudad de México, Junio 2016.



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA
Unidad Iztapalapa

Fecha : 27/06/2016

Página : 1/1

CONSTANCIA DE PRESENTACION DE EXAMEN DE GRADO

La Universidad Autónoma Metropolitana extiende la presente CONSTANCIA DE PRESENTACION DE EXAMEN DE GRADO de MAESTRA EN CIENCIAS (MATEMÁTICAS APLICADAS E INDUSTRIALES) de la alumna MICHELLE CARREON MIRANDA, matrícula 2133802714, quien cumplió con los 147 créditos correspondientes a las unidades de enseñanza aprendizaje del plan de estudio. Con fecha veintisiete de junio del 2016 presentó la DEFENSA de su EXAMEN DE GRADO cuya denominación es:

MODELACION MATEMATICAS DE LA FUERZA DE MORTALIDAD Y SU APLICACION A FINANZAS

Cabe mencionar que la aprobación tiene un valor de 60 créditos y el programa consta de 207 créditos.

El jurado del examen ha tenido a bien otorgarle la calificación de:

APROBADA

JURADO

Presidenta

Secretaria

DRA. MARIA ASUNCION BEGOÑA FERNANDEZ
FERNANDEZ

DRA. PATRICIA SAAVEDRA BARRERA

Vocal

DR. GABRIEL NUÑEZ ANTONIO

Agradecimientos

El presente trabajo representa el termino de una etapa más en mi vida profesional, en el cual participaron varias personas a las que me gustaría agradecer.

Empezaré agradeciendo a mi tutora la Dra. Patricia Saavedra Barrera, por la confianza depositada en mi persona para la elaboración de este trabajo, su compromiso y motivación. A la Dra. Begoña Fernández y el Dr. Gabriel Núñez, por aceptar ser sinodales de este trabajo.

A mis padres Naty y Felipe por su amor y apoyo incondicional. A mis hermanos Kenia y Miguel por su cariño y disposición para hacer divertidos mis fines de semana.

A mis amigos Alejandra, Carlos, Isabel y Belem por brindarme su amistad y ser un apoyo emocional durante el proceso de investigación y escritura de esta tesis.

Al Consejo Nacional de Ciencia y Tecnología por su apoyo económico.

¡A todos ustedes GRACIAS!

Introducción

Como se sabe los sistemas actuales de pensiones en México y en el mundo no son sostenibles. Representan un pasivo financiero enorme para los gobiernos y el problema no deja de crecer debido al incremento de la esperanza de vida. Por primera vez, en el actual sexenio, se ha planteado estructurar un plan de pensiones universales lo cual puede agravar aún más el problema. Por lo tanto, es importante estudiar y modelar la evolución de la mortalidad en México.

En general, para las compañías de seguros el problema consiste en construir una reserva financiera asociada a los seguros de vida o a las pensiones que garantice su solvencia. De hecho, nuevas reglamentaciones de la Comisión Nacional de Seguros y Finanzas exige hoy en día a las empresas de seguros requerimientos de capital mucho mayores que en el pasado. Estos deben tomar en cuenta el incremento en la esperanza de vida. Cambios sensibles en la esperanza de vida pueden significar para una compañía de seguros pérdidas millonarias en sus productos tanto de pensiones como en seguros de vida. El riesgo al que se expone una compañía de seguros cuando la esperanza de vida se incrementa o reduce debido a avances médicos, pandemias, entre otros, se denomina riesgo de longevidad.

Una posible solución a los problemas de cobertura de la mortalidad es la titularización: derivados financieros cuyo subyacente sean los índices de mortalidad o supervivencia a futuro. Los primeros derivados de este tipo fueron los bonos de longevidad propuestos en 2001 por

Blake y Barrows. Los pagos, tanto de cupones como la devolución del principal, se condicionó a la supervivencia de un cohorte poblacional de referencia a lo largo de un periodo de tiempo determinado. En diciembre del 2003, Swiss Re emitió el primer bono con estas características, cuyo pago principal estaba relacionado con los escenarios adversos que pueden derivarse del riesgo de mortalidad. El vencimiento de la emisión era de 4 años y los cupones eran variables, pagados a los inversionistas del LIBOR más 135 puntos básicos. Este tipo de cupón resulta elevado si se compara con los cupones pagados por bonos con la misma calificación crediticia, pero compensa el hecho de que la devolución del principal se condicionó a que la media ponderada de la mortalidad general de una población extraída de cinco países de referencia (EE.UU., Reino Unido, Francia, Italia y Suiza) no excediera el 130 % del nivel de mortalidad base de 2002. Además de los bonos, existen otras formas posibles de derivados de supervivencia, por ejemplo, los *swaps* de supervivencia descritos en el trabajo de Pérez F. [16]. Si existiera un mercado de bonos de longevidad, entonces las aseguradoras tendrían la posibilidad de cubrirse del riesgo de longevidad.

Antes de definir el objetivo de la tesis cabe mencionar que actualmente no existe un mercado de bonos de longevidad con la profundidad necesaria para considerar que el mercado es completo. También es importante señalar que hasta ahora, los instrumentos que se han propuesto son a corto plazo debido a los desafíos que se presentan en la estimación de instrumentos a largo plazo que involucran la estimación tanto de la curva de estructura de tasas de interés como la de mortalidad. Es por ello que decidimos en este trabajo abordar únicamente la valuación de instrumentos a corto plazo pues implica la estimación tanto de las tasas de interés como las tasas de mortalidad *spot*.

En el presente trabajo, se valuó un instrumento financiero: un bono cupón cero de longevidad para un género y un cohorte de edad específico. Cabe señalar que a diferencia de las compañías de seguros que cuentan con una amplia base de datos que les permite hacer inferencias y fijar precios de manera específica a cada uno de sus productos, se utilizó la base

de datos de la CONAPO del periodo de 1990 a 2009 para estimar el riesgo de longevidad.

Objetivo de la tesis:

Estudiar los modelos más usados para describir la evolución del índice de supervivencia y su aplicación en la valuación de instrumentos financieros que puedan servir para cubrirse del riesgo de longevidad a corto plazo.

El contenido de esta tesis es el siguiente: en el capítulo 1 se hace una presentación de los conceptos básicos que se han estudiado desde el punto de vista actuarial para definir la probabilidad de supervivencia en términos de la llamada fuerza de mortalidad. Este último concepto será de gran importancia para el desarrollo de este trabajo. También en este capítulo se estudian los modelos de tasa corta de interés. En el capítulo 2, se hace una revisión en la literatura de dos de los modelos matemáticos más utilizados para modelar la fuerza de mortalidad. El primero es el modelo Gompertz-Makeham, el cual es un modelo paramétrico que solo depende de la edad de un individuo. El segundo es el modelo Lee-Carter que tiene como característica principal el proporcionar tablas de vida dinámicas en el tiempo. A diferencia del Modelo Gompertz-Makeham, este modelo considera como parámetros tanto la edad de un individuo como el tiempo en el calendario. Se implementó la metodología propuesta para ambos modelos y se obtuvo estimaciones para la probabilidad de supervivencia de la población mexicana.

Las propuestas más recientes de modelar la fuerza de mortalidad se basan en suponer que su dinámica se describe por un proceso estocástico. Esto sustentado en el hecho de que la mortalidad de las personas depende de eventos aleatorios tales como pandemias, guerras civiles o desastres naturales. Por ello, en el capítulo 2 también se estudia un marco teórico para tratar a la fuerza de mortalidad como un proceso estocástico.

En el capítulo 3, se estudiará la metodología a seguir para valorar un bono de longevidad a corto plazo, además se estima un tercer modelo para la fuerza de mortalidad cuando ésta

sigue un modelo estocástico tipo CIR. Cabe mencionar que la estimación de la fuerza de mortalidad para los tres modelos no tiene como objetivo obtener un modelo que describa la dinámica de la fuerza de mortalidad para fines demográficos, tampoco proyectar la futura evolución de ésta; solo será utilizada con el fin de valorar un bono de longevidad para un género y un cohorte generacional específico.

En el capítulo 4, se calcula numéricamente la probabilidad de supervivencia a diferentes edades para cada uno de los tres modelos. También se calculó el precio de un contrato actuarial bajo el marco de los tres modelos.

En el capítulo 5, se presentan las conclusiones de este trabajo y las futuras investigaciones que de éste se pueden derivar. Por último, se citan las referencias bibliográficas que sustentan el desarrollo de este trabajo.

Índice general

Agradecimientos	3
Introducción	5
1. Antecedentes	11
1.1. Mortalidad	11
1.1.1. El tiempo hasta la muerte de una persona de edad x	13
1.1.2. La fuerza de mortalidad	16
1.1.3. Tablas de vida y su relación con la función de supervivencia	21
1.2. Modelos de tasas de interés	24
1.2.1. Bonos cupón cero	24
1.2.2. Modelo de Vasicek	26
1.2.3. Modelo Cox, Ingersoll y Ross (CIR)	28
1.2.4. Estimación de parámetros para modelos de tasa de interés	29
1.2.4.1. Modelo Vasicek	30
1.2.4.2. Modelo CIR	32
2. Modelación de la fuerza de mortalidad	33
2.1. Modelo de Gompertz-Makeham para la fuerza de mortalidad	34

2.1.1.	Graduación de q_x para la población mexicana	39
2.2.	Modelo de Lee-Carter	51
2.2.1.	Modelo Lee-Carter para la población mexicana	56
2.3.	Modelos estocásticos para la fuerza de mortalidad	68
2.3.1.	Marco de modelación	69
2.3.2.	Tiempos de paro y el tiempo restante de vida de un individuo de edad x	70
2.3.3.	La fuerza de mortalidad y el índice de supervivencia	70
3.	Valuación de bonos de longevidad a corto plazo	75
3.1.	El índice medio de supervivencia y la fuerza de mortalidad <i>forward</i>	76
3.2.	Valuación de bonos cupón cero de longevidad a corto plazo	82
3.3.	Estimación de la fuerza de mortalidad estocástica para la población mexicana	84
4.	Resultados numéricos	89
5.	Conclusiones y perspectivas	99
	Bibliografía	101

Capítulo 1

Antecedentes

En este capítulo estudiaremos los conceptos de mortalidad y tasas de interés, que serán fundamentales para exponer en los próximos capítulos la fuerza de mortalidad estocástica y bonos de longevidad, que serán objeto de estudio en el presente trabajo. Para escribir este capítulo se siguieron las siguientes referencias: Bowers [1] y Carstens [3] para la parte actuarial y para las tasas de interés el libro *Financial markets in continuous time* de Rose Anne Dano y Monique Jeanblanc [7].

1.1. Mortalidad

Una compañía aseguradora se enfrenta a varios tipos de riesgo al vender sus contratos. En lo que a seguros de vida se refiere, uno de los principales factores de riesgo que es necesario considerar es la mortalidad. Si se conoce con cierta exactitud la futura evolución de ésta, se puede entonces calcular las primas y reservas que permitan a la aseguradora cubrirse de riesgos derivados de los cambios en la mortalidad. Por lo general, dicho cálculo se realiza al suponer que la mortalidad es una variable aleatoria estática, es decir, se considera que solo depende de la edad del asegurado y no del tiempo, sin embargo, en los últimos años, esta manera de calcular ha sido cuestionada, pues es un hecho que la esperanza de vida

ha aumentado a lo largo del tiempo y es erróneo suponer que la mortalidad se mantenga constante a lo largo del tiempo.

Los conceptos que definiremos a continuación, serán fundamentales para el desarrollo del presente trabajo, por tal motivo se definirán en términos actuariales según Carstens [3] las siguientes variables aleatorias.

Consideremos un espacio completo de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, y sea X una variable aleatoria continua definida sobre ese espacio a la cual llamaremos **edad hasta la muerte** de un recién nacido. Denotaremos como $F_X(x)$ a la función de distribución de X , es decir

$$F_X = \mathbb{P}(X \leq x), \quad x \geq 0.$$

Definición 1.1. La función $s: [0, \infty) \rightarrow [0, 1]$ definida por la fórmula

$$s(x) = 1 - F_X(x) = \mathbb{P}(X > x), \quad x \geq 0, \quad (1.1)$$

es llamada **función de supervivencia** y representa la probabilidad de que un recién nacido alcance la edad x , para $x \geq 0$.

Supondremos lo siguiente:

1. Asumiremos que $F_X(0) = 0$ y por lo tanto $s(0) = 1$.
2. La función de distribución es absolutamente continua. Si $f_X(x)$ denota la función de densidad de X , entonces:

$$f_X(x) = \frac{dF_X(x)}{dx} = F'_X(x) = -s'_X(x). \quad (1.2)$$

3. La función $F_X(x)$ es creciente, lo cual implica que la función $s(x)$ es decreciente para todo x .

La función de supervivencia y la función de distribución tienen una estrecha relación y ambas pueden definir la distribución de la variable X . Consideremos la probabilidad de que

un recién nacido muera entre la edades x y z ($x < z$), esto es:

$$\begin{aligned}\mathbb{P}(x < X \leq z) &= \mathbb{P}(X \leq z) - \mathbb{P}(X \leq x) \\ &= F_X(z) - F_X(x) \\ &= s(x) - s(z).\end{aligned}\tag{1.3}$$

Ahora, consideremos la probabilidad condicional de que un recién nacido muera durante el intervalo $(x, z]$ dado que ha sobrevivido hasta la edad x . En términos de la v.a. X y usando (1.1) y (1.3) tenemos

$$\mathbb{P}(x < X \leq z | X > x) = \frac{\mathbb{P}(x < X \leq z)}{\mathbb{P}(X > x)} = \frac{s(x) - s(z)}{s(x)},\tag{1.4}$$

siempre que $s(x) \neq 0$.

Desde el punto de vista estadístico o probabilista, la función de distribución es utilizada como un punto de partida para futuras investigaciones. Desde el punto de vista actuarial, este papel lo toma la función de supervivencia.

1.1.1. El tiempo hasta la muerte de una persona de edad x

En esta sección, consideremos a una población en la cual todos sus elementos tienen edad x . Seleccionemos aleatoriamente a uno de ellos y denotaremos a este individuo por el símbolo (x) . A (x) también se le conoce como “una vida de edad x ”. Definiremos la variable aleatoria continua τ_x como **el tiempo de supervivencia** de (x) .

Nota 1.1. La distribución de la v.a. τ_x es la misma que la distribución condicional del tiempo de vida de un recién nacido dado que éste ha alcanzado la edad x . En otras palabras, $\tau_x = X - x$ en distribución dado $X > x$. En comparación con la sección anterior, el subíndice x de la v.a. τ_x es un valor fijo que no debe confundirse con los valores que puede tomar la v.a. τ_x .

Definición 1.2. La probabilidad de que (x) muera dentro de t años se denota con el símbolo ${}_tq_x$ y se define como

$${}_tq_x = \mathbb{P}(\tau_x \leq t) \quad t \geq 0, \quad (1.5)$$

esto es, ${}_tq_x$ es la función de distribución de τ_x . La probabilidad de que (x) sobreviva t años se denota ${}_tp_x$ y se define como

$${}_tp_x = 1 - {}_tq_x = \mathbb{P}(\tau_x > t) \quad t \geq 0. \quad (1.6)$$

Como un caso especial, consideremos una vida de edad 0, es decir $\tau_0 = X$, entonces

$${}_xp_0 = s(x) \quad x \geq 0. \quad (1.7)$$

En otras palabras, la probabilidad de sobrevivir de un recién nacido, puede ser representada por su función de supervivencia hasta la edad x .

Definición 1.3. La probabilidad de que (x) muera entre las edades $x+t$ y $x+t+u$ se denota con el símbolo ${}_{t|u}q_x$ y se define como

$$\begin{aligned} {}_{t|u}q_x &= \mathbb{P}(t < \tau_x \leq t+u) \\ &= \mathbb{P}(\tau_x \leq t+u) - \mathbb{P}(\tau_x \leq t) \\ &= {}_{t+u}q_x - {}_tq_x \\ &= {}_tp_x - {}_{t+u}p_x, \end{aligned} \quad (1.8)$$

${}_{t|u}q_x$ también puede ser interpretado como la probabilidad de que (x) alcanzará la edad $x+t$ y morirá dentro de los siguientes u años.

Por notación, cuando $t = 1$ el prefijo en los símbolos definidos en (1.5) y (1.6) se omitirán. Lo mismo pasará cuando $u = 1$ en la ecuación (1.8).

Hasta este momento, hemos dado dos definiciones para la probabilidad de que (x) muera dentro del intervalo $(x, x+u]$. La primera, está dada por la expresión (1.4) haciendo $z = x+u$ y la segunda está dada por la expresión (1.8) con $t = 0$. Recuerde que la fórmula (1.4) representa

la probabilidad condicional de que un recién nacido muera dentro del intervalo $(x, x + u]$. Mientras que, el caso especial de la fórmula (1.8) con $t = 0$ representa la probabilidad de que una vida de edad x muera dentro del intervalo $(x, x + u]$. En el primer caso, la única información que conocemos sobre el recién nacido es que éste ha sobrevivido hasta la edad (x) . Por lo tanto, el cálculo de la probabilidad depende de la distribución condicional de los recién nacidos. Sin embargo, en el segundo caso, la observación de una vida de edad x podría incluir más información que simplemente la supervivencia. Por simplicidad, se asumirá que la información en cada uno de los casos es la misma y, por lo tanto, no habrá distinciones entre (1.4) y (1.8), es decir, siempre que $s(x) \neq 0$ se tiene:

$${}_t p_x = \frac{{}_{x+t} p_0}{{}_x p_0} = \frac{s(x+t)}{s(x)} \quad (1.9)$$

$${}_t q_x = 1 - {}_t p_x = \frac{s(x) - s(x+t)}{s(x)}. \quad (1.10)$$

Proposición 1.1

i) Sea $p_x = {}_1 p_x$, entonces para cualquier $t \geq 1$,

$${}_t p_x = p_x \cdot {}_{t-1} p_{x+1}.$$

ii) Sea $p_x = {}_1 p_x$, entonces para los enteros $t = 1, 2, \dots$, se cumple

$${}_t p_x = p_x \cdot p_{x+1} \cdot \dots \cdot p_{x+t-1}.$$

iii)

$${}_{t|u} q_x = {}_t p_x \cdot {}_u q_{x+t}.$$

Demostración

Ver Carstens [3].

1.1.2. La fuerza de mortalidad

Recuérdese que la expresión (1.4) define, en términos de la función de densidad de probabilidad o en términos de la función de supervivencia, la probabilidad de que un recién nacido muera entre la edad x y z , condicionado a que el recién nacido ha alcanzado la edad x . De esta manera, la expresión (1.4) puede ser utilizada para calcular la probabilidad de muerte instantánea para un recién nacido, dado que ha sobrevivido hasta la edad x . Haciendo $z = x + dx$,

$$\begin{aligned} \mathbb{P}(x < X \leq x + dx | X > x) &= \frac{s(x) - s(x + dx)}{s(x)} \\ &= \frac{F_X(x + dx) - F_X(x)}{1 - F_X(x)} \\ &\approx \frac{f_X(x)dx}{1 - F_X(x)}, \end{aligned}$$

siempre que $F_X(x) \neq 1$.

La función $\frac{f_X(x)}{1 - F_X(x)}$, puede ser interpretada como la función de densidad de probabilidad condicional de X a la edad exacta x , dado que X ha sobrevivido a la edad x , ésta es denotada por $\mu(x)$. Así, usando (1.1) y (1.2) tenemos:

$$\mu(x) = \frac{f_X(x)}{1 - F_X(x)} = \frac{F'_X(x)}{1 - F_X(x)} = \frac{-s'(x)}{s(x)}, \quad (1.11)$$

para $s(x) \neq 0$. Debido a las propiedades de $f_X(x)$ y $s(x)$, la función $\mu(x)$ es no negativa. La función $\mu(x)$ es llamada **tasa de riesgo** en la teoría general de probabilidad. Sin embargo, en las ciencias actuariales la función $\mu(x)$ es llamada **la fuerza de mortalidad**. La siguiente proposición muestra la importancia de la fuerza de mortalidad, pues la distribución de X puede ser justificada utilizando este concepto.

Proposición 1.2. Sea $\mu(x)$ la fuerza de mortalidad de “la edad hasta la muerte” de la v.a. X , entonces

$${}_t p_x = \exp \left\{ - \int_x^{x+t} \mu(z) dz \right\}, \quad (1.12)$$

o equivalentemente

$${}_t p_x = \exp \left\{ - \int_0^t \mu(x + z) dz \right\}. \quad (1.13)$$

En particular, la función de supervivencia $s(x)$ esta dada por

$$s(x) = {}_x p_0 = \exp \left\{ - \int_0^x \mu(z) dz \right\}. \quad (1.14)$$

Demostración

Solo consideraremos el caso cuando $s(u) > 0; \forall u \in [x, x+t]$, $s'(u)$ existe y es continua en $[x, x+t]$. Notemos que la ecuación (1.11) puede expresarse como:

$$\begin{aligned} \mu(x) &= -\frac{s'(x)}{s(x)} \\ &= -\frac{d}{dx} \ln s(x). \end{aligned}$$

En la ecuación anterior, cambiando la variable x por y obtenemos

$$-\mu(y)dy = d \ln s(y).$$

Después, integrando esta expresión de x a $x+t$, tenemos

$$\begin{aligned} - \int_x^{x+t} \mu(y)dy &= \ln \left(\frac{s(x+t)}{s(x)} \right) \\ &= \ln {}_t p_x, \end{aligned}$$

tomando exponencial en ambos lados de la ecuación anterior obtenemos

$${}_t p_x = \exp \left\{ - \int_x^{x+t} \mu(y)dy \right\}. \quad (1.15)$$

Es conveniente hacer el cambio de variables $z = y - x$, así (1.15) se puede expresar como

$${}_t p_x = \exp \left\{ - \int_0^t \mu(x+z)dz \right\}.$$

En particular, si consideramos “una vida de edad cero” y un cambio en el tiempo de supervivencia dado por x , es decir, ajustando $x = 0$ y $t = x$ en (1.13) y utilizando la ecuación (1.7), se puede establecer una relación entre la fuerza de mortalidad y la función de supervivencia dada por

$${}_x p_0 = s(x) = \exp \left\{ - \int_0^x \mu(z)dz \right\}.$$

□

Corolario 1.1. Sea $\mu(x)$ la fuerza de mortalidad de X . Entonces se siguen los siguientes resultados:

i) La función de distribución de $F_X(x)$ está dada por

$$F_X(x) = 1 - \exp \left\{ - \int_0^x \mu(z) dz \right\},$$

ii) Una expresión para la función de densidad $f_X(x)$ de X en términos de la fuerza de mortalidad está dada por

$$f_X(x) = s(x) \cdot \mu(x). \quad (1.16)$$

Demostración:

i) De las ecuaciones (1.1) y (1.14), se sigue

$$F_X(x) = 1 - s(x) = 1 - \exp \left\{ - \int_0^x \mu(z) dz \right\}.$$

ii) De las ecuaciones (1.2) y (1.14),

$$\begin{aligned} f_X(x) &= F'_X(x) = -s'(x) = -\frac{d}{dx} \exp \left\{ - \int_0^x \mu(z) dz \right\} \\ &= \exp \left\{ - \int_0^x \mu(z) dz \right\} \mu(x) \\ &= s(x) \mu(x). \end{aligned}$$

□

Nota 1.2. Si se considera una fuerza de mortalidad constante en la ecuación (1.14), digamos $\mu(x) \equiv \mu, \forall x > 0$, entonces la función de supervivencia se expresa como

$$s(x) = \exp \left\{ - \int_0^x \mu dz \right\} = e^{-\mu x}$$

En otras palabras, la variable aleatoria X esta distribuida exponencialmente con parámetro μ . Observe además que si se sustituye (1.14) en (1.16) y se compara el resultado con la función de densidad de una v.a. distribuida exponencialmente con parámetro μ , es decir $f_X(x) = \mu e^{-\mu x}$, podemos concluir que se tiene una generalización de una distribución exponencial.

Para la v.a. “tiempo de supervivencia” τ_x , también será conveniente encontrar una expresión para la fuerza de mortalidad y la función de densidad de τ_x .

Definición 1.4. Dada una v.a. “tiempo de supervivencia” τ_x con función de distribución $F_{\tau_x}(t)$ y una función de densidad $f_{\tau_x}(t)$, la fuerza de mortalidad de τ_x está definida como la función

$$\mu_x(t) = \frac{f_{\tau_x}(t)}{1 - F_{\tau_x}(t)},$$

para $F_{\tau_x}(t) < 1$.

Definición 1.5. Dada la fuerza de mortalidad de τ_x , se define el **índice de supervivencia** $S(t, x)$ para (x) y $t > 0$ como

$$S(t, x) = \exp \left\{ - \int_0^t \mu_x(z) dz \right\}. \quad (1.17)$$

Proposición 1.3. Consideremos la fuerza de mortalidad de la v.a. τ_x . Si $f_{\tau_x}(t)$ y $\mu_x(t)$ denotan la función de densidad y la fuerza de mortalidad, respectivamente, entonces

$$\begin{aligned} f_{\tau_x}(t) &= {}_t p_x \mu(x+t), \\ y \\ \mu_x(t) &= \mu(x+t). \end{aligned}$$

Demostración. Notemos que de (1.5) $F_{\tau_x}(t) = \mathbb{P}(\tau_x \leq t) = {}_t q_x$, por definición de la función de densidad tenemos

$$\begin{aligned} f_{\tau_x}(t) &= \frac{d}{dt} {}_t q_x \\ &= \frac{d}{dt} (1 - {}_t p_x) \\ &= - \frac{d}{dt} {}_t p_x \\ &= - \frac{d}{dt} \exp \left\{ - \int_x^{x+t} \mu(z) dz \right\} \\ &= \mu(x+t) \exp \left\{ - \int_x^{x+t} \mu(z) dz \right\} \\ &= \mu(x+t) {}_t p_x, \end{aligned}$$

donde en el cuarto paso se utilizó la ecuación (1.12), en el quinto paso se diferencié la ecuación (1.12) con respecto de t y en el último paso, se utilizó (1.12) una vez más. Por otro lado, utilizando la definición de tasa de riesgo, se sigue que

$$\begin{aligned}
 \mu_x(t) &= \frac{f_{\tau_x}(t)}{1 - F_{\tau_x}(t)} \\
 &= \frac{f_{\tau_x}(t)}{f_{\tau_x}(t)} \\
 &= \frac{{}_t p_x}{\mu(x+t) {}_t p_x} \\
 &= \frac{{}_t p_x}{{}_t p_x} \\
 &= \mu(x+t).
 \end{aligned}$$

□

Corolario 1.2. Se sigue que

$${}_t p_x = S(t, x). \quad (1.18)$$

Proposición 1.4. Para $t \leq T$, la probabilidad de que (x) alcanzará la edad $x + T$ dado que (x) ha alcanzado la edad $x + t$ es igual a

$$\begin{aligned}
 \mathbb{P}(\tau_x > T | \tau_x > t) &= \mathbb{P}(\tau_{x+t} > T - t) \\
 &= {}_{T-t} p_{x+t}.
 \end{aligned} \quad (1.19)$$

Demostración. Por definición, $\mathbb{P}(\tau_{x+t} > T - t) = {}_{T-t} p_{x+t}$. Por otro lado, utilizando la definición de probabilidad condicional y la ecuación (1.9) tenemos

$$\begin{aligned}
 \mathbb{P}(\tau_x > T | \tau_x > t) &= \frac{\mathbb{P}(\tau_x > T)}{\mathbb{P}(\tau_x > t)} \\
 &= \frac{{}_T p_x}{{}_t p_x} \\
 &= \frac{s(x+T)}{s(x+t)} \\
 &= \frac{s(x+t+T-t)}{s(x+t)} \\
 &= {}_{T-t} p_{x+t}.
 \end{aligned}$$

□

Sin embargo, utilizando la definición de fuerza de mortalidad, ${}_{T-t}p_{x+t}$ se puede expresar como

$${}_{T-t}p_{x+t} = \exp \left\{ - \int_t^T \mu_x(z) dz \right\} = \exp \left\{ - \int_0^T \mu(x+z) dz \right\}, \quad (1.20)$$

o alternativamente, en términos del índice de supervivencia S o de la función de supervivencia s , se tiene la siguiente expresión

$${}_{T-t}p_{x+t} = \frac{S(T, x)}{S(t, x)} = \frac{s(x+T)}{s(x+t)}. \quad (1.21)$$

1.1.3. Tablas de vida y su relación con la función de supervivencia

Las tablas de vida, representan un componente fundamental de muchos modelos matemáticos en las ciencias actuariales. En el caso del presente trabajo, la importancia de las tablas de vida radica en que en ellas se pueden resumir las distribuciones de las variables aleatorias “el tiempo hasta la muerte” de X y “el tiempo de supervivencia” τ_x .

De la ecuación (1.10) recordemos que la probabilidad condicional de que (x) muera dentro de los siguientes t años está dada por la expresión

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

y como un caso especial, cuando $t = 1$ tenemos:

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

Ahora, denotemos por l_0 el número de recién nacidos de un grupo. Para cada recién nacido en el grupo, la v.a. “la edad hasta la muerte” se distribuye de acuerdo a la función de supervivencia $s(x)$. Además, sea $L(x)$ el número de sobrevivientes de edad x del grupo y consideremos que todos los recién nacidos están indexados por $j = 1, 2, \dots, l_0$. Sea I_j la función indicadora para la supervivencia de los recién nacidos j , es decir

$$I_j = \begin{cases} 1 & \text{si } j \text{ sobrevive a la edad } x, \\ 0 & \text{otro caso.} \end{cases}$$

Con la definición anterior, podemos decir que I_j tiene distribución Bernoulli con parámetros $s(x)$ para $j = 1, 2, \dots, l_0$. Adicionalmente, notemos que

$$L(x) = \sum_{j=1}^{l_0} I_j,$$

y como además $\mathbb{E}[I_j] = s(x)$, entonces

$$\mathbb{E}[L(x)] = \sum_{j=1}^{l_0} \mathbb{E}[I_j] = l_0 s(x).$$

Denotaremos por l_x el número esperado de supervivientes a la edad x , entonces

$$l_x = \mathbb{E}[L(x)] = l_0 s(x). \quad (1.22)$$

Note que de (1.22) podemos obtener un estimador para la función de supervivencia, esto es

$$s(x) = \frac{l_x}{l_0}. \quad (1.23)$$

Si asumimos que las funciones indicadoras I_j son mutuamente independientes, entonces $L(x)$ tiene distribución Binomial con parámetros $n := l_0$ y $p := s(x)$.

Ahora, denotemos por ${}_tD_x$ el número de muertes que ocurrieron durante el intervalo $(x, x + t]$ de los l_0 recién nacidos. Sea $\mathbb{E}[{}_tD_x]$ el número esperado de muertes en ese intervalo que denotaremos por ${}_td_x$. Como la probabilidad de que un recién nacido muera durante el intervalo de tiempo $(x, x + t]$ esta dada por $s(x) - s(x + t)$, entonces

$$\begin{aligned} {}_td_x &= \mathbb{E}[{}_tD_x] = l_0 (s(x) - s(x + t)) \\ &= l_x - l_{x+t}. \end{aligned}$$

En particular, cuando $t = 1$ el prefijo en ${}_tD_x$ y ${}_td_x$ se omite.

Usando (1.22) y (1.2) busquemos una expresión para ${}_tq_x$ y ${}_tp_x$ en términos de l_x . Derivando la ecuación (1.22) con respecto de x tenemos

$$\begin{aligned} -\frac{1}{l_x} \frac{dl_x}{dx} &= -\frac{1}{s(x)} \frac{ds(x)}{dx} \\ &= \mu(x), \end{aligned}$$

o análogamente

$$-dl_x = l_x \mu_x(x) dx.$$

Finalmente, utilizando (1.23) y (1.9) tenemos expresiones para ${}_tq_x$ y ${}_tp_x$ en términos de los estimadores de $s(x)$

$${}_tp_x = \frac{s(x+t)}{s(x)} = \frac{l_{x+t}/l_0}{l_x/l_0} = \frac{l_{x+t}}{l_x},$$

y

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

Otra información de gran utilidad que se puede obtener de una tabla de vida es, el número promedio de años de vida restante al inicio del intervalo de edad que se espera viva una persona de edad x . El valor esperado de la v.a. τ_x es llamada la “esperanza completa de vida” y es denotada por $\overset{\circ}{e}_x$ que se define como:

$$\begin{aligned} \overset{\circ}{e}_x &= \mathbb{E}[\tau_x] = \int_0^\infty t {}_tp_x \mu(x+t) dt, \\ &= \int_0^\infty t d(-{}_td_x), \\ &= t(-{}_tp_x) \Big|_0^\infty + \int_0^\infty {}_tp_x dt, \end{aligned} \tag{1.24}$$

donde la segunda integral es calculada por partes. La existencia de $\mathbb{E}[\tau_x]$ implica que $\lim_{t \rightarrow \infty} t(-{}_tp_x) = 0$. Así

$$\overset{\circ}{e}_x = \int_0^\infty {}_tp_x dt.$$

Utilizando integración por partes una vez más en la ecuación anterior, se obtiene una expresión para $\mathbb{E}[\tau_x^2]$ como sigue

$$\begin{aligned} \mathbb{E}[\tau_x^2] &= \int_0^\infty t^2 {}_tp_x \mu(x+t) dt \\ &= 2 \int_0^\infty t {}_tp_x dt. \end{aligned} \tag{1.25}$$

De las ecuaciones (1.24) y (1.25) se puede obtener una expresión para la $\text{Var}[\tau_x]$ como sigue

$$\text{Var}[\tau_x] = \mathbb{E}[\tau_x^2] - \mathbb{E}[\tau_x]^2 = 2 \int_0^\infty t {}_tp_x dt - \overset{\circ}{e}_x^2, \tag{1.26}$$

siempre que $\mathbb{E}[\tau_x]$ y $\mathbb{E}[\tau_x^2]$ existan.

La esperanza completa de vida, a menudo se utiliza para comparar los niveles de salud pública entre las diferentes poblaciones.

1.2. Modelos de tasas de interés

1.2.1. Bonos cupón cero

En general, los bonos son instrumentos financieros de deuda, que se pueden negociar en el mercado secundario, por lo tanto quien es propietario del bono tiene el derecho de recibir el valor nominal en la fecha de vencimiento y cuando es el caso, los intereses en forma de cupones. Dependiendo de la duración del bono, pueden subdividirse en bonos a corto plazo (de 1 a 3 años) o a largo plazo. El precio del bono depende de la tasa de interés pactada y del riesgo de incumplimiento, mientras mayor es el riesgo crece la posibilidad de que el tenedor no reciba el pago acordado. Si la tasa de interés del mercado es muy alta, la institución que emite el bono debe ser consciente de su exposición al riesgo para que el bono sea aceptado por el mercado. Un bono cupón cero es un instrumento financiero que no tiene pagos a lo largo de su duración, pues hay un solo pago al final del tiempo pactado.

Sea $B(t; T)$ el precio al tiempo t de un bono cupón cero con madurez T y, por simplicidad, supongamos que al término del contrato paga un peso, es decir, $B(T; T) = 1$.

Definición 1.6. El rendimiento al vencimiento $R(t; T)$ al tiempo t del bono descontado $B(t; T)$ es la tasa de rendimiento compuesta continuamente que hace que el precio del bono sea \$1 al tiempo T , esto es

$$B(t; T)e^{(T-t)R(t; T)} = 1,$$

o bien, en términos del rendimiento

$$R(t; T) = -\frac{\ln(B(t; T))}{T - t}.$$

Cuando fijamos t y aumentamos T , $R(t; T)$ determina la estructura temporal de las tasas de interés, por lo que en este caso la curva de rendimientos es igual a la estructura temporal.

Algunos de los modelos basados en arbitraje asumen que un solo factor puede explicar la estructura temporal y suponen que la tasa *spot* instantánea es la única variable de estado. Ahora, denotemos por $V(t)$ el valor al tiempo t del reclamo contingente que depende de la tasa *spot* con madurez t . Supongamos que bajo la medida objetiva \mathbb{P} , la dinámica de la tasa *spot* esta dada por la ecuación

$$dr(t) = a(t, r(t))dt + \rho(t, r(t))dW(t), \quad (1.27)$$

donde $W(t)$ es un movimiento browniano estándar.

El precio del bono depende de la tasa de interés, es decir $B(t; T) = B(t, r(t), T) \in C^{1,1,2}(\mathbb{R}^+ \times \mathbb{R}^+, \mathbb{R})$ sobre $(\omega, \mathcal{G}, \mathbb{P})$ \mathcal{G} -filtración del movimiento Browniano. Por lo tanto este precio junto con todos sus derivados estan completamente determinado por la ecuación general:

$$\frac{\partial B(t, r(t), T)}{\partial t} + (a(t, r(t)) - \rho(t, r(t))q(t)) \frac{\partial B(t, r(t), T)}{\partial r} + \frac{1}{2} \rho^2 \frac{\partial^2 B(t, r(t), T)}{\partial r^2} - rB(t, r(t), T) = 0. \quad (1.28)$$

con condición terminal

$$B(T, T) = 1, \quad (1.29)$$

donde $q(t) = \frac{a(t, r(t)) - r(t)}{\rho(t, r(t))}$ es el valor del riesgo de mercado. El término $a(t, r(t)) - \rho(t, r(t))q(t)$ es el término de tendencia (o *drift*) de la tasa *spot* bajo la medida martingala \mathbb{Q} .

Sea $V^*(t)$ el precio relativo al activo V bajo el numerario N al tiempo t , como $V^*(t) = \frac{V(t)}{N(t)}$, se tiene la siguiente proposición.

Proposición 1.5. *Bajo ciertas condiciones de regularidad, un mercado es libre de arbitraje si existe una medida martingala equivalente \mathbb{Q} , tal que el proceso de precios relativo de cualquier valor es una \mathbb{Q} -martingala, es decir, para $t \leq T$*

$$\mathbb{E}^{\mathbb{Q}} [V^*(T) | \mathcal{G}_t] = V^*(t).$$

Demostración. Ver Karatzas y Shreve [10], página 232.

Tomemos como un ejemplo de numerario el precio $\beta(t)$ al tiempo t de una unidad monetaria reinvertida continuamente a la tasa *spot* dado un tiempo inicial fijo 0. Entonces

$$\beta(t) = e^{\int_0^t r(s)ds}.$$

Bajo este numerario, el precio relativo del activo V esta dado por

$$V^*(t) = \frac{V(t)}{\beta(t)}.$$

En la proposición anterior, sea $V(t) = B(t, T)$ un bono cupón cero con maduración T cuya numeraria es $\beta(t)$, por ser martingala

$$\mathbb{E}^{\mathbb{Q}} [B^*(T, T) | \mathcal{G}_t] = \mathbb{E}^{\mathbb{Q}} \left[\frac{B(T; T)}{\beta(T)} | \mathcal{G}_t \right] = \frac{B(t; T)}{\beta(t)}, \quad (1.30)$$

pero como $B(T; T) = 1$ se tiene

$$\mathbb{E}^{\mathbb{Q}} = \left[e^{-\int_t^T r(s)ds} | \mathcal{G}_t \right] = B(t; T). \quad (1.31)$$

A continuación, vamos a presentar dos modelos en los cuales $e^{\int_t^T r(s)ds}$ tiene distribución gaussiana. Uno de ellos es propuesto por Vasicek en 1977 y el otro es el modelo CIR propuesto en 1985.

Recuerde que una variables aleatoria X con distribución gaussiana $N(\mu, \sigma^2)$ tiene función generadora de momentos

$$E(e^{aX}) = e^{a\mu + \frac{1}{2}\sigma^2 a^2}.$$

A estos se les llama **modelos afines** y para ellos se tiene un expresión analítica para calcular el precio de un bono.

1.2.2. Modelo de Vasicek

Propone modelar la tasa de interés *spot* con un proceso tipo Ornstein-Uhlenbeck como sigue:

$$dr(t) = k(\theta - r(t))dt + \sigma dW(t), \quad (1.32)$$

k , θ y σ son constantes positivas y $W(t)$ es un movimiento browniano estándar. La solución a la ecuación (1.32) es

$$r(t) = \theta + (r(s) - \theta)e^{-k(t-s)} + \sigma \int_s^t e^{-k(t-u)} dW(u), \quad \forall s \leq t. \quad (1.33)$$

Este modelo define una caminata aleatoria alrededor de una tendencia, por lo tanto tiene reversión a la media.

Dada la información al tiempo s , $r(t)$ se distribuye como una gaussiana con media condicional

$$\mathbb{E}[r(t)|\mathcal{G}_s] = \theta + (r(s) - \theta)e^{-k(t-s)}, \quad (1.34)$$

y varianza condicional

$$\begin{aligned} \text{Var}[r(t)|\mathcal{G}_s] &= \sigma^2 \int_s^t e^{-2k(t-u)} du, \\ &= \frac{\sigma^2}{2k} (1 - e^{-2k(t-s)}). \end{aligned} \quad (1.35)$$

Este modelo tiene un inconveniente, pues permite que las tasas de interés puedan ser negativas. Por otro lado, notemos que para valores grandes de t , el valor esperado y la varianza de la tasa *spot* son θ y $\frac{\sigma^2}{2k}$ respectivamente, lo que quiere decir que el proceso de reversión a la media evita que estos dos valores diverjan.

Sea $I(t; T) = \int_t^T r(s) ds$, entonces $I(t; T)$ tiene distribución gaussiana con media condicional

$$\mathbb{E}[I(t; T)|\mathcal{G}_t] = \theta(T - t) + (r(t) - \theta) \left(\frac{1 - e^{-k(T-t)}}{k} \right), \quad (1.36)$$

y varianza condicionada

$$\begin{aligned} \text{Var}[I(t; T)|\mathcal{G}_t] &= \sigma^2 \int_t^T \left(\frac{1 - e^{-k(T-u)}}{k} \right)^2 du, \\ &= \frac{\sigma^2}{k^2} \left(T - t - 2 \left(\frac{1 - e^{-k(T-t)}}{k} \right) + \frac{1}{2k} (1 - e^{-2k(T-t)}) \right). \end{aligned} \quad (1.37)$$

Entonces el precio de un bono cupón cero que paga un peso al término del contrato satisface:

$$B(t, r(t), T) = \mathbb{E}[e^{-I(t; T)}|\mathcal{G}_t] = e^{a(t, T) - r(t)b(t, T)}, \quad (1.38)$$

donde $b(t, T) = \frac{1-e^{-k(T-t)}}{k}$ y $a(t, T) = \frac{1}{k^2} (b(t, T) - T + t) (k^2\theta - \frac{1}{2}\sigma^2) - \frac{\sigma^2 b^2(t, T)}{4k}$. Por lo tanto, la estructura temporal de la tasa de interés está dada por

$$\begin{aligned} R(t, T) &= -\frac{\ln B(t, T)}{T-t} \\ &= \frac{r(t)b(t, T) - a(t, T)}{T-t} \\ &= r(t)\frac{1-e^{-k(T-t)}}{k(T-t)} - \left(\frac{1-e^{-k(T-t)}}{k(T-t)} - 1\right) \left(\theta - \frac{\sigma^2}{2k^2}\right) + \frac{\sigma^2(1-e^{-k(T-t)})^2}{4k^3(T-t)}. \end{aligned} \quad (1.39)$$

Nota 1.3. La tasa de interés con madurez infinita es:

$$\lim_{T \rightarrow \infty} R(t; T) = \theta - \frac{\sigma^2}{2k^2}. \quad (1.40)$$

1.2.3. Modelo Cox, Ingersoll y Ross (CIR)

Este modelo describe el comportamiento de la tasa *spot* con reversión a la media como sigue:

$$dr(t) = k(\theta - r(t))dt + \sigma\sqrt{r(t)}dW^2(t), \quad (1.41)$$

donde k , θ y σ son constantes positivas, $W^2(t)$ es un movimiento browniano estándar. Este modelo a diferencia del de Vasicek no permite valores negativos de la tasa de interés. Es conocido como **proceso de raíz cuadrada**, el cual tiene varianza $\sigma^2 t$ por unidad de tiempo.

Si aplicamos el teorema de Itô al proceso $X(t) = r(t)e^{kt}$ se obtiene la solución de la ecuación diferencial estocástica (1.41)

$$r(t) = \theta + (r(s) - \theta)e^{-k(t-s)} + \sigma \int_s^t e^{-k(t-u)} \sqrt{r(u)} dW^2(u), \quad \forall t \geq s. \quad (1.42)$$

La tasa *spot* de este modelo se distribuye como una ji-cuadrada (χ^2). Por lo tanto, el valor esperado de $r(t)$ y su varianza condicionadas a la información que se tiene hasta el tiempo s

son

$$\mathbb{E}[r(t)|\mathcal{G}_s] = \theta + (r(s) - \theta)e^{-k(t-s)}, \quad (1.43)$$

y

$$Var[r(t)|\mathcal{G}_s] = \frac{\sigma^2\theta}{2k} (1 - e^{-k(t-s)})^2 + \frac{\sigma^2 r(s)}{k} (e^{-k(t-s)} - e^{-2k(t-s)}). \quad (1.44)$$

Cuando $t \rightarrow \infty$ en las ecuaciones (1.43) y (1.44) la esperanza es θ y la varianza es $\theta \frac{\sigma^2}{2k}$.

Bajo la medida neutra al riesgo \mathbb{Q} , el valor de un bono descontado es martingala y el riesgo de mercado está dado por $q(t, r) = \frac{q_1}{\sqrt{r}} + q_2\sqrt{r}$. El precio de un bono cupón cero que paga un peso al término de la maduración es

$$\begin{aligned} B(t, r(t), T) &= \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} | \mathcal{G}_t \right] \\ &= A(t, T) e^{-r(t)D(t, T)}, \end{aligned} \quad (1.45)$$

donde $a = k + q_2\sigma$, $b = \frac{k\theta - q_1\sigma}{a}$, $\gamma = \sqrt{a^2 + 2\sigma^2}$, $A(t, T) = \left(\frac{2\gamma e^{-(\gamma-a)(T-t)/2}}{(\gamma+a)(1-e^{-\gamma(T-t)}) + 2\gamma e^{-\gamma(T-t)}} \right)^{\frac{2ab}{\sigma^2}}$ y $D(t, T) = \frac{2(1-e^{-\gamma(T-t)})}{(\gamma+a)(1-e^{-\gamma(T-t)}) + 2\gamma e^{-\gamma(T-t)}}$.

La estructura de la tasa de interés esta dada por

$$\begin{aligned} R(t, T) &= -\frac{\ln(B(t, T))}{T-t} \\ &= \frac{r(t)D(t, T) - \ln(A(t, T))}{T-t}. \end{aligned} \quad (1.46)$$

1.2.4. Estimación de parámetros para modelos de tasa de interés

La estimación de parámetros es una herramienta importante cuando, para problemas reales no se pueden establecer los parámetros del modelo a priori. Éstos, por lo general, se estiman al analizar el comportamiento histórico de datos temporales. En los modelos de tasa de interés, cuando se asume que los coeficientes del proceso son constantes, existen diferentes técnicas de estimación, entre las cuales están: el método de los momentos, mínimos cuadrados ordinarios y máxima verosimilitud. A continuación, se estimarán los parámetros de los modelos Vasicek y CIR por el método de máxima verosimilitud.

1.2.4.1. Modelo Vasicek

Consideremos que la tasa *spot* sigue una dinámica dada por la ecuación (1.32), así los parámetros a estimar son: k , θ y σ . Dado que $r(t)$ sigue un proceso gaussiano donde la media y la varianza están dadas por las ecuaciones (1.34) y (1.35) respectivamente, su función de probabilidad de transición p para la tasa $r(t)$ dada una tasa inicial $r(s)$ está dada por la ecuación:

$$p(t, y; s, x) = \sqrt{\frac{k}{\pi\sigma^2(1 - e^{-2k(t-s)})}} \exp\left(\frac{-k[(y - \theta) - (x - \theta)e^{-k(t-s)}]^2}{\sigma^2(1 - e^{-2k(t-s)})}\right), \quad (1.47)$$

donde por simplicidad en la notación, hemos hecho $y = r(t)$ y $x = r(s)$.

Sean

$$\begin{aligned} \alpha &= \frac{k}{\sigma^2(1 - e^{-2k(t-s)})} \\ \beta &= 1 - e^{-k(t-s)}. \end{aligned} \quad (1.48)$$

Utilizando (1.47) y (1.48) definamos

$$P := \ln(p) = -\ln(\sqrt{\pi}) + \frac{1}{2} \ln(\alpha) - \alpha [x - y + \beta(\theta - x)]^2. \quad (1.49)$$

Diferenciando la ecuación (1.49) con respecto a α , β y θ respectivamente tenemos:

$$\begin{aligned} \frac{\partial P}{\partial \alpha} &= \frac{1}{2\alpha} - [(x - y) + \beta(\theta - x)]^2, \\ \frac{\partial P}{\partial \beta} &= -2\alpha [(x - y) + \beta(\theta - x)] (\theta - x), \\ \frac{\partial P}{\partial \theta} &= -2\alpha [(x - y) + \beta(\theta - x)] \beta. \end{aligned} \quad (1.50)$$

Supongamos que podemos observar al tiempo t_i la tasa *spot* de interés $r(t_i)$, $i = 1, \dots, n$. con densidades de transición dadas por (1.47). Además, supongamos que las observaciones tienen incrementos de tiempos iguales, es decir, $\Delta t_i = t_{i+1} - t_i = t$. Definamos $X = (r(t_0), \dots, r(t_{n-1}))$ y $Y = (r(t_1), \dots, r(t_n))$. Entonces, utilizando la definición de máxima vero-

similitud tenemos que encontrar α , β y θ tal que

$$\begin{aligned} \frac{n}{2\alpha} - \sum_{i=1}^n \{[(X_i - Y_i) + \beta(\theta - X_i)]^2\} &= 0, \\ \sum_{i=1}^n \{[(X_i - Y_i) + \beta(\theta - X_i)](\theta - X_i)\} &= 0, \\ \sum_{i=1}^n [(X_i - Y_i) + \beta(\theta - X_i)] &= 0. \end{aligned} \quad (1.51)$$

Cuyas soluciones son:

$$\begin{aligned} \alpha &= -\frac{n}{2} \frac{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i\right)^2}{-2 \left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right) \left(\sum_{i=1}^n X_i Y_i\right) + n \left(\sum_{i=1}^n X_i Y_i\right)^2 + \left(\sum_{i=1}^n Y_i\right)^2 \left(\sum_{i=1}^n X_i^2\right) + \left(\sum_{i=1}^n Y_i^2\right) \left(\sum_{i=1}^n X_i\right)^2 - n \left(\sum_{i=1}^n Y_i^2\right) \left(\sum_{i=1}^n X_i^2\right)}, \\ \beta &= \frac{\left(\sum_{i=1}^n X_i\right)^2 - \left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right) - n \sum_{i=1}^n X_i^2 - n \sum_{i=1}^n (X_i Y_i)}{\left(\sum_{i=1}^n X_i\right)^2 - n \sum_{i=1}^n X_i^2}, \\ \theta &= \frac{\left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n X_i Y_i\right) - \left(\sum_{i=1}^n Y_i\right) \left(\sum_{i=1}^n X_i^2\right)}{\left(\sum_{i=1}^n X_i\right)^2 - \left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right) - n \sum_{i=1}^n X_i + n \sum_{i=1}^n X_i Y_i}. \end{aligned}$$

Por lo tanto, sustituyendo las soluciones anteriores en las ecuaciones (1.48), obtenemos que las estimaciones para los parámetros originales k , θ y σ^2 son:

$$\begin{aligned} k &= \frac{1}{(t-s)} \ln \left(\frac{n \sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right)}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i\right)^2} \right), \\ \theta &= \frac{\left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n X_i Y_i\right) - \left(\sum_{i=1}^n Y_i\right) \left(\sum_{i=1}^n X_i^2\right)}{\left(\sum_{i=1}^n X_i\right)^2 - \left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right) - n \sum_{i=1}^n X_i + n \sum_{i=1}^n X_i Y_i}, \\ \sigma^2 &= \frac{(1-\beta)^2 \ln(1-\beta)}{\alpha \beta (t-s) (\beta-2)}. \end{aligned}$$

1.2.4.2. Modelo CIR

El modelo CIR es un caso de un proceso difusivo, en el cuál la función de probabilidad de transición tiene una forma cerrada. Es decir, dada una tasa r_t al tiempo t , la densidad de $r_{t+\Delta t}$ al tiempo $t + \Delta t$ tiene la forma:

$$P(r_{t+\Delta t}|r_t; k, \theta, \sigma, \Delta t) = ce^{-u-v} \left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}),$$

donde

$$\begin{aligned} c &= \frac{2k}{\sigma^2(1 - e^{-k\Delta t})}, \\ u &= cr_t e^{-k\Delta t}, \\ v &= cr_t + \Delta t, \\ q &= \frac{2k\theta}{\sigma^2} - 1, \end{aligned}$$

$I_q(\cdot)$ es la función modificada de Bessel de primer tipo y de orden q . La función de verosimilitud para la tasa de interés de una serie de tiempo de N observaciones esta dada por:

$$L(k, \theta, \sigma) = \prod_{i=1}^{N-1} P(r_{t+\Delta t}|r_t; k, \theta, \sigma, \Delta t),$$

y por lo tanto, la función log-verosimilitud es:

$$\ln L(k, \theta, \sigma) = (N - 1) \ln c + \sum_{i=1}^{N-1} \left\{ -u_{t_i} - v_{t_{i+1}} + \frac{1}{2} q \ln \left(\frac{v_{t_{i+1}}}{u_{t_i}} \right) + \ln [I_q(2\sqrt{u_{t_i} v_{t_{i+1}}})] \right\},$$

donde $u_{t_i} = cr_{t_i} e^{-k\Delta t}$ y $v_{t_{i+1}} = cr_{t_{i+1}}$.

Maximizar la función de verosimilitud es análogo a maximizar la función log-verosimilitud, ya que la función logaritmo es monótona creciente y esta última es mas fácil de calcular. La solución entonces tiene la forma:

$$(\hat{k}, \hat{\theta}, \hat{\sigma}) = \arg \max_{(k, \theta, \sigma)} \ln L(k, \theta, \sigma).$$

Modelación de la fuerza de mortalidad

Cuando una compañía de seguros de vida celebra un contrato se ve expuesta a dos tipos de riesgo: uno es el **riesgo específico o diversificable**, es el riesgo que se presenta por ejemplo, cuando un asegurado en particular se desvía de la media, el cual se puede eliminar suponiendo que se venden suficientes pólizas de seguros y por la ley de los grandes números el promedio de las realizaciones convergen a la media teórica. El otro tipo riesgo se conoce como **riesgo sistémico**, es el riesgo de sobrestimar o subestimar una proyección anticipada de la evolución de la mortalidad que se observará en el futuro. Es importante mencionar que no existe una manera estándar de calcular una prima que cubra a una aseguradora del riesgo sistémico, pero lo que si es un hecho es que debido a la necesidad de predecir este riesgo se han desarrollado varios marcos de modelación que expondremos a continuación.

En este capítulo se presentan dos de los modelos más utilizados en la literatura para la fuerza de mortalidad. Uno de ellos es el modelo Gompertz-Makeham que se utilizará análogamente al trabajo hecho por Debón [8], el segundo es el modelo Lee-Carter [12]. Además, se define un marco de modelación según Carstens [3], Carnis, Blake and Dowd [2], Dahl M. [6] para tratar a la fuerza de mortalidad como un proceso estocástico.

2.1. Modelo de Gompertz-Makeham para la fuerza de mortalidad

Históricamente la ciencia actuarial ha trabajado con los datos de mortalidad de una población. El primer paso y quizá una de las partes fundamentales en las que interviene la estadística es la graduación de los mismos. Vamos a entender por graduación la aplicación de los principios y métodos por medio de los cuales un conjunto de probabilidades observadas se ajustan para proporcionar un modelo suavizado que nos permita hacer inferencias y además cálculos prácticos de primas, reservas, etc.

Modelos estáticos en el tiempo para la fuerza de mortalidad han sido estudiados y utilizados para dar solución al problema de predicción. En la tesis doctoral de Debón [8] por ejemplo, cuyo objetivo es una presentación exhaustiva de los métodos disponibles para ajustar tablas de mortalidad, se realiza la graduación de datos de mortalidad de la Comunidad de Valencia mediante métodos paramétricos, lo que supone ajustar una función paramétrica que depende solo de la edad de los individuos.

El modelar la v.a. τ_x mediante la búsqueda de su función de distribución ${}_tq_x$ (ver def. 1.2), tiene la ventaja de permitir su estimación mediante un reducido número de parámetros, ventaja deseable cuando se dispone de pocos datos. A lo largo del tiempo diversos autores han propuesto modelos para el comportamiento probabilístico de τ_x . Uno de ellos es el modelo Gompertz-Makeham cuya definición y utilidad se expondrá a continuación para modelar q_x .

Primero, vamos a graduar la probabilidad de que una persona de edad x muera en el transcurso de un año, esto es q_x . Supongamos que E_x personas entran en observación a la edad exacta x y continúan en observación hasta que sobreviven a una edad exacta $x + 1$ o mueren antes. Es decir, E_x son los individuos iniciales expuestos al riesgo y determinan por tanto el número de individuos en estudio. Además, se supone que la muerte o supervivencia de uno es independiente de la muerte o supervivencia de los otros. Si llamamos D_x a la variable aleatoria que representa el número de muertes que ocurren en el transcurso del año,

entonces $D_x \sim Bin(E_x, q_x)$.

Si el valor de q_x es desconocido, pero hemos observado D_x muertes de E_x personas, entonces $E_x - D_x$ de ellos sobreviven, luego el estimador de máxima verosimilitud de q_x es

$$\hat{q}_x = \frac{D_x}{E_x}. \quad (2.1)$$

A las estimaciones (2.1) las denominaremos probabilidades brutas.

Si suponemos que no hay cambios bruscos en las probabilidades de muerte q_x de una edad a otra, entonces podemos ajustar una función de probabilidad que describa su relación con la edad. En nuestro caso, vamos a utilizar la función Gompertz-Makeham (G-M) y su transformada logit, las cuales se definen como:

$$GM(r, s) = \sum_{i=1}^r \alpha_i x^{i-1} + \exp \left\{ \sum_{j=1}^s \alpha_j x^{j-1} \right\}, \quad (2.2)$$

y

$$LGM(r, s) = \frac{GM(r, s)}{1 + GM(r, s)}. \quad (2.3)$$

Las ecuaciones anteriores, constan de una parte polinomial con r parámetros y una exponencial con s , de tal manera que tenemos que determinar $r + s$ parámetros.

Si hemos de estimar los parámetros de las funciones G-M que mejor ajustan las probabilidades brutas con dicha distribución, no podemos recurrir a modelos lineales clásicos por motivos obvios de no linealidad en las funciones presentadas. Por ello vamos a recurrir a los conceptos de los Modelos Lineales Generalizados (GLM).

Un GLM es una generalización flexible de la regresión lineal ordinaria. Relaciona la variable dependiente en el experimento (la función de distribución) con la parte independiente (no aleatoria o predictor) a través de una función llamada función de enlace. Fueron formulados como una manera de unificar varios modelos estadísticos, incluyendo regresión lineal, regresión logística de Poisson, bajo un solo marco teórico. Esto permitió desarrollar un algoritmo general para la estimación por máxima verosimilitud en todos estos modelos. Los componentes de un GLM son:

1. Una función de distribución f perteneciente a la familia exponencial.
2. Un predictor lineal $\eta = X\beta$.
3. Una función de enlace g tal que $E(Y) = \mu = g^{-1}(\eta)$.

En nuestro caso, la variable dependiente es q_x y la independiente es la edad x . Buscamos ahora una función enlace que nos permita obtener una ecuación lineal para la variable dependiente. En el caso de la familia binomial, la función enlace más utilizada es el logit cuya definición es la siguiente:

$$\text{logit}(q_x) = \log\left(\frac{q_x}{1 - q_x}\right). \quad (2.4)$$

Vamos a suponer que $q_x = LGM(0, s)$, es decir vamos a prescindir de la parte polinomial de la ecu.(2.3). Sustituyendo lo anterior en la ecu.(2.4) obtenemos una expresión lineal en términos del predictor lineal α_j como sigue:

$$\begin{aligned} \text{logit}(q_x) &= \log\left(\frac{LGM(0, s)}{1 - LGM(0, s)}\right) = \log\left(\exp\left(\sum_{j=1}^s \alpha_j x^{j-1}\right)\right) \\ &= \log(GM(0, s)) \\ &= \sum_{j=1}^s \alpha_j x^{j-1}. \end{aligned} \quad (2.5)$$

Tomando exponencial en ambos lados de la ecuación (2.5) obtenemos la igualdad

$$\frac{1}{1 - q_x} = 1 + GM(0, s). \quad (2.6)$$

Estimación de parámetros

Para estimar los parámetros, primero se maximiza la función de verosimilitud del modelo probabilista:

$$L(q) = \prod_{x=n}^m \binom{E_x}{D_x} q_x^{D_x} (1 - q_x)^{(E_x - D_x)}, \quad (2.7)$$

y su log-verosimilitud

$$\log L(q) = \sum_{x=n}^m \left(\log \binom{E_x}{D_x} + D_x \log(q_x) + (E_x - D_x) \log(1 - q_x) \right). \quad (2.8)$$

Luego, de la ecuación (2.8) y al hacer uso de la igualdad (2.6) se obtiene una expresión para la función log-verosimilitud en términos de la función enlace como sigue:

$$\log L(q) = \sum_{x=n}^m \left(\log \left(\frac{E_x}{D_x} \right) + D_x \log(GM(0, s)) - E_x \log(1 + GM(0, s)) \right). \quad (2.9)$$

En el caso de los modelos lineales generalizados, la estimación de los parámetros se hace mediante medidas de bondad de ajuste entre los datos observados y los valores predichos por el modelo. Dos de las más populares, dada su disponibilidad en los distintos *softwares*, son el test basado en la Devianza D y el estadístico χ^2 .

Definición 2.1. La Devianza se define como:

$$D(q) = 2 \log(L(\hat{q})) - 2 \log(L(q)), \quad (2.10)$$

donde $L(q)$ y $L(\hat{q})$ son las funciones de log-verosimilitud de los q_x observados y estimados, respectivamente.

La Devianza, mide la discrepancia entre la log-verosimilitud de los datos y la del modelo utilizado, en consecuencia, maximizar la verosimilitud equivale a minimizar $D(q)$. El estadístico así construido tiene distribución asintótica χ^2 , con grados de libertad la diferencia del espacio paramétrico y la dimensión de este espacio bajo la hipótesis nula, es decir, $D(q) \sim \chi_{n-p}^2$ donde n es el número de datos y p es el número de parámetros ajustados bajo H_0 , que es la hipótesis del modelo.

El estadístico χ_{n-p}^2 también puede calcularse como la suma de los cuadrados

$$\chi^2 = \sum_{x=n}^m r_x^2,$$

donde

$$r_x = \frac{D_x - E_x \hat{q}_x}{\sqrt{E_x \hat{q}_x (1 - \hat{q}_x)}},$$

que fueron denominados por Hosmer como residuos de Pearson.

La Devianza es más usada para comparar dos modelos que como medida absoluta de bondad de ajuste. Por ejemplo, si queremos contrastar si añadir una nueva covariable mejora el modelo, si H_0 es el modelo a validar, H_1 la extensión del modelo conteniendo una covariable adicional y \hat{q}_x^0 y \hat{q}_x^1 los correspondientes valores ajustados, la diferencia de Devianza,

$$D(\hat{q}_x^0) - D(\hat{q}_x^1) = \log L(\hat{q}_x^0) - \log L(\hat{q}_x^1)$$

es el estadístico utilizado para contrastar H_0 contra H_1 y se distribuye aproximadamente χ^2 con 1 grado de libertad y, en general, con tantos grados de libertad como el número de parámetros añadidos. Así, la hipótesis nula será rechazada para el nivel de significancia α cuando $D \geq \chi_{1,\alpha}^2$, que es equivalente a que el p -valor del contraste sea menor que el nivel de α fijado.

El problema de la modelación requiere un equilibrio entre la mejora del modelo y el nada deseable incremento de la complejidad de éste al ir añadiendo cada vez más parámetros. Para esto utilizaremos el Criterio de Información de Akaike (AIC por sus siglas en inglés). Este criterio es una medida de la calidad relativa de modelos estadísticos para un conjunto dado de datos. Dada una colección de modelos candidatos para ajustar los datos, AIC compara la calidad de cada modelo en relación con cada uno de los otros modelos. Por lo tanto, AIC proporciona un medio para la selección del modelo.

Suponga que se tiene un modelo estadístico de algunos datos. Sea L el valor maximizado de la función de verosimilitud para ese modelo; Sea p el número de los parámetros estimados del modelo. Entonces, el valor AIC del modelo es:

$$AIC = 2p - 2 \ln L. \tag{2.11}$$

Por lo tanto, el modelo preferido es el que tiene el valor mínimo AIC.

En la siguiente sección, al emplear datos de México vamos a analizar tiempos de supervivencia, es decir, la duración τ_x hasta que se produce la muerte. Se pretende pues, utilizar las herramientas anteriormente expuestas para describir cómo se comporta la supervivencia

de un mexicano y además pronosticar una proyección futura.

2.1.1. Graduación de q_x para la población mexicana

Sin duda alguna, las compañías aseguradoras cuentan con una amplia base de datos históricos de sus poblaciones objetivo, con las cuales, dependiendo del producto que deseen vender hacen estimaciones futuras del comportamiento de la mortalidad para cubrirse de posibles pérdidas. En ausencia de una base de datos tan específica como la de estas empresas, se optó por recurrir a una base de datos más general como la proporcionada por el Consejo Nacional de Población (CONAPO). Este tipo de información será de importancia, por ejemplo, para establecer un plan de pensiones universales como las que se pretenden impulsar en el actual sexenio en nuestro país. Sin embargo, uno de los inconvenientes que encontramos al elegir esta base de datos fue que es limitada, pues únicamente contiene información del periodo de tiempo comprendido entre 1990 y 2009.

Se utilizó la base de datos de la CONAPO [20], que contiene información del número de personas muertas D_x^t y vivas E_x^t del periodo de $t=1990$ al 2009 de las edades 20 a 89. Se asumirá la hipótesis de uniformidad para las muertes a lo largo de cada año para obtener las estimaciones de \hat{q}_x^t (ecuación (2.1)) para $t=1990, 1995, 2000$ y 2009 (el super índice t indica el año en el calendario). Se hará diferencias entre hombres y mujeres, pues se ha observado que las mujeres son más longevas que los hombres.

La función *glm* de *R-project* implementa la metodología de los modelos lineales generalizados que se expuso anteriormente. Una descripción detallada de la estructura de esta función en el lenguaje R puede encontrarse en Verrall [19]. Se utilizó dicha función para ajustar un predictor lineal a la información que se tiene de cada uno de los años $t=1990, 1995, 2000$ y 2009. Primero, se ajustaron las funciones $q_x^t = LGM(0, s)$ con $s = 1, \dots, 14$, es decir, para cada año vamos a graduar polinomios de ajuste de hasta grado 13 para x . Las tablas 2.1, 2.2, 2.3, 2.4 y 2.5 muestran los valores de las medidas de bondad de ajuste para cada año. Como era de esperar, se observa una mejoría paulatina de la bondad del ajuste de la

devianza a medida que aumenta el número de parámetros, pero necesitamos contrastar si la disminución de la devianza es significativa y si se mejora el ajuste en relación al incremento de complejidad del modelo.

En las tablas 2.6, 2.7, 2.8, 2.9 y 2.10 se reporta para cada año, el test basado en la devianza y su respectivo p -valor para comparar los modelos al ir aumentando el grado del polinomio de ajuste y también el AIC. Se observa lo siguiente: en el año 1990, para mujeres se acepta un polinomio de grado 5 y para hombres un polinomio de grado 6. En el año 1995 se acepta para mujeres un polinomio de grado 3, mientras que para los hombres uno de grado 4. Para el 2000 se acepta para mujeres un polinomio de grado 5 y para los hombres uno de grado 6. Por último, para el año 2009 se acepta un polinomio de grado 5 para las mujeres y uno de 3 para los hombres.

Las gráficas 2.1, 2.2, 2.3, 2.4 y 2.5 muestran el ajuste de los datos para diferentes grados en el polinomio que depende de la edad y cada una para un periodo de tiempo diferente.

Note que todos los cálculos anteriores son hechos para cada año, lo cual implica que tenemos cálculos estacionarios en el tiempo. Además solo ajustamos para edades adultas, pues la mortalidad infantil es decreciente y ajustar una curva para todas las edades aumentaría la complejidad del modelo, ya que necesitaríamos graduar un polinomio de grado mayor. También note que el grado del polinomio de ajuste para hombres y mujeres es diferente, esto muestra la necesidad de hacer una diferencia para estimar la mortalidad de hombres y mujeres.

Con la anterior elección del grado del polinomio de ajuste, tanto para hombres como para mujeres, se calculó el valor del predictor lineal para los datos del año 2009, estos se indican en la tabla 2.11.

Por último, al utilizar los valores de la Tabla 2.11 y despejar q_x de la ecuación (2.6), se obtiene un ajuste a la probabilidad de muerte para el año 2009. En realidad, lo que nos interesa estudiar es la supervivencia de los asegurados, es por esto que si se utiliza la relación descrita en la ecuación (1.10), a partir de la probabilidad de muerte, se consigue la

HOMBRES 1990							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	270257.4634	4350.0679	746.9260	294.6735	231.5372	229.9795	132.0033
gl	69	68	67	66	65	64	63
Log-ver	-964441.0532	-829159.4542	-826878.8917	-826733.8989	-826707.9984	-826707.3342	-826655.4320
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	112.7410	89.4024	60.8727	46.6223	38.3005	36.5208	36.4760
gl	62	61	60	59	58	57	56
Log-ver	-826643.9748	-826630.4604	-826614.2343	-826605.8275	-826600.8594	-826599.7039	-826599.6601
MUJERES 1990							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	301975.3389	306.3492	258.3740	183.6304	182.6620	127.7757	119.0398
gl	69	68	67	66	65	64	63
Log-ver	-716417.1014	-563391.3004	-563319.9717	-563306.4839	-563305.2590	-563278.1175	-563273.1135
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	91.1628	69.9819	50.1137	37.2594	33.5394	31.6943	31.6943
gl	62	61	60	59	58	57	56
Log-ver	-563257.8024	-563245.8884	-563234.6772	-563227.2891	-563225.0123	-563223.8868	-563223.8865

Tabla 2.1: Medidas de bondad de ajuste para hombres y mujeres de LGM con diferentes grados de x .

HOMBRES 1995							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	291174.5974	4357.3005	354.8939	63.5053	46.0840	15.6280	5.1101
gl	69	68	67	66	65	64	63
Log-ver	-1007011.2109	-861346.9624	-858858.6585	-858782.8277	-858778.5552	-858761.9877	-858756.9696
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	2.0946	1.9192	1.7791	1.5999	1.4133	1.0071	0.7742
gl	62	61	60	59	58	57	56
Log-ver	-858755.4636	-858755.3608	-858755.2991	-858755.2045	-858755.1126	-858754.8964	-858754.7631
MUJERES 1995							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	322295.3273	250.1161	116.5876	39.3199	16.3451	4.7344	1.5739
gl	69	68	67	66	65	64	63
Log-ver	-767698.1882	-604553.2005	-604408.8171	-604395.8987	-604379.9739	-604375.0477	-604373.4058
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	1.5689	1.5392	1.5380	1.5141	1.3536	0.8362	0.7301
gl	62	61	60	59	58	57	56
Log-ver	-604373.4014	-604373.3839	-604373.3831	-604373.3714	-604373.2874	-604373.0110	-604372.9468

Tabla 2.2: Medidas de bondad de ajuste para hombres y mujeres de LGM con diferentes grados de x .

HOMBRES 2000							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	332359.9552	4372.3503	352.0183	36.1073	23.6143	9.526	1.4592
gl	69	68	67	66	65	64	63
Log-ver	-1096353.8484	-929827.3022	-927312.7766	-927231.1853	-927229.0452	-927221.0618	-927217.1436
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	0.8872	0.8189	0.6728	0.6617	0.5639	0.4605	0.3771
gl	62	61	60	59	58	57	56
Log-ver	-927216.8844	-927216.8464	-927216.7666	-927216.7591	-927216.7118	-927216.6560	-927216.6096
MUJERES 2000							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	365789.2021	357.0254	173.2895	47.9310	20.9477	2.5366	1.1768
gl	69	68	67	66	65	64	63
Log-ver	-866587.7152	-681514.6745	-681328.4030	-681300.6459	-681282.0757	-681273.9726	-681273.2904
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	1.0119	0.9212	0.9180	0.9083	0.8147	0.5956	0.5350
gl	62	61	60	59	58	57	56
Log-ver	-681273.2033	-681273.1528	-681273.1520	-681273.1464	-681273.0994	-681272.9828	-681272.9473

Tabla 2.3: Medidas de bondad de ajuste para hombres y mujeres de LGM con diferentes grados de x .

HOMBRES 2005							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	383898.8321	4422.1736	365.1800	25.9531	22.6671	4.8608	1.8238
gl	69	68	67	66	65	64	63
Log-ver	-1217944.7138	-1025202.7457	-1022627.6578	-1022542.1468	-1022542.8437	-1022532.4980	-1022531.1818
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	0.8539	0.8136	0.6100	0.5930	0.2969	0.2744	0.2676
gl	62	61	60	59	58	57	56
Log-ver	-1022530.6700	-1022530.6438	-1022530.5289	-1022530.5231	-1022530.3728	-1022530.3588	-1022530.35423
MUJERES 2005							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	425384.2249	466.7871	227.4044	68.3973	18.9540	5.0397	1.5550
gl	69	68	67	66	65	64	63
Log-ver	-998152.9355	-782882.1084	-782646.0792	-782609.4360	-782577.0070	-782571.3154	-782569.4588
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	1.4271	1.4113	1.1158	0.7350	0.4133	0.4109	0.4098
gl	62	61	60	59	58	57	56
Log-ver	-782569.3799	-782569.3683	-782569.2065	-782569.0232	-782568.8536	-782568.8514	-782568.8503

Tabla 2.4: Medidas de bondad de ajuste para hombres y mujeres de LGM con diferentes grados de x .

HOMBRES 2009							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	405279.6125	9150.5381	969.884	133.0698	111.6291	13.1611	2.9086
gl	69	68	67	66	65	64	63
Log-ver	-1369962.883	-1168718.243	-1163807.393	-1163536.303	-1163531.5	-1163478.885	-1163474.378
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	2.78	2.4441	2.2222	2.1235	1.6568	1.3002	1.2763
gl	62	61	60	59	58	57	56
Log-ver	-1163474.323	-1163474.155	-1163474.035	-1163473.978	-1163473.749	-1163473.571	-1163473.556
MUJERES 2009							
	LGM(0,1)	LGM(0,2)	LGM(0,3)	LGM(0,4)	LGM(0,5)	LGM(0,6)	LGM(0,7)
Devianza	473937.5812	832.0491	458.2614	136.1454	26.1228	7.0316	4.2240
gl	69	68	67	66	65	64	63
Log-ver	-1119217.349	-879462.84	-879120.86	-879024.62	-878957.673	-878949.806	-878948.281
	LGM(0,8)	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)	LGM(0,13)	LGM(0,14)
Devianza	3.6830	2.9407	1.4576	1.4575	0.9131	0.0538	0.0481
gl	62	61	60	59	58	57	56
Log-ver	-878947.976	-878947.629	-878946.863	-878946.863	-878946.594	-878946.407	-878946.378

Tabla 2.5: Medidas de bondad de ajuste para hombres y mujeres de LGM con diferentes grados de x .

	MUJERES			HOMBRES		
	Devianza	p-valor	AIC	Devianza	pvalor	AIC
LGM(0,1) vs LGM(0,2)	301668.9897	0	958.4197	265907.3956	0	5038.5805
LGM(0,2) vs LGM(0,3)	47.9752	0	912.4445	3603.1418	0	1437.4387
LGM(0,3) vs LGM(0,4)	74.7436	0	839.7009	452.2526	0	987.1861
LGM(0,4) vs LGM(0,5)	0.9684	0.3251	840.7324	63.1362	0	926.0499
LGM(0,5) vs LGM(0,6)	54.8863	0	787.8462	1.5578	0.2119	926.4921
LGM(0,6) vs LGM(0,7)	8.7359	0.0031	781.1103	97.9762	0	830.5159
LGM(0,7) vs LGM(0,8)	27.8770	0	755.2332	19.2623	0	813.2537
LGM(0,8) vs LGM(0,9)	21.1809	0	736.0523	23.3386	0	791.9151
LGM(0,9) vs LGM(0,10)	19.8682	0	718.1841	28.5297	0	765.3853
LGM(0,10) vs LGM(0,11)	12.8543	0.0003	707.3298	14.2504	0.0001	753.1330
LGM(0,11) vs LGM(0,12)	3.7200	0.0538	705.6099	8.3218	0.0039	746.8132
LGM(0,12) vs LGM(0,13)	1.8451	0.1744	705.7648	1.7797	0.1822	747.0334
LGM(0,13) vs LGM(0,14)	5.6035×10^{-06}	0.99811	707.7648	0.0448	0.83229	748.9886

Tabla 2.6: AIC hombres y mujeres para el año 1990.

	MUJERES			HOMBRES		
	Devianza	pvalor	AIC	Devianza	pvalor	AIC
LGM(0,1) vs LGM(0,2)	322045.2112	0	905.4349	286817.2969	0	5046.7959
LGM(0,2) vs LGM(0,3)	133.5286	0	773.9064	4002.4067	0	1046.3892
LGM(0,3) vs LGM(0,4)	77.2676	0	698.6387	291.3885	0	757.0007
LGM(0,4) vs LGM(0,5)	22.9749	0	677.6639	17.4213	0	741.5793
LGM(0,5) vs LGM(0,6)	11.6106	0.0007	668.0532	30.4560	0	713.1234
LGM(0,6) vs LGM(0,7)	3.1605	0.0754	666.8927	10.5180	0.0012	704.6054
LGM(0,7) vs LGM(0,8)	0.0050	0.9436	668.8876	3.0154	0.0825	703.5900
LGM(0,8) vs LGM(0,9)	0.0297	0.8631	670.8580	0.1754	0.6753	705.4145
LGM(0,9) vs LGM(0,10)	0.0012	0.9726	672.8568	0.1401	0.7082	707.2745
LGM(0,10) vs LGM(0,11)	0.0239	0.8772	674.8330	0.1792	0.6721	709.0953
LGM(0,11) vs LGM(0,12)	0.1605	0.6887	676.6724	0.1866	0.6657	710.9086
LGM(0,12) vs LGM(0,13)	0.5174	0.4720	678.1550	0.4062	0.5239	712.5025
LGM(0,13) vs LGM(0,14)	0.1061	0.7446	680.0490	0.2329	0.6294	714.2696

Tabla 2.7: AIC hombres y mujeres para el año 1995.

	MUJERES			HOMBRES		
	Devianza	pvalor	AIC	Devianza	p-valor	AIC
LGM(0,1) vs LGM(0,2)	365432.1767	0	1019.105	327987.6050	0	5066.1830
LGM(0,2) vs LGM(0,3)	183.7359	0	837.3693	4020.3319	0	1047.8511
LGM(0,3) vs LGM(0,4)	125.3584	0	714.0109	315.9110	0	733.9401
LGM(0,4) vs LGM(0,5)	26.9833	0	689.0276	12.4930	0.0004	723.4471
LGM(0,5) vs LGM(0,6)	18.4112	0	672.6164	14.0880	0.0002	711.3591
LGM(0,6) vs LGM(0,7)	1.3597	0.2436	673.2567	8.0671	0.0045	705.2920
LGM(0,7) vs LGM(0,8)	0.1650	0.6846	675.0917	0.5721	0.4494	706.7200
LGM(0,8) vs LGM(0,9)	0.0906	0.7634	677.0011	0.0683	0.7938	708.6517
LGM(0,9) vs LGM(0,10)	0.0032	0.9546	678.9979	0.1461	0.7023	710.5056
LGM(0,10) vs LGM(0,11)	0.0097	0.9217	680.9882	0.0111	0.9161	712.4945
LGM(0,11) vs LGM(0,12)	0.0936	0.7597	682.8946	0.0978	0.7545	714.3967
LGM(0,12) vs LGM(0,13)	0.2192	0.6397	684.6754	0.1033	0.7479	716.2933
LGM(0,13) vs LGM(0,14)	0.0606	0.8055	686.6148	0.0835	0.7727	718.2099

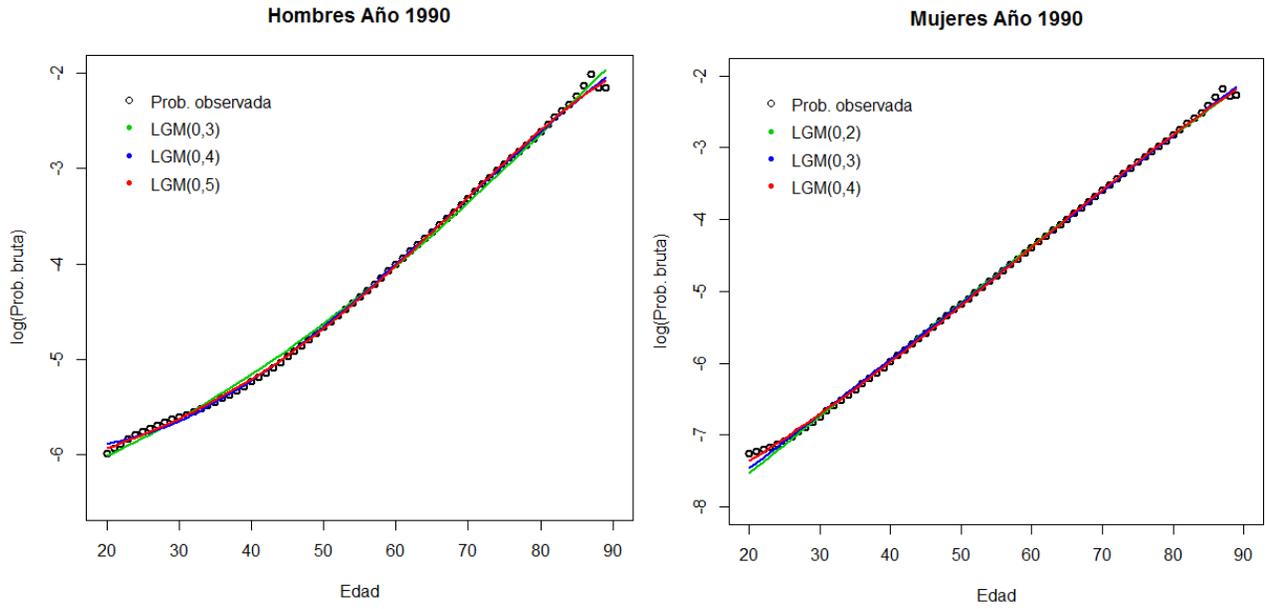
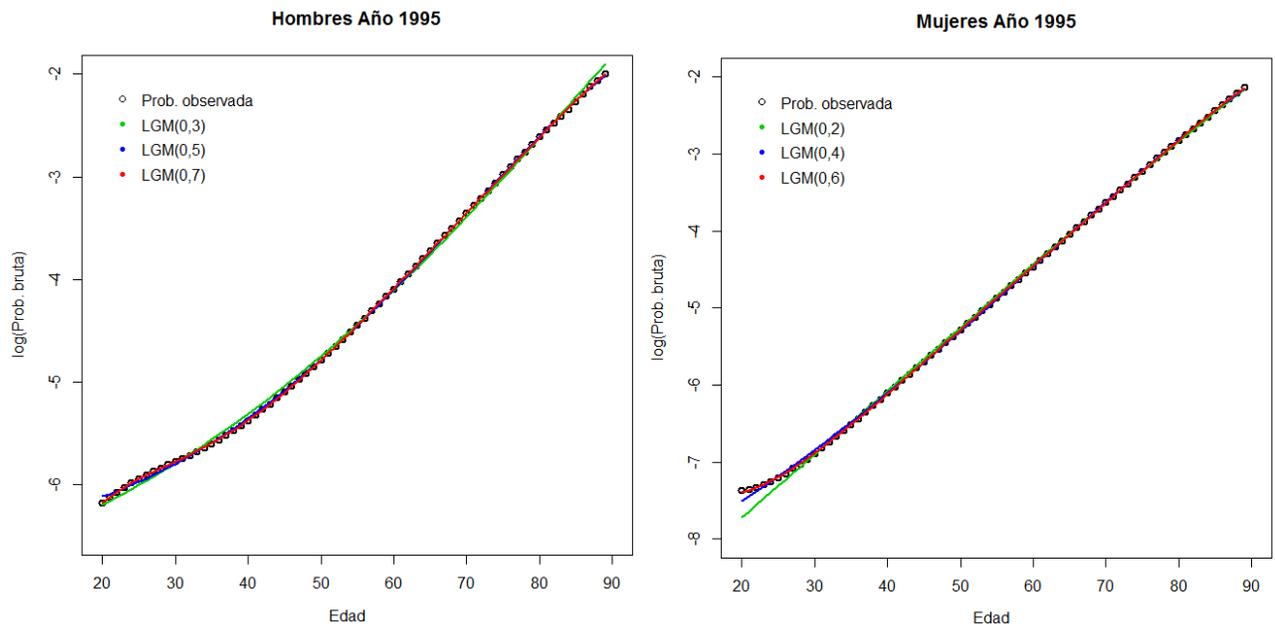
Tabla 2.8: AIC hombres y mujeres para el año 2000.

	MUJERES			HOMBRES		
	Devianza	pvalor	AIC	Devianza	pvalor	AIC
LGM(0,1) vs LGM(0,2)	424917.4378	0	1137.2274	379476.6584	0	5122.3686
LGM(0,2) vs LGM(0,3)	239.3828	0	899.8447	4056.9936	0	1067.3749
LGM(0,3) vs LGM(0,4)	159.0071	0	742.8376	339.2269	0	730.1480
LGM(0,4) vs LGM(0,5)	49.4433	0	695.3943	3.2860	0.0699	728.8620
LGM(0,5) vs LGM(0,6)	13.9143	0.0002	683.4799	17.8062	0	713.0558
LGM(0,6) vs LGM(0,7)	3.4846	0.0619	681.9953	3.0369	0.0814	712.0188
LGM(0,7) vs LGM(0,8)	0.1279	0.7206	683.8674	0.9699	0.3247	713.0489
LGM(0,8) vs LGM(0,9)	0.0158	0.8999	685.8516	0.0403	0.8408	715.0085
LGM(0,9) vs LGM(0,10)	0.2955	0.5867	687.5561	0.2036	0.6518	716.8049
LGM(0,10) vs LGM(0,11)	0.3808	0.5372	689.1753	0.0169	0.8963	718.7879
LGM(0,11) vs LGM(0,12)	0.3217	0.5706	690.8536	0.2961	0.5863	720.4918
LGM(0,12) vs LGM(0,13)	0.0024	0.9609	692.8512	0.0224	0.8810	722.4694
LGM(0,13) vs LGM(0,14)	0.0011	0.9735	694.8501	0.0069	0.9339	724.4625

Tabla 2.9: AIC hombres y mujeres para el año 2005.

	MUJERES			HOMBRES		
	Devianza	p-valor	AIC	Devianza	p-valor	AIC
LGM(0,1) vs LGM(0,2)	473105.5321	0	1509.937	396129.074	0	9860.573
LGM(0,2) vs LGM(0,3)	373.7877	0	1138.149	8180.655	0	1681.919
LGM(0,3) vs LGM(0,4)	322.1159	0	818.034	836.814	0	847.105
LGM(0,4) vs LGM(0,5)	110.0226	0	710.011	21.441	0	827.664
LGM(0,5) vs LGM(0,6)	19.0912	0	692.919	98.468	0	731.196
LGM(0,6) vs LGM(0,7)	2.8076	0.094	692.112	10.253	0.00136	722.944
LGM(0,7) vs LGM(0,8)	0.541	0.462	693.571	0.1285	0.7199	724.815
LGM(0,8) vs LGM(0,9)	0.7423	0.389	694.828	0.3359	0.5621	726.479
LGM(0,9) vs LGM(0,10)	1.4831	0.223	695.345	0.2219	0.6376	728.257
LGM(0,10) vs LGM(0,11)	5.6684×10^{-05}	0.994	697.345	0.0987	0.7533	730.159
LGM(0,11) vs LGM(0,12)	0.5444	0.4606	698.801	0.4667	0.4945	731.692
LGM(0,12) vs LGM(0,13)	0.3754	0.5401	700.425	0.3566	0.5504	733.336
LGM(0,13) vs LGM(0,14)	0.0564	0.812	702.369	0.0238	0.8773	735.312

Tabla 2.10: AIC hombres y mujeres para el año 2009.

Figura 2.1: Gráfica de la graduación de q_x .Figura 2.2: Gráfica de la graduación de q_x .

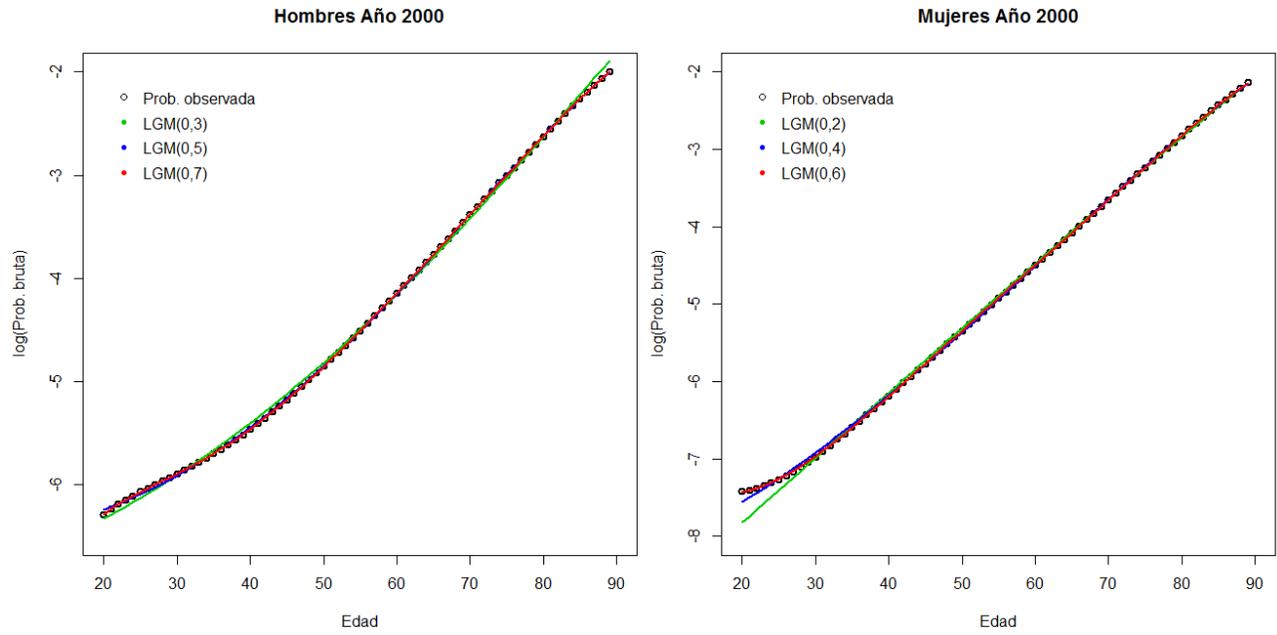


Figura 2.3: Gráfica de la graduación de q_x .

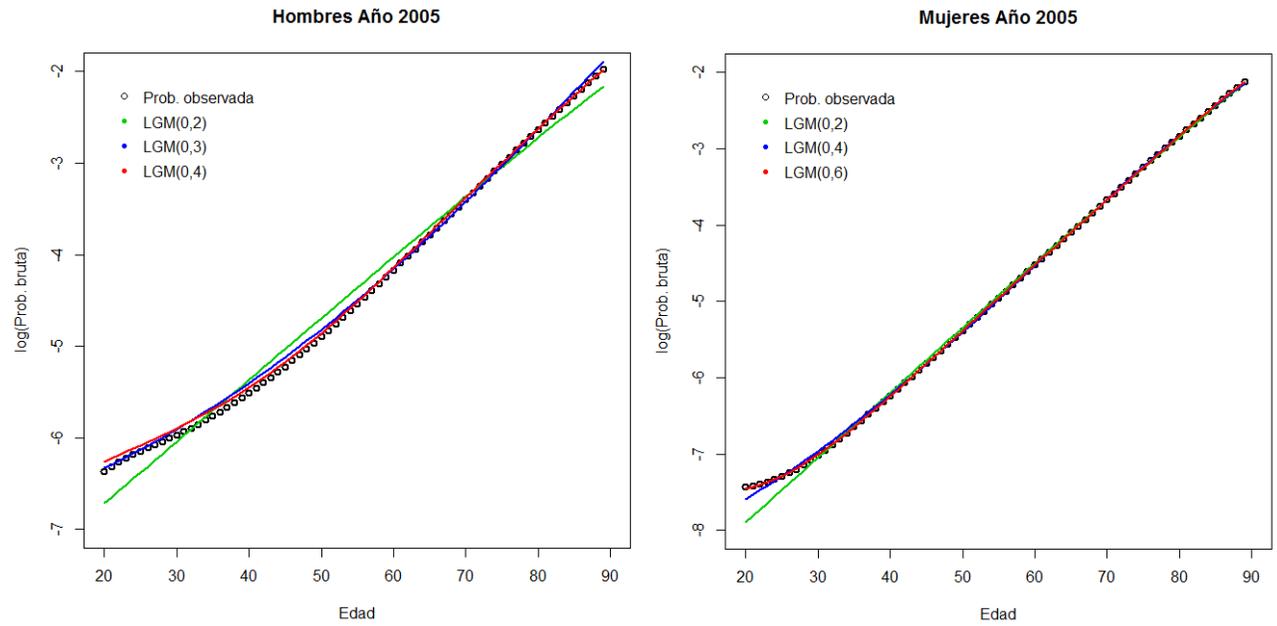


Figura 2.4: Gráfica de la graduación de q_x .

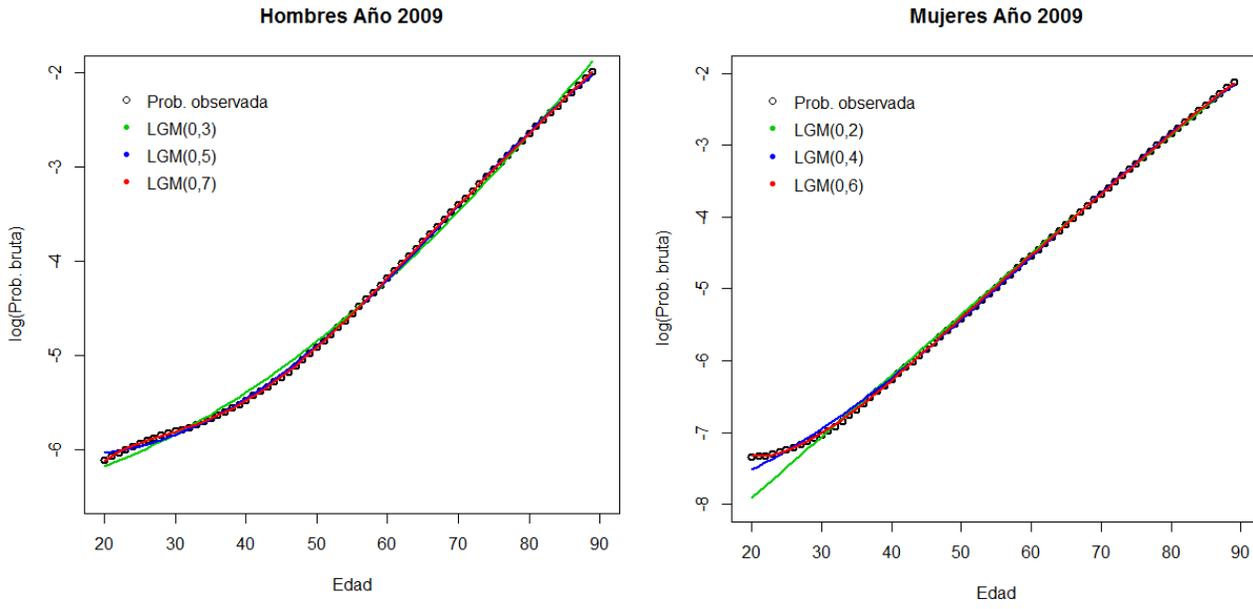


Figura 2.5: Gráfica de la graduación de q_x .

HOMBRES						
α_1	α_2	α_3	α_4	α_5	α_6	α_7
-11.5853	0.6896	-0.0340	8.4920×10^{-04}	-1.0933×10^{-05}	7.1022×10^{-08}	-1.8477×10^{-06}
MUJERES						
-4.3213	-0.3638	0.0146	2.3310×10^{-04}	1.8248×10^{-06}	-5.6279×10^{-09}	—

Tabla 2.11: Valores del predictor lineal para hombres y mujeres del año 2009.

probabilidad de supervivencia para hombres y mujeres de 20 a 89 años, la cual se muestra en la Tabla 2.12.

Se aplicó el test Kolmogorov-Smirnov para validar el ajuste de los datos a cada uno de los polinomios elegidos para cada año, lo anterior se reporta en la tabla 2.13.

2.2. Modelo de Lee-Carter

Después de haber estudiado en la sección anterior la graduación de datos de mortalidad con funciones paramétricas que solo se centran en la influencia que la edad tiene sobre éstas, se pretende ahora analizar el efecto del tiempo del calendario. Se ha observado que la probabilidad de muerte a una determinada edad calculada en periodos diferentes cambia, por lo tanto la hipótesis de estacionariedad no se cumple.

Siguiendo con la idea de desarrollar métodos que dependan ahora también del tiempo, tenemos el desarrollado por Lee y Carter (1992) [12]. Este modelo trata de ajustar una función lineal a los logaritmos de los radios centrales de mortalidad observados para cada edad específica. A continuación se presentará dicho modelo.

Definición 2.2. El radio central de mortalidad m_x^t es el índice de muerte medio sobre el intervalo de edad x a $x + 1$ en el año cronológico t que relaciona las muertes con la población en el centro del año. Además puede relacionarse con q_x^t a través de la expresión,

$$m_x^t = \frac{2q_x^t}{2 - q_x^t}. \quad (2.12)$$

Lee-Carter proponen construir una matriz formada por los radios centrales de muerte ordenados por filas de acuerdo a la edad x y columnas según el año t . A esta matriz, se ajusta la función

$$m_x^t = \exp(a_x + b_x k_t) + \epsilon_x^t, \quad (2.13)$$

o, equivalentemente, la siguiente función a la matriz de los logaritmos,

$$\ln(m_x^t) = a_x + b_x k_t + \epsilon_x^t. \quad (2.14)$$

Edad	Hombres	Mujeres	Edad	Hombres	Mujeres
20	0.9978114647	0.9993328276	55	0.9895347439	0.9930775719
21	0.9977008899	0.9993396309	56	0.9887249371	0.9924476635
22	0.9976009100	0.9993389593	57	0.9878424820	0.9917608181
23	0.9975103620	0.9993314970	58	0.9868819107	0.9910119906
24	0.9974278016	0.9993176834	59	0.9858374376	0.9901957029
25	0.9973516046	0.9992977679	60	0.9847029415	0.9893060146
26	0.9972800497	0.9992718478	61	0.9834719444	0.9883364936
27	0.9972113815	0.9992398959	62	0.9821375845	0.9872801861
28	0.9971438557	0.9992017797	63	0.9806925850	0.9861295872
29	0.9970757671	0.9991572749	64	0.9791292155	0.9848766110
30	0.9970054646	0.9991060750	65	0.9774392473	0.9835125622
31	0.9969313552	0.9990477975	66	0.9756139005	0.9820281072
32	0.9968518994	0.9989819880	67	0.9736437845	0.9804132472
33	0.9967656016	0.9989081221	68	0.9715188289	0.9786572900
34	0.9966709952	0.9988256061	69	0.9692282087	0.9767488235
35	0.9965666262	0.9987337761	70	0.9667602612	0.9746756885
36	0.9964510345	0.9986318964	71	0.9641023989	0.9724249495
37	0.9963227351	0.9985191564	72	0.9612410175	0.9699828636
38	0.9961801986	0.9983946663	73	0.9581614044	0.9673348449
39	0.9960218327	0.9982574530	74	0.9548476495	0.9644654224
40	0.9958459637	0.9981064534	75	0.9512825645	0.9613581874
41	0.9956508185	0.9979405077	76	0.9474476156	0.9579957278
42	0.9954345078	0.9977583514	77	0.9433228792	0.9543595442
43	0.9951950100	0.9975586060	78	0.9388870283	0.9504299412
44	0.9949301555	0.9973397692	79	0.9341173626	0.9461858891
45	0.9946376121	0.9971002028	80	0.9289898961	0.9416048461
46	0.9943148715	0.9968381202	81	0.9234795182	0.9366625342
47	0.9939592359	0.9965515725	82	0.9175602476	0.9313326574
48	0.9935678059	0.9962384331	83	0.9112056014	0.9255865525
49	0.9931374693	0.9958963808	84	0.9043891018	0.9193927599
50	0.9926648905	0.9955228822	85	0.8970849461	0.9127165028
51	0.9921465000	0.9951151715	86	0.8892688661	0.9055190619
52	0.9915784857	0.9946702304	87	0.8809192003	0.8977570327
53	0.9909567833	0.9941847653	88	0.8720181980	0.8893814547
54	0.9902770678	0.9936551843	89	0.8625535708	0.8803368024

Tabla 2.12: Probabilidad de supervivencia del año 2009 para hombres y mujeres de 20 a 89 años, modelo Gompertz-Makeham.

Año	Polinomio de ajuste	Kolmogorov Mujeres	p-valor Mujeres	Polinomio de ajuste	Kolmogorov Hombres	p-valor Hombres
1990	LGM(0,6)	0.02857	1	LGM(0,7)	0.02857	1
1995	LGM(0,4)	0.04286	0.999	LGM(0,5)	0.02857	1
2000	LGM(0,6)	0.01429	1	LGM(0,7)	0.01429	1
2005	LGM(0,6)	0.01428	1	LGM(0,7)	0.01428	1
2009	LGM(0,6)	0.01428	1	LGM(0,4)	0.04286	0.999

Tabla 2.13: Prueba Kolmogorov de bondad de ajuste.

Cabe señalar que por tratarse de una función exponencial las predicciones de los radios centrales de mortalidad no pueden tomar valores negativos, lo que por otra parte sería totalmente absurdo. La interpretación de los parámetros es la que sigue:

- a_x es una constante que depende de la edad y describe el perfil general a lo largo de la edad del esquema de mortalidad.
- b_x es otra constante que también depende de la edad y refleja lo rápido que decrecen los índices en respuesta a los cambios de k_t $\left(\frac{d \ln m_{xt}}{dt} = b_x \frac{dk_t}{dt}\right)$.
- Los valores k_t representan la tendencia de la mortalidad a lo largo del periodo t .
- Los errores ϵ_x^t , con media cero y varianza σ_ϵ^2 , reflejan influencias históricas particulares de cada edad específica que no son explicadas por el modelo.

Dada una matriz de radios m_x^t , para estimar los parámetros del modelo en uno de los pasos vamos a recurrir a la solución por mínimos cuadrados de la ecuación (2.14). Sin embargo, este modelo es indeterminado puesto que dada una solución (a, b, k) , son soluciones también $(a - bc, b, k + c)$ y $(a, bc, k/c)$, donde c es un escalar cualquiera. Así pues, para que la solución sea única, es necesario normalizar b_x para que sumen 1 y k_t para sumar 0, lo cual implica que los a_x son simplemente las medias a lo largo del tiempo de $\ln(m_x^t)$. En efecto,

$$\begin{aligned}
\sum_t \sum_x \ln(m_x^t) &= \sum_t \sum_x (a_x + b_x k_t) \\
&= \sum_t \sum_x a_x + \sum_t \sum_x b_x k_t \\
&= \sum_x a_x \sum_t 1 + \sum_t k_t \sum_x b_x.
\end{aligned} \tag{2.15}$$

Si $\sum_t k_t = 0$ y $\sum_x b_x \neq 0$, desaparece el segundo sumando de la ecuación (2.15) y entonces,

$$\sum_t \sum_x \ln(m_{xt}) = T \sum_x a_x. \quad (2.16)$$

Luego,

$$a_x = \frac{\sum_t \ln(m_{xt})}{T}, \quad (2.17)$$

cumpliría con la con la restricción. Entonces,

$$\sum_x a_x = \sum_x \frac{\sum_t \ln(m_{xt})}{T} = \frac{\sum_t \sum_x \ln(m_{xt})}{T}.$$

Según Lee y Carter [12] los b_x podrían ser negativos indicando que la mortalidad a ciertas edades tiende a subir mientras que para otras cae, pero en la práctica esto no ocurre a lo largo de un periodo largo de años.

El modelo (2.14) no puede ser ajustado por métodos de regresión habituales, pues a su izquierda tiene la variable dependiente $\ln(m_x^t)$ pero a su derecha tiene una función cuyos parámetros acompañan a un índice k_t y no a una variable independiente con valores observables. Por ello, aplicaremos el método de descomposición de SVD a la matriz de los logaritmos de los radios centrales de mortalidad (2.14) menos los valores de a_x (ecuación 2.17) para encontrar la solución por mínimos cuadrados. Tenemos pues que resolver el sistema

$$\ln(m_x^t) - \frac{\sum_t \ln(m_x^t)}{T} = b_x k_t. \quad (2.18)$$

Sea $Z_x^t = \ln(m_x^t) - \frac{\sum_t \ln(m_x^t)}{T}$ entonces podemos estimar, respectivamente, k_t y b_x como los primeros vectores derecho e izquierdo correspondientes al principal valor de la SVD de la matriz Z_x^t , o bien, estimar k_t como $\hat{k}_t = \sum_x Z_x^t$ para después determinar b_x a partir de la regresión de (2.18), una vez sustituidos los valores \hat{k}_t .

Los radios centrales de mortalidad calculados de esta manera no estiman el número real de muertes para el total de edades, puesto que k_t ha sido estimada minimizando los errores en los logaritmos y no en los radios de muerte. Por ello, se utiliza la ecuación

$$D^t = \sum_x (N_x^t \exp(a_x + k_t b_x)); \quad (2.19)$$

donde D^t es el número total de muertes y N_x^t es la población de edad x en el año t .

Así reestimamos k_t sustituyendo los valores que ya teníamos de a_x en la ecuación (2.4) y los de b_x que obtuvimos al emplear la descomposición SVD. Estos valores de k_t se encuentran mediante búsqueda iterativa y difieren de la estimaciones directas mediante SVD.

Una vez desarrollado y ajustado el modelo demográfico, vamos a centrarnos en el problema de la predicción. El último paso del método Lee-Carter es encontrar un modelo para la predicción de los valores de los índices de mortalidad, k_t . Se hace uso de la metodología Box-Jenkins que se aplica a los modelos autorregresivos de media móvil ARMA para encontrar el mejor ajuste de una serie temporal de valores, a fin de que los pronósticos sean más acertados. Los autores llegan a la siguiente expresión

$$k_t = c + k_{t-1} + u_t, \quad (2.20)$$

donde c es una constante y u_t un ruido blanco. Con este modelo, la predicción de k_t varía linealmente y cada radio de muerte predicho varía con radio exponencial constante.

Note que este modelo tiene un parámetro que depende del tiempo y por lo tanto es dinámico. Tiene sencillez en cuanto a los parámetros que utiliza y puede aplicarse aún cuando se tengan pocos datos. Es ampliamente utilizado por lo que está implementado en diferentes *softwares*.

Proporciona tablas de vida a partir de la información pasada para predecir la mortalidad hasta 75 años futuros como se ha hecho en Australia [9].

2.2.1. Modelo Lee-Carter para la población mexicana

El paquete *Demography* implementado en el lenguaje R está basado en la metodología propuesta por Lee-Carter que describimos anteriormente. Éste fue lanzado en el 2010 y diseñado por Rob J. Hyndman en colaboración con Heather Bothoo, Tickle Leonie y Maindonald doctores e investigadores del Departamento de Econometría y Estadística Empresarial de la Universidad de Monash, Victoria, Australia. Una de las principales fortalezas de este paquete es la construcción tablas de vida dinámicas haciendo uso de este modelo.

Esta vez, se usó la base proporcionada por la CONAPO [20] del número de población viva a mitad de cada año del periodo comprendido del año 1990 al 2009 y la población muerta al final del año de dicho periodo. Para las edades de 0 a 89 se graduó la probabilidad de muerte mediante el uso de la ecuación (2.12), también una predicción para un futuro comportamiento de la mortalidad.

Para hacer el análisis, al igual que en el modelo anterior, se hizo diferencia entre hombres y mujeres. En las figuras 2.6 a) y 2.7 a) se observa la graduación de la probabilidad de muerte de la información histórica de hombres y mujeres respectivamente, mientras que en las figuras 2.6 b) y 2.7 b) se ilustra una predicción según el modelo Lee-Carter para una probabilidad de muerte futura. En comparación con el de Gompertz-Makeham, este modelo ajusta la probabilidad de muerte en edades infantiles sin necesidad de utilizar más parámetros y por lo tanto complicar el modelo.

El comportamiento de los coeficientes a_x , b_x y k_t se ilustra en las figuras 2.8 y 2.9 para hombres y mujeres. En ambos casos, el coeficiente k_t es decreciente. Esto de acuerdo con el hecho de que la mortalidad disminuye a lo largo del tiempo, lo cual conduce a un aumento en la esperanza de vida. En Tabla 2.14 se registró el valor de k_t para cada uno de los años 1990 a 2009 mientras que la tabla 2.15 contiene información del valor de las predicciones hasta el año 2015 de k_t así como su intervalo de error. Observe que el intervalo de error crece a medida que el tiempo transcurre, lo cual hace que para periodos en el futuro más largos la predicción sea más inexacta. La tabla 2.16 contiene información de los valores de los coeficientes a_x y

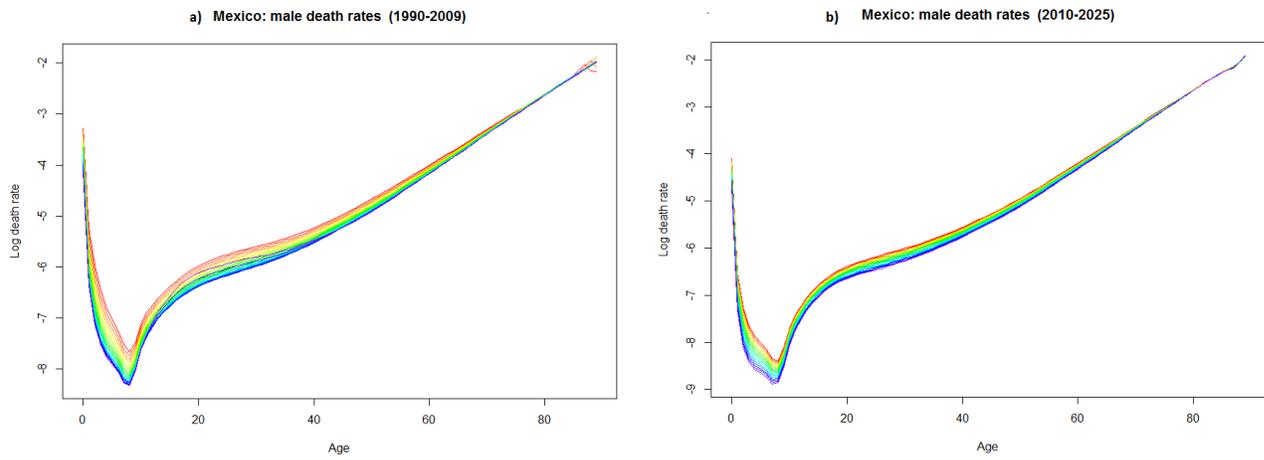


Figura 2.6: Las curvas de arriba hacia abajo representan, de menor a mayor año el logaritmo de la probabilidad de muerte.

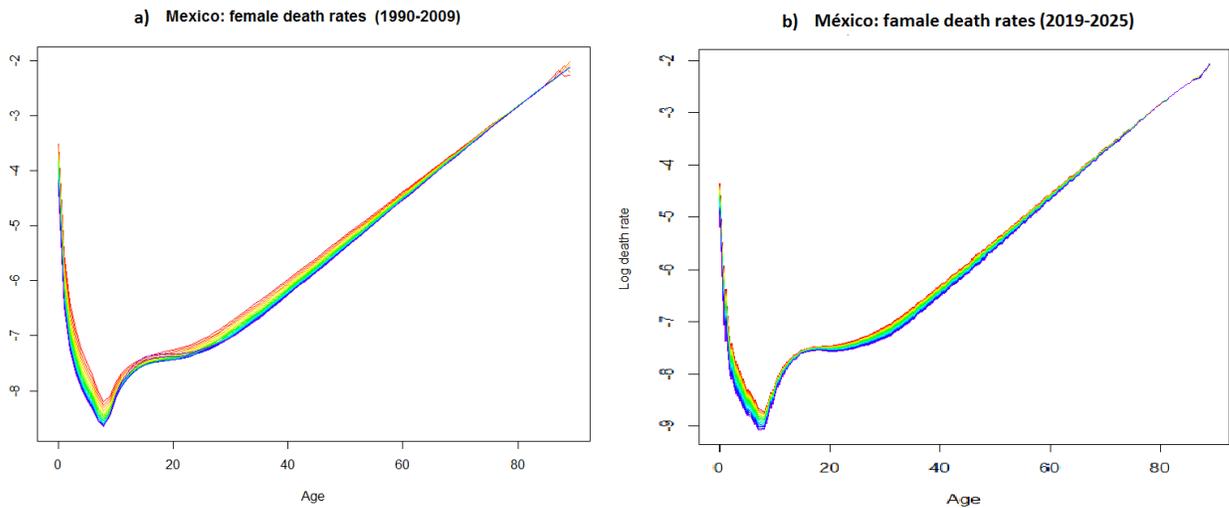


Figura 2.7: Las curvas de arriba hacia abajo representan, de menor a mayor año el logaritmo de la probabilidad de muerte.

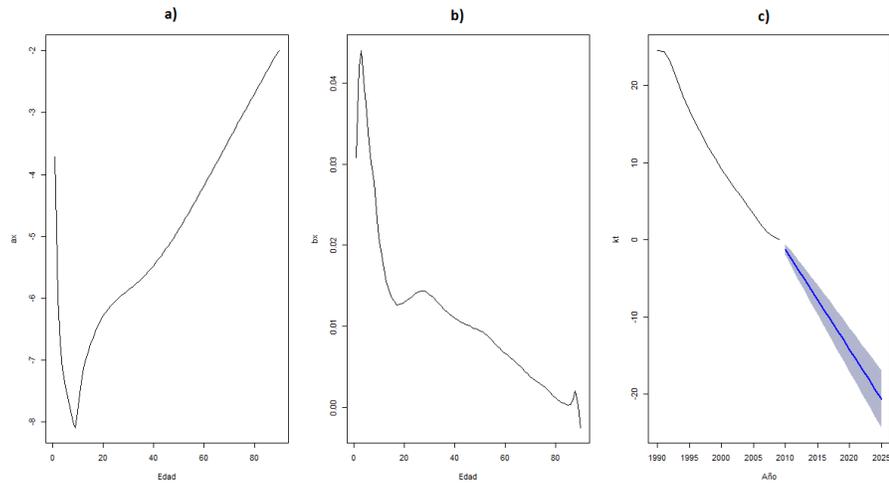


Figura 2.8: Hombres mexicanos. a) comportamiento con respecto a la edad del coeficiente a_x , b) comportamiento con respecto a la edad del coeficiente b_x y c) comportamiento con respecto del tiempo de k_t , además una predicción futura y un intervalo de error.

b_x para ambos sexos.

Con la información de las tablas anteriores y al utilizar la relación definida en la ecuación (2.12) entre la probabilidad de muerte y el radio central de mortalidad, se construyeron las tablas 2.19 y 2.20 que son estimaciones de la probabilidad de supervivencia para el periodo 2010-2014 de hombres y mujeres respectivamente.

Para medir el ajuste de los datos al modelo propuesto por Lee-Carter, se compararon los valores predichos mediante las siguientes medidas de ajuste:

- Error cuadrático medio

$$ECM = \frac{1}{N} \sum_{t=1}^N \varepsilon_t^2.$$

- Error absoluto medio

	HOMBRES	MUJERES
Año	coeficiente k_t	coeficiente k_t
1990	24.5481252	18.3314329
1991	24.4037243	18.7663708
1992	23.0898101	18.0439539
1993	20.7637104	16.1386577
1994	18.6201545	14.4351189
1995	16.7094241	13.0876169
1996	15.0203636	11.9709862
1997	13.4623587	10.9543497
1998	11.9445715	9.9242337
1999	10.5066008	8.9979196
2000	9.2007393	7.9715994
2001	7.9571610	7.2090589
2002	6.7841594	6.4417727
2003	5.6346014	5.6923508
2004	4.5214654	4.9254400
2005	3.3162896	4.1105463
2006	2.1754842	3.0510904
2007	1.0583809	1.9782058
2008	0.5329322	0.9429035
2009	0.0000000	0.0000000

Tabla 2.14: Valores del ajuste del coeficiente k_t de 1990 a 2009 .

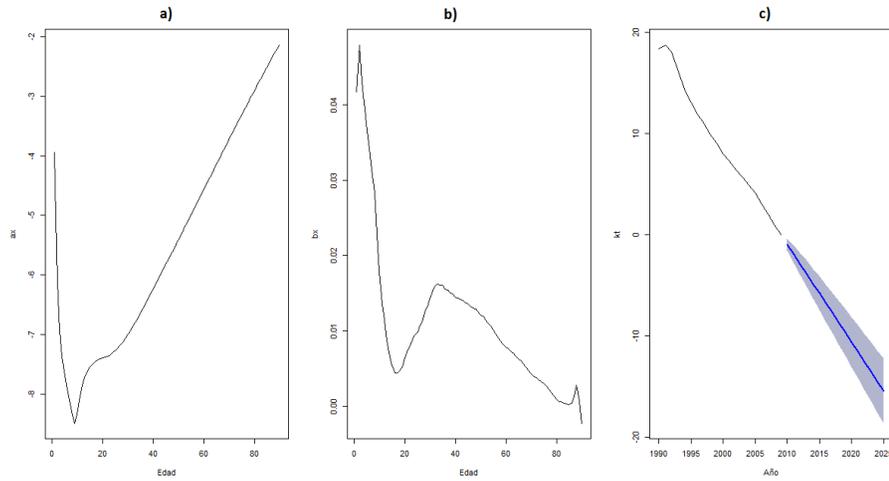


Figura 2.9: Mujeres mexicanas. a) comportamiento con respecto a la edad del coeficiente a_x , b) comportamiento con respecto a la edad del coeficiente b_x y c) comportamiento con respecto del tiempo de k_t , además una predicción futura y un intervalo de error.

Año	HOMBRES			MUJERES		
	coeficiente k_t	Límite inferior	Límite superior	coeficiente k_t	Límite inferior	Límite superior
2010	-1.292007	-1.991148	-0.592865	-0.9648123	-1.572433	-0.3571917
2011	-2.584013	-3.597166	-1.570861	-1.9296245	-2.810150	-1.0490986
2012	-3.876020	-5.146074	-2.605966	-2.8944368	-3.998234	-1.7906392
2013	-5.168026	-6.667518	-3.668535	-3.8592490	-5.162450	-2.5560481
2014	-6.460033	-8.172573	-4.747493	-4.8240613	-6.312422	-3.3357009
2015	-7.752040	-9.666718	-5.837361	-5.7888735	-7.452911	-4.1248360
2016	-9.044046	-11.153091	-6.935001	-6.7536858	-8.586647	-4.9207250
2017	-10.336053	-12.633667	-8.038439	-7.7184980	-9.715343	-5.7216530
2018	-11.628059	-14.109766	-9.146353	-8.6833103	-10.840149	-6.5264712
2019	-12.920066	-15.582318	-10.257814	-9.6481226	-11.961873	-7.3343723
2020	-14.212072	-17.051999	-11.372146	-10.6129348	-13.081101	-8.1447683
2021	-15.504079	-18.519317	-12.488841	-11.5777471	-14.198276	-8.9572182
2022	-16.796086	-19.984662	-13.607509	-12.5425593	-15.313736	-9.7713831
2023	-18.088092	-21.448339	-14.727846	-13.5073716	-16.427746	-10.5869976
2024	-19.380099	-22.910591	-15.849607	-14.4721838	-17.540517	-11.4038503
2025	-20.672105	-24.371615	-16.972596	-15.4369961	-18.652222	-12.2217702

Tabla 2.15: Valores del coeficiente k_t y su respectivo intervalo de error.

Edad	HOMBRES		MUJERES	
	Coficiente	Coficiente	Coficiente	Coficiente
	a_x	b_x	a_x	b_x
0	-3.718093	0.0308275393	-3.941184	0.0416676783
1	-5.986964	0.0419752813	-6.107660	0.0479590069
2	-6.721347	0.0440095703	-6.937169	0.0430646814
3	-7.148953	0.0403776560	-7.375029	0.0403198430
4	-7.422056	0.0363408619	-7.695913	0.0365366060
5	-7.602709	0.0327509080	-7.922543	0.0341753585
6	-7.770669	0.0301593501	-8.105180	0.0313621457
7	-8.001295	0.0277088629	-8.367529	0.0283717230
8	-8.094344	0.0241991037	-8.499569	0.0228087046
9	-7.846687	0.0210194154	-8.337032	0.0178089862
10	-7.449128	0.0188456369	-8.050516	0.0135578931
11	-7.189826	0.0171947445	-7.847373	0.0113755495
12	-7.020819	0.0155139101	-7.720357	0.0087500166
13	-6.875663	0.0144069283	-7.629457	0.0066152907
14	-6.745875	0.0135759195	-7.551842	0.0054006723
15	-6.631205	0.0130602901	-7.497334	0.0044128122
16	-6.527192	0.0126475648	-7.461892	0.0043733621
17	-6.436509	0.0127142959	-7.436467	0.0046766997
18	-6.354923	0.0127845046	-7.415031	0.0053027448
19	-6.288061	0.0130185425	-7.397788	0.0063987949
20	-6.229450	0.0132097848	-7.384684	0.0072604952
21	-6.175594	0.0134651945	-7.366934	0.0079860991
22	-6.127696	0.0136787422	-7.343310	0.0088529451
23	-6.084113	0.0139465771	-7.310829	0.0094378018
24	-6.042638	0.0142272836	-7.273391	0.0098319697
25	-6.004047	0.0143045421	-7.230969	0.0107415937
26	-5.968219	0.0144024754	-7.180194	0.0112368148
27	-5.934477	0.0144299425	-7.126779	0.0127368154
28	-5.901784	0.0142312148	-7.067259	0.0132693897
29	-5.868541	0.0139648582	-7.005303	0.0142458800
30	-5.836905	0.0137730922	-6.938849	0.0153094696
31	-5.804742	0.0134223506	-6.867863	0.0158985245
32	-5.772036	0.0131185699	-6.794306	0.0162044100
33	-5.735428	0.0127132847	-6.718240	0.0160399264
34	-5.695671	0.0124189250	-6.642119	0.0159891799
35	-5.655463	0.0121132478	-6.564017	0.0155782644
36	-5.613865	0.0118068826	-6.484865	0.0154036427
37	-5.569136	0.0115546237	-6.402237	0.0150663718
38	-5.523061	0.0112899108	-6.320340	0.0149256500
39	-5.474624	0.0110869136	-6.239771	0.0144435607
40	-5.423640	0.0108781069	-6.157107	0.0143769175
41	-5.371337	0.0107030386	-6.072933	0.0141927144
42	-5.317341	0.0105106438	-5.989814	0.0140122965
43	-5.261801	0.0103496125	-5.907731	0.0137419981
44	-5.204540	0.0101923027	-5.824879	0.0135863141
45	-5.145521	0.0101046961	-5.742957	0.0132266046
46	-5.085437	0.0098627249	-5.659233	0.0130472748
47	-5.022733	0.0097704232	-5.574418	0.0129107356
48	-4.958348	0.0096755597	-5.491602	0.0125355414
49	-4.893006	0.0095022473	-5.408101	0.0121724300

50	-4.827064	0.0093425062	-5.324321	0.0119543018
51	-4.760124	0.0090695384	-5.240160	0.0114581894
52	-4.691243	0.0088832824	-5.156053	0.0110669506
53	-4.622213	0.0085151623	-5.072064	0.0106595974
54	-4.552122	0.0082586856	-4.988110	0.0100783349
55	-4.481868	0.0079194594	-4.904521	0.0095918649
56	-4.410226	0.0075482066	-4.820812	0.0090539082
57	-4.338741	0.0073028351	-4.736150	0.0085743292
58	-4.266679	0.0069553981	-4.652057	0.0082487485
59	-4.193661	0.0067246158	-4.567930	0.0078517633
60	-4.120442	0.0065036662	-4.484035	0.0076511475
61	-4.046615	0.0062136478	-4.399845	0.0073178935
62	-3.972217	0.0059702086	-4.316030	0.0069783267
63	-3.897611	0.0056785553	-4.231826	0.0066555536
64	-3.822885	0.0054239019	-4.148156	0.0063192040
65	-3.748065	0.0051276352	-4.064275	0.0059975615
66	-3.673127	0.0048079583	-3.980602	0.0056376657
67	-3.597896	0.0044996027	-3.897066	0.0051684767
68	-3.522652	0.0041535663	-3.812912	0.0047277855
69	-3.447226	0.0038656375	-3.729171	0.0043544786
70	-3.371635	0.0036211037	-3.645610	0.0040122439
71	-3.296011	0.0033724101	-3.562000	0.0037691338
72	-3.220244	0.0031406890	-3.478418	0.0034671678
73	-3.144632	0.0029628227	-3.395302	0.0032706584
74	-3.068986	0.0027089472	-3.312161	0.0029596159
75	-2.993529	0.0025121798	-3.229885	0.0026651141
76	-2.918658	0.0021924562	-3.147736	0.0022493175
77	-2.844649	0.0019119018	-3.066327	0.0017905914
78	-2.771188	0.0015415988	-2.985781	0.0013631000
79	-2.697647	0.0012363925	-2.905676	0.0008762799
80	-2.625028	0.0009300133	-2.826302	0.0006469176
81	-2.552220	0.0007462270	-2.747052	0.0005670319
82	-2.479733	0.0006082353	-2.668289	0.0003904582
83	-2.408777	0.0004793749	-2.589971	0.0003376714
84	-2.337874	0.0003379373	-2.511838	0.0001623716
85	-2.265349	0.0003880715	-2.432647	0.0003721520
86	-2.191743	0.0009511080	-2.352045	0.0013026376
87	-2.115854	0.0020628044	-2.269459	0.0028423618
88	-2.050809	0.0001608320	-2.196509	0.0006991063
89	-1.990697	-0.0024678331	-2.127967	-0.0022902134

Tabla 2.16: Valores de los coeficientes a_x y b_x para cada sexo.

$$MAE = \frac{1}{N} \sum_{t=1}^N |\varepsilon_t|.$$

- Error porcentual absoluto medio

$$MAPE = \frac{100}{N} \sum_{t=1}^N \frac{|\varepsilon_t|}{X_t}.$$

- Error medio

$$ME = \frac{1}{N} \sum_{t=1}^N \varepsilon_t.$$

Donde:

N es el número total de observaciones.

ε_t es la diferencia entre los valores observados y los predichos por el modelo en el año t .

X_t es la serie ajustada por el modelo.

Los criterios que se utilizarán para considerar un resultado aceptable son los siguientes (ver Romero R. [18]):

- $0 < ECM < 0,04$
- $0 < MAE < 0,03$
- $0 < ME < 0,03$; $-0,03 < ME < 0$
- $0\% < MAPE < 8\%$

La tabla 2.17 muestra los valores obtenidos de las medidas de ajuste para los hombres de 0 a 89 años del año 1990 al 2009. Podemos observar que el ECM más grande es $6,16 \times 10^{-6}$ que corresponde al año 1990 y está dentro del nivel apropiado. El MAE más alto es

Año	ECM	MAE	MAPE (%)	ME
1990	6.162435×10^{-6}	7.508462×10^{-4}	4.8609420	-6.592270×10^{-5}
1991	2.671682×10^{-6}	3.676055×10^{-4}	1.5270703	-1.155165×10^{-4}
1992	4.686791×10^{-6}	3.857229×10^{-4}	0.6553135	-1.905003×10^{-4}
1993	1.681197×10^{-6}	2.623366×10^{-4}	0.8206537	-3.615416×10^{-5}
1994	4.379522×10^{-7}	2.000487×10^{-4}	1.0229388	4.749700×10^{-5}
1995	1.461536×10^{-7}	1.702954×10^{-4}	1.2558237	7.450778×10^{-5}
1996	1.530721×10^{-7}	1.800476×10^{-4}	1.5516498	7.857767×10^{-5}
1997	1.546345×10^{-7}	1.882945×10^{-4}	1.7844949	7.752471×10^{-5}
1998	1.222627×10^{-7}	1.732232×10^{-4}	1.8749497	7.504974×10^{-5}
1999	1.029605×10^{-7}	1.693363×10^{-4}	1.8673671	7.020254×10^{-5}
2000	7.938609×10^{-8}	1.545414×10^{-4}	1.8615736	5.791043×10^{-5}
2001	5.683938×10^{-8}	1.313645×10^{-4}	1.7356481	3.836897×10^{-5}
2002	4.566792×10^{-8}	1.148710×10^{-4}	1.5750772	1.973728×10^{-5}
2003	4.881448×10^{-8}	1.004435×10^{-4}	1.3279365	5.614361×10^{-6}
2004	5.036607×10^{-8}	8.988877×10^{-5}	1.0648948	-4.006344×10^{-6}
2005	3.973775×10^{-8}	8.501068×10^{-5}	1.0808989	-9.704571×10^{-6}
2006	4.454371×10^{-8}	1.043593×10^{-4}	1.3102050	-1.037744×10^{-5}
2007	7.213679×10^{-8}	1.369931×10^{-4}	1.5915196	-1.149282×10^{-5}
2008	1.165250×10^{-7}	2.149354×10^{-4}	4.5274192	-6.608227×10^{-5}
2009	1.977797×10^{-7}	3.085880×10^{-4}	7.8003335	-1.056359×10^{-4}

Tabla 2.17: Medidas de ajuste hombres de 0 a 89 años. Modelo Lee-Carter

$7,508462 \times 10^{-4}$ que también corresponde al año 1990 y es aceptable. El *ME* más alto es $7,857767 \times 10^{-5}$ del año 1996 pero es aceptable. El *MAPE* más alto es 7.8% y corresponde al año 2009, sin embargo es aceptable.

La tabla 2.18 muestra los valores obtenidos de las medidas de ajuste para las mujeres de 0 a 89 años del año 1990 al 2009. Podemos observar que el *ECM* y el *MAE* más grandes son 3.669006×10^{-6} y 5.670784×10^{-4} respectivamente, que corresponden al año 1990 y están dentro del nivel aceptable. El *ME* más alto es 6.554202×10^{-5} del año 1995 pero es aceptable. El *MAPE* más alto es 4.9% y corresponde al año 1900 y también es aceptable.

Año	ECM	MAE	MAPE	ME
1990	3.669006×10^{-6}	5.670784×10^{-4}	4.9054899	-6.351270×10^{-5}
1991	1.824144×10^{-6}	2.852485×10^{-4}	1.4717457	-8.709512×10^{-5}
1992	3.178751×10^{-6}	3.361517×10^{-4}	0.8314188	-1.416644×10^{-4}
1993	1.093280×10^{-6}	2.046263×10^{-4}	0.7434178	-1.894264×10^{-5}
1994	2.551125×10^{-7}	1.507523×10^{-4}	0.7738982	4.765357×10^{-5}
1995	1.006497×10^{-7}	1.169207×10^{-4}	1.0370088	6.554202×10^{-5}
1996	1.050095×10^{-7}	1.269005×10^{-4}	1.2679174	6.219055×10^{-5}
1997	1.003212×10^{-7}	1.279258×10^{-4}	1.5176807	5.659733×10^{-5}
1998	9.234594×10^{-8}	1.222377×10^{-4}	1.4890829	4.930849×10^{-5}
1999	7.784153×10^{-8}	1.197103×10^{-4}	1.4926527	3.862339×10^{-5}
2000	7.983961×10^{-8}	1.117685×10^{-4}	1.2564059	3.321994×10^{-5}
2001	4.389155×10^{-8}	9.218422×10^{-5}	1.1978933	1.817578×10^{-5}
2002	2.637567×10^{-8}	7.653316×10^{-5}	1.1387559	6.186408×10^{-6}
2003	2.557568×10^{-8}	7.355630×10^{-5}	1.0104050	-5.177698×10^{-6}
2004	3.381634×10^{-8}	6.865692×10^{-5}	0.7893662	-1.190050×10^{-5}
2005	3.168433×10^{-8}	6.370169×10^{-5}	0.5268831	-1.544501×10^{-5}
2006	3.152806×10^{-8}	6.279622×10^{-5}	0.5047958	-1.682936×10^{-5}
2007	4.798295×10^{-8}	9.013722×10^{-5}	0.9491704	-1.895202×10^{-5}
2008	5.982926×10^{-8}	1.214965×10^{-4}	2.7872991	-2.835983×10^{-5}
2009	7.844598×10^{-8}	1.511718×10^{-4}	4.4666308	-3.622480×10^{-5}

Tabla 2.18: Medidas de ajuste mujeres de 0 a 89 años. Modelo Lee-Carter

Edad	2010	2011	2012	2013	2014	Edad	2010	2011	2012	2013	2014
0	0.9812186	0.9818873	0.9825324	0.9831547	0.9837551	46	0.9943969	0.9944632	0.9945287	0.9945934	0.9946573
1	0.9982101	0.9982971	0.9983799	0.9984587	0.9985337	47	0.9940337	0.9941040	0.9941735	0.9942422	0.9943100
2	0.9991553	0.9991984	0.9992392	0.9992780	0.9993148	48	0.9936359	0.9937105	0.9937843	0.9938572	0.9939293
3	0.9994331	0.9994597	0.9994851	0.9995093	0.9995323	49	0.9932008	0.9932795	0.9933574	0.9934343	0.9935104
4	0.9995547	0.9995736	0.9995917	0.9996090	0.9996256	50	0.9927325	0.9928158	0.9928982	0.9929797	0.9930602
5	0.9996176	0.9996322	0.9996464	0.9996599	0.9996730	51	0.9922162	0.9923035	0.9923898	0.9924752	0.9925596
6	0.9996700	0.9996817	0.9996930	0.9997039	0.9997144	52	0.9916509	0.9917431	0.9918343	0.9919244	0.9920136
7	0.9997329	0.9997417	0.9997502	0.9997583	0.9997662	53	0.9910283	0.9911235	0.9912178	0.9913111	0.9914034
8	0.9997498	0.9997570	0.9997640	0.9997708	0.9997774	54	0.9903569	0.9904566	0.9905552	0.9906528	0.9907494
9	0.9996714	0.9996797	0.9996877	0.9996956	0.9997032	55	0.9896258	0.9897289	0.9898310	0.9899321	0.9900322
10	0.9995025	0.9995138	0.9995248	0.9995356	0.9995461	56	0.9888185	0.9889243	0.9890292	0.9891330	0.9892359
11	0.9993466	0.9993602	0.9993734	0.9993864	0.9993991	57	0.9879599	0.9880692	0.9881776	0.9882849	0.9883913
12	0.9992156	0.9992302	0.9992446	0.9992588	0.9992726	58	0.9870166	0.9871276	0.9872376	0.9873467	0.9874548
13	0.9990845	0.9991004	0.9991160	0.9991313	0.9991464	59	0.9860008	0.9861148	0.9862278	0.9863400	0.9864512
14	0.9989502	0.9989673	0.9989841	0.9990006	0.9990169	60	0.9849073	0.9850245	0.9851407	0.9852560	0.9853705
15	0.9988170	0.9988354	0.9988536	0.9988714	0.9988890	61	0.9837138	0.9838330	0.9839514	0.9840689	0.9841856
16	0.9986819	0.9987016	0.9987211	0.9987402	0.9987591	62	0.9824267	0.9825491	0.9826707	0.9827914	0.9829113
17	0.9985566	0.9985782	0.9985994	0.9986203	0.9986409	63	0.9810280	0.9811522	0.9812756	0.9813981	0.9815199
18	0.9984338	0.9984571	0.9984800	0.9985026	0.9985249	64	0.9795247	0.9796510	0.9797766	0.9799014	0.9800255
19	0.9983274	0.9983525	0.9983772	0.9984015	0.9984255	65	0.9778983	0.9780260	0.9781529	0.9782792	0.9784047
20	0.9982282	0.9982549	0.9982812	0.9983071	0.9983327	66	0.9761406	0.9762692	0.9763970	0.9765242	0.9766506
21	0.9981337	0.9981622	0.9981903	0.9982180	0.9982453	67	0.9742399	0.9743691	0.9744976	0.9746256	0.9747529
22	0.9980461	0.9980764	0.9981062	0.9981355	0.9981644	68	0.9721754	0.9723014	0.9724268	0.9725516	0.9726759
23	0.9979646	0.9979967	0.9980283	0.9980594	0.9980900	69	0.9699577	0.9700797	0.9702013	0.9703224	0.9704430
24	0.9978843	0.9979183	0.9979518	0.9979847	0.9980171	70	0.9675802	0.9676991	0.9678175	0.9679355	0.9680531
25	0.9978038	0.9978393	0.9978742	0.9979085	0.9979423	71	0.9650244	0.9651401	0.9652554	0.9653703	0.9654848
26	0.9977277	0.9977647	0.9978011	0.9978369	0.9978722	72	0.9622750	0.9623874	0.9624994	0.9626111	0.9627225
27	0.9976532	0.9976917	0.9977295	0.9977667	0.9978033	73	0.9593488	0.9594606	0.9595720	0.9596832	0.9597940
28	0.9975747	0.9976141	0.9976528	0.9976910	0.9977285	74	0.9561945	0.9563056	0.9564165	0.9565271	0.9566374
29	0.9974907	0.9975309	0.9975704	0.9976093	0.9976476	75	0.9528472	0.9529651	0.9530827	0.9532000	0.9533170
30	0.9974091	0.9974502	0.9974907	0.9975305	0.9975697	76	0.9492381	0.9493614	0.9494845	0.9496072	0.9497297
31	0.9973195	0.9973611	0.9974020	0.9974423	0.9974820	77	0.9454303	0.9455671	0.9457035	0.9458397	0.9459755
32	0.9972261	0.9972683	0.9973097	0.9973506	0.9973909	78	0.9413303	0.9414798	0.9416290	0.9417778	0.9419262
33	0.9971161	0.9971586	0.9972005	0.9972417	0.9972824	79	0.9369324	0.9370983	0.9372638	0.9374288	0.9375934
34	0.9969950	0.9970383	0.9970810	0.9971231	0.9971646	80	0.9322080	0.9323773	0.9325462	0.9327147	0.9328828
35	0.9968668	0.9969109	0.9969544	0.9969972	0.9970395	81	0.9271577	0.9273310	0.9275040	0.9276766	0.9278488
36	0.9967282	0.9967731	0.9968174	0.9968610	0.9969040	82	0.9217746	0.9219536	0.9221322	0.9223104	0.9224882
37	0.9965740	0.9966199	0.9966653	0.9967100	0.9967541	83	0.9161594	0.9163527	0.9165456	0.9167380	0.9169301
38	0.9964070	0.9964540	0.9965004	0.9965462	0.9965914	84	0.9102050	0.9104306	0.9106557	0.9108802	0.9111041
39	0.9962246	0.9962731	0.9963210	0.9963683	0.9964149	85	0.9035912	0.9038059	0.9040201	0.9042339	0.9044472
40	0.9960228	0.9960730	0.9961225	0.9961715	0.9962198	86	0.8963752	0.8965022	0.8966292	0.8967560	0.8968826
41	0.9958060	0.9958583	0.9959100	0.9959610	0.9960113	87	0.8886766	0.8888913	0.8887060	0.8887206	0.8887353
42	0.9955694	0.9956239	0.9956778	0.9957310	0.9957836	88	0.8819360	0.8822704	0.8826038	0.8829363	0.8832680
43	0.9953131	0.9953702	0.9954267	0.9954825	0.9955376	89	0.8714673	0.8711586	0.8708492	0.8705392	0.8702284
44	0.9950334	0.9950934	0.9951526	0.9952111	0.9952689						
45	0.9947307	0.9947941	0.9948568	0.9949187	0.9949799						

Tabla 2.19: Probabilidad de supervivencia p_x para hombres utilizando el modelo Lee-Carter para el periodo de 2010-2014.

Edad	2010	2011	2012	2013	Edad	2010	2011	2012	2013
0	0.9864350	0.9869666	0.9874774	0.9879684	45	0.9971928	0.9972289	0.9972646	0.9972998
1	0.9985178	0.9985848	0.9986488	0.9987099	46	0.9969452	0.9969843	0.9970228	0.9970609
2	0.9993257	0.9993531	0.9993794	0.9994046	47	0.9966733	0.9967156	0.9967573	0.9967985
3	0.9995547	0.9995717	0.9995881	0.9996038	48	0.9963767	0.9964216	0.9964660	0.9965098
4	0.9996666	0.9996782	0.9996893	0.9997001	49	0.9960514	0.9960992	0.9961464	0.9961930
5	0.9997289	0.9997377	0.9997463	0.9997545	50	0.9957011	0.9957525	0.9958032	0.9958534
6	0.9997688	0.9997757	0.9997824	0.9997889	51	0.9953062	0.9953603	0.9954139	0.9954667
7	0.9998177	0.9998226	0.9998274	0.9998321	52	0.9948783	0.9949356	0.9949922	0.9950483
8	0.9998326	0.9998363	0.9998399	0.9998434	53	0.9944099	0.9944703	0.9945301	0.9945892
9	0.9997947	0.9997982	0.9998017	0.9998051	54	0.9938893	0.9939519	0.9940139	0.9940752
10	0.9997166	0.9997203	0.9997240	0.9997277	55	0.9933273	0.9933925	0.9934571	0.9935210
11	0.9996463	0.9996502	0.9996541	0.9996579	56	0.9927077	0.9927749	0.9928415	0.9929074
12	0.9995893	0.9995928	0.9995963	0.9995997	57	0.9920244	0.9920934	0.9921618	0.9922297
13	0.9995419	0.9995449	0.9995478	0.9995507	58	0.9912933	0.9913649	0.9914360	0.9915064
14	0.9994998	0.9995025	0.9995050	0.9995076	59	0.9904904	0.9905639	0.9906368	0.9907092
15	0.9994673	0.9994695	0.9994718	0.9994740	60	0.9896353	0.9897122	0.9897886	0.9898644
16	0.9994476	0.9994499	0.9994522	0.9994544	61	0.9886901	0.9887694	0.9888481	0.9889262
17	0.9994346	0.9994370	0.9994395	0.9994419	62	0.9876672	0.9877489	0.9878300	0.9879106
18	0.9994250	0.9994278	0.9994306	0.9994333	63	0.9865506	0.9866349	0.9867186	0.9868018
19	0.9994200	0.9994234	0.9994267	0.9994299	64	0.9853407	0.9854271	0.9855130	0.9855984
20	0.9994163	0.9994200	0.9994238	0.9994275	65	0.9840249	0.9841137	0.9842020	0.9842898
21	0.9994093	0.9994134	0.9994176	0.9994217	66	0.9825915	0.9826821	0.9827722	0.9828619
22	0.9993998	0.9994045	0.9994092	0.9994138	67	0.9810122	0.9811019	0.9811912	0.9812800
23	0.9993835	0.9993886	0.9993938	0.9993988	68	0.9792773	0.9793646	0.9794516	0.9795382
24	0.9993624	0.9993680	0.9993735	0.9993790	69	0.9774017	0.9774852	0.9775684	0.9776514
25	0.9993404	0.9993467	0.9993530	0.9993592	70	0.9753758	0.9754550	0.9755340	0.9756127
26	0.9993097	0.9993167	0.9993236	0.9993305	71	0.9732063	0.9732843	0.9733620	0.9734395
27	0.9992822	0.9992905	0.9992988	0.9993069	72	0.9708461	0.9709219	0.9709975	0.9710729
28	0.9992428	0.9992521	0.9992613	0.9992703	73	0.9683324	0.9684090	0.9684853	0.9685614
29	0.9992024	0.9992130	0.9992234	0.9992337	74	0.9655974	0.9656741	0.9657507	0.9658270
30	0.9991564	0.9991685	0.9991804	0.9991922	75	0.9626982	0.9627803	0.9628622	0.9629440
31	0.9990997	0.9991131	0.9991264	0.9991395	76	0.9595376	0.9596250	0.9597122	0.9597992
32	0.9990342	0.9990489	0.9990635	0.9990777	77	0.9561467	0.9562422	0.9563375	0.9564326
33	0.9989573	0.9989730	0.9989886	0.9990039	78	0.9525317	0.9526400	0.9527480	0.9528559
34	0.9988754	0.9988924	0.9989091	0.9989255	79	0.9485928	0.9487101	0.9488272	0.9489440
35	0.9987808	0.9987987	0.9988164	0.9988338	80	0.9443857	0.9445101	0.9446342	0.9447580
36	0.9986794	0.9986986	0.9987175	0.9987362	81	0.9398662	0.9399972	0.9401279	0.9402584
37	0.9985624	0.9985829	0.9986030	0.9986229	82	0.9349713	0.9351050	0.9352383	0.9353714
38	0.9984388	0.9984608	0.9984825	0.9985039	83	0.9298374	0.9299892	0.9301406	0.9302917
39	0.9983019	0.9983251	0.9983480	0.9983705	84	0.9243125	0.9244882	0.9246635	0.9248385
40	0.9981559	0.9981810	0.9982058	0.9982302	85	0.9181397	0.9183030	0.9184659	0.9186285
41	0.9979923	0.9980195	0.9980462	0.9980726	86	0.9111343	0.9112073	0.9112803	0.9113533
42	0.9978167	0.9978460	0.9978749	0.9979034	87	0.9035446	0.9034994	0.9034542	0.9034089
43	0.9976264	0.9976578	0.9976888	0.9977194	88	0.8976735	0.8978976	0.8981212	0.8983444
44	0.9974197	0.9974537	0.9974871	0.9975202	89	0.8867879	0.8864997	0.8862109	0.8859214

Tabla 2.20: Probabilidad de supervivencia p_x para mujeres utilizando el modelo Lee-Carter para el periodo de 2010-2013.

2.3. Modelos estocásticos para la fuerza de mortalidad

Si analizamos la tendencia histórica de las tasas de mortalidad, se puede observar que aparentemente no siguen una tendencia definida, debido a que las mejoras en la mortalidad son en efecto estocásticas, es decir dependen de eventos impredecibles tales como avances médicos, condiciones ambientales, desastres naturales, accidentes, etc. La incertidumbre de las previsiones de la mortalidad, es ilustrada en el trabajo hecho por Currie et. al. [5], quienes analizaron las tendencias históricas en la mortalidad de una población de UK utilizando P-splines. En términos generales, ellos encontraron que la tasas de mejoras han variado significativamente a lo largo del tiempo y que además varían substancialmente entre diferentes grupos de edades, concluyendo que el futuro de las mejoras en la mortalidad no puede ser pronosticada con ningún grado de precisión.

Los trabajos recientes que han buscado modelar la fuerza de mortalidad con el fin de crear instrumentos financieros para cubrirse del riesgo de longevidad y no con fines demográficos han propuesto modelarla como una tasa de interés dada la relación entre el índice de supervivencia y la fuerza de mortalidad, véase la expresión (1.17). Para ello, se proponen diversos enfoques que están relacionados con el tipo de instrumento financiero que se desea implementar. No es lo mismo trabajar con un seguro de vida que se negocia anualmente, que con un derivado asociado a un índice de supervivencia con una vigencia de tres a cinco años que con una pensión vitalicia con una duración de al menos veinte años. Todos estos instrumentos pueden modelarse como bonos con incumplimiento en el que el incumplimiento es la muerte del tenedor del bono. Para cada uno de ellos la fuerza de mortalidad debe modelarse de manera distinta: por ejemplo usar los modelos de tasa *spot* para los primeros dos casos y los modelos de tasa *forward* que conlleva la estimación de la curva de estructura de tasas para el caso de las pensiones como lo propone Cairns [2].

Recordemos que buscamos un modelo, ahora estocástico, que nos permita hacer frente al riesgo sistémico. El objetivo será volver a definir los conceptos vistos en el capítulo 1 que nos permitan considerar que la mortalidad está definida por un proceso estocástico, y que

además depende tanto de la edad del asegurado como del tiempo.

Según Cairns et. al. [2] si se asume que la mortalidad es un proceso estocástico entonces, cualquier modelo admisible tiene que satisfacer algunos criterios importantes como son:

- (i) La fuerza de mortalidad debe ser positiva.
- (ii) El modelo debe ser consistente con respecto al comportamiento histórico de los datos.
- (iii) El modelo debe ser biológicamente razonable a largo plazo (por ejemplo, la tasa de mortalidad de los ancianos debe incrementarse con la edad).
- (iv) Mejoras en la mortalidad a largo plazo pueden no ser reversibles a la media a un nivel determinista predeterminado. Por ejemplo, si en el pasado las mejoras en la mortalidad se debieron a avances médicos, entonces sería difícil decir que en el futuro los avances médicos tendrán al menos el mismo impacto que en el pasado con el fin de causar mejoras en la mortalidad.
- (v) Por otro lado, desviaciones a corto plazo de la tendencia causada por fluctuaciones ambientales locales deberían ser reversibles a la media alrededor de la tendencia estocástica a largo plazo.

2.3.1. Marco de modelación

Sea $T^* > 0$ un horizonte temporal fijo. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad equipado con una filtración $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ que satisface las condiciones usuales: continuidad por la derecha y completos. \mathbb{P} es una medida de probabilidad observada en la vida real. En todo, \mathcal{F}_t representa la información total disponible hasta el tiempo t .

2.3.2. Tiempos de paro y el tiempo restante de vida de un individuo de edad x

Definición 2.3. Una variable aleatoria $\tau : \Omega \rightarrow [0, \infty]$ es llamada **un tiempo de paro** con respecto a la filtración \mathbb{F} si para cada t , $0 \leq t \leq \infty$, el evento

$$\{\tau \leq t\} := \{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

Notemos que un tiempo de paro esta caracterizado por el hecho de que cuenta con toda la información hasta el tiempo t . Es decir, para todo t , se puede decidir si τ ha ocurrido o no. Recordemos que en el capítulo anterior, la variable aleatoria continua τ_x definida en $(\Omega, \mathcal{F}, \mathbb{P})$, representa el futuro tiempo restante de vida de un individuo de edad x . Ahora, vamos a ver cómo se define esta variable cuando también depende del tiempo.

Definamos la filtración $\mathbb{H} = (\mathcal{H}_t)_{0 \leq t \leq T}$ como el \mathbb{P} -aumento de la filtración natural generada por τ_x . Esto es, $\mathcal{H}_t = \mathcal{H}_t^+ \vee \mathcal{N}$, donde \mathcal{N} es la σ -álgebra generada por todos los conjuntos \mathbb{P} -nulos y

$$\mathcal{H}_t^+ = \sigma(\{\tau_x > u\}, u \leq t). \quad (2.21)$$

Entonces, τ_x es un tiempo de paro respecto a la filtración \mathbb{H} .

2.3.3. La fuerza de mortalidad y el índice de supervivencia

Ahora, vamos a adaptar la metodología utilizada para los modelos de estructura de tasas de interés a un modelo de estructura en términos de la mortalidad. Esto se puede hacer, reemplazando la tasa de interés por la fuerza de mortalidad. Recordemos que $\mu(t, x)$ representa la fuerza de mortalidad al tiempo t de un individuo de edad x . Como vimos en las secciones anteriores, muchos modelos de mortalidad asumen una fuerza de mortalidad estacionaria con respecto al tiempo, esto es, $\mu(t, x) \equiv \mu(x)$ para todo t y x ; otros modelos de la mortalidad asumen que la fuerza de mortalidad $\mu(t, x)$ es una función determinista con respecto a t y x . Sin embargo, de ahora en adelante, la fuerza de mortalidad $\mu(t, x)$ será llamada **fuerza de mortalidad *spot*** y será necesario que satisfaga las siguientes suposiciones.

Suposición 2.1. Para un valor fijo pero arbitrario de x , se asume:

- La fuerza de mortalidad *spot* $\mu(t, x)$ será modelada como un proceso estocástico adaptado sobre el espacio filtrado de probabilidad $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Es decir, $\mu(\cdot, x) = (\mu(t, x))_{t \geq 0}$.
- Las trayectorias de la muestra de $\mu(\cdot, x)$ son casi seguramente (c. s.) integrables respecto a la medida de Lebesgue sobre el intervalo $[0, T^*]$.

Una vez más, consideremos un valor fijo pero arbitrario de x y sea $S(\cdot, x) = (S(t, x))_{t \geq 0}$ un proceso estocástico adaptado de variación finita y con trayectorias continuas de la muestra sobre el mismo espacio filtrado de probabilidad $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Definición 2.4. Para todo $0 \leq s \leq t \leq T^*$, el proceso del **índice de supervivencia** esta definido por:

$$S(t, x + s) = e^{-\int_0^t \mu(u+s, x+s+u) du}, \quad (2.22)$$

o equivalentemente por la dinámica de

$$dS(t, x + s) = -\mu(t + s, x + s + t)S(t, x + s)dt,$$

con condición inicial $S(t, x + s) = 1$.

En particular, cuando $s = 0$ resulta:

$$S(t, x) = e^{-\int_0^t \mu(u, x+u) du}, \quad \forall t \in [0, T^*], \quad (2.23)$$

o equivalentemente definido por la dinámica

$$dS(t, x) = -\mu(t, x + t)S(t, x)dt,$$

con condición inicial $S(0, x) = 1$.

Frecuentemente se usa el caso especial $s=0$ en al proceso del índice de supervivencia, es decir, $S(t, x)$. Sin embargo es necesaria una definición general para futuras investigaciones.

En el trabajo de Carstens [3] se han definido la fuerza de mortalidad *spot* y el índice de supervivencia como procesos estocásticos adaptados, de esta manera, mirando hacia delante desde el tiempo 0, $S(t, x)$ será una variable aleatoria. Observe que estando parados en el tiempo $u > 0$, $S(u, t)$ puede ser considerada todavía como una probabilidad de supervivencia. Sin embargo, puede extraerse una probabilidad desde el tiempo 0 en adelante tomando una esperanza condicional de funciones indicadoras que mostramos a continuación.

Se define la filtración $\mathbb{I} = (\mathcal{I}_t)_{0 \leq t \leq T}$ como un \mathbb{P} -aumento de la filtración natural generada por $\mu(t, x + t)$. Esto es, $\mathcal{I}_t = \mathcal{I}_t^+ \vee \mathcal{N}$, donde \mathcal{N} esta definida como antes y

$$\mathcal{I}_t^+ = \sigma \{ \mu(u, x + u), u \leq t \}. \quad (2.24)$$

En otras palabras, \mathcal{I}_t representa el flujo de información generada por las observaciones de la evolución de la fuerza de mortalidad *spot*, $\mu(t, x + t)$ hasta el tiempo t .

Considere la filtración $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T^*}$ introducida anteriormente y asuma que esta definida y es igual a

$$\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{I}_t.$$

Además, considere a un individuo de edad x al tiempo 0, cuyo futuro tiempo restante de vida esta dado por τ_x . Defínase el proceso $(I_{\{\tau_x > t\}})_{t \geq 0}$ como una cadena continua de Markov sobre $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, la cual toma los siguientes valores

$$1_{\{\tau_x > t\}} = \begin{cases} 1 & \text{si } (x) \text{ esta vivo al tiempo } t; \\ 0 & \text{otro caso.} \end{cases}$$

Observe que dada la información \mathcal{F}_t al tiempo t , $1_{\{\tau_x > t\}}$ es una función indicadora de una variable aleatoria de Bernoulli con parámetro $p := S(t, x)$. Así, la probabilidad al tiempo 0 de que (x) alcanzará la edad $x + T$ bajo la probabilidad objetiva \mathbb{P} , esta dado por (compare con el Corolario 1.2)

$$\begin{aligned} {}_T p_x(0; T) &= \mathbb{P}(\tau_x > T) \\ &= \mathbb{E}^{\mathbb{P}}[1_{\{\tau_x > T\}}] \\ &= \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[1_{\{\tau_x > T\}} | \mathcal{F}_T]] \\ &= \mathbb{E}^{\mathbb{P}}[S(T, x)]. \end{aligned} \quad (2.25)$$

Nota 2.1. Por la desigualdad de Jensen, $\mathbb{E}^{\mathbb{P}}[S(t, x)] > \exp[-\int_0^t \mathbb{E}^{\mathbb{P}}(\mu(u, x + u))du]$. Además, si $\hat{\mu}(t, x + t)$ es el mejor estimador determinista al tiempo 0 para una mortalidad futura (por ejemplo la media) usualmente tendremos que $\mathbb{E}^{\mathbb{P}}[S(t, x)] \neq \exp[-\int_0^t \mathbb{E}^{\mathbb{P}}(\hat{\mu}(u, x + u))du]$.

De manera general, la probabilidad de supervivencia al tiempo t de un individuo de edad x al tiempo 0, quien ha alcanzado la edad $x + t$ al actual tiempo t y sobrevivirá hasta el tiempo posterior T , bajo la probabilidad objetiva \mathbb{P} , se define como (compare con ecuación 1.21)

$$\begin{aligned}
{}_{T-t}p_{x+t}(t; T) &= \mathbb{P}(\tau_x > T | \mathcal{F}_t) \\
&= \mathbb{P}(\tau_x > T | \mathcal{I}_t, \mathcal{H}_t) \\
&= \mathbb{E}^{\mathbb{P}}[1_{\{\tau_x > T\}} | \mathcal{I}_t, \tau_x > t] \\
&= \frac{\mathbb{E}^{\mathbb{P}}[1_{\{\tau_x > T\}} | \mathcal{I}_t]}{\mathbb{P}(\tau_x > t | \mathcal{I}_t)} \\
&= \frac{\mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[1_{\{\tau_x > T\}} | \mathcal{I}_T] | \mathcal{I}_t]}{\mathbb{E}^{\mathbb{P}}[1_{\{\tau_x > t\}} | \mathcal{I}_t]} \\
&= \frac{\mathbb{E}^{\mathbb{P}}[S(T, x) | \mathcal{I}_t]}{S(t, x)} \\
&= \mathbb{E}^{\mathbb{P}} \left[\frac{S(T, x)}{S(t, x)} \middle| \mathcal{I}_t \right], \quad \text{para } t < T.
\end{aligned} \tag{2.26}$$

La ecuación (2.26) solo necesita de la información dada por la filtración $(\mathcal{I}_u)_{u \geq 0}$ hasta el tiempo t , es decir, la información generada por el proceso $\mu(t, x)$. Sin embargo, como $\mathcal{I}_t \subset \mathcal{F}_t$ para toda $t \leq T^*$, entonces si se incorpora información adicional, no se altera el resultado. Por lo tanto si se incorpora la información dada por $(\mathcal{H}_u)_{u \geq 0}$ hasta el tiempo t ,

$${}_{T-t}p_{x+t}(t; T) = \mathbb{P}(\tau_x > T | \mathcal{F}_t) = \mathbb{E}^{\mathbb{P}} \left[\frac{S(T, x)}{S(t, x)} \middle| \mathcal{F}_t \right], \quad \text{para } t < T. \tag{2.27}$$

Cada uno de los enfoques vistos se aplicaran en relación con el instrumento financiero. En el siguiente capítulo proponemos la valuación de un bono de longevidad a corto plazo por lo que usaremos el índice de mortalidad bajo la probabilidad de riesgo neutro.

Valuación de bonos de longevidad a corto plazo

Un instrumento financiero, según las Normas Internacionales de Información Financiera (NIIF), es un contrato que da origen a un activo financiero en una empresa y un pasivo financiero o instrumento de patrimonio en otra. Algunos ejemplos de instrumentos financieros son las acciones, ordinarias o preferentes, que constituyen el capital de una empresa; las obligaciones, ya sean simples o convertibles; los bonos o pagarés emitidos por el estado y adquiridos por las empresas; los préstamos que una empresa concede a las empresas de su grupo o a sus trabajadores; las cuentas a cobrar de los clientes e incluso el efectivo mantenido en caja o en cuentas corrientes bancarias. Por otra parte, los instrumentos financieros se pueden clasificar por clase de activos, dependiendo de si son basados en equidad (lo que refleja la propiedad de la entidad emisora) o basados en la deuda (lo que refleja un préstamo que el inversor ha hecho a la entidad emisora). Si se trata de la deuda, puede clasificarse en a corto plazo (menos de tres años) o a largo plazo.

Es importante mencionar que los modelo descritos en el capítulo 2 serán utilizados con el objetivo de valorar un instrumento financiero. No se pretende obtener un modelo demográfico ni un modelo de proyección para la fuerza de mortalidad.

Este capítulo tiene como objetivo valorar un bono de longevidad a corto plazo. Un bono de longevidad es un instrumento cuyo valor depende de dos factores de riesgo: la fuerza

de mortalidad y la tasa de interés. Como varios autores lo han indicado estos factores son independientes por lo que los modelaremos como dos procesos estocásticos independientes.

Para garantizar que hay ausencia de arbitraje en el proceso de valuación se va a trabajar bajo la medida neutra al riesgo por lo que vamos a dar un marco teórico para usar el Teorema de Girsanov multidimensional que nos permite definir una medida equivalente bajo la cual los bonos de longevidad descontados son martingala.

En el mercado financiero cuando se fija un “precio justo” a algún derivado, generalmente se hace bajo el marco de una **medida neutra al riesgo** \mathbb{Q} equivalente a la medida objetiva \mathbb{P} . La medida \mathbb{Q} tiene como principal característica que los precios descontados (traídos a valor presente) son martingalas y con ello, por el Primer Teorema Fundamental de la Valuación de Activos conduce a la ausencia de arbitraje. Aunque la existencia de \mathbb{Q} garantice la ausencia de arbitraje en el mercado, se necesita que sea única para que conduzca a un mercado completo. Análogamente a la metodología utilizada para fijar un precio a un derivado en el mercado financiero, lo que buscamos es fijar el precio del riesgo sistémico derivado de la mortalidad con la ayuda de una medida de probabilidad neutra al riesgo \mathbb{Q} .

La correspondiente probabilidad de supervivencia bajo una medida martingala \mathbb{Q}_1 se define como

$$\begin{aligned} {}_{T-t}\bar{p}_{x+t}(t; T) &= \mathbb{Q}_1(\tau_x > T | \mathcal{F}_t) \\ &= \mathbb{E}^{\mathbb{Q}_1}[1_{\{\tau_x > T\}} | \mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}_1}\left[\frac{S(T, x)}{S(t, x)} \middle| \mathcal{F}_t\right], \quad \text{para toda } t < T. \end{aligned} \tag{3.1}$$

En la expresión (3.1) se supone que el riesgo específico que se deriva de la mortalidad no implica una prima.

3.1. El índice medio de supervivencia y la fuerza de mortalidad *forward*

Suposición 3.1 . Existe una medida de probabilidad \mathbb{Q}_1 martingala sobre (Ω, \mathcal{F}) equivalente

a la medida de probabilidad objetiva \mathbb{P} tal que para cualquier valor arbitrario pero fijo de x se asume:

- La fuerza de mortalidad *spot* $\bar{\mu}(t, x + t)$ es modelada como un proceso adaptado sobre el espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}_1)$, $\bar{\mu}(\cdot, x) = (\bar{\mu}(t, x + t))_{t \geq 0}$.
- Las trayectorias de la muestra de $\bar{\mu}(\cdot, x)$ son integrables casi seguramente con respecto a la medida de Lebesgue sobre el intervalo $[0, T^*]$.
- La expresión dada por (3.1) se sigue para todo $t, T \in [0, T^*]$, $t \leq T$.

Definición 3.1. Un proceso estocástico estrictamente positivo adaptado ${}_T\bar{p}_x(\cdot; T) = ({}_T\bar{p}_x(t; T))_{t \geq 0}$ sobre $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q}_1)$ definido por la fórmula

$${}_T\bar{p}_x(t; T) = \mathbb{E}^{\mathbb{Q}_1}[S(T, x) | \mathcal{F}_t], \quad \forall t \in [0, T], \quad (3.2)$$

es llamado **índice medio de supervivencia** para un individuo de edad x al tiempo cero y para un horizonte fijo de tiempo T .

Se asume que el proceso es estrictamente positivo para descartar la posibilidad de eventos catastróficos que elimine a toda la población.

Lema 3.1. El proceso del índice medio de supervivencia es una \mathbb{Q}_1 -martingala.

Demostración. El proceso ${}_T\bar{p}_x(t; T)$ es adaptado pues está definido por una esperanza condicional. Además, ${}_T\bar{p}_x(t; T)$ es integrable, esto es

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_1}[|{}_T\bar{p}_x(t; T)|] &= \mathbb{E}^{\mathbb{Q}_1}[|\mathbb{E}^{\mathbb{Q}_1}[S(T, x) | \mathcal{F}_t]|] \\ &\leq \mathbb{E}^{\mathbb{Q}_1}[\mathbb{E}^{\mathbb{Q}_1}[|S(T, x)| | \mathcal{F}_t]] \\ &= \mathbb{E}^{\mathbb{Q}_1}[S(T, x)] \\ &= {}_T\bar{p}_x(0; T), \end{aligned}$$

donde ${}_T\bar{p}_x(0; T) \in [0, 1]$. Al usar la suposición 3.1, se tiene que $S(t, x)$ y ${}_T\bar{p}_x(0; T)$ están bien definidos. Por último, para $t \leq u \leq T$

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_1} [{}_T\bar{p}_x(u; T) | \mathcal{F}_t] &= \mathbb{E}^{\mathbb{Q}_1} [\mathbb{E}^{\mathbb{Q}_1} [S(T, x) | \mathcal{F}_u] | \mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}_1} [S(T, x) | \mathcal{F}_t] \\ &= {}_T\bar{p}_x(t; T). \end{aligned}$$

□

Bajo la medida \mathbb{Q}_1 martingala, la probabilidad de supervivencia al tiempo t de un individuo de edad x al tiempo 0 que ha alcanzado la edad $x + t$ al actual tiempo t y que sobrevivirá hasta el tiempo T , ${}_{T-t}\bar{p}_{x+t}(t; T)$, puede ser expresada en términos de los procesos ${}_T\bar{p}_x(t; T)$ y ${}_t\bar{p}_x(t; t)$ para $t \leq T$ al utilizar la ecuación (2.27) de la siguiente manera:

$$\begin{aligned} {}_{T-t}\bar{p}_{x+t}(t; T) &= \mathbb{E}^{\mathbb{Q}_1} \left[\frac{S(T, x)}{S(t, x)} \middle| \mathcal{F}_t \right] \\ &= \frac{\mathbb{E}^{\mathbb{Q}_1} [S(T, x) | \mathcal{F}_t]}{S(t, x)} \\ &= \frac{\mathbb{E}^{\mathbb{Q}_1} [S(T, x) | \mathcal{F}_t]}{\mathbb{E}^{\mathbb{Q}_1} [S(t, x) | \mathcal{F}_t]} \\ &= \frac{{}_T\bar{p}_x(t; T)}{{}_t\bar{p}_x(t; t)}. \end{aligned} \tag{3.3}$$

Como ${}_T\bar{p}_x(t; T)$ y ${}_t\bar{p}_x(t; t)$ pueden ser considerados como estimadores *spot*, el término ${}_{T-t}\bar{p}_{x+t}(t; T)$ puede ser visto como una **probabilidad de supervivencia spot**.

Definición 3.2. El proceso ${}_T\bar{p}_x(t; T)$ define la **fuerza de mortalidad forward**, es decir

$$\bar{\mu}(t, T, x + T) = -\frac{\partial}{\partial T} \log_T \bar{p}_x(t; T). \tag{3.4}$$

Por otro lado, el conocer la fuerza de mortalidad *forward* nos permite calcular

Lema 3.2 . Para $0 \leq t \leq T$ y un x fijo, tenemos

$$\frac{{}_T\bar{p}_x(t; T)}{{}_t\bar{p}_x(t; t)} = \exp \left(- \int_t^T \bar{\mu}(t, s, x + s) ds \right). \tag{3.5}$$

Demostración. Comenzamos utilizando (3.4), multiplicando por -1 e integrando con respecto a la variable T sobre el intervalo $[t, T]$, obtenemos

$$\log_T \bar{p}_x(t; T) - \log_t \bar{p}_x(t; t) = - \int_t^T \bar{\mu}(t, u, x + u) du.$$

Lo siguiente es tomar función exponencial en ambos lados de la ecuación anterior y entonces se obtiene (3.5). □

Sin embargo, combinando las ecuaciones (3.3) y (3.5), podemos obtener una expresión para las probabilidades de supervivencia *spot* en términos de la fuerza de mortalidad *forward* de la siguiente manera

$${}_{T-t}\bar{p}_{x+t}(t; T) = \exp \left(- \int_t^T \bar{\mu}(t, s, x + s) ds \right). \quad (3.6)$$

Además, la fuerza de mortalidad $\mu(t, x + t)$ se puede obtener a partir de conocer la fuerza de mortalidad *forward* $\bar{\mu}(t, T, x + T)$ tomando $T = t$, es decir

$$\begin{aligned} \bar{\mu}(t, t, x + t) &= - \frac{\partial}{\partial t} \log_t \bar{p}_x(t; t) \\ &= - \frac{\partial}{\partial t} \log \mathbb{E}^{\mathbb{Q}_1} [S(t, x) | \mathcal{F}_t] \\ &= - \frac{\partial}{\partial t} \log S(t, x) \\ &= - \frac{\partial}{\partial t} \log \left[\exp \left(- \int_0^t \bar{\mu}(s, x + s) ds \right) \right] \\ &= - \frac{\partial}{\partial t} \left(- \int_0^t \bar{\mu}(s, x + s) ds \right) \\ &= \bar{\mu}(t, x + t). \end{aligned} \quad (3.7)$$

La probabilidad de supervivencia *forward* puede ser definida como sigue.

Definición 3.3. Dada la información disponible \mathcal{F}_t al tiempo t , ${}_{T_2-T_1}\bar{p}_{x+T_1}(t; T_2)$ es llamado el proceso de **probabilidad de supervivencia forward** y denota la probabilidad de que un individuo de edad x al tiempo 0 que sigue vivo al tiempo T_1 sobrevivirá al tiempo T_2 , para $t \leq T_1 \leq T_2$. Esta probabilidad esta dada por

$$\begin{aligned} {}_{T_2-T_1}\bar{p}_{x+T_1}(t; T_2) &= \exp \left(- \int_{T_1}^{T_2} \bar{\mu}(t, u, x + u) du \right) \\ &= \frac{{}_{T_2}\bar{p}_x(t; T_2)}{{}_{T_1}\bar{p}_x(t; T_1)}. \end{aligned} \quad (3.8)$$

La estructura de la tasa de interés, se puede estudiar desde dos diferentes marcos de modelado: a corto o largo plazo. Con el propósito de modelar la dinámica del proceso ${}_T\bar{p}_x(t; T)$, se pueden considerar emplear los dos diferentes marcos anteriores, sustituyendo el término correspondiente a la tasa de interés por el de la fuerza de mortalidad. Estos son:

1. Modelado de tasa *spot* (tasa cota) para la dinámica de la fuerza de mortalidad $\bar{\mu}(t, x)$ (que corresponde a los modelos de tasa corta interés $r(t)$, por ejemplo Vasicek o el modelo CIR).
2. Marco de modelado de la mortalidad *forward* para la dinámica de las tasas de mortalidad *forward* $\bar{\mu}(t, T, x + T)$.

Asumiendo que la tasa de interés es cero, sea $B(t, T, x)$ el precio al tiempo t de un instrumento financiero que paga un monto M al tiempo T a un asegurado que al tiempo 0 tiene x años. Para fijar el precio de estos instrumentos bajo la medida neutra al riesgo, necesitamos usar la fuerza de mortalidad *forward* (ecuación (3.4)). Entonces, el precio de este instrumento se define de la siguiente manera:

$$\bar{\mu}(t, T, x + T) = -\frac{\partial}{\partial T} \log B(t, T, x). \quad (3.9)$$

En la literatura, a $\bar{\mu}(t, T, x + T)$ se le conoce como **superficie de la fuerza de mortalidad *forward*** (ver por ejemplo Milevsky [15] y Dahl [6]).

Si se conoce toda la superficie de la fuerza de mortalidad *forward* entonces podemos fijar el precio del instrumento como:

$$\frac{B(t, T, x)}{B(t, t, x)} = \exp \left(\int_t^T \bar{\mu}(t, s, x + s) ds \right). \quad (3.10)$$

Note la relación que existe entre la ecuación anterior y la ecuación de la fuerza de mortalidad *forward* (3.5).

Si hacemos $T = t$ en la ecuación (3.10), obtenemos la fuerza de mortalidad *spot*:

$$\bar{\mu}(t, x + t) = \bar{\mu}(t, t, x + t). \quad (3.11)$$

Hacer una estimación de toda una superficie de mortalidad que dependa de la edad de cada individuo y para cualquier fecha de madurez, es una tarea muy complicada. Por ello primero vamos a estudiar instrumentos a corto plazo para los cuales vamos a fijar la edad del asegurado. Esto nos lleva a hacer uso de la fuerza de mortalidad *spot* para modelar curvas de supervivencia asociadas a un cohorte generacional específico. Antes de comenzar con la obtención de las diferentes curvas de supervivencia, es importante mencionar que éstas se definen dependiendo el enfoque del instrumento financiero que se quiera utilizar. En general, supongamos que tenemos observaciones a lo largo del tiempo de individuos a ciertas edades. Si acomodamos la información en una tabla en la cuál las columnas sean las observaciones en el tiempo y las filas sean la edad de cada individuo, hemos creado una herramienta que en demografía se conoce como diagrama de Lexis. Con dicho diagrama, se pueden hacer estudios transversales o longitudinales de los datos. En nuestro caso, el objetivo es estudiar cohortes generacionales, por lo tanto vamos a hacer uso de un estudio longitudinal. Así, por ejemplo, supongamos que tenemos una tabla con observaciones de un periodo de 1990 a 2009 de individuos que tienen de 0 a 89 años, es decir, contamos con información de 20 columnas a lo largo del tiempo y 90 filas que representan la edad de los individuos. Supongamos que nos interesa estudiar el comportamiento del cohorte de individuos que en 1990 tienen 50 años. Entonces el dato que tenemos que considerar para el año 1991 es la observación de los individuos que en este año tiene 51 y así sucesivamente hasta considerar la observación en el 2009 de individuos que tienen la edad 59. Observe que hemos utilizado la información de la tabla en forma diagonal.

A continuación, vamos a definir un bono cupón cero de longevidad a corto plazo B_L , cuyos factores de riesgo son la fuerza de mortalidad y la tasa de interés. Se utilizarán los modelos de tasa *spot*: esto es, se modela la tasa de mortalidad *spot* $q(t, x)$ o la fuerza de mortalidad *spot* $\mu(t, x)$. En la literatura se han propuesto modelos de este tipo, uno de ellos es el modelo de Milevsky y Promislow [15] para la evolución de la fuerza de mortalidad en tiempo continuo, asumen que la fuerza de mortalidad $\mu(t, x)$ sigue un modelo Gompertz $\xi_0(t) \exp(\xi_1 x)$, donde

el término $\xi_0(t)$ es modelado usando un proceso de difusión con reversión a la media. Otro, es el propuesto por Menoncin [14]. Presentan una solución cerrada para el precio de un bono de longevidad cuando la fuerza de mortalidad sigue un proceso estocástico de raíz cuadrada cuyo valor esperado coincide, en cualquier momento, con la fuerza de la mortalidad dada por la llamada densidad de Gompertz-Makeham.

En el capítulo 4 describiremos cómo utilizaremos la fuerza de mortalidad asociada a un cohorte generacional para la valuación de un instrumento actuarial.

3.2. Valuación de bonos cupón cero de longevidad a corto plazo

Del capítulo 2 recordar que: τ_x es el futuro tiempo restante de vida de un individuo de edad x al tiempo 0 que genera la filtración \mathbb{I}_t (ecuación 2.24). Se definió también la filtración \mathbb{H}_t (ecuación 2.21) generada por $\mu(t, x+t)$ y que $\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{I}_t$. Recuerde además que la tasa de interés genera la filtración \mathcal{G}_t . Vamos a suponer que las tasas de interés y la fuerza de mortalidad son independientes, en otras palabras, la fuerza de mortalidad de una población no tiene dependencia alguna con el comportamiento del mercado.

Consideremos los procesos:

$$\begin{aligned} d\mu_x(t) &= \hat{\alpha}(t, \mu_x)dt + \hat{\sigma}(t, \mu_x)dW^1(t), \quad \mu_x(0) = \mu_0, \\ dr(t) &= k(\theta - r(t))dt + \sigma\sqrt{r(t)}dW^2(t), \quad r(0) = r_0. \end{aligned}$$

con $W^1(t)$ y $W^2(t)$ procesos brownianos independientes.

Definición 3.4. Un bono de longevidad cupón cero con vencimiento T es un instrumento financiero que paga a un inversionista un monto M si sobrevive hasta el tiempo T . Denotemos por $B_L(t; T)$ el valor del bono al tiempo t con $0 \leq t \leq T$.

Teorema 3.1. Por el Teorema de Girsanov multidimensional existe una probabilidad \mathbb{Q} neutra al riesgo equivalente a \mathbb{P} bajo la cual el precio descontado de un bono de longevidad es martingala.

Demostración: Ver en el capítulo 3 de Dahl [6].

Teorema 3.2. El valor de un bono de longevidad cupón cero con vencimiento T al tiempo t es

$$B_L(t; T) = \mathbb{E}^{\mathbb{Q}} \left[M 1_{\{\tau_x > T\}} e^{-\int_t^T r(u) du} | \mathcal{F}_t \vee \mathcal{G}_t \right]. \quad (3.12)$$

Demostración: Esto es resultado de que bajo \mathbb{Q} los bonos de longevidad descontados son martingala.

Observemos que usando la independencia de τ_x y de la tasa de interés $r(t)$

$$\begin{aligned} B_L(t; T) &= M \mathbb{E}^{\mathbb{Q}} [1_{\{\tau_x > T\}} | \mathcal{F}_t] \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r(u) du} | \mathcal{G}_t \right], \\ &= M {}_{T-t}p_{x+t}(t; T) \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r(u) du} | \mathcal{G}_t \right]. \end{aligned} \quad (3.13)$$

donde ${}_{T-t}p_{x+t}(t; T)$ la definimos en (2.27), $r(t)$ es la tasa de interés del mercado y \mathcal{G}_t la filtración generada por ella.

Nota 3.1. Dado que no existe un mercado de bonos de longevidad, la probabilidad de riesgo neutra no es única, asumiremos en este caso que el precio del riesgo de mercado de longevidad es cero, tal como lo sugiere Dahl [6]

En la Sección 1.2.3 se estudió el modelo CIR para la tasa *spot* $r(t)$, entonces la esperanza condicional que depende la tasa de interés en la ecuación (3.13) se puede calcular explícitamente utilizando la ecuación (1.45).

Para calcular ${}_{T-t}p_{x+t}(t; T)$, podemos utilizar las tablas de probabilidades de supervivencia que obtuvimos con los modelos Gompertz-Makeham y Lee-Carter en el capítulo anterior, pero no tendríamos un proceso estocástico para la mortalidad. A continuación, estudiaremos una

de las propuestas para considerar la fuerza de mortalidad como una tasa de interés estocástica que sigue un modelo tipo CIR.

Vamos a suponer que la fuerza de mortalidad para un cohorte generacional específico μ_x , donde el subíndice x es un valor fijo, sigue un proceso estocástico bajo la probabilidad neutra al riesgo dado por la ecuación:

$$\begin{aligned} d\mu_x(t) &= \alpha_x(\beta_x - \mu_x(t)) + \delta_x \sqrt{\mu_x(t)} dW^1(t), \\ \mu_x(0) &= -\ln({}_0p_x), \end{aligned} \quad (3.14)$$

donde α_x , β_x y δ_x son constantes definidas para cada cohorte generacional que satisfacen $2\alpha_x\beta_x > \delta_x^2$ para garantizar que son estrictamente positivos, $W^1(t)$ es el movimiento browniano estándar. Ahora, el problema es estimar las constantes en la ecuación anterior, para ello, recordemos que la fuerza de mortalidad, por definición es:

$$\mu_x(t) = -\frac{d}{dt} \ln({}_t p_x) \approx -\lim_{\Delta t \rightarrow \infty} \left(\frac{\ln_{t_{i+1}} p_x - \ln_{t_i} p_x}{t_{i+1} - t_i} \right), \quad (3.15)$$

donde $i = 1 \dots N$ son observaciones de la probabilidad de supervivencia de una persona de edad x . Lo siguiente es usar la metodología descrita en la sección 1.2.4.2 reemplazando la tasa de interés por la fuerza de mortalidad dada por la ecuación (3.15). Así, la probabilidad de supervivencia del bono de longevidad de la ecuación (3.13) es:

$${}_{T-t}p_{x+t}(t; T) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T \mu_x(s) ds} | \mathcal{F}_t \right] = a(t, T) e^{-\mu_x(t)b(t, T)}, \quad (3.16)$$

donde $\gamma = \sqrt{\alpha^2 + 2\delta^2}$, $a(t, T) = \left(\frac{2\gamma e^{(\alpha+\gamma)((T-t)/2)}}{(\gamma+\alpha)(e^{\gamma(T-t)}-1)+2\gamma} \right)^{\frac{2\alpha\beta}{\delta^2}}$ y $b(t, T) = \frac{2(e^{\gamma(T-t)}-1)}{(\gamma+\alpha)(e^{\gamma(T-t)}-1)+2\gamma}$.

Por lo tanto podemos calcular explícitamente un bono de longevidad a corto plazo.

3.3. Estimación de la fuerza de mortalidad estocástica para la población mexicana

Para calcular la probabilidad de supervivencia en la ecuación (3.16) cuando la fuerza de mortalidad sigue un proceso estocástico dado por la ecuación (3.14), debemos primero

1990		2008	2009
x-19	←		
		x-1	
			x

Figura 3.1: Graduación de los parámetros del modelo de fuerza de mortalidad tipo CIR.

estimar los parámetros α_x , β_x y δ_x . Para ello, hemos tomado los datos proporcionados por LA CONAPO [20] del número de defunciones $D_x^{t_i}$ y el número de personas vivas $E_x^{t_i}$, donde t_i denota los años comprendidos de 1990 a 2009 y x la edad, para aproximar la fuerza de mortalidad definida en la ecuación (3.15), es decir, para una edad fija x :

$$\mu_x(t_i) \approx -\frac{\ln\left(1 - \frac{D_x^{t_{i+1}}}{E_x^{t_{i+1}}}\right) - \ln\left(1 - \frac{D_x^{t_i}}{E_x^{t_i}}\right)}{t_{i+1} - t_i}, \quad i = 1 \dots 19. \tag{3.17}$$

Los parámetros del modelo CIR para tasas de interés, se estiman a partir de datos históricos. Análogamente, se calcularon los parámetros del modelo CIR para la fuerza de mortalidad estocástica α_x , β_x y δ_x con ayuda de la base de datos histórica que describimos en el párrafo anterior, pero siguiendo la idea de cohorte generacional pasado. Es decir, si en el 2009 queremos estudiar la probabilidad de supervivencia de una persona de edad x , se considerara como cohorte generacional histórico a las personas que en el 2008 tienen $x - 1$ y así sucesivamente siguiendo una línea diagonal hacia atrás hasta considerar a los individuos que en 1990 tienen $x - 19$ años, tal como se ilustra en la figura 3.1.

En la literatura, se han propuestos diferentes modelos de la fuerza de mortalidad desde el marco de modelación a corto plazo. Por ejemplo:

- Menoncin F. 2009 [14] propone que la fuerza de mortalidad sigue la dinámica de la

ecuación:

$$\begin{aligned}d\mu(t) &= a_\mu(\gamma_\mu - \mu(t))dt + \sigma_\mu\sqrt{\mu(t)}dW_\mu(t), \\ \mu(t_0) &= \phi + \left(\frac{1}{b}e^{\frac{t_0-m}{b}}\right),\end{aligned}$$

donde

$$\gamma_\mu(t) \equiv \phi + \left(\frac{1}{a_\mu b} + 1\right) \frac{1}{b}e^{\frac{t-m}{b}},$$

a_μ y b_μ son dos parámetros constantes. Observe que el modelo propone como condición inicial una función Gompertz-Makeham.

- En el trabajo hecho por Luciano E. y Vigna E. 2008 [13], se modela la fuerza de mortalidad bajo 3 modelos diferentes. Éstos son:

1. $d\mu(t) = a\mu(t)dt + \sigma dW(t)$.

2. $d\mu(t) = a\mu(t)dt + \sigma dW(t) + dJ(t)$,

donde J es un es un proceso de salto de Poisson.

- 3.

$$d\mu(t) = a\mu(t)dt + \sigma\sqrt{\mu(t)}dW(t). \quad (3.18)$$

Note que la ecuación anterior es un proceso de raíz cuadrada sin reversión a la media.

Estos modelos se calibraron para diferentes cohortes generacionales utilizando información de la población de UK. Los parámetros a y σ son constantes y se estiman a partir de la serie histórica de los datos.

- Milevsky y Promislow [15] proponen modelar la fuerza de mortalidad h_t como un proceso de difusión en tiempo continuo que puede servir como un proceso estocástico análogo a la función de supervivencia de Gompertz. Eligieron el llamado proceso Browniano

Gompertz de reversión a la media (MRBG por sus siglas en inglés):

$$\begin{aligned}h_t &= h_0 e^{gt + \sigma Y_t}, & g, \sigma, h_0 > 0 \\dY_t &= -bY_t dt + dB_t^h, & Y_0 = 0, b \geq 0,\end{aligned}\tag{3.19}$$

donde B_t es un movimiento browniano estándar. Cuando $b = 0$, el proceso converge a un movimiento browniano geométrico ya que Y_t se convierte en B_t . El inconveniente de este modelo es que no es afin, por lo tanto el cálculo es más complicado.

Resultados numéricos

En esta sección, se utilizarán los modelos para la fuerza de mortalidad que se estudiaron en el capítulo anterior para valorar numéricamente un instrumento actuarial de longevidad con madurez de uno y tres años. Dicho cálculo se realizará asociado a un género y a un cohorte generacional específico. Observe que si se supone que la tasa de interés del mercado y la fuerza de mortalidad son independientes, el valor del bono de longevidad se puede calcular como el producto de un bono cupón cero que solo depende de la tasa de interés del mercado multiplicado por la probabilidad de supervivencia. Con respecto a la tasa de interés, vamos a suponer que sigue un proceso CIR dado por la ecuación (1.41). Para la estimación de los parámetros de este proceso se utilizó el paquete SMFI5 del *software* R. El paquete está basado en la metodología propuesta en el capítulo 5 del libro *Statistical Methods for Financial Engineering* [17].

Primero calcularemos el valor de un bono cupón cero, después compararemos el cálculo de la probabilidad de supervivencia dada por los tres modelos de la fuerza de mortalidad para finalmente calcular el valor del bono de longevidad y comparar los resultados que se obtienen con los tres modelos.

Precio de un bono cupón cero

El valor al tiempo presente, es decir $t = 0$, de un bono cupón cero que promete pagar \$1

con una madurez T cuando la tasa de interés sigue un modelo tipo CIR bajo una medida martingala \mathbb{Q}_2 , está dado por la ecuación (1.45). A continuación, haremos el cálculo del precio de un bono con madurez de 1 y 3 años. Para ello, hemos utilizado los valores de la tasa de rendimientos de CETES a 28 días proporcionados por el Banco de México (BANXICO) [21] del periodo de 06/01/2005 al 31/12/2009 (5 años). Se utilizó este periodo de información de tal manera que coincida con el periodo de la base de datos que hemos utilizado para la estimación de la fuerza de mortalidad. Con la ayuda de la función *est.cir* del paquete SMFI5 en R, se obtuvo que los valores de los parámetros del modelo CIR para el bono cupón cero son:

$$k=0.9221892, \theta=0.014127578, \sigma=0.7082809, q_1=1.8394235 \text{ y } q_2=-1.2984093.$$

Por lo tanto, el precio de un bono cupón cero al tiempo $t = 0$ que paga un peso con madurez de 1 y 3 años respectivamente es:

$$\begin{aligned} B(0, 1) &= 0.9592301, \\ B(0, 3) &= 0.9218557. \end{aligned} \tag{4.1}$$

Cálculo de la probabilidad de supervivencia

Vamos a pensar que estamos situados en el año 2009 y que queremos saber la probabilidad de que una persona de x años sobreviva los próximos t años. Según lo que hemos estudiado, tenemos que hacer el cálculo de la ecuación (1.9). En las tablas 4.1, 4.2 y 4.3 hemos calculado la probabilidad de supervivencia de los siguientes 10 años de individuos que en el año 2009 tienen 60, 70 y 80 años para hombres y mujeres, respectivamente bajo los tres modelos estudiados en el capítulo 2. Estos resultados se obtuvieron de la siguiente manera:

Primero, recordemos de la ecuación (1.14) la relación que existe entre la función de supervivencia de una persona de edad x al tiempo t con la probabilidad de supervivencia, esto es:

$$s_x(t) = {}_t p_x = \exp \left\{ - \int_x^{x+t} \mu(z) dz \right\}. \tag{4.2}$$

Bajo el modelo Gompertz-Makeham, en la Tabla 2.12 calculamos la probabilidad de supervivencia de hombres y mujeres en el año 2009. Al suponer que estas probabilidades de supervivencia son constantes a lo largo del tiempo, haciendo uso de la Proposición 1.1 y la ecuación anterior, calculamos la probabilidad de que un individuo de x años en el 2009 sobreviva t años de la siguiente manera:

$$s_x(t) = {}_t p_x^{2009} = p_x^{2009} p_{x+1}^{2009} \cdots p_{x+t}^{2009}.$$

Observe que para calcular la probabilidad de supervivencia bajo este modelo, sólo se utiliza la información del ajuste que se obtuvo con los datos del 2009, lo que supone que la mortalidad es constante a lo largo del tiempo.

Con el modelo Lee-Carter, se construyeron las Tablas 2.19 y 2.20 que son dinámicas en el tiempo, entonces el cálculo de la función de supervivencia en términos de la probabilidad de que un individuo de x años sobreviva de los próximos t años parados en 2009 se calculó como sigue:

$$s_x(t) = {}_t p_x^{2009} = p_x^{2010} p_{x+1}^{2011} \cdots p_{x+t}^{2010+(t-1)}.$$

El cálculo anterior se realiza de forma longitudinal, de tal manera que si en el 2009 se fija una edad para un asegurado, estaríamos calculando la probabilidad de supervivencia de su cohorte generacional.

En el capítulo anterior, en la sección 3.3, se presentó la metodología para estimar la fuerza de mortalidad cuando sigue un modelo estocástico. Después de hacer esta estimación, se hizo uso de la función *est.feller* del paquete SFMI5 de R, la cual estima los parámetros α_x , β_x y δ_x de la ecuación (3.14). El cálculo de estos coeficientes se realizó bajo la probabilidad objetiva \mathbb{P} ya los métodos que disponemos para hacer estos cálculos suponen que el mercado es incompleto. Observe que la probabilidad de supervivencia en este caso, puede ser calculada como un bono cupón cero sustituyendo la tasa de interés por la fuerza de mortalidad. Por lo tanto, se calculó la probabilidad de que una persona de x años sobreviva durante t años haciendo uso del modelo afin que definimos en la ecuación (3.16).

t	HOMBRES			MUJERES		
	G-M	L-C	CIR	G-M	L-C	CIR
1	0.9850918	0.9849073	0.9846540	0.9894839	0.9896353	0.9892940
2	0.9693422	0.9689843	0.9695003	0.9781109	0.9785210	0.9786712
3	0.9527063	0.9521925	0.9545372	0.9658228	0.9666124	0.9681312
4	0.9351388	0.9344799	0.9397633	0.9525601	0.9538549	0.9576738
5	0.9165946	0.9158141	0.9251768	0.9382632	0.9401988	0.9472990
6	0.8970293	0.8961511	0.9107761	0.9228728	0.9255919	0.9370064
7	0.8764002	0.8754514	0.8965597	0.9063314	0.9099754	0.9267958
8	0.8546670	0.8536798	0.8825260	0.8885839	0.8932604	0.9166671
9	0.8317929	0.8307735	0.8686732	0.8695797	0.8753647	0.9066198
10	0.8077461	0.8067133	0.8549999	0.8492736	0.8562314	0.8966539

Tabla 4.1: Probabilidad de supervivencia ${}_t p_{60}$.

t	HOMBRES			MUJERES		
	G-M	L-C	CIR	G-M	L-C	CIR
1	0.9687467	0.9675802	0.9667312	0.9745131	0.975375	0.9746218
2	0.9359893	0.9338505	0.9344820	0.9474162	0.9493179	0.9498125
3	0.9017191	0.8988305	0.9032244	0.9186905	0.9217853	0.9255616
4	0.8659432	0.8625925	0.8729309	0.8883334	0.8928057	0.9018586
5	0.8286869	0.8251883	0.8435749	0.8563619	0.8623639	0.8786932
6	0.7899974	0.7867624	0.8151303	0.8228153	0.8305487	0.8560553
7	0.7499469	0.7474037	0.7875717	0.7877578	0.7973758	0.8339348
8	0.7086362	0.7073286	0.7608741	0.7512811	0.7629379	0.8123220
9	0.6661983	0.6666686	0.7350133	0.7135067	0.7273784	0.7912069
10	0.6228010	0.6256087	0.7099656	0.6745870	0.6907472	0.7705801

Tabla 4.2: Probabilidad de supervivencia ${}_t p_{70}$.

t	HOMBRES			MUJERES		
	G-M	L-C	CIR	G-M	L-C	CIR
1	0.9291086	0.9322080	0.9287545	0.9408816	0.9443857	0.9413951
2	0.8573993	0.8644654	0.8624106	0.8806593	0.8877199	0.8860622
3	0.7853284	0.7971514	0.8006440	0.8196980	0.8302297	0.8338287
4	0.7134138	0.7307790	0.7431513	0.7584039	0.7723558	0.7845306
5	0.6422295	0.6658157	0.6896480	0.6972177	0.7144391	0.7380121
6	0.5723965	0.6023369	0.6398678	0.6366063	0.6565358	0.6941250
7	0.5045692	0.5403778	0.5935615	0.5770526	0.5984795	0.6527284
8	0.4394166	0.4802765	0.5504957	0.5190432	0.5405634	0.6136887
9	0.3775994	0.4248458	0.5104521	0.4630543	0.4862115	0.5768787
10	0.3197429	0.3690483	0.4732264	0.4095371	0.4298934	0.5421778

Tabla 4.3: Probabilidad de supervivencia ${}_t p_{80}$.

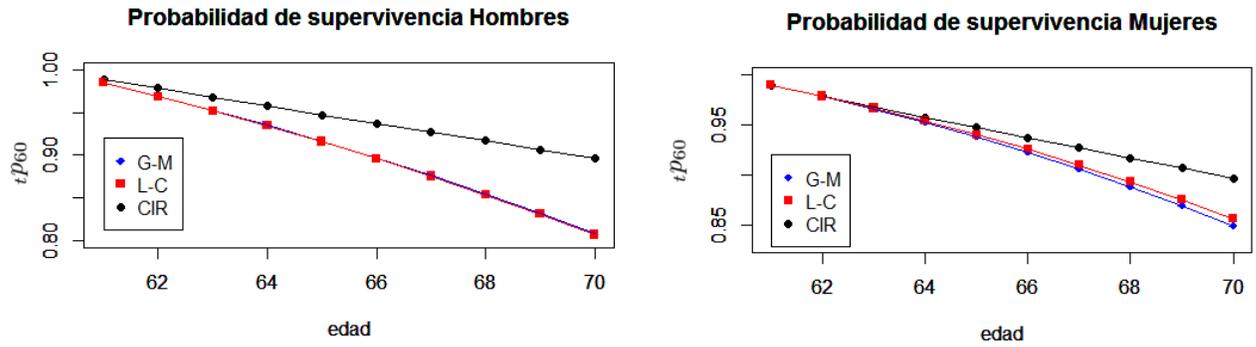


Figura 4.1: Curvas de supervivencia de individuos que en el 2009 tienen 60 años.

En las figuras 4.1, 4.2 y 4.3 se muestra una comparación entre las curvas de supervivencia $s_x(t)$ de cada modelo estudiado. En las tres figuras se puede observar que la función de supervivencia tiene un comportamiento muy diferente entre hombres y mujeres; las curvas de supervivencia de las mujeres tienen un decrecimiento más lento en comparación con los hombres; además la probabilidad de supervivencia es más alta en las mujeres que en los hombres. En la figura 4.1 observe que los modelos Gompertz-Makeham y Lee-Carter predicen la probabilidad de supervivencia muy parecido en los hombres, mientras que en las mujeres el modelo G-M queda por debajo de los otros dos conforme el tiempo avanza. El modelo CIR decrece más lento en comparación con los otros dos modelos, esto de acuerdo a la idea de que la esperanza de vida aumenta con el tiempo y por lo tanto la probabilidad de supervivencia decrece menos al avanzar el tiempo. Un comportamiento muy parecido se puede observar en la figura 4.2, con la diferencia de que el rango del valor de la probabilidad de supervivencia es menor a los 70 años en comparación con la de los 60 años, como era de esperarse. En la figura 4.3 se observa que las estimaciones de las curvas de supervivencia para los tres modelos se empiezan a separar, el rango de la probabilidad es más pequeño en comparación con los anteriores cohortes. El modelo G-M queda por debajo de los otros dos, mientras que el modelo CIR predice un valor más alto de la probabilidad de supervivencia.

Hechos todos los cálculos anteriores, podemos ahora conocer el valor al tiempo $t = 0$, es

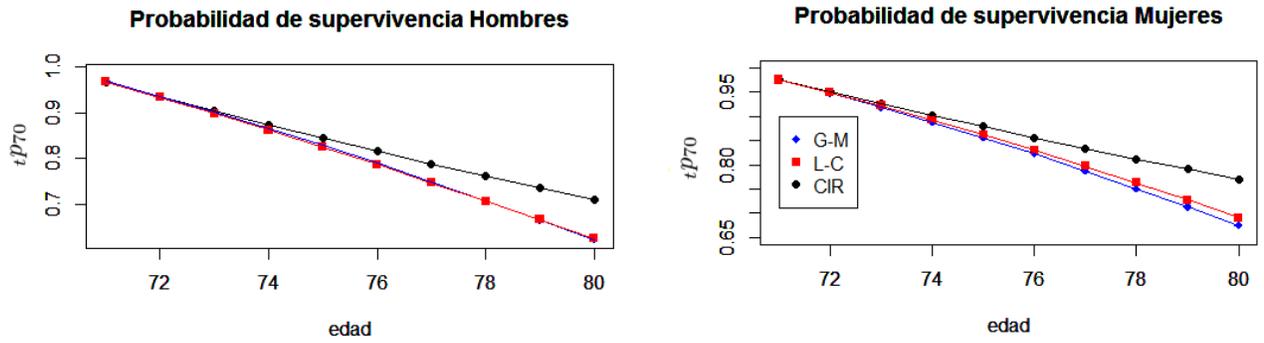


Figura 4.2: Curvas de supervivencia de individuos que en el 2009 tienen 70 años.

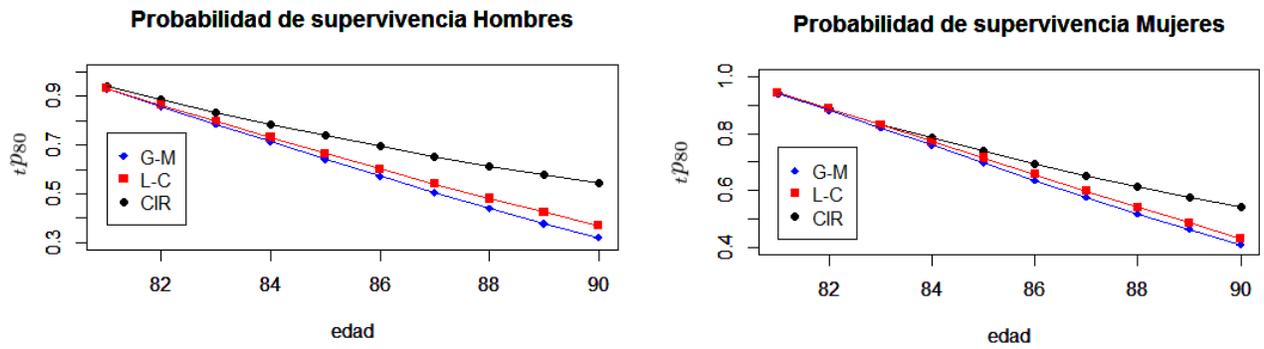


Figura 4.3: Curvas de supervivencia de individuos que en el 2009 tienen 80 años.

decir en el año 2009, de un bono cupón cero de longevidad que se calculó multiplicando el bono cupón cero que calculamos en la ecuación (4.1) por la probabilidad de supervivencia. Se calculó la probabilidad de supervivencia de 1 y 3 años de personas que en el 2009 tienen 50, 60, 70, 80 y 85 años, respectivamente considerando los tres modelos propuestos para la fuerza de mortalidad.

En las figuras 4.4, 4.5, 4.6 y 4.7 se puede observar el valor de un bono de longevidad con madurez de un año y tres años considerando diferentes edades de un asegurado hombre o mujer. Las figuras 4.4 y 4.6 ilustran el valor de un bono de longevidad con madurez de un año, se puede observar que para edades de 50 a 60 años, no se nota mucha diferencia entre los valores de los precios del bono estimado por los tres diferentes modelos; sin embargo, para edades mayores de 70 a 85 años se nota una diferencia entre los precios del bono estimado por cada uno de los modelos. En los hombres, para las edades mayores, el modelo G-M estima un precio mas bajo, mientras que el modelo L-C un precio más alto para valor del bono de longevidad. Para las mujeres, no se aprecia una diferencia entre los precios estimados por los modelos G-M y CIR, pero es también el modelo L-C el que estima un precio mas alto en comparación con los otros dos modelos. En las figuras 4.5 y 4.7 se ilustra el precio de un bono de longevidad con una madurez de 3 años para hombres y mujeres, respectivamente. Observe que para ambos, el modelo G-M estima un precio mas bajo para el bono y el modelo CIR un precio más alto.

Con el objetivo de comparar con datos reales de mortalidad todos los resultados anteriores, se planteó realizar un *back-testing*. Como se mencionó anteriormente, de la base de datos de CONAPO sólo se pudo recopilar información de defunciones y número total de la población para el periodo de 1990 a 2009. Se hizo entonces una búsqueda de información en las bases de datos del Instituto Nacional de Estadística y Geografía (INEGI). Sin embargo, la información que se encontró para realizar esta tarea no fue adecuada, pues encontramos información del número de defunciones del periodo 2010-2015, pero agrupada en bloques de edades por cada 5 años. Además, el número total de la población al inicio de año sólo está publicada cada 10

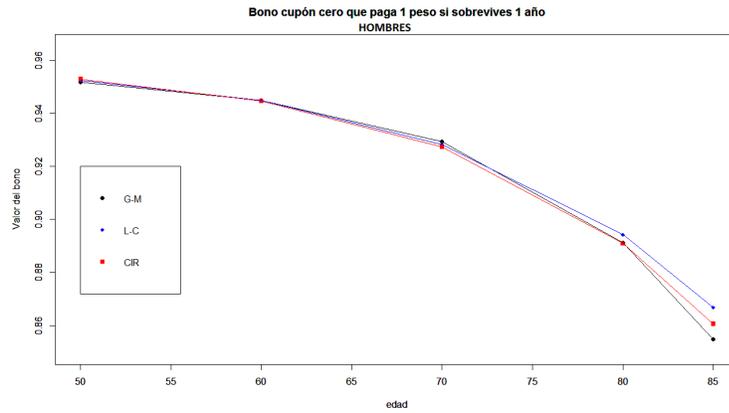


Figura 4.4: Valor de un bono cupón cero de longevidad para una madurez de un año para diferentes edades.

años. Por lo tanto, debido a estas dificultades no se realizó la prueba. Sería muy interesante en el futuro firmar un convenio con una compañía de seguros que nos pudiera dar acceso a su base de datos para realizar un ejercicio similar.

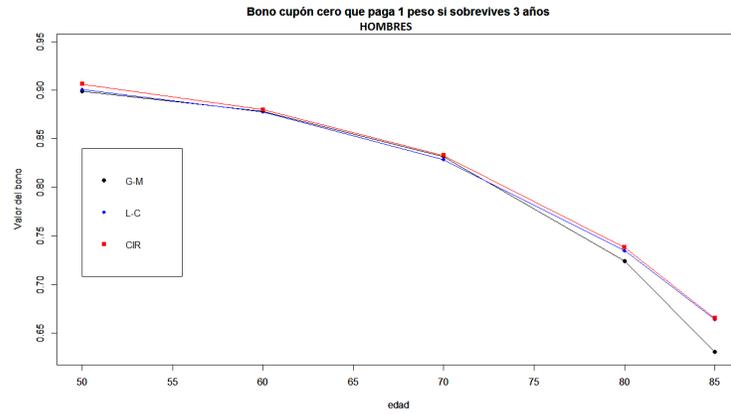


Figura 4.5: Valor de un bono cupón cero de longevidad para una madurez de tres años para diferentes edades.

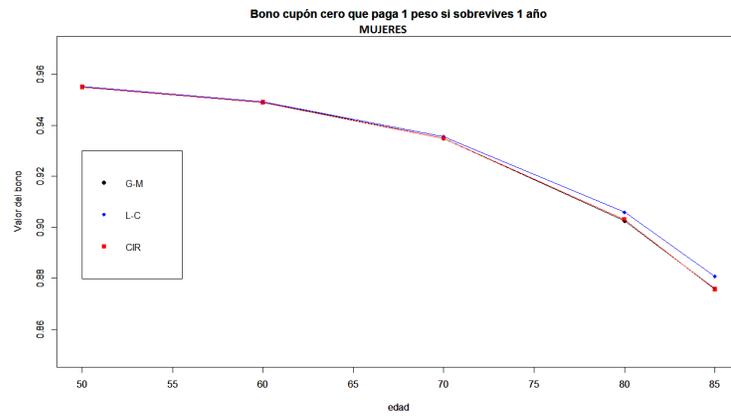


Figura 4.6: Valor de un bono cupón cero de longevidad para una madurez de un año para diferentes edades.

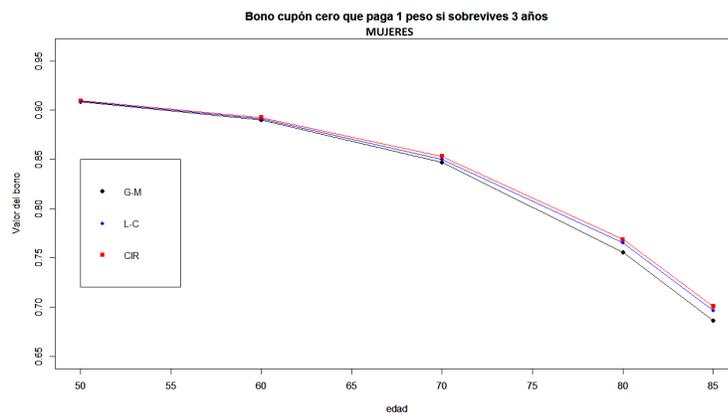


Figura 4.7: Valor de un bono cupón cero de longevidad para una madurez de tres años para diferentes edades.

Conclusiones y perspectivas

Se estudiaron dos de los modelos más utilizados en la literatura para describir el comportamiento de la fuerza de mortalidad. Uno de ellos es el modelo Gompertz-Makeham que estima la probabilidad de supervivencia de un individuo a través de la construcción de una función de distribución. El segundo es el modelo Lee-Carter, el cual a partir del ajuste de datos históricos de mortalidad, hace predicciones para un futuro comportamiento de la fuerza de mortalidad. Cabe señalar, que en el trabajo de Romero R. [18] se hizo una comparación entre las tablas dinámicas de mortalidad hechas con este modelo y las tablas de mortalidad estáticas publicadas en el Diario Oficial de la Federación (DOF). Las estimaciones fueron hechas hasta el 2025 para la población mexicana, tomando como base la información del periodo de 1990 a 2009. Llegaron a la conclusión de que las tablas estáticas están sobre estimadas, lo cual encarece los productos o servicios que ofrecen compañías de seguros o pensiones que las utilizan. Ambos son modelos paramétricos, pero la diferencia entre uno y otro radica en que el primero solo considera parámetros que dependen de la edad de un individuo, mientras que en el segundo se incluye un parámetro que cambia dependiendo del tiempo; lo anterior basado en que la mortalidad sufre cambios al pasar el tiempo. Se seleccionó de la literatura especializada en tasas de interés, un modelo estocástico con reversión a la media tipo CIR para describir la dinámica de la fuerza de mortalidad. Esto substituyendo la tasa de interés por

la fuerza de mortalidad. La elección de un modelo estocástico se sustenta en la idea de que la mortalidad depende de sucesos aleatorios como avances médicos, estilos de vida, epidemias, guerras, cambios bruscos en el clima, entre otros.

En la construcción de las curvas de supervivencia para diferentes cohortes generacionales, se puede notar una diferencia entre la estimación de la probabilidad de supervivencia para hombres y mujeres. Los tres modelos logran capturar el hecho de que las mujeres son más longevas que los hombres y por lo tanto las curvas de supervivencia decrecen más lento en comparación con la de los hombres.

El modelo Gompertz-Makeham tiene la desventaja de ser estacionario en el tiempo, lo cual repercute en subestimar la probabilidad de supervivencia de un asegurado, esto se observa en las figuras 4.1, 4.2 y 4.3 en las cuales este modelo queda por debajo de los otros dos. Esto sería un problema muy grave para una empresa de seguros de vida, pues si se subestima la mortalidad corren el riesgo de vender planes de seguros pocos rentables.

El modelo estocástico con reversión a la media que se estudió, cumple con los requerimientos propuestos enunciados en la sección 2.3 para la fuerza de mortalidad. Es decir, este modelo garantiza que la fuerza de mortalidad es positiva para todo tiempo, además la fuerza de mortalidad para los ancianos aumenta a lo largo del tiempo, lo cual se observó al tener un decrecimiento en las curvas de supervivencia a medida que el tiempo pasa. Sin embargo, es discutible el supuesto de que la mortalidad tenga reversión a la media, pues no está garantizado que la mortalidad oscile alrededor de un valor medio. Una posible solución a este inconveniente, que quedará como un trabajo futuro, es tratar de modelar a la fuerza de mortalidad con un modelo estocástico de raíz cuadrada sin reversión a la media tal como se plantea en la ecuación (3.18) propuesto por Luciano E. y Vigna E. 2008 [13]. Además, dado que la probabilidad neutra al riesgo no es única, sería más conveniente valuar los instrumentos financieros con la metodología propuesta para mercados incompletos.

Bibliografía

- [1] Bowers N., Gerber H., Hickman J., Jones D. and Nesbitt C.; *Actuarial Mathematics, The Society of Actuaries*, 1997.
- [2] Cairns, Blake and Dowd, *Pricing Death: frameworks for the valuation and securitization of mortality. Center for Risk and Insurance Studies, University of Nottingham*, 2006.
- [3] Carstens C., Documento interno de la Comisión Nacional de Seguros y Finanzas, 2013.
- [4] Cox, J. C., Ingersoll, J.E. y Ross, S.A. *A theory of term structure of interest rates. Econometrica*,53,385-407, 1985.
- [5] Currie I.D., Durban M. and Eilers, *Smoothing and forecasting mortality rates.*, *Statistical Modelling*, **4**, 297-298, 2004.
- [6] Dahl M., *Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. Insurance: Mathematics and Economics*, **35**, 113-136. 2004.
- [7] Dana R. A., Jeanblanc M., *Financial markets in continuous time*. Springer, 2003.
- [8] Debón Aucejo A., Graduación de tablas de mortalidad. Aplicaciones actuariales, Universidad de Valencia. Tesis Doctoral, España 2003.

- [9] Hyndman Rob J., *Robust forecasting of mortality and fertility rates: a functional data approach*. Computational Statistics and Data Analysis. Volume 51, Issue 10, Pages 4942–4956, 15 June 2007.
- [10] Karatzas, I. y Shreve, S. *Brownian motion and stochastic calculus*. Springer Verlag N.Y. 1998.
- [11] Lay, D. C.. *Algebra Lineal y sus Aplicaciones*. Prentice-Hall, Mexico, 2a edición, 2001.
- [12] Lee, R. y Carter, L. *Modelling and forecasting U.S. mortality*. *Journal of the American Statistical Association*, 1992.
- [13] Luciano E. and Vignam E., *Mortality risk via affine stochastic intensities: calibration and empirical relevance*, MPRA Paper No. 59627, posted 4. November 2014 05:37 UTC.
- [14] Menoncin F. *Death bonds with stochastic force of mortality*, 2009.
- [15] Milevsky, M.A., y Promislow S.D., *Mortality derivatives and the option to annuitise*. *Insurance: Mathematics and Economics*, 2001.
- [16] Pérez F. Ma. J., *Cobertura alternativa del riesgo de longevidad a través de bonos y swaps de los mercados de capital*. Universidad Carlos III de Madrid. 2006.
- [17] Remillard B., *Statistical Methods for Financial Engineering*, 2013.
- [18] Romero Rangel L. M., *Predicción de tablas de vida dinámicas hasta el año 2025 para México*. Tesis de Maestría, Colegio de Postgraduados, 2012.
- [19] Verrall, R.J., *A unified framework for graduation*, Actuarial Research Paper No.91. Department of Actuarial Science and Statistics, City University London, November 1996, ISBN 1 874 770 99 9.
- [20] http://www.conapo.gob.mx/es/CONAPO/Proyecciones_Datos

[21] <http://www.banxico.org.mx/SieInternet/consultarDirectorioInternetAction>