



UNIVERSIDAD DE LA REPÚBLICA
Facultad de Ciencias Económicas y de Administración
Licenciatura en Estadística
Informe de Pasantía

**Tabla de Mortalidad
para empleados del BROU**

Lara Blanco

Tutores:

Ramón Alvarez

Gabriel Camaño

Montevideo, 25 de Febrero de 2016.

UNIVERSIDAD DE LA REPÚBLICA

FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN

El tribunal docente integrado por los abajo firmantes aprueba el trabajo de

Pasantía:

Tabla de Mortalidad para empleados del BROU

Lara Blanco

Tutores:

Ramón Alvarez

Gabriel Camaño

Licenciatura en Estadística

Puntaje

Tribunal

Profesor.....(nombre y firma).

Profesor.....(nombre y firma).

Profesor.....(nombre y firma).

Fecha.....

Agradecimientos

En primer lugar quiero agradecerle a Juan Cirielli¹, quien pensó en mí para la realización de este estudio como trabajo de pasantía. A Carla Angele-
ro², Tania Steffen³ y al BROU (donde actualmente me encuentro trabajando)
por la oportunidad y la confianza otorgada.

A mis tutores, en especial a Ramón Álvarez, que a pesar de haber pasado un año complicado en relación a su salud, me ha ayudado continuamente.

A mi familia y amigas que siempre están presentes con su apoyo infinito y tolerancia en mis momentos de estrés, siempre alentándome a seguir adelante.

A mis compañeros del grupo de viaje que me bancaron la ausencia, en el momento más crucial de comprar vuelos y reservar hospedaje, para que yo pudiera concluir este trabajo.

¹Controlador Estadístico de la Oficina de Políticas y Control de Riesgo del BROU

²Gerente del Área Contabilidad y Control del BROU

³Analista Contable de la Oficina de Análisis Contable y Gestión del BROU

Resumen

El objetivo principal del presente trabajo es construir una tabla de mortalidad específica para los funcionarios del Banco República Oriental del Uruguay. Motiva dicho trabajo el considerar que éstos podrían llegar a tener una mortalidad menor a la del resto de la población uruguaya, es decir, se esperaría que los mismos presentaran una esperanza de vida superior a la de la población en general debido a la calidad de vida de los funcionarios, ya que los mismos tienen una jornada laboral de lunes a viernes de 6 hora, 30 minutos cuando la mayoría de la población trabaja 8 horas o más, lo cual les permite contar con más tiempo libre para descansar y desarrollar otras actividades como puede ser, actividad física. Actividades que pueden acceder también por sus buenos sueldos y otros beneficios que les brinda el banco a sus funcionarios.

Los datos utilizados fueron brindados por el BROU. Los mismos se conforman de toda la plantilla de funcionarios activos al 31/07/2014, de los funcionarios que se fueron del banco a igual fecha por alguna causal, entre ellas, la jubilación y también se brindó información de los funcionarios fallecidos (datos que le brindó la Caja de Jubilaciones y Pensiones Bancarias al BROU).

Para comparar la tabla de mortalidad obtenida se construyen las tablas de mortalidad de la población uruguaya a partir de las tasas de mortalidad del Instituto Nacional de Estadística y las del Banco Central del Uruguay.

Dada la población objeto de estudio, se simulan los fallecimientos de la misma a través de la tasa de mortalidad del BCU mediante el método Monte Carlo pretendiendo ver que tan diferente se comporta la mortalidad de los funcionarios del BROU al utilizar otra tasa de mortalidad.

Por último se proyecta la mortalidad para los próximos cinco años utilizando el modelo de Lee-Carter mediante dos métodos diferentes de estimación de sus parámetros.

Palabras clave: *BROU, métodos demográficos, método de Lee-Carter, población bancaria, proyección de la mortalidad, Tabla de mortalidad*

Índice general

Índice general	IX
Índice de Cuadros	XIII
Índice de figuras	XVII
Introducción	3
1. Metodología	11
1.1. La Tabla de Mortalidad	11
1.1.1. Tipos de tablas	12
1.1.2. Estructura	14
1.1.3. Construcción de la Tabla de Mortalidad	26
1.2. Descripción de la mortalidad en el tiempo	29
1.2.1. Modelo de Lee-Carter	29
1.2.2. Estimación de los parámetros del modelo	31
1.2.3. Diagnóstico y bondad de ajuste del modelo	35
1.2.4. Proyección	38
2. Datos de Aplicación	41

2.1. Análisis Descriptivo	42
2.1.1. Activos	42
2.1.2. Jubilados	44
2.1.3. Expuestos	45
2.1.4. Fallecidos	47
2.2. Datos de aplicación	48
3. Resultados	51
3.1. Tabla de mortalidad BROU	52
3.1.1. Tablas de mortalidad a partir de los datos del INE y del BCU	59
3.1.2. Tabla empírica simulada	63
3.2. Descripción y Proyección de la mortalidad	67
3.2.1. Estimación del modelo mediante SVD y proyección . . .	69
3.2.2. Estimación del modelo mediante MV y proyección . . .	74
4. Conclusiones	81
Bibliografía	87
A. Resultados	93
A.1. Datos BROU	93
A.2. Datos INE	97
A.3. Datos BCU	100
A.4. Tabla empírica simulada	100
A.5. Lee-Carter	102

Índice de Cuadros

3.1. Expuestos al riesgo de morir por año y edades agrupadas . . .	53
3.2. Fallecidos por año y edades agrupadas	53
3.3. Tasas centrales de mortalidad observada por año para edades agrupadas	54
3.4. Estimación de q_x por los métodos Lineal, Exponencial, Reed- Merrell, Greville y Keyfitz	54
3.5. Tabla de mortalidad para los funcionarios del BROU	56
3.6. Tabla de mortalidad para los funcionarios del BROU suavizada	58
3.7. Tabla de mortalidad abreviada BCU	63
3.8. Promedio de los fallecimientos simulados y fallecimientos ob- servados de la muestra del BROU	66
3.9. Estimaciones a_x y b_x de los modelos de Lee-Carter con librerías lca e ilc para los grupos de edades 40-44 años a 75-79 años . .	78
3.10. Estimaciones k_t de los modelos de Lee-Carter con librerías lca e ilc para los grupos de edades 40-44 años a 75-79 años	79
4.1. Esperanza de vida e_x según BCU y para el BROU respectiva- mente	82

A.1. Cantidad de expuesto al riesgo de morir por edad simple durante los años 1995 a 2013	94
A.2. Cantidad de fallecidos por edad simple durante los años 1995 a 2013	95
A.3. Esperanzas de vida e_x por grupos de edades calculadas a partir de los distintos métodos de estimación de q_x : Lineal, Exponencial, Reed-Merrel, Greville y Keyfitz	96
A.4. Esperanzas de vida e_x por grupos de edades para los años 2011 a 2013	96
A.5. Tasas de mortalidad q_x a partir de la m_x del INE utilizando los métodos: Lineal, Exponencial, Reed-Merrell y Greville . . .	98
A.6. Tabla de mortalidad generada con la q_x hallada por el método lineal a partir de la m_x del INE y e_x calculada mediante el método planteado en (Alho, 2005)	99
A.7. Tabla de mortalidad calculada a partir de la q_x del BCU, para la población femenina y la masculina respectivamente	101
A.8. Tasa central de mortalidad simulada	102
A.9. p-valores de los tests Jarque-Bera y Ljung-Box, modelo Lee-Carter con librería <i>demography</i> para todos los grupos de edades de 25-29 a 80 años y más	102
A.10. p-valores de los tests Jarque-Bera y Ljung-Box, modelo Lee-Carter con librería <i>demography</i> para los grupos de edades de 35-39 a 75-79 años	103

A.11.p-valores de los tests Jarque-Bera y Ljung-Box, modelo Lee-Carter con librería <i>demography</i> para los grupos de edades de 40-44 a 75-79 años	103
A.12.Estimaciones de a_x y b_x del modelo de Lee-Carter con librería <i>demography</i> para los grupos de edades 40-44 a 75-79 años . . .	104
A.13.Tendencia de la mortalidad estimada k_t del modelo de Lee-Carter con librería <i>demography</i> para los grupos de edades 40-44 a 75-79 años	104
A.14.Tendencia de la mortalidad k_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería <i>demography</i> para los grupos de edades 40-44 a 75-79 años e intervalos de confianza	104
A.15.Tasa de mortalidad m_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería <i>demography</i> para los grupos de edades 40-44 a 75-79 años	105
A.16.Estimaciones de a_x y b_x del modelo de Lee-Carter con librería <i>ilc</i> para los grupos de edades 40-44 años a 75-79 años	105
A.17.Tendencia de la mortalidad estimada k_t del modelo de Lee-Carter con librería <i>ilc</i> para los grupos de edades 40-44 años a 75-79 años	105
A.18.Tendencia de la mortalidad k_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería <i>ilc</i> para los grupos de edades 40-44 años a 75-79 años e intervalos de confianza . .	106
A.19.Tasa de mortalidad m_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería <i>ilc</i> para los grupos de edades 40-44 años a 75-79 años	106

A.20. Estimaciones de a_x y b_x del modelo de Lee-Carter con librería ilc para todos los grupos de edades desde los 25-29 años a 80 y más	107
A.21. Tendencia de la mortalidad estimada k_t del modelo de Lee-Carter con librería ilc para todos los grupos de edades desde los 25-29 años a 80 y más	107
A.22. Tendencia de la mortalidad k_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería ilc para todos los grupos de edades desde los 25-29 años a 80 y más e intervalos de confianza	107
A.23. Tasa de mortalidad m_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería ilc para todos los grupos de edades desde los 25-29 años a 80 y más	108

Índice de figuras

1.1. Funciones l_x , d_x y q_x de la población uruguaya	15
2.1. Distribuciones por edad y sexo de los funcionarios activos . . .	42
2.2. Pirámide poblacional de funcionarios activos	43
2.3. Distribuciones por edad y sexo de los funcionarios jubilados . .	44
2.4. Pirámide poblacional de jubilados	45
2.5. Distribuciones por edad y sexo de la población expuesta al riesgo de morir	46
2.6. Pirámide poblacional de expuestos al riesgo de morir	46
2.7. Distribuciones por edad y sexo de los funcionarios fallecidos .	47
2.8. Pirámide poblacional de fallecidos	47
2.9. Distribuciones por edad de los hombres expuestos al riesgo de morir y de los hombres fallecidos	49
3.1. Tasas de mortalidad q_x	55
3.2. Funciones l_x , d_x y q_x de los funcionarios del BROU	57
3.3. Esperanzas de vida e_x para los distintos tramos de edades durante los años 2011 a 2013	58
3.4. Funciones l_x , d_x y q_x con <i>smooth</i>	59

3.5. Tasa de mortalidad q_x según datos del INE calculadas a partir de los métodos: Lineal, Exponencial, Reed-Merrell y Greville	60
3.6. Tasas centrales de mortalidad m_x , simuladas y observadas respectivamente, para los años 2007 a 2013	66
3.7. m_x y $\log(m_x)$ por edad para los años 1995 a 2013 y $\log(m_x)$ por año para las distintas edades	68
3.8. d_x y $\log(d_x)$ por edad para los años 1995 a 2013	69
3.9. Residuos del modelo de Lee-Carter estimado para los grupos de edades 40-44 a 75-79 años durante los años 1995 a 2013, mediante la librería <i>demography</i>	71
3.10. Estimaciones de a_x , b_x y k_t mediante SVD	71
3.11. Proyección de k_t y $\log(m_x)$ por edad para los años 2014 a 2018	73
3.12. Tasas de mortalidad proyectadas para los años 2014 a 2018	74
3.13. Estimaciones a_x , b_x y k_t mediante la librería <i>ilc</i>	75
3.14. Estimaciones a_x , b_x y k_t para todos los grupos de edades mediante MV	76
3.15. Proyección de k_t y $\log(m_x)$ por edad para los años 2014 a 2018	77
3.16. Tasas de mortalidad proyectadas para los años 2014 a 2018 de todos los grupos de edades	78
A.1. Funciones l_x , d_x y q_x construidas a partir de la m_x del INE	97
A.2. Funciones l_x , d_x y q_x construidas a partir de la m_x del BCU	100
A.3. Residuos del modelo de Lee-Carter, estimado mediante librería <i>demography</i> para todos los grupos de edades de 25-29 a 80 años y más	103

A.4. Residuos del modelo Lee-Carter estimado con la librería <i>ilc</i> para los grupos de edades 40-44 años a 74-79 años	106
A.5. Residuos del modelo Lee-Carter estimado con la librería <i>ilc</i> para todos los grupos de edades desde los 25-29 años a 80 y más	108

Introducción

Introducción

La Tabla de Mortalidad, también llamada Tabla de Vida, es uno de los métodos más utilizados en demografía para el análisis de la mortalidad y la supervivencia de una población. Mide la probabilidad de vida o de muerte de dicha población en función de su edad. Eso permite responder a preguntas como ¿cuál es la probabilidad que un hombre de 30 años sobreviva hasta su edad de jubilación a los 60? o ¿cuántos años se espera que viva una persona que ya alcanzó los 65 años de edad?

Siendo su estudio de interés tanto para los demógrafos como para los actuarios y los diversos profesionales vinculados a los temas de salud pública y planificación, en una gran variedad de problemas, entre los cuales pueden mencionarse, la estimación del nivel y la tendencia de la mortalidad, los análisis de la mortalidad por causas, los estudios de fecundidad, estructura dinámica y crecimiento de la población, y el análisis de diversas características socioeconómicas, tales como, la composición de la fuerza de trabajo y la regulación de los sistemas de jubilaciones y pensiones para las personas que pasan a la edad de retiro.

Antecedentes

Nivel local

Dada la importancia de las tablas de mortalidad para el cálculo de las previsiones actuariales es que se busca información, a nivel local, en las siguientes cajas previsionales: CJPB, CJPPU, CN, SRPFFAA, DNASSP, BPS y AFAPs.⁴

La única caja que se encontró que usa una tabla de mortalidad específica para su población afiliada es la CJPPU⁵, la cual utiliza tablas dinámicas. Siendo las tablas dinámicas las que contemplan en sus formulaciones, no solo el tiempo biológico o edad de los individuos, sino también el tiempo cronológico.

En la página web del BPS se encuentran varios estudios referentes a la tasa de mortalidad:

- Estimación de las Tasas de Mortalidad específicas para los Jubilados por Vejez e Invalidez del Régimen Previsional Contributivo Uruguayo de (Lazo, 2010). En dicho estudio se utilizan los datos del stock de jubilados por vejez e invalidez para los años 2006 y 2008 del Centro de Desarrollo de Servicios Informáticos de Prestaciones del BPS y se procedió al análisis y cálculo de la correspondiente probabilidad de sobrevivencia bi – anual para cada uno de los colectivos por sexo y edad simple. Luego se calculan las tasas de mortalidad y se aplica un método logit para su graduación.

⁴El significado de todas las siglas utilizadas en el presente trabajo se encuentran descriptas en la página 91

⁵el dato fue obtenido de los estados contables 2012 y 2013 de la mencionada caja.

- Análisis del Equilibrio Financiero Individual de un Sistema de Prestación Definida, computando mejoras futuras en las Tasas de Mortalidad de (Camacho, 2010). Aquí se compara la ecuación de equilibrio financiero, a través de la utilización de tablas de mortalidad estáticas y dinámicas, concluyendo en la importancia de la utilización de tablas dinámicas y dado que en la práctica en muchos casos no se utilizan tablas de mortalidad dinámicas, se destaca la importancia de la actualización de las tasas de mortalidad.
- Estimación de las tasas de mortalidad futuras para su aplicación en las proyecciones financieras del régimen previsional de (Camacho, 2009), donde también destaca la importancia de la utilización de tablas dinámicas.

Las AFAPs utilizan la tabla de mortalidad que dispone el BCU en la circular N^o 2111 del 27 de junio de 2012. La misma se extiende hasta los 110 años y ha sido motivo de reclamo del PIT-CNT⁶, ya que afirman que ese mecanismo, reduce las prestaciones que se podría cobrar y reclaman que se permita a las aseguradoras calcular las jubilaciones en función de las expectativas de vida planteadas por el INE, que van hasta los 95 años⁷. Desde un estudio de la Super Intendencia del Sistema Financiero (SSF) del BCU-Dassattia and Natalia Mariño, 2014, se concluye que en base a modelos de otros países con mayores flexibilidades en la participación de los agentes del mercado, se debería permitir el cálculo de tablas por parte de las asegurado-

⁶Comunicado a la opinión pública de fecha 09/09/2015

⁷<http://www.noticiasbyo.org/2012/05/07/controversia-por-forma-de-calculo-de-jubilacion-a-traves-de-las-afap/>

ras y su validación por parte de las SSF. También se nombra la importancia en la actualización y metodología de cálculo de la tabla de mortalidad, introduciendo tablas de mortalidad dinámicas o la revisión de los parámetros utilizados en la metodología vigente por parte de la SSF.

Para el colectivo de los funcionarios bancarios, no se encontró que existan tablas de mortalidad específicas. Tampoco se encontraron publicaciones de otros países para poblaciones específicas de empleados bancarios.

Nivel internacional

Se investiga a nivel internacional, para tener una referencia de otros países, cual es la metodología utilizada en la construcción de las tablas de mortalidad por parte de los Institutos de Estadística Nacionales de Argentina, España y Chile. Lo que marca la diferencia en la construcción, es la manera con la que se estima la tasa de mortalidad.

De la publicación realizada por el INDEC de Argentina, acerca de la estimación y proyección de la población 2010-2040 (INDEC, 2013), se construyeron tablas de mortalidad nacionales y provinciales para el año 2009 con el promedio de defunciones de 2008, 2009 y 2010 y una población estimada para mitad del año 2009, que se obtuvo mediante interpolación.

El INE de España en el informe de Metodología de tablas de mortalidad (INE, 2015), estima la tasa específica de mortalidad a la edad x observada sobre la población en estudio, m_x , bajo la hipótesis de distribución uniforme de los cumpleaños de todos los individuos de la población que no mueren a lo largo del año con una determinada edad y de distribución también uniforme

a lo largo del año del día de llegada de los individuos que se incorporan a la población en estudio y del día de salida de los individuos que emigran de dicha población durante el año de observación, mediante la expresión:

$$m_x = \frac{D(t, x, s)}{\frac{P(t, x, s) - D_2(t, x, s)}{2} + \sum_{i=1}^{D_2(t, x, s)} b_2(t, x, s, i) + \frac{P(t+1, x, s)}{2} + \sum_{i=1}^{D_1(t, x, s)} b_1(t, x, s, i)}$$

Siendo:

- t , el año o período de observación
- x , la edad o años cumplidos, con $x = 0, 1, \dots, 99$
- s , el sexo, que puede tomar los atributos varón, mujer o ambos sexos
- $P(t, x, s)$, es el stock de población residente a 1^o de enero del año t , con edad x y sexo s
- $D(t, x, s)$, es el número de fallecidos en el año t , con edad x y sexo s
- $D_1(t, x, s)$, es el número de fallecidos en el año t , con edad x y sexo s , que cumple x años en a lo largo del año t
- $D_2(t, x, s)$, es el número de fallecidos en el año t , con edad x y sexo s , que cumplió x años en a lo largo del año $t - 1$
- $b_1(t, x, s, i)$, se define como la diferencia (en años) entre la fecha de defunción y la fecha de cumpleaños (en el año t) de cada individuo i de sexo s fallecido durante el año t , con edad x y que cumplió los x años a lo largo de t

- $b_2(t, x, s, i)$, se define como la diferencia (en años) entre la fecha de defunción y el 1^o de enero del año t para cada individuo i , de sexo s , fallecido durante el año t , con edad x y que cumplió los x años a lo largo de $t - 1$

Asimilando los valores estimados de las tasas específicas de mortalidad de la población en estudio con los correspondientes a las tasas específicas de mortalidad en cada edad x de una cohorte ficticia de 100.000 individuos, la probabilidad o riesgo de muerte a la edad x , q_x , de dicha cohorte de individuos, que presenta la misma incidencia de la mortalidad a cada edad que la población observada en el año de referencia, se estima por la expresión:

$$q_x = \frac{m_x}{1 + (1 + c_x)m_x}$$

donde c_x , es el promedio de años vividos en el último año de vida por aquellos individuos de la cohorte ficticia que mueren con edad cumplida x . Para el grupo abierto (100 o más años) de edad, considerando que el suceso de muerte es un suceso seguro, se tiene: $q_{100+} = 1$

El INE de Chile construye tablas de mortalidad abreviadas con los datos del censo del año 2002 (INE, 2004). El cálculo de la tasa central de mortalidad para los grupos de edades mayores a 5 años, se calcula del siguiente modo:

$${}_5m_x = \frac{\frac{1}{2}({}_5D(t, x) + {}_5D(t + 1, x))}{{}_5N(t + 1, x)}$$

donde el numerador, corresponde al promedio de las defunciones de los años $t(2001)$ y $t + 1(2002)$ ocurridas a personas de un sexo y con edades entre los

x y $x + 5$ años de edad y el denominador, corresponde a la población del mismo sexo y edades que para el caso de las defunciones, pero referida al 1^o de enero del año $t + 1$ (2002). Las tablas de mortalidad, se calculan utilizando la tasa de mortalidad, q_x , con metodología de Reed-Merrell, que se detalla más adelante.

Como se observa, se pueden encontrar diferentes maneras de estimar q_x para construir las tablas de mortalidad.

Objetivos

El objetivo general del presente trabajo es realizar una tabla de mortalidad de los funcionarios del Banco República (BROU) para luego ser utilizada por éste en el cumplimiento de la Norma de Contabilidad Internacional N° 19 referente a los Beneficios de los Empleados. La Norma requiere que la entidad reconozca:

- (a) un pasivo, cuando el empleado ha prestado servicios a cambio de beneficios a los empleados a pagar en el futuro; y
- (b) un gasto, cuando la entidad consume el beneficio económico procedente del servicio prestado por el empleado a cambio de los beneficios a los empleados.

Para realizar estos cálculos y dado que hay beneficios que el banco sigue pagando luego de que el funcionario se jubila, es que se precisa estimar hasta qué edad viven los funcionarios bancarios. O sea, su edad de fallecimiento, momento en el cual se deja de pagar por tales beneficios.

En función de éste, surgen los siguientes objetivos específicos:

- Determinar si la población bancaria tiene mayor esperanza de vida a la de la población general uruguaya.
- Determinar la diferencia entre la esperanza de vida masculina y femenina, ya que es una regla general que las mujeres vivan más años que los hombres.
- Proyectar la mortalidad de los funcionarios del BROU para los próximos años.

En el próximo capítulo, se desarrollan las metodologías a aplicar referente a la construcción de tablas de mortalidad y a la proyección de la mortalidad. En el capítulo 2, se describen y preparan los datos para ser analizados mediante la metodología planteada. En el capítulo 3, se presentan los resultados obtenidos de aplicar la citada metodología a los datos de análisis. El informe termina en el capítulo 4, donde se plantean las conclusiones a las que se llega al finalizar el estudio, las limitaciones que se tuvieron y algunas recomendaciones para futuros análisis.

Capítulo 1

Metodología

En este capítulo se desarrollan las metodologías utilizadas para el análisis de los datos de los funcionarios del BROU. El capítulo se divide en dos grandes secciones. La primera hace referencia a la construcción de las tablas de mortalidad y la segunda al estudio de la mortalidad en el tiempo y su proyección futura basada en el método de Lee-Carter.

Para el desarrollo de la construcción de tablas de mortalidad, se sigue básicamente a los autores (Hinde, 2014) y (Caselli et al., 2006). En la segunda sección, referente a la descripción de la mortalidad en el tiempo, se toma como referencia lo propuesto por (Lee and Carter, 1992) y por (Wilmoth, 1993).

1.1. La Tabla de Mortalidad

La tabla de mortalidad es una herramienta para el análisis de la mortalidad de una población. La mortalidad muestra variaciones significativas en relación a ciertas características, siendo la edad, la variable demográfica más

importante en el análisis de la mortalidad debido a la relación estrecha que hay entre ésta y el riesgo de muerte.

Las características innatas más agudamente definidas son el sexo y la raza. Siendo el sexo, una característica de primordial importancia en el estudio de la mortalidad, ya que está comprobado de que mortalidad masculina es más alta que la femenina.

Otras características demográficas en el análisis de la mortalidad son el estado civil, el nivel socioeconómico y el lugar de residencia (urbano o rural).

La mortalidad también varía según la comunidad y el entorno físico en el cual se vive. Estas características incluyen el clima, la altitud, la calidad de la asistencia médica, las condiciones medioambientales tales como el tipo de suministro de agua, el grado de polución del aire y la calidad y cantidad de comida a la que se accede, así como otras condiciones socioeconómicas.

En este estudio las características que se toman en cuenta son únicamente la edad y el sexo dado que al ser un grupo muy específico de la población ya se está asumiendo que todos cuentan con ciertas características en común que podrían influir en la mortalidad del colectivo y diferenciarla de la mortalidad del resto de la población uruguaya.

1.1.1. Tipos de tablas

Las tablas de mortalidad o de vida, se pueden clasificar teniendo en cuenta varios aspectos. Una primera distinción se puede hacer según el año de referencia de la tabla, teniendo: *la tabla de mortalidad de momento o de período* y *la tabla de mortalidad de cohorte o generación*. La primera tabla, se

basa en la experiencia sobre un período corto de tiempo, tal como 1 o 3 años, o un período intercensal, donde la mortalidad se mantiene sustancialmente igual. Este tipo de tablas, representa la experiencia de la mortalidad por conjuntos de edades de la población, en un período particular del tiempo; ésta no representa la mortalidad de la cohorte actual. En su lugar, se asume una cohorte hipotética, que es sujeta a las tasas de muerte específicas por edad observadas en el período particular. Luego, la tabla de mortalidad de período puede ser vista como una foto de la mortalidad actual. Y es una excelente descripción resumida de la mortalidad en ese período corto de tiempo.

El otro tipo de tablas de mortalidad, las tablas de mortalidad de generación o cohorte, se basan en las tasas de mortalidad experimentadas por una cohorte particular. Acorde con este tipo de tablas, la mortalidad experimentada por las personas en la cohorte serían observaciones desde el momento de su nacimiento, pasando por cada edad consecutiva en sucesivos años hasta que todas las personas objeto de estudio fallezcan. Obviamente, se necesita un largo período de años para completar una tabla y por eso se buscan otros métodos.

Otra clasificación para las tablas de vida, es según el largo del intervalo de las edades en que los datos son presentados. Éstas pueden ser, *completas* o *abreviadas*. La tabla de mortalidad completa, contiene los datos para cada edad singular de edad (0, 1, 2, 3, 5 años, etc). La tabla de mortalidad abreviada, contiene los datos en intervalos de 5 o 10 años para la mayor parte del rango de edades, salvo las primeras edades y el grupo de edades más avanzadas, por ejemplo: 0-1, 1-4, 5-9, ..., 75-79, 80-84 años y el último grupo que contiene todas las personas mayores de 85 años. La división del primer

intervalo de edades, se debe a la alta mortalidad infantil y la extensión del último grupo, es debido a que son pocas las personas que sobreviven a los 85 años en este ejemplo.

1.1.2. Estructura

En esencia, la tabla de mortalidad se deriva de una cohorte de personas nacidas en un mismo año, la cual muestra la evolución y constante decrecimiento de la cohorte, midiendo la proporción de personas que continúan vivas en cada edad hasta que fallece la última. En la práctica, las tablas de mortalidad se basan en la tasa de mortalidad q_x calculada para un grupo de personas de la misma edad¹.

Su estructura, usualmente, contiene 7 columnas encabezadas por las siguientes funciones biométricas: x , l_x , d_x , q_x , L_x , T_x y e_x .

x , representa la edad del individuo, $0 \leq x \leq \omega$, siendo ω la edad límite.

l_x , representa el número de individuos que sobreviven a la edad x .

d_x , representa el número de muertes entre las edades x y $x + 1$,

$$d_x = l_x - l_{x+1} \tag{1.1}$$

q_x , es la proporción de individuos que fallecen entre la edad x y la edad $x + 1$,

$$q_x = \frac{d_x}{l_x} \tag{1.2}$$

¹Cabe aclarar que a lo largo de este trabajo se llama a q_x : tasa de mortalidad. Ya que, así la llaman (Hinde, 2014) y (Alho, 2005), a quienes se sigue para la descripción de la metodología, siendo en verdad la probabilidad de muerte, como la llama (Keyfitz, 2005).

Estas funciones presentadas anteriormente, se suelen graficar y mantienen una forma que se repite en todas las poblaciones con las variaciones específicas de cada población. A continuación, se grafican las funciones l_x , d_x y q_x de la población uruguaya a partir de la q_x del BCU. En Apéndice A, cuadro A.7, se encuentran las tablas de mortalidad completas de la población.

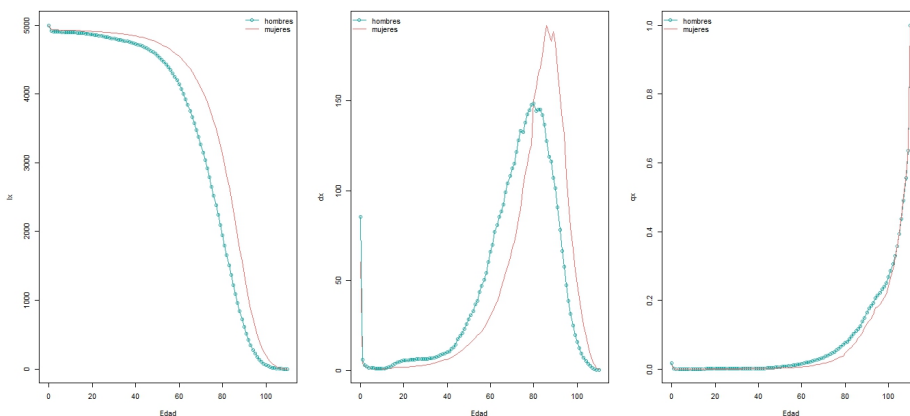


Figura 1.1: Funciones l_x , d_x y q_x de la población uruguaya

En las gráficas, se observa claramente la diferencia entre la mortalidad femenina y la masculina. De la gráfica l_x , se infiere que partiendo de la misma cantidad de hombres que de mujeres, las mujeres sobreviven en mayor cantidad a los hombres en todas las edades. De la gráfica para d_x , se observa una elevada cantidad de muertes para los primeros meses de vida, cayendo rápidamente para las edades siguientes al año y comenzando a aumentar de manera más rápida, a partir de los 40 años. También se observa que las mujeres fallecen a edades más avanzadas que los hombres. Lo anterior hace que la q_x masculina, sea mayor a la femenina para todas las edades, con una marcada diferencia a partir de los 50 años de edad.

Continuando con las funciones de la tabla,

e_x , es la esperanza de vida residual a la edad x , representa los años que le restan por vivir a un individuo que alcanzó la edad x ,

$$e_x = \frac{T_x}{l_x} \quad (1.3)$$

donde T_x , es el total de años que todos los individuos que sobreviven a la edad x esperan vivir,

$$T_x = \sum_{i=1}^{\omega} L_i \quad (1.4)$$

Asumiendo que las muertes se distribuyen uniformemente a través de cada año de vida, se define L_x como el número de años-persona vividos entre la edad exacta x y la edad exacta $x + 1$. Un año-persona es una persona que vive durante un año. Dos personas que viven 6 meses cada una, representan un año-persona.

En otras palabras, el número de años-persona vividos entre la edad exacta x y la edad exacta $x + 1$, es igual al promedio del número de personas vivas a la edad exacta x y el número de personas vivas a la edad exacta $x + 1$.

$$L_x = \frac{1}{2}(l_x + l_{x+1}) \quad (1.5)$$

De otra manera, cada persona que sobreviva a la edad $x + 1$, vivió un año completo entre su cumpleaños x y su cumpleaños $x + 1$. Asumiendo que las muertes se distribuyen de manera uniforme entre las edades exactas x y $x + 1$, cada persona que sobreviva a la edad exacta x pero que muera antes

de su próximo cumpleaños $x + 1$ vive, en promedio, medio año-persona. Así,

$$L_x = l_{x+1} + \frac{1}{2}d_x = l_{x+1} + \frac{1}{2}(l_x - l_{x+1}) = \frac{1}{2}(l_x + l_{x+1}) \quad (1.6)$$

Ocasionalmente, otra cantidad se incluye en la tabla que es la proporción de personas que sobreviven desde su cumpleaños x hasta su cumpleaños $x + 1$ y se simboliza como:

$$p_x = \frac{l_{x+1}}{l_x} = 1 - q_x \quad (1.7)$$

Función de supervivencia

Tal como lo presenta (Débon Aucejo et al., 2008), siendo x la edad de un individuo, con $x \in [0, \omega]$, T representa su tiempo futuro de supervivencia. La función de distribución de probabilidad de T ,

$$G(t) = P(T \leq t), t \geq 0,$$

representa la probabilidad que el individuo tiene de morir dentro de los t años siguientes. A partir de $G(t)$ se define la función de supervivencia:

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

siendo ésta la probabilidad que tiene una persona de sobrevivir t años. De su definición se derivan las siguientes propiedades:

- es una función no creciente
- en los extremos del intervalo de supervivencia, toma los valores $s(0) =$

1, puesto que $G(0) = 0$, y $s(\omega) = 0$, por tratarse de la edad máxima alcanzable.

Para un individuo de edad x , la probabilidad de sobrevivir al menos t años es:

$$s(t) = {}_t p_x \quad (1.8)$$

Tasa de mortalidad m_x

Como plantea (Cox, 2008), para muchos propósitos es útil considerar, no el número de personas vivas a la edad exacta x sino el número promedio de personas vivas a la edad x en su último cumpleaños, esto es L_x y de ella se deriva la tasa ‘central’ de mortalidad:

$$m_x = \frac{d_x}{L_x} \quad (1.9)$$

Como L_x es aproximadamente igual a $l_x - \frac{1}{2}d_x$, se puede escribir como:

$$m_x \doteq \frac{d_x}{l_x - \frac{1}{2}d_x}$$

y así $\frac{1}{m_x} \doteq \frac{1}{q} - \frac{1}{2}$. Entonces,

$$m_x \doteq \frac{2q_x}{2 - q_x} \quad (1.10)$$

o

$$q_x \doteq \frac{2m_x}{2 + m_x} \quad (1.11)$$

Esta relación no se mantiene, en el caso de los primeros años de edad y para edades muy avanzadas.

Tablas de mortalidad abreviadas

Las tablas de mortalidad abreviadas contienen la información, usualmente, en grupos de edad de cinco años. Se indica el tamaño de grupo de edad en años mediante el símbolo n . De este modo,

${}_n d_x$, es el número de muertes ocurridas entre la edad x y la edad $x + n$;

${}_n q_x$, es la proporción de personas que llegaron a su cumpleaños x y murieron antes de su cumpleaños $x + n$;

${}_n L_x$, es el número de años-persona vividos entre las edades exactas x y $x + n$.

Las fórmulas que conectan las cantidades, son exactamente las mismas en las tablas de mortalidad abreviadas y en las completas, con una excepción, que surge de la relación entre l_x y L_x . En la tabla de mortalidad completa, se asume que las muertes se distribuyen uniformemente a través de cada año de vida según la ecuación (1.5). En la tabla de mortalidad abreviada, se asume que las muertes se distribuyen uniformemente dentro de cada grupo de edad.

Dado este supuesto, la ecuación equivalente a la ecuación (1.5) en la tabla de mortalidad abreviada es

$${}_n L_x = \frac{n}{2}(l_x + l_{x+n}) \quad (1.12)$$

y la equivalente a la ecuación (1.4) es

$$T_x = \sum_{i=1}^{\infty} {}_n L_i \quad (1.13)$$

donde i toma los valores x , $x + n$, $x + 2n$ y así sucesivamente.

Hay dos complicaciones con las tablas de mortalidad abreviada. El supuesto de que las muertes se distribuyen uniformemente a través del intervalo, no se cumple para las edades más pequeñas (0-4 años), donde la mayoría de las muertes ocurren antes del primer año. Para contemplar esto, es usual que la tabla de mortalidad abreviada divida éste grupo de edades en dos partes: los menores de 1 año y de 1-4 años.

De este modo, la tabla de mortalidad abreviada queda dividida en los siguientes factores de separación: 0-1, 1-4, 5-9, 10-14, 15-19, 20-24 y así sucesivamente hasta el último grupo de edades el cual queda abierto, siendo por ejemplo, 90 años y más.

Para las edades más avanzadas, el problema es que no se conoce el ancho del intervalo, entonces no es claro a qué edad la persona más vieja sobrevive. Se pueden hacer varios supuestos para calcular ${}_nL_x$ para este último grupo de edades.

Una opción, es asumir que nadie sobrevive a la edad más grande. Si a esta edad la llamamos ω , entonces $l_\omega = 0$. El problema con este supuesto es que es poco probable que las muertes se distribuyan uniformemente sobre el rango de edad, entre la edad más chica del intervalo de edad avanzada y la edad más grande ω .

Otra alternativa, es hacer un supuesto acerca del número promedio de años que le quedan por vivir, a una persona que llega al inicio del grupo de mayor edad. Por ejemplo, si el grupo de edad más avanzada es el de personas de 90 años y más, se hace un supuesto sobre e_{90} y luego se calcula ${}_nL_{90} = l_{90}e_{90}$.

Una tercera alternativa, es usar el hecho de que lógicamente ${}_nq_x$ para el

grupo más grande de edad, es igual a 1 y dado el valor observado de ${}_n m_x$ para el grupo de edad más avanzada, calcular n . Esto asegura que las muertes se distribuyen uniformemente para éste grupo de edad. Siempre que la atención no se dirija específicamente a la mortalidad en las edades más avanzadas, proporciona una aproximación práctica aceptable. Ésta aproximación es equivalente a usar la ecuación:

$${}_n L_x = \frac{{}_n d_x}{{}_n m_x} \quad (1.14)$$

para el último grupo de edad.

Fuerza de mortalidad

La tabla de mortalidad, se describe considerando sucesivos valores de q_x , o sea, la proporción de personas que alcanzaron la edad x y fallecieron antes de cumplir $x + 1$. Esta división de personas vivas dentro de los años de edad, es solamente por conveniencia analítica, implicando que el riesgo de muerte cambie abruptamente a cada edad. Lo que, claramente, no es cierto para muchas personas, la mortalidad realmente está cambiando continuamente con la edad.

Considerando una cohorte de personas l_0 . A la edad exacta x , hay l_x que continúan vivas. A alguna edad cercana mayor, $x + dx$, hay l_{x+dx} aún vivas. El número de muertes entre las edades x y $x + dx$ es $l_x - l_{x+dx}$. La intensidad de la mortalidad, depende de la “velocidad” a que esas muertes ocurren con respecto a la edad. A su vez, la “velocidad” a la que las personas van muriendo, depende del largo del intervalo de edad dx .

La tasa a la que las muertes ocurren por año de edad es $= \frac{l_x - l_{x+dx}}{dx}$. El impacto de la mortalidad representado por las muertes $l_x - l_{x+dx}$, depende del número de personas vivas. Luego, la intensidad de la mortalidad es una combinación de la tasa a la que las personas mueren por año de edad y la proporción de personas que mueren:

$$\text{Intensidad de la mortalidad en el intervalo de edad } x \text{ a } dx = \frac{l_x - l_{x+dx}}{l_x dx}$$

Si se supone el largo del intervalo dx , bien pequeño, se puede pensar en el límite de la intensidad de la mortalidad cuando $dx \rightarrow 0$. En demografía, esto se conoce como μ_x , y se la llama Fuerza de mortalidad a la edad x .

$$\mu_x = \lim_{dx \rightarrow 0} \frac{l_x - l_{x+dx}}{l_x dx} = -\frac{1}{l_x} \lim_{dx \rightarrow 0} \frac{l_{x+dx} - l_x}{dx} = -\frac{1}{l_x} \frac{d}{dx} l_x = -\frac{d}{dx} \ln(l_x)$$

Dentro de un intervalo anual, la fuerza de la mortalidad mide la mortalidad a cada instante. En este sentido, esta tasa puede ser considerada como una clase de promedio de todas las tasas instantáneas del intervalo. De este modo, la tasa central de mortalidad m_x , es un buen estimador de la fuerza de mortalidad en la mitad del intervalo:

$$\mu_{x+0,5} \cong m_x \tag{1.15}$$

Si la relación anterior fuera constante a lo largo del intervalo, el error sería nulo.

La fuerza de mortalidad y la probabilidad de morir

Si se analiza la función de supervivencia, se puede estudiar la relación entre la fuerza de mortalidad y la probabilidad de morir. Integrando $\mu_x = [-dl_x/dx]/l_x$ se obtiene:

$$l_x = l_0 \exp \left[- \int_0^x \mu(u) du \right]$$

Si $l_0 = 1$, l_x es igual a la probabilidad de sobrevivir desde el nacimiento hasta la edad x :

$$p_0 = \exp \left[- \int_0^x \mu(u) du \right]$$

Luego la probabilidad condicional de sobrevivir hasta la edad $x + 1$ habiendo alcanzado la edad x es:

$$p_x = \exp \left[- \int_x^{x+1} \mu(u) du \right] \quad (1.16)$$

De este modo la probabilidad condicional de morir antes de la edad $x + 1$ teniendo la edad x , es:

$$q_x = 1 - \exp \left[- \int_x^{x+1} \mu(u) du \right] \quad (1.17)$$

Como un intervalo de edad comprende una infinidad de edades discretas, la fuerza de la mortalidad en el medio del intervalo discreto de edad $(x, x + 1)$ es considerado, por conveniencia, como igual a la probabilidad de morir en

el modo de cálculo discreto:

$$q_x \cong 1 - \exp[-\mu_{x+0,5}]$$

Esta relación muestra que, cuando la fuerza de mortalidad $\mu_{x+0,5}$ tiende a infinito, q_x tiende a 1.

Usando la relación $\mu_{x+0,5} \cong m_x$ se obtiene:

$$q_x \cong 1 - \exp[-m_x] \tag{1.18}$$

Entonces,

$$m_x \cong -\ln(1 - q_x)$$

Estas relaciones muestran que, si se define el año de edad en una unidad, se pueden usar generalmente la tasa, la probabilidad o la fuerza de mortalidad indistintamente. Pero cualquiera sea el largo del intervalo de edad, estas ecuaciones también se hacen posibles cuando se conoce alguno de los indicadores que determine los otros dos.

Expectativa de vida en términos continuos

Expresando L_x en modo continuo, el número de años-persona vividos entre dos edades subsecuentes, x y $x + n$, y considerando l_0 :

$${}_nL_x = \int_x^{x+n} l(u) du \tag{1.19}$$

Esta aproximación para L_x , requiere asumir que hay distribución uniforme

de la muerte dentro de cada grupo de edad, lo que equivale a asumir que la probabilidad de supervivencia para alguna edad específica se distribuye linealmente entre las edades x y $x + n$.

El número de años-persona vividos para edades mayores a x , es dado por la ecuación:

$$T_x = {}_{\omega-x}L_x = \int_x^{\omega} l(u) du$$

Así la expresión para la expectativa de vida a la edad x , es dada por:

$$e_x = \frac{\int_x^{x+n} l(u) du}{l_x} \quad (1.20)$$

Otro cálculo de la expectativa de vida

Una manera diferente de calcular la expectativa de vida es presentada por (Alho, 2005) donde define $I(t)$ como indicador de procesos de tiempo de espera X , siendo $I(t) = 1$ si $X > t$ y $I(t) = 0$ de otro modo. De este modo:

$$X = \int_0^{\infty} I(t) dt$$

Nótese que la probabilidad de que $X > t$ es igual a $p(t) = E[I(t)]$, entonces:

$$E(X) = \int_0^{\infty} p(t) dt$$

Y siendo $e_x = E[X - x \mid X > x]$ expectativa de vida, dada la supervivencia hasta la edad x , y $p(x+t) = p(x+t)/p(x)$ la probabilidad condicional de

sobrevivir hasta la edad $x + t$ dado que se sobrevivió hasta la edad x , luego:

$$e_x = \int_0^{\infty} p(x+t)/p(x) dt \quad (1.21)$$

La aproximación más común asume la linealidad de $p(t)$ en el intervalo $[x, x+1]$, esto es equivalente al así llamado método trapezoidal de integración numérica, lo que conduce a la fórmula aproximada:

$$e_x \approx \frac{1}{2} + \sum_{t=1}^{\infty} p(x+t)/p(x) \quad (1.22)$$

1.1.3. Construcción de la Tabla de Mortalidad

Para construir la tabla de mortalidad, lo primero que se hace es tomar el criterio para calcular la tasa central de mortalidad observada m_x y a partir de ella, calcular la tasa de mortalidad q_x . Luego, partiendo de una cohorte hipotética, se calculan el resto de las funciones biométricas de la tabla.

Estimación de m_x

Existen diversas formas de calcular la tasa central de mortalidad observada. Convencionalmente, se calcula mediante la siguiente fórmula:

$${}_5m_x = \frac{{}_5D_x}{{}_5N_x} = \frac{\# \text{ de defunciones entre las edades } x \text{ y } x+5}{\# \text{ de expuestos al riesgo de morir entre las edades } x \text{ y } x+5}$$

Métodos de estimación de q_x

Existen diversos procedimientos para estimar el cociente de mortalidad a partir de la tasa observada. Algunos de ellos son: método lineal, método

exponencial, método de Reed-Merrell, método de Greville y método de Keyfitz. Todos ellos arrojan resultados muy parecidos, por lo que usualmente se utiliza aquel cuya expresión es más sencilla (método lineal).

Método lineal Partiendo de que la tasa central de mortalidad se puede definir como la relación del número de eventos observados y el número de años-persona y suponiendo que las muertes se distribuyen uniformemente a lo largo del intervalo de edad, se llega a la relación que fue explicada anteriormente en (1.1.2):

$${}_nq_x = \frac{2n_n m_x}{2 + n_n m_x} \quad (1.23)$$

Método exponencial Asumiendo que l_x se comporta de modo exponencial se llega a lo visto anteriormente en (1.1.2):

$${}_nq_x = 1 - e^{-n_n m_x} \quad (1.24)$$

Método de Reed-Merrell En el método de Reed y Merrell (1939)² las tasas de mortalidad se leen de un conjunto de tablas de conversión estándar, que muestran las tasas de mortalidad asociadas con varias tasas centrales de mortalidad observadas. Las tablas estándar para ${}_3m_x$, ${}_5m_x$ y ${}_{10}m_x$, fueron preparadas asumiendo la siguiente función exponencial:

$${}_nq_x = 1 - e^{-n_n m_x - a n^3 {}_n m_x^2} \quad (1.25)$$

²para éste método y el siguiente (método de Greville) se sigue a (Kintner, 2004)

donde n es el largo del intervalo, ${}_n m_x$ es la tasa central de mortalidad y a es una constante. Reed y Merrell encontraron que el valor de $a = 0,008$ producía resultados aceptables. La conversión de ${}_n m_x$ a ${}_n q_x$ mediante el uso de tablas de Reed-Merrell es aplicado usualmente a datos agrupados en 5 o 10 años.

Método de Greville El método sugerido por Greville (1943) para convertir las tasas centrales de mortalidad en probabilidades de muerte es:

$${}_n q_x = \frac{{}_n m_x}{\frac{1}{n} + {}_n m_x \left[\frac{1}{2} + \frac{n}{12} ({}_n m_x - \log C) \right]} \quad (1.26)$$

donde C viene de asumir que los valores de ${}_n m_x$, siguen una curva exponencial. Se asume el valor de $\log C = 0,095$.

Método de Keyfitz Siguiendo a (Keyfitz, 2005), proponen agregar un término C de corrección a la tasa de mortalidad observada y utilizando el método exponencial para calcular la tasa de muerte. Siendo:

$$C = \frac{({}_n N_{x-n} - {}_n N_{x+n})({}_n m_{x+n} - {}_n m_{x-n})}{48 {}_n N_x}$$

donde ${}_n N_x$, es la población observada en el intervalo de edad $x + n$.

De este modo:

$${}_n q_x = 1 - e^{-n({}_n m_x + C)} \quad (1.27)$$

Finalizado el proceso de construcción de la Tabla de Mortalidad se pasa

a desarrollar, en la siguiente sección, la metodología a utilizar para describir y proyectar la mortalidad en el tiempo.

1.2. Descripción de la mortalidad en el tiempo

Siguiendo a (Giroso and King, 2007), se describe el modelo propuesto por Lee y Carter. Dicho modelo, fue realizado para describir la mortalidad de los Estados Unidos y proyectarla. Actualmente es aplicado para poblaciones de diferentes países como ser Argentina³, México⁴, Costa Rica⁵ y España⁶ entre otros.

1.2.1. Modelo de Lee-Carter

(Lee and Carter, 1992), desarrollan el siguiente modelo para describir la mortalidad a través del tiempo:

$$m_{x,t} = e^{a_x + b_x k_t + \varepsilon_{x,t}} \quad (1.28)$$

Donde $m_{x,t}$ es la tasa central de muerte para la edad x en el año t . Debido a que intervienen la variable edad y la variable tiempo, se denomina a este modelo como bivariable.

a_x , refiere a la estructura de la mortalidad y representa cómo se comporta la mortalidad a través de las edades.

³(Belliard and Williams, 2013), (Andreozzi and Blaconá, 2011)

⁴(García Guerrero and Ordorica Mellado, 2012)

⁵(Aguilar Fernandez, 2013)

⁶(Débon Aucejo et al., 2008)

b_x , representa el patrón de cambio en la mortalidad o la “velocidad” con que ésta varía a cada edad o grupo de edades, cuando varía el nivel general (tendencia) de mortalidad.

k_t , explica la tendencia de la mortalidad en el tiempo.

$\varepsilon_{x,t}$, es un término de error que depende del tiempo y la edad, con media 0 y varianza σ_ε^2 , que refleja los efectos históricos no capturados en el modelo en cada edad o grupo de edades.

Este modelo permite proyectar hacia el futuro la mortalidad estimada a partir de los datos históricos. Los parámetros a_x y b_x , capturan la información histórica de la mortalidad por la edad de la persona y el parámetro k_t , la evolución de la mortalidad histórica en el transcurso del tiempo. El modelo combina un enfoque paramétrico con una utilización del método estadístico de series temporales.

Para facilitar la estimación de los parámetros, se linealiza la expresión anterior:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \quad (1.29)$$

El término $b_x k_t$, implica infinitos valores arbitrarios para ambos parámetros lo que hace que el modelo sea indeterminado. Suponiendo que los vectores a, b, k son una solución, luego para cualquier constante c , también $a - bc$, $b, k + c$ son una solución. También es claro que si a, b, k son una solución, luego $a, bc, k/c$ también son una solución. Esto no es un obstáculo, simplemente significa que la verosimilitud asociada con el modelo tiene infinitos máximos equivalentes, cada uno de los cuales produce iguales estimaciones. En la práctica, y siguiendo a Lee y Carter, se escoge una arbitraria pero

suficiente parametrización consistente para su identificación. Esto se puede hacer mediante la imposición de dos restricciones: $\sum_x b_x = 1$ y $\sum_t k_t = 0$.

1.2.2. Estimación de los parámetros del modelo

El modelo original de Lee y Carter supone la utilización de Descomposición de Valores Singulares (SVD) para la estimación de los parámetros del modelo. El mismo presenta la desventaja de asumir homocedasticidad en los errores (es decir que poseen la misma varianza a través de todas las edades), supuesto que no se cumple con datos de tablas de vida, en los cuales, en las edades avanzadas se tienen pocos datos y por tal motivo, la varianza de los estimadores es grande. Éste problema fue resuelto por (Wilmoth, 1993), quien propone la versión Log-Binomial-Poisson del modelo de Lee-Carter, basado en suponer una distribución Poisson para la variable aleatoria número de defunciones, lo que permite tener en cuenta la presencia de heterocedasticidad. Wilmoth estima los parámetros mediante el método de Máxima Verosimilitud (MV).

A continuación, se presentan ambos métodos para estimar los parámetros de modelo.

Estimación de los parámetros mediante Descomposición de Valores Singulares

El modelo no puede ajustarse a través de un método de regresión usual, ya que no existe una variable regresora observable, en el lado derecho de la ecuación (1.29), se tienen sólo los parámetros a ser estimados y el índice

desconocido k_t . Lo que proponen Lee y Carter es, una vez obtenida la estimación de a_x , utilizar el método SVD para encontrar la solución de los restantes parámetros a partir de la minimización de la matriz:

$$Z = \sum_{x,t} [\ln(m_{x,t}) - \hat{a}_x - b_x k_t]^2 \quad (1.30)$$

Desarrollando:

$$\frac{\partial Z}{\partial a_x} = \frac{\partial Z}{\partial b_x} = \frac{\partial Z}{\partial k_t} = 0$$

$$\frac{\partial Z}{\partial a_x} = 0 \Rightarrow \sum_t \ln(m_{x,t}) = \sum_t a_x + b_x k_t$$

y teniendo en cuenta la restricción $\sum_t k_t = 0$, se obtiene $\hat{a}_x = \overline{\ln m_x}$, esta restricción implica que los valores estimados de a_x son el promedio de los logaritmos de las tasas observadas.

Luego a la matriz:

$$Z_{x,t} = \ln m_{x,t} - \hat{a}_x \quad (1.31)$$

se le aplica el método SVD:

$$PdQ' = SVD(Z_{x,t}) = \sum d_j * P_{x,j} * Q_{t,j}$$

Donde la matriz P representa la componente edad, d representa los valores singulares y Q representa la componente tiempo.

Del sistema anterior se obtienen los valores de \hat{b}_x y \hat{k}_t . Siendo $\hat{b}_x = P_{x,1}$ y $\hat{k}_t = d_t * Q_{t,1}$. De este modo $\hat{Z}_{x,t} = \hat{b}_x \hat{k}_t$ y por consiguiente el logaritmo de la

tasa central de muerte estimada es:

$$\ln(\hat{m}_{x,t}) = \hat{a}_x + \hat{Z}_{x,t} \quad (1.32)$$

Las tasas de muerte derivadas de éste procedimiento, generalmente, no producen el número de muertes reales cuando se aplica a la distribución por edades de la población dada. Además k , es estimada más bien para minimizar los errores en los logaritmos de las tasas de muerte y no de las tasas de muerte en sí mismas. Usando la ecuación (1.29), se puede reestimar k_t en un segundo paso, tomando los valores estimados de a_x y b_x del primer paso. De este modo se encuentra una nueva estimación para k tal que, para cada año, dada la distribución por edad de la población real, el número implícito de muertes sea igual al número de muertes reales.

Estimación de los parámetros mediante Máxima Verosimilitud

Una forma alternativa de ajustar el modelo de Lee-Carter, fue propuesto por (Wilmoth, 1993), quien propone especificar un modelo probabilístico cuyos parámetros son estimados mediante el método de MV.

Siendo $D_{x,t}$, la variable aleatoria que representa las muertes a la edad x en el período de tiempo t y siendo $d_{x,t}$, el correspondiente número de muertes observado, $D_{x,t}$ puede ser satisfactoriamente aproximado mediante una distribución Poisson con media $\lambda_{x,t}$, donde $\lambda_{x,t} = m_{x,t}E_{x,t}$ y $E_{x,t}$ representa los expuestos al riesgo de morir a la edad x en el tiempo t .

Abandonando los subíndices temporalmente, la función de verosimilitud

para una única combinación de edad-tiempo se puede escribir:

$$L(d; \lambda) = \frac{\lambda^d e^{-\lambda}}{d!} \quad (1.33)$$

Del mismo modo, la función log-verosimilitud en este caso es:

$$l(d; \lambda) = d \ln(\lambda) - \lambda - \ln(d!) \quad (1.34)$$

Asumiendo la independencia de las observaciones, se suma a través de de las distintas edades y tiempos y se obtiene la log-verosimilitud total:

$$l = \sum_{x,t} (d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t} - \ln(d_{x,t}!)) \quad (1.35)$$

Las estimaciones máximo verosímiles son los valores de $\lambda_{x,t}$ que maximizan la ecuación (1.35). Dado que el tercer término de la ecuación no depende de $\lambda_{x,t}$, resulta suficiente maximizar la ecuación:

$$l = \sum_{x,t} (d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t}) \quad (1.36)$$

Si no hay restricciones sobre $\lambda_{x,t}$, se verifica que la ecuación (1.36) alcanza su máximo valor cuando $\lambda_{x,t} = d_{x,t}$ ⁷. Por otro lado, el modelo de Lee-Carter requiere que $\lambda_{x,t}$ satisfaga la ecuación:

$$\lambda_{x,t} = m_{x,t} E_{x,t} = e^{a_x + b_x k_t} E_{x,t} \quad (1.37)$$

⁷Nótese que $\hat{m}_{x,t} = d_{x,t}/E_{x,t}$ es la estimación máximo verosímil de $m_{x,t}$

Así, las estimaciones máximo verosímiles de los parámetros del modelo de Lee-Carter, se encuentran sustituyendo $\lambda_{x,t}$ por $e^{a_x+b_x k_t} E_{x,t}$ en la ecuación (1.36) y maximizándola con respecto a a_x , b_x y k_t .

1.2.3. Diagnóstico y bondad de ajuste del modelo

En esta sección se presentan los tests utilizados para corroborar si los residuos del modelo estimado cumplen las hipótesis básicas realizadas sobre los mismos y que tan bien se ajusta el modelo estimado a los datos de la realidad.⁸

El modelo original de Lee-Carter supone que los residuos de modelo tienen:

- media 0
- varianza constante
- incorrelación
- distribución Normal

Para evaluar los primeros dos supuestos se grafican los residuos y se observa si los mismos tienen media 0 y varianza constante.

Contraste de Ljung-Box

El supuesto de incorrelación de los residuos se contrasta con este test, el cual consiste en un contraste global de que las primeras h autocorrelaciones

⁸Para la descripción de estos métodos se siguen las notas de clase de las materias: Series Cronológicas I y Modelos Lineales.

son cero. Si los residuos siguen un proceso tipo ruido blanco, los coeficientes de correlación estimado, ρ_k^2 , son asintóticamente normales, con media cero y varianza $(T - k)/T(T + 2)$.

Siendo: $H_0 : \rho_1 = \dots = \rho_k = 0$,

el estadístico:

$$Q(h) = T(T + 2) \sum_{j=1}^h \frac{\rho_j^2}{T - j} \sim \chi^2$$

donde $\rho_k = \frac{\sum_{j=k+1}^T \epsilon_j \epsilon_{j-k}}{\sum_{j=1}^T \epsilon_j^2}$, se distribuye, asintóticamente, como una χ^2 con grados de libertad igual al número de coeficientes en la suma (h) menos el número de parámetros estimados n . Se concluye que el modelo es inadecuado, si el valor de $Q(h)$ es mayor que el percentil 0.95 de la distribución χ_{h-n}^2 .

Contraste Jarque-Bera

El estadístico planteado por Jarque Bera, mide la diferencia de los coeficientes de asimetría (cs) y kurtosis (ck) de la serie con la de una serie con distribución normal. El test plantea las siguientes hipótesis para la prueba conjunta:

$$H_0 : cs = ck = 0$$

$$H_1 : cs \neq 0 \quad ck \neq 0$$

Se calculan los coeficientes de asimetría y kurtosis de los residuos y bajo la hipótesis de normalidad:

$$X = \frac{T(cs)^2}{6} + \frac{T(ck - 3)^2}{24} \sim \chi_2^2$$

El coeficiente de asimetría de una variable con distribución Normal, es 0 y el coeficiente de kurtosis de una variable con distribución Normal, es 3. Si el coeficiente excede el valor de 3, la distribución es más empinada que la Normal (leptocúrtica) y si el coeficiente es menor que 3, la distribución es más chata que la normal (platicúrtica).

Se rechaza H_0 , si $X > \chi_{\alpha,2}^2$, siendo α el nivel de significación.

Coefficiente de determinación R^2

El modelo de Lee-Carter parte del supuesto que existe una relación lineal entre el logaritmo de las tasas centrales de mortalidad $m_{x,t}$ y los factores que la determinan: la edad x y el tiempo t .

La recta de regresión, tiene carácter de línea media (promedio), tratando por lo tanto de resumir la información suministrada por los datos. Para saber que tan buena es la recta, se debe tener una medida de dispersión que tenga en cuenta la dispersión de cada observación, con respecto a la recta. Es decir, se debe evaluar la distancia vertical a la recta, es decir, los errores residuales del modelo.

Si las dispersiones son pequeñas, la recta será un buen representante de la dispersión de los datos, o sea, la bondad de ajuste del modelo será alta. Si la dispersión es grande, la bondad de ajuste será baja.

Una forma de medir dicha bondad, mediante el coeficiente de determinación, expresado como:

$$R^2 = 1 - \frac{SCR}{SCT} = 1 - \frac{\sum_t \sum_x \epsilon_{x,t}^2}{\sum_t \sum_x [\ln(m_{x,t}) - a_x]^2}$$

donde SCT, es la suma de las desviaciones cuadráticas de cada valor con respecto a la media y se la denomina Suma de Cuadrados Total. SCR, es la Suma de Cuadrados de los Residuos y mide la dispersión no explicada por el modelo.

Ésta proporción mide la variabilidad total explicada por el modelo de regresión planteado. Se espera que esta proporción sea alta y cercana al 100 % y solo una pequeña parte sea debido al error.

Devianza

Para el método de estimación propuesto por (Wilmoth, 1993), donde $D_{x,t}$, son las defunciones totales de individuos a la edad x , ocurridas durante el año t y $D'_{x,t}$, las defunciones anuales estimadas a cada edad, otra medida para determinar el ajuste del modelo es la Devianza:

$$devianza_t = 2 \sum_x D_{x,t} \ln \left(\frac{D_{x,t}}{D'_{x,t}} \right) - (D_{x,t} - D'_{x,t})$$

El criterio es que cuanto menor sea la devianza, mejor será el ajuste del modelo a los datos de las defunciones.

1.2.4. Proyección

Para producir proyecciones de la mortalidad, o sea, estimar la mortalidad futura, Lee y Carter asumen constantes los errores de b_x a lo largo del tiempo y usan las proyecciones de k_t para un modelo estándar univariado de series

de tiempo. Luego de testear muchas especificaciones de modelos ARIMA⁹, encontraron que una caminata aleatoria con deriva¹⁰, era el modelo más apropiado para sus datos. Especificaron que otros modelos ARIMA podrían ser preferibles para otros conjuntos de datos, pero en la práctica, el modelo de caminata aleatoria con deriva para k_t se ha utilizado casi exclusivamente. El modelo es el siguiente:

$$\hat{k}_t = \hat{k}_{t-1} + \theta + \xi_t \quad \text{con} \quad \xi_t \sim N(0, \sigma_{RW}^2) \quad (1.38)$$

donde θ es el parámetro de desvío y su valor estimado por máxima verosimilitud es: $\hat{\theta} = (\hat{k}_t - \hat{k}_1)/(T - 1)$, que solo depende de la primer y última estimación de k y siendo T , el último año de la serie. Luego para proyectar dos períodos hacia adelante, se coloca la estimación del parámetro de deriva $\hat{\theta}$ y se sustituye por la definición de \hat{k}_{t-1} corrido en el tiempo un período:

$$\hat{k}_t = \hat{k}_{t-1} + \hat{\theta} + \xi_t \quad (1.39)$$

$$= (\hat{k}_{t-2} + \hat{\theta} + \xi_{t-1}) + \theta + \xi_t \quad (1.40)$$

$$= \hat{k}_{t-2} + 2\hat{\theta} + (\xi_{t-1} + \xi_t) \quad (1.41)$$

Para proyectar \hat{k}_t hasta el tiempo $T + (\Delta t)$ con los datos disponibles hasta el período T , se sigue el mismo procedimiento iterativamente (Δt) veces y se

⁹se trata de modelos de series temporales, es decir, las estimaciones futuras vienen explicadas por los datos del pasado. Se suelen expresar como ARIMA(p,d,q) donde los parámetros p, d y q son números enteros no negativos que indican el orden de las distintas componentes del modelo (respectivamente, las componentes autorregresiva, integrada y de media móvil).

¹⁰El modelo de caminata aleatoria con deriva (Random Walk with drift) es un modelo ARIMA(0,1,0)

obtiene:

$$\hat{k}_{T+(\Delta t)} = \hat{k}_T + (\Delta t)\hat{\theta} + \sum_{i=1}^{\Delta t} \xi_{T+i-1} \quad (1.42)$$

$$= \hat{k}_T + (\Delta t)\hat{\theta} + \sqrt{(\Delta t)}\xi_t \quad (1.43)$$

donde la segunda línea es una simplificación posible, por el hecho que las variables aleatorias ξ_t son asumidas en el modelo como independientes con la misma varianza. Esta segunda línea de la ecuación, está indicando que los errores estándar condicionados para las proyecciones aumentan con la raíz cuadrada de la distancia al horizonte proyectado (Δt).

A partir de este modelo se obtienen las estimaciones puntuales proyectadas que siguen una línea recta como función de (Δt), con pendiente $\hat{\theta}$:

$$E(\hat{k}_{T+(\Delta t)} | \hat{k}_1, \dots, \hat{k}_T) \equiv \mu_{T+(\Delta t)} = \hat{k}_T + (\Delta t)\hat{\theta} \quad (1.44)$$

El modelo de Lee-Carter para los \hat{k} es muy simple: extrapola a partir de una línea recta que pasa por el primer punto \hat{k}_1 por el último punto \hat{k}_T . Todos los demás \hat{k} son ignorados.

Ahora se utiliza ésta última expresión para obtener la proyección de la estimación puntual de los logaritmos de la mortalidad:

$$\mu_{T+(\Delta t)} = \overline{\ln(m)} + \hat{b}\hat{k}_{T+(\Delta t)} \quad (1.45)$$

$$= \overline{\ln(m)} + \hat{b} \left[\hat{k}_T + (\Delta t)\hat{\theta} \right] \quad (1.46)$$

Capítulo 2

Datos de Aplicación

Los datos utilizados fueron brindados por el BROU, los mismos se conforman de toda la plantilla de funcionarios activos y que egresaron del banco al 31/07/2014. También se brindó información de los funcionarios fallecidos (datos que aportó la Caja de Jubilaciones y Pensiones Bancarias al BROU).

No son tomados en cuenta para este estudio los funcionarios cuya categoría de ingreso fuera: “Becario”, “Contrato a término”, “Designación directa por PE” (es el Director del Banco) y “En comisión en BROU”. El criterio se toma en base a que estos funcionarios, ya sea que se encuentren trabajando o que sean egresos del banco, no alcanzan una antigüedad que les permita jubilarse por la Caja de Jubilaciones y Pensiones Bancarias (CJPB) y por tal motivo el banco no deberá afrontar ningún gasto por concepto de beneficios por los mismos.

Los datos de los egresos del banco se componen de los fallecimientos, los funcionarios jubilados y de otros egresos (como ser renuncia, destitución, no confirmación, etc.). Siguiendo el criterio anterior, los funcionarios que entran

en el estudio son los que ya pertenecen a la CJPB y que el banco debe seguir pagando por sus beneficios adquiridos hasta su fallecimiento y los funcionarios ya fallecidos. Quedan fuera del análisis los que egresaron del banco por otras causas y que no alcanzaron la antigüedad de 30 años para poder jubilarse por la CJPB.

2.1. Análisis Descriptivo

2.1.1. Activos

Los funcionarios activos que se analizan, comprenden un total de 4167 y se encuentran entre los 20 y los 65 años de edad. Los mismos ingresaron a trabajar al banco entre los años 1969 y el 2014. De ellos 2413 son hombres y 1754 son mujeres.

A continuación se presentan: histograma, distribución sexo y pirámide poblacional. De los gráficos, se observa la brecha generacional entre los fun-

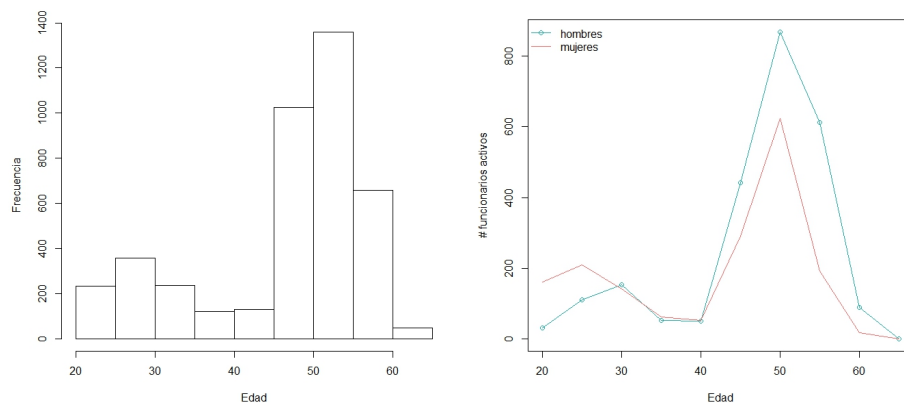


Figura 2.1: Distribuciones por edad y sexo de los funcionarios activos

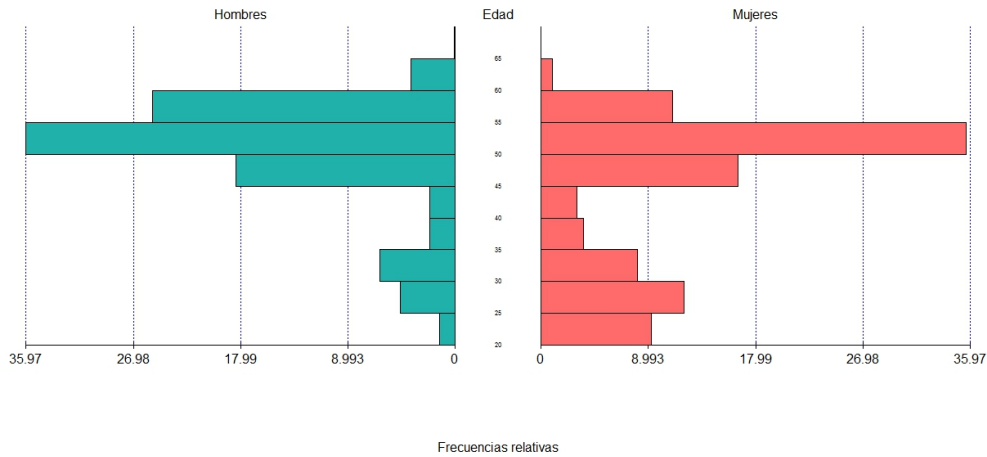


Figura 2.2: Pirámide poblacional de funcionarios activos

cionarios más jóvenes y los más viejos dada la escasez de funcionarios entre los 35 y 45 años de edad. Esto es debido a la falta de ingresos de nuevos funcionarios al banco por concurso en los años noventa y hasta el 2007. También se observa que la mayor concentración de la población tiene entre 45 y 55 años, y por lo tanto, se encuentra muy cercana a la jubilación.

Del gráfico de distribución por sexo, se ve una mayor cantidad de mujeres que de hombres en las edades más jóvenes, esto es entre los 20 y los 30 años, siendo muy pareja la cantidad entre los 30 y los 40 años de edad, y una mayor cantidad de hombres que de mujeres para las edades siguientes.

De la pirámide poblacional se observa, dentro del total de cada sexo, una mayor proporción de mujeres que de hombres en las edades más jóvenes y hasta los 45 años, siendo muy pareja la proporción para el tramo de edades de los 50 a los 55 años y habiendo una mayor proporción de hombres que de mujeres para el resto de los tramos etarios.

La edad de jubilación se alcanza a los 60 años, pudiendo el funcionario quedarse en la institución hasta los 65 años. Es por eso que se observan muy pocas personas en este último grupo de edades.

2.1.2. Jubilados

Se cuenta con 1959 datos de jubilados conformado por 1803 hombres y 156 mujeres. La poca cantidad de mujeres jubiladas es debido a que antes los funcionarios bancarios eran únicamente hombres. Recién llegando al año 1960 comienzan a ingresar las primeras funcionarias mujeres, que en proporción a los hombres eran muy pocas. De la distribución por sexo y de la pirámide

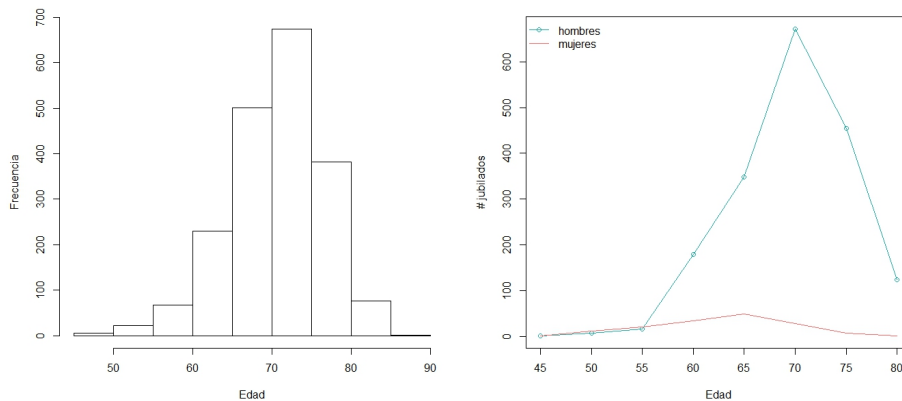


Figura 2.3: Distribuciones por edad y sexo de los funcionarios jubilados

poblacional, se observa un corrimiento en el máximo de la distribución de hombres y mujeres, siendo que para los hombres, el máximo se alcanza entre los 70 y los 75 años y en las mujeres, entre los 65 y los 70 años.

El grueso de los jubilados hombres, se concentra a partir de los 60 años de edad ya que es la edad de jubilación establecida. La reglamentación para la

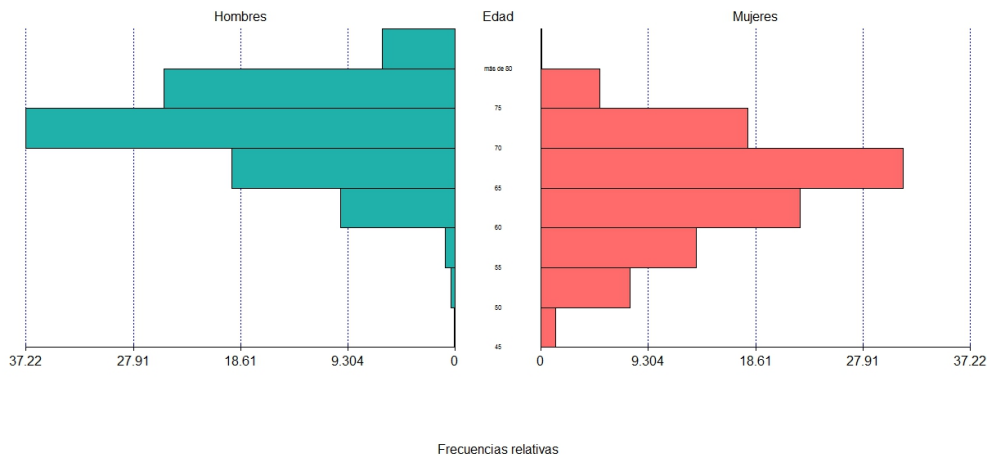


Figura 2.4: Pirámide poblacional de jubilados

jubilación de las mujeres, es la misma que para los hombres en este momento, pero hasta hace unos 8 años aproximadamente, la edad de jubilación de las mujeres era de 55 años de edad y por eso es que se observa una jubilación más temprana de las mujeres, aunque llama la atención el porcentaje de mujeres jubiladas antes de los 55 años de edad.

2.1.3. Expuestos

Para el estudio de la mortalidad de los funcionarios del banco, se analizan juntos los datos de los funcionarios activos y de los funcionarios jubilados, ya que en su conjunto forman los expuestos al riesgo de morir.

El total de expuestos son 6126, conformado de 4216 hombres y 1910 mujeres. La mayor cantidad de expuestos, se concentra entre los 45 y los 55 años de edad aproximadamente, según muestran los gráficos que se presentan a continuación. También se observa que las mujeres expuestas al riesgo de

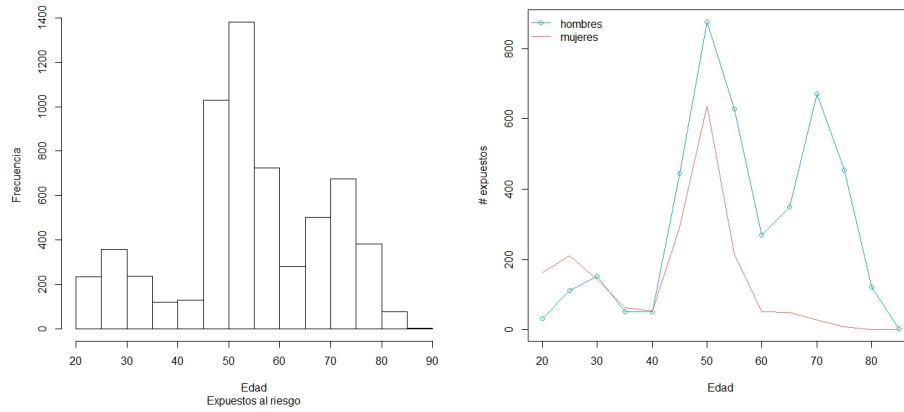


Figura 2.5: Distribuciones por edad y sexo de la población expuesta al riesgo de morir

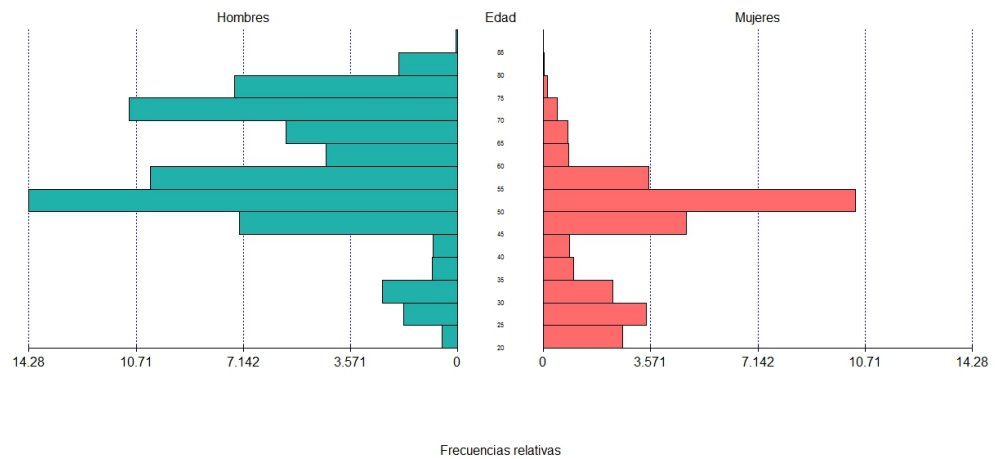


Figura 2.6: Pirámide poblacional de expuestos al riesgo de morir

muerte, son muy pocas a partir de los 60 años de edad, cosa que ya se había observado en la población de mujeres jubiladas, ésto como ya se explicó, es porque antes los trabajadores bancarios eran solo hombres. La poca cantidad de hombres entre los 60 y los 70 años de edad, seguramente se explique por falta de ingresos al banco hace unos 40 años.

2.1.4. Fallecidos

Se cuenta con los datos de 646 fallecimientos, recabados desde el año 1995 hasta el 31/07/2014 y algunos datos para los años 1992 al 1994. De ellos, 599 son fallecimientos de hombres y sólo 47 datos son fallecimientos de mujeres.

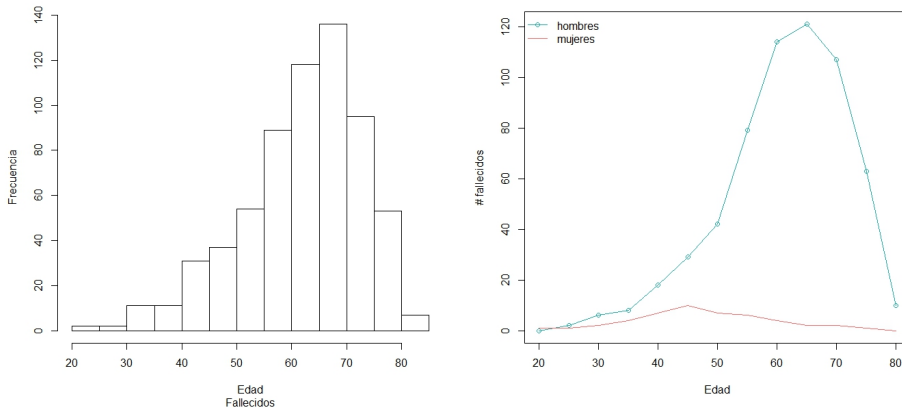


Figura 2.7: Distribuciones por edad y sexo de los funcionarios fallecidos

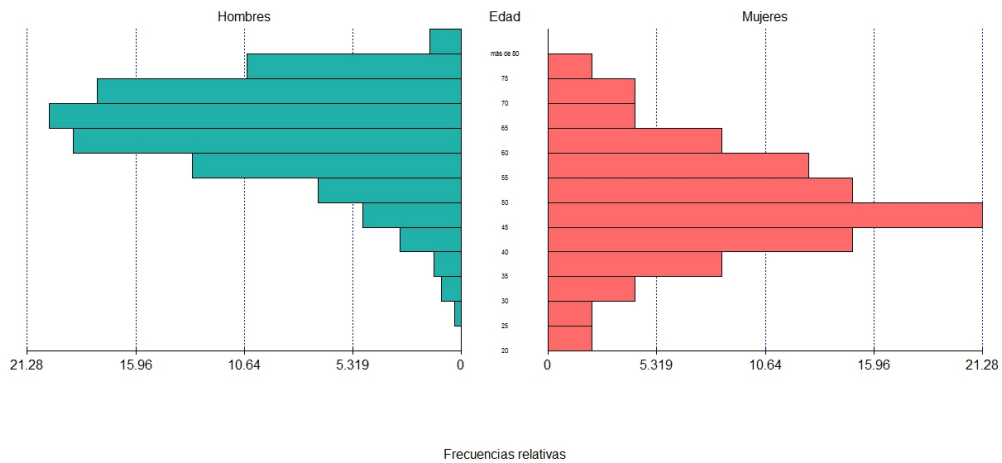


Figura 2.8: Pirámide poblacional de fallecidos

De las gráficas para las muertes femeninas, no se pueden sacar conclusiones,

ya que son muy pocos los datos con los que se cuenta. Si bien se observa que el mayor porcentaje de muertes femeninas se da entre los 45 y los 55 años de edad, esto es debido a que la mayor cantidad de mujeres expuestas rondan esa edad (ver Figuras 2.1 y 2.2).

La distribución de los fallecimientos masculinos que se observan en la pirámide, es igual a como se distribuye el total de la población fallecida antes descrita, ya que las muertes femeninas, al ser tan pocas, no tiene peso en la misma. Del resto de las gráficas se observa una tendencia de crecimiento exponencial en la distribución a medida que aumentan los años hasta los 70, luego decrece, ya que son menos personas las que alcanzan edades mayores, sobre todo edades mayores a 80 años.

2.2. Datos de aplicación

Dado que la muestra de mujeres con la que se cuenta es muy pequeña para poder analizarla, es que se decide dejarlas fuera del presente trabajo y seguir adelante con la población masculina.

A modo de resumen, se cuenta con 4216 hombres expuestos a riesgo de morir, comprendidos entre los 20 y los 86 años de edad y con 599 fallecimientos registrados entre los 24 y los 84 años de edad, que se distribuyen de la siguiente manera:

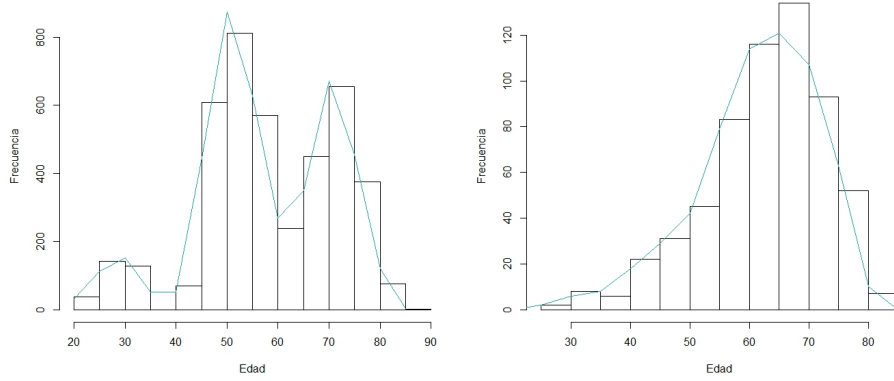


Figura 2.9: Distribuciones por edad de los hombres expuestos al riesgo de morir y de los hombres fallecidos

Capítulo 3

Resultados

Los resultados obtenidos, así como el análisis de los datos, son realizados mediante la utilización del software estadístico R-project (R Core Team, 2015). Para la creación de la tabla de mortalidad se utiliza la librería *lubridate* (Grolemund and Wickham, 2011), para poder operar con las fechas de nacimiento y fallecimiento de la población objeto de estudio. Para la estimación de los parámetros del modelo de Lee-Carter se utilizan dos técnicas, SVD y MV, cuya implementación en R es realizada mediante las librerías *demography* (with contributions from Heather Booth et al., 2014) e *ilc* (Butt et al., 2014) respectivamente. Luego, para corroborar que se cumplan los supuestos de los modelos de Lee-Carter, se utiliza la librería *tseries* (Trapletti and Hornik, 2015) y *car* (Fox and Weisberg, 2011). Y para proyectar la mortalidad, se utiliza la librería *forecasting* (Hyndman and Khandakar, 2008).

3.1. Tabla de mortalidad para los funcionarios del BROU

Como ya se aclaró en el capítulo anterior, el análisis se hace únicamente para la población masculina del BROU. Lo primero que se realiza, es el conteo de la cantidad de expuestos y fallecidos en las distintas edades y en los diferentes años, desde 1995 al 2013. Se dejan fuera del estudio los años 1992 al 1994, ya que los fallecimientos no fueron contabilizados durante todos los meses del año, lo mismo para el año 2014 que se deja fuera del análisis, ya que sólo se cuenta con los datos hasta el 31/07/2014 y para el cálculo de las tasas de mortalidad, se precisan los datos de los años completos.

Dado que la cantidad de fallecidos es de 575, al distribuirlos en forma anual y por edades, los mismos quedan muy dispersos y la matriz queda con muchos ceros (ver Apéndice A, cuadro A.2). Por tal motivo, se agrupan los datos por edades de a 5 años, comenzado con la edad más joven a los 20 años. De este modo, los datos se agrupan de 20 a 24 años, de 25 a 29 años y así sucesivamente hasta el último tramo de edades, donde se agrupan todos los mayores de 80 años. No se crean los grupo de edades: "80 a 84 años" y "85 años y más", dado que sólo hay 3 expuestos con la edad de 85 años y 1 con la edad de 86, ésto en lo referente al año 2013. Si se observa la tabla por edades simples, en el Anéxo A, cuadro A.1, se puede observar que, para los primeros años no se cuenta con datos de hombres expuestos al riesgo de morir con más de 70 años de edad, lo cual no es lógico y hace pensar que los datos de los fallecidos brindados por la CJPB no están completos.

A continuación, se presentan las matrices que contienen la cantidad de

3.1. Tabla de mortalidad BROU

expuestos al riesgo de morir y de fallecidos por año para los distintos tramos de edades.

Edades	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
20-24	31	34	34	29	18	6	0	0	0	0	0	0	86	78	72	44	37	13	30
25-29	201	121	57	31	22	31	36	36	30	18	7	1	50	88	129	145	160	141	120
30-34	850	780	681	519	352	203	124	59	33	25	35	39	40	41	32	31	73	112	139
35-39	562	611	662	749	833	858	790	688	526	354	210	136	71	48	42	46	46	49	53
40-44	333	386	454	503	532	564	620	674	762	848	881	809	701	544	373	226	145	78	53
45-49	249	234	224	245	285	339	398	466	513	545	579	664	720	835	901	912	830	717	556
50-54	654	584	512	431	323	254	240	228	252	292	351	420	500	606	642	674	724	775	867
55-59	559	606	668	673	683	676	604	528	446	332	261	244	232	263	323	390	461	543	620
60-64	191	286	333	422	510	579	626	689	700	725	717	635	560	463	352	281	258	249	278
65-69	16	27	59	102	156	206	308	360	449	537	609	655	722	745	773	767	686	603	501
70-74	0	0	0	1	3	17	31	69	116	176	234	355	410	493	579	648	692	762	805
75-79	0	0	0	0	0	0	0	0	1	4	21	38	85	140	209	280	414	467	545
80 y más	0	0	0	0	0	0	0	0	0	0	0	1	2	5	8	26	44	96	164

Cuadro 3.1: Expuestos al riesgo de morir por año y edades agrupadas

Edades	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
20-24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25-29	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
30-34	3	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35-39	1	0	0	0	2	2	1	2	0	0	0	0	0	0	0	0	0	0	0
40-44	1	1	1	1	1	1	0	1	4	2	1	0	0	0	1	1	0	0	0
45-49	1	1	1	2	2	2	1	1	2	3	1	1	1	2	1	3	0	2	2
50-54	2	7	4	0	2	2	0	0	1	2	2	1	1	3	0	2	2	2	4
55-59	6	4	6	2	5	7	7	8	4	2	3	1	3	4	3	2	1	6	3
60-64	3	1	8	2	2	6	1	9	8	8	7	13	10	11	7	4	3	6	2
65-69	0	0	0	2	2	5	9	8	5	9	7	8	4	7	8	13	12	11	9
70-74	0	0	0	0	0	1	1	2	3	1	12	11	12	11	11	6	11	11	10
75-79	0	0	0	0	0	0	0	0	1	1	1	2	1	2	4	11	4	14	18
80 y más	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	4

Cuadro 3.2: Fallecidos por año y edades agrupadas

Con los datos de los expuestos y fallecidos, por año y grupos de edades, se calculan las tasas centrales de mortalidad observada. Tomando las tasas observadas para el año 2013, se calculan las tasas de mortalidad aplicando los diferentes métodos mencionados en la sección 1.1.3.

Los valores "NaN" que se observan en el cuadro 3.3, es debido a que se tienen cero datos de fallecidos y de expuestos y por lo tanto su cociente da indeterminado. El software estadístico R expresa este valor mediante la sigla "NaN", que significa "Not a Number".

Como se puede ver en el cuadro 3.4 y en la figura 3.1, las tasas de mortalidad q_X , sea cual sea el método que se utilice para su estimación, dan

Capítulo 3. Resultados

Edades	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
20-24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	NaN	NaN	NaN	NaN	NaN	NaN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
25-29	0,0000	0,0083	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0062	0,0000	0,0000
30-34	0,0035	0,0000	0,0015	0,0039	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
35-39	0,0018	0,0000	0,0000	0,0000	0,0024	0,0023	0,0013	0,0029	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
40-44	0,0030	0,0026	0,0022	0,0020	0,0019	0,0018	0,0000	0,0015	0,0052	0,0024	0,0011	0,0000	0,0000	0,0000	0,0027	0,0044	0,0000	0,0000	0,0000
45-49	0,0040	0,0043	0,0045	0,0082	0,0070	0,0059	0,0025	0,0021	0,0039	0,0055	0,0017	0,0015	0,0014	0,0024	0,0011	0,0033	0,0000	0,0028	0,0036
50-54	0,0031	0,0120	0,0078	0,0000	0,0062	0,0079	0,0000	0,0000	0,0040	0,0068	0,0057	0,0024	0,0020	0,0050	0,0000	0,0030	0,0028	0,0026	0,0046
55-59	0,0107	0,0066	0,0090	0,0030	0,0073	0,0104	0,0116	0,0152	0,0090	0,0060	0,0115	0,0041	0,0129	0,0152	0,0093	0,0051	0,0022	0,0110	0,0048
60-64	0,0157	0,0035	0,0240	0,0047	0,0039	0,0104	0,0016	0,0131	0,0114	0,0110	0,0098	0,0205	0,0179	0,0238	0,0199	0,0142	0,0116	0,0241	0,0072
65-69	0,0000	0,0000	0,0000	0,0196	0,0128	0,0243	0,0292	0,0222	0,0111	0,0168	0,0115	0,0122	0,0055	0,0094	0,0103	0,0169	0,0175	0,0182	0,0180
70-74	NaN	NaN	NaN	0,0000	0,0000	0,0588	0,0323	0,0290	0,0259	0,0057	0,0513	0,0310	0,0293	0,0223	0,0190	0,0093	0,0159	0,0144	0,0124
75-79	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1,0000	0,2500	0,0476	0,0526	0,0118	0,0143	0,0191	0,0393	0,0097	0,0300	0,0330
80 y más	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0,0000	0,0000	0,0000	0,0000	0,0000	0,0227	0,0312	0,0244

Cuadro 3.3: Tasas centrales de mortalidad observada por año para edades agrupadas

Edades	Nx	Dx	Mx	qx1	qx2	qx3	qx4	qx5
20-24	30	0	0,000000	0,000000	0,000000	0,000000	0,000000	-
25-29	120	0	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
30-34	139	0	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
35-39	53	0	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
40-44	53	0	0,000000	0,000000	0,000000	0,000000	0,000000	-0,003562
45-49	556	2	0,003597	0,017825	0,017825	0,017838	0,017837	0,017134
50-54	867	4	0,004614	0,022805	0,022804	0,022825	0,022825	0,022795
55-59	620	3	0,004839	0,023904	0,023903	0,023926	0,023926	0,024152
60-64	278	2	0,007194	0,035336	0,035332	0,035382	0,035381	0,035896
65-69	501	9	0,017964	0,085960	0,085905	0,086200	0,086198	0,085381
70-74	805	10	0,012422	0,060241	0,060222	0,060367	0,060366	0,060142
75-79	545	18	0,033028	0,152542	0,152223	0,153147	0,153146	0,153465
80 y más	164	4	0,024390	0,114943	0,114808	0,115335	0,115333	-

Cuadro 3.4: Estimación de q_x por los métodos Lineal, Exponencial, Reed-Merrell, Greville y Keyfitz

todas muy similares. Por lo tanto, lógicamente, si se generan las tablas de mortalidad a partir de cada una de las q_x , se llega a e_x muy similares (ver cuadro A.3). Se observa también, que el término de corrección al método exponencial sugerido por Keyfitz, no genera grandes diferencias al resto de las estimaciones y que para el grupo de edades 40-44 años, genera un valor negativo que es inconsistente con la definición de q_x .

Dado todo lo anterior, se elige trabajar con la tasa de mortalidad estimada por el método lineal para la construcción de la tabla de mortalidad de los empleados del BROU, partiendo de una población hipotética de 5000 hombres y bajo el supuesto de una población cerrada, en la cual no hay ingresos y los únicos egresos que ocurren, son por la causal: fallecimiento.

Para el último grupo de edades, se calcula L_{80y+} , según la ecuación (1.14).

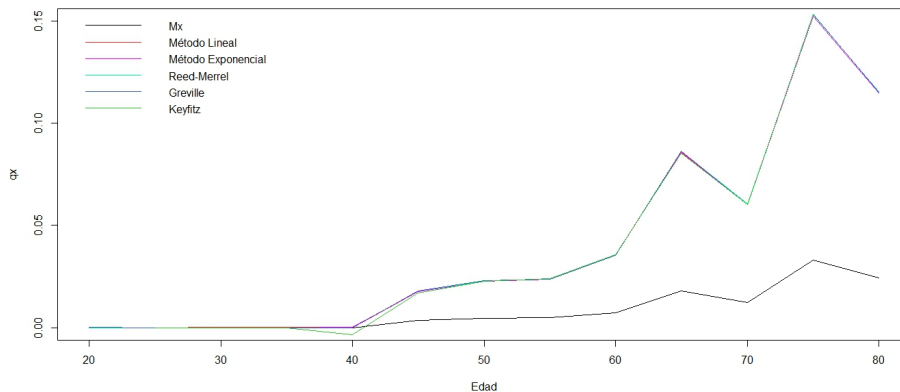


Figura 3.1: Tasas de mortalidad q_x

Se observa de la tabla de mortalidad (cuadro 3.5) y de la figura 3.2, que los fallecimientos y el descenso de la población, comienza a partir de los 45 años de edad, dado que no se observaron fallecimientos para las edades anteriores. Comparando las formas de las funciones biométricas l_x , d_x y q_x y las de la

Capítulo 3. Resultados

Edades	qx_1	lx	dx	Lx	Tx	ex
20-24	0,0000	5000,00	0,00	25000,00	295845,87	59,17
25-29	0,0000	5000,00	0,00	25000,00	270845,87	54,17
30-34	0,0000	5000,00	0,00	25000,00	245845,87	49,17
35-39	0,0000	5000,00	0,00	25000,00	220845,87	44,17
40-44	0,0000	5000,00	0,00	25000,00	195845,87	39,17
45-49	0,0178	5000,00	89,13	24777,18	170845,87	34,17
50-54	0,0228	4910,87	111,99	24274,39	146068,69	29,74
55-59	0,0239	4798,88	114,71	23707,62	121794,30	25,38
60-64	0,0353	4684,17	165,52	23007,04	98086,68	20,94
65-69	0,0860	4518,65	388,42	21622,19	75079,65	16,62
70-74	0,0602	4130,23	248,81	20029,11	53457,46	12,94
75-79	0,1525	3881,42	592,08	17926,88	33428,35	8,61
80 y más	0,1149	3289,34	378,08	15501,47	15501,47	4,71

Cuadro 3.5: Tabla de mortalidad para los funcionarios del BROU

población uruguaya presentadas en la figura 1.1, llama la atención el descenso de los fallecimientos entre los 70 y los 74 años de edad, lo que luego impacta en la estimación de la tasas de mortalidad, que se esperaría fuera de forma creciente a medida que avanzan las edades. La caída de los fallecimientos a partir de los 80 años de edad, se debe a la menor cantidad de hombres que alcanzan esas edades, pero cuando se calcula la tasa de mortalidad, se esperaría que ésta fuera mayor que la del tramo de edad anterior y no que cayera, como lo que ocurre con la tasa de mortalidad para los funcionarios del BROU.

La otra función biométrica de suma importancia que brinda la tabla de mortalidad, es la esperanza de vida para las distintas edades e_x . Los funcionarios más jóvenes, cuya edad es de 20 años, se espera que vivan 59 años más. Un funcionario que alcanza la edad de 60 años y que está en condiciones de jubilarse, se espera que viva casi 21 años más. Un funcionario jubilado que alcanza los 80 años de edad se espera que viva casi 5 años más.

Como la construcción de la tabla está basada únicamente en los datos del último año completo, 2013, es que se estudia la evolución de la esperanza

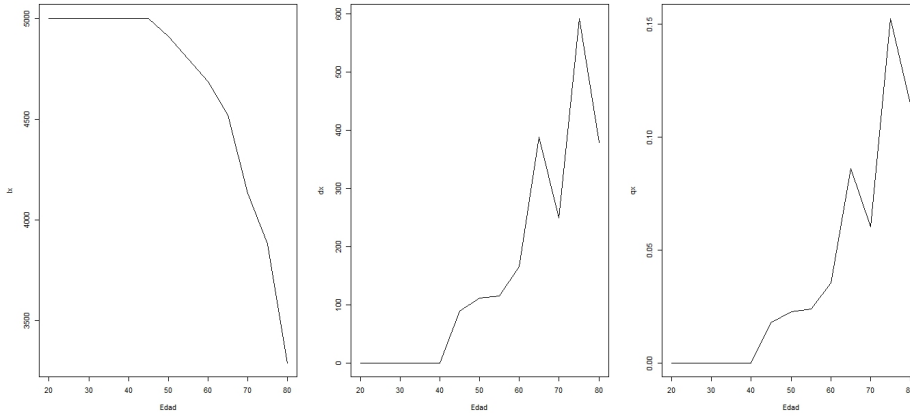


Figura 3.2: Funciones l_x , d_x y q_x de los funcionarios del BROU

de vida e_x para los distintos tramos de edades en los últimos tres años, del 2011 al 2013, y poder decir que la misma se ha mantenido constante (ver figura 3.3). Se elige un horizonte de tres años por dos razones, la primera es que debido a los ceros y los "NaN" que aparecen en el cuadro de las tasas centrales de mortalidad (ver cuadro 3.3) y los ceros de la tabla de fallecidos (ver cuadro 3.2) y dada la fórmula de cálculo de L_x para el último tramo de edades (ver ecuación 1.14), ésta da valores indeterminados. La otra razón, es que la mortalidad varía a lo largo de los años por el avance de la medicina y la mejora de la calidad de vida de las personas, lo que hace que aumente la esperanza de vida; por lo que tomar horizontes demasiado largos, no es conveniente para ver que las esperanzas de vida se mantienen constantes.

Suavizado de q_x

Para corregir las irregularidades que se observaron en la figura 3.2 respecto a la forma de d_x y de q_x , se suaviza la tasa de mortalidad q_x y se vuelven a calcular las funciones biométricas que conforman la tabla de mortalidad.

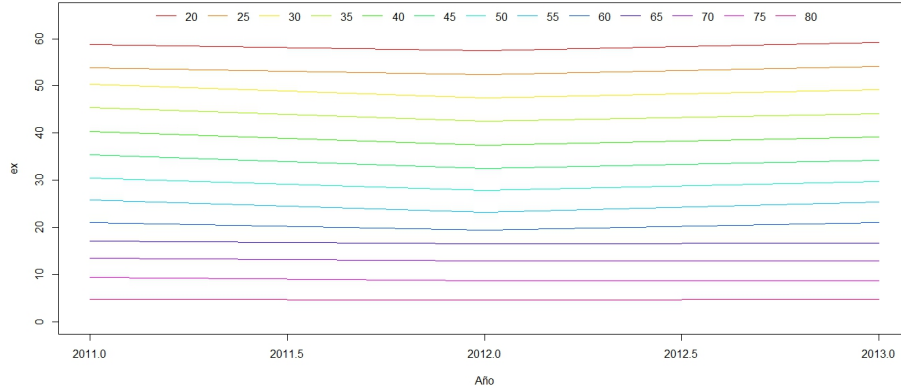


Figura 3.3: Esperanzas de vida e_x para los distintos tramos de edades durante los años 2011 a 2013

Para dicho suavizado se utiliza la función *smooth*, la cual utiliza el algoritmo no lineal de Tukey, que suaviza los datos mediante el uso de una combinación de suavizado de la mediana y filtrado lineal.

Edades	qx.smooth	lx	dx	Lx	Tx	ex
20-24	0,0000	5000,00	0,00	25000,00	297479,57	59,50
25-29	0,0000	5000,00	0,00	25000,00	272479,57	54,50
30-34	0,0000	5000,00	0,00	25000,00	247479,57	49,50
35-39	0,0000	5000,00	0,00	25000,00	222479,57	44,50
40-44	0,0000	5000,00	0,00	25000,00	197479,57	39,50
45-49	0,0178	5000,00	89,13	24777,18	172479,57	34,50
50-54	0,0228	4910,87	111,99	24274,39	147702,38	30,08
55-59	0,0239	4798,88	114,71	23707,62	123428,00	25,72
60-64	0,0353	4684,17	165,52	23007,04	99720,38	21,29
65-69	0,0602	4518,65	272,21	21912,72	76713,34	16,98
70-74	0,0860	4246,44	365,02	20319,64	54800,62	12,91
75-79	0,1149	3881,42	446,14	18291,74	34480,97	8,88
80 y más	0,1149	3435,28	394,86	16189,24	16189,24	4,71

Cuadro 3.6: Tabla de mortalidad para los funcionarios del BROU suavizada

De la figura 3.4, se observa que se logró corregir la caída de la tasa de mortalidad en los tramos de edades 70-74 años y 80 y más, siendo la misma creciente en todo su recorrido. También se corrige la disminución de los fallecimientos que se observaba para el tramo de edades 70-74 años. De este

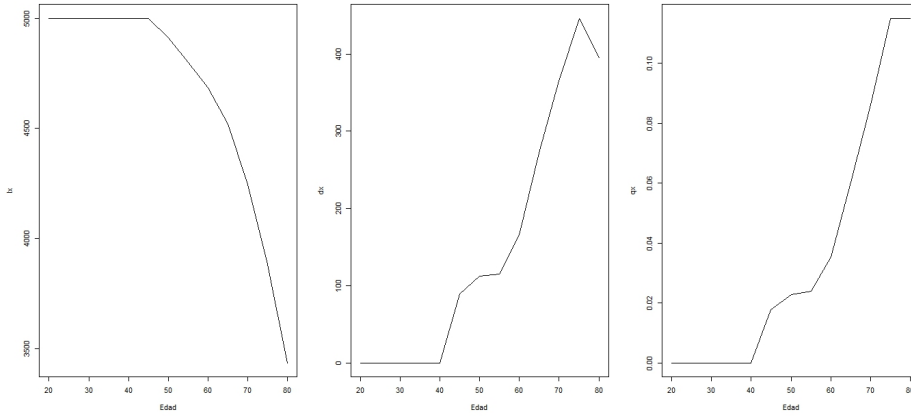


Figura 3.4: Funciones l_x , d_x y q_x con *smooth*

modo, las funciones biométricas l_x , d_x y q_x , se comportan de forma similar a las de la población uruguaya.

Sin embargo, no se encuentran grandes diferencias entre las esperanzas de vida obtenidas con la q_x suavizada y la q_x sin suavizar (ver cuadro 3.5).

3.1.1. Tablas de mortalidad a partir de los datos del INE y del BCU

A modo de comparar los resultados obtenidos para la población del BROU, se analizan los resultados de las estimaciones de q_x con los datos del INE y se construye la correspondiente tabla de mortalidad para la población masculina uruguaya. También se construyen las tablas de mortalidad agrupada y por edades simples con la tasa de mortalidad del BCU.

Tabla de mortalidad con datos del INE

Se toma del INE, la tasa central de mortalidad observada por edades simples m_x , cabe mencionar que la misma va hasta los 100 años de edad. Partiendo de ella, se calcula la tasa de mortalidad q_x , mediante los diferentes métodos presentados en la sección 1.1.3.

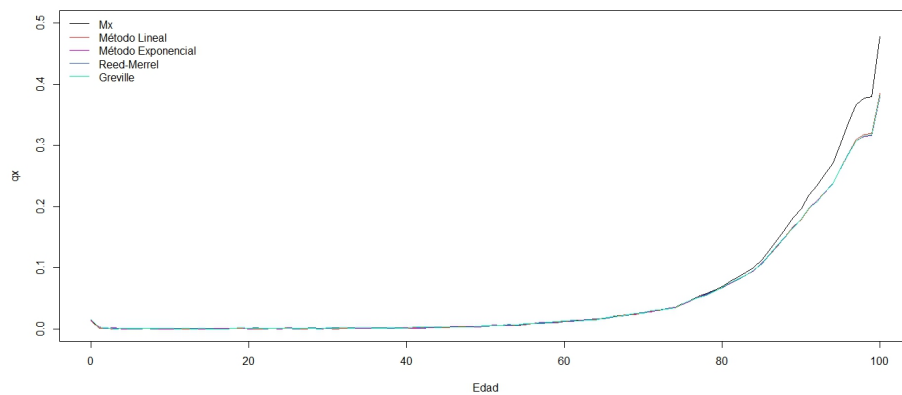


Figura 3.5: Tasa de mortalidad q_x según datos del INE calculadas a partir de los métodos: Lineal, Exponencial, Reed-Merrell y Greville

Al igual que con los datos del BROU y como se observa en la figura 3.5, no se encuentran grandes diferencias entre las distintas estimaciones de q_x . Los valores de dichas estimaciones, se pueden ver en el cuadro A.5.

Comparando las gráficas de las tasas de mortalidad obtenidas con los datos del BROU (figura 3.1) y las obtenidas con los datos del INE, se observa que éstas últimas son más parsimoniosas, esto debido a que las tasas centrales de mortalidad observadas por el INE, son construidas en base a todos los hombres del Uruguay y las del BROU, en base a los hombres que pertenecen a ésta institución, siendo su número muy reducido en comparación con la población masculina total de Uruguay. También, las q_x obtenidas con los

datos del INE, son crecientes a lo largo de todo el recorrido, cosa que no ocurre con las tasas de mortalidad del BROU, que se muestran decrecientes para las edades de 70 a 74 años y con los mayores de 80.

Otra diferencia importante, se observa entre la relación de la m_x y las q_x de cada población. Observando las estimaciones del INE, sus q_x son casi iguales a la m_x y recién se despega para las últimas edades, siendo su brecha pequeña. Además, la m_x del INE queda por arriba de las q_x , a diferencia de lo que pasa con las estimaciones del BROU, donde la m_x queda por debajo de las estimaciones de las q_x con una brecha mayor.

Para la construcción de la tabla de mortalidad con los datos del INE y dado que las tasas de mortalidad estimadas mediante los diferentes métodos ya mencionados dan muy similares, se toma la estimación de q_x , calculada por el método lineal. Dicha tabla se obtiene, partiendo de una población hipotética de 5.000 hombres. También se calcula la esperanza de vida e_x , a partir de la ecuación (1.22), presentada en (Alho, 2005) para observar que tan buena es ésta estimación. La tabla se puede consultar en el cuadro A.6.

La esperanza de vida calculada con la ecuación (1.22), se observa que sobreestima los valores de e_x . Para las edades 20, 60 y 80 años, los valores de e_x son: 76, casi 37 y 18,5 años respectivamente por lo que hay una sobreestimación de 19, 16 y 10 años en dichas edades, en comparación con la e_x para el INE, calculada por el método convencional. Por lo tanto se descarta el cálculo de la esperanza de vida utilizando este método.

Tabla de mortalidad con datos del BCU

Como se mencionó en los antecedentes (Capítulo), el BCU construye una tasa de mortalidad hasta los 110 años para regular los cálculos de las reservas técnicas de las aseguradoras. Se toma la tasa de mortalidad para los hombres y se construye la tabla de mortalidad, partiendo de una población hipotética de 5000 personas, la misma se puede ver en el cuadro A.7.

Comparando la esperanza de vida obtenida en la tabla de mortalidad, con la tabla de mortalidad del INE para las edades 20, 60 y 80, se observa una diferencia de 2 años y medio para la edad de 20 años, donde el INE espera que ésta viva casi 57 años más y el BCU espera que viva 54 años y medio más. A medida que avanzan las edades, la brecha de la esperanza de vida entre una y otra tabla se achica. Una persona de 60 años de edad, se espera que viva 20 años y medio más según el INE o 19 años más según el BCU. Para una persona que alcanza los 80 años de edad, ambas entidades esperan que ésta persona viva 8 años más.

Se construye la tabla de mortalidad abreviada del BCU para poder compararla con la tabla de mortalidad obtenida para los funcionarios del BROU. Dado que para el último tramo de edades la q_x del BCU es igual a 1, L_{80} y + se calcula bajo el supuesto de que $e_{80} = 4,71$ y $L_{80} = l_{80}e_{80}$, siendo el valor de e_{80} , el que se obtuvo para la población del BROU. Se calcula de diferente manera, ya que al valer 1 la tasa de mortalidad, el valor de L_{80} que se obtiene utilizando el mismo criterio que para el cálculo de L_{80} del BROU (mediante la ecuación 1.14), da negativo.

Como se observa en el Cuadro 3.7, la esperanza de vida que se obtiene

3.1. Tabla de mortalidad BROU

Edades	q_x .bcu	l_x	dx	L_x	T_x	e_x
20-24	0,0060	5000,00	29,93	24925,19	266269,09	53,25
25-29	0,0065	4970,07	32,32	24769,57	241343,91	48,56
30-34	0,0070	4937,75	34,64	24602,17	216574,34	43,86
35-39	0,0090	4903,11	44,26	24404,90	191972,17	39,15
40-44	0,0127	4858,85	61,83	24139,68	167567,27	34,49
45-49	0,0228	4797,02	109,56	23711,21	143427,59	29,90
50-54	0,0368	4687,46	172,31	23006,54	119716,38	25,54
55-59	0,0583	4515,16	263,37	21917,36	96709,84	21,42
60-64	0,0918	4251,79	390,13	20283,62	74792,48	17,59
65-69	0,1311	3861,66	506,14	18042,95	54508,87	14,12
70-74	0,1870	3355,52	627,59	15208,61	36465,92	10,87
75-79	0,2660	2727,92	725,50	11825,88	21257,31	7,79
80 y más	1,0000	2002,43	2002,43	9431,43	9431,43	4,71

Cuadro 3.7: Tabla de mortalidad abreviada BCU

mediante la q_x del BCU es menor a la del BROU para todos los tramos de edades.

3.1.2. Tabla empírica simulada

Utilizando la tasa de mortalidad del BCU, se simula, mediante el método Monte Carlo, los fallecimientos de los expuestos al riesgo de morir del BROU en los distintos tramos etarios en cada año.

Como se observa en la matriz de expuestos al riesgo de morir (cuadro 3.1), la misma tiene muchos ceros en el último tramo de edades durante los años 1995 a 2008 y ceros para el tramo de edades más jóvenes entre 2001 y 2006. Por este motivo, se decide tomar para la simulación los años 2007 a 2013.

Lo que se busca con esta simulación, es ver que tanto se asemeja o se diferencia el comportamiento de los fallecimientos estimados mediante la tasa de mortalidad del BCU, del que realmente se observa en la muestra de fallecidos del BROU. Esto permitirá evaluar la importancia, o no, de contar con una tasa de mortalidad específica para esta población o utilizar la del BCU sin

mayores esfuerzos.

La simulación, es una "Técnica numérica que permite conducir experimentos en una computadora, para lo cual se precisan Modelos Lógicos Matemáticos que describan la conducta del negocio, fenómeno o realidad que se intenta reproducir, a través de periodos de tiempo real". Se recurre a la simulación, cuando se desea modificar algunos de los componentes que integran el problema de análisis, para verificar algún resultado o estudiar la sensibilidad a cambios que se desean introducir. Ese sistema o modelo que se quiere reproducir, se puede implementar desde la perspectiva determinística o en condiciones de incertidumbre. La simulación que involucra aspectos aleatorios, generalmente se engloba en lo que se llama Simulación Monte Carlo (Álvarez and Massa, 2015).

El método consiste en la generación de observaciones de una variable aleatoria, en este caso, de una variable aleatoria uniforme. La variable aleatoria que se simula es $D \sim Bin(n, q_x)$, que es la suma de n variables aleatorias independientes $d \sim Ber(q_x)$, siendo n el número de expuestos en el tramo de edad x y q_x la probabilidad de morir en ese tramo de edad. Las variables aleatorias d , toman el valor $d = 1$ con probabilidad q_x y $d = 0$ con probabilidad $1 - q_x$. Para simular d , se utilizan los valores de las observaciones generadas $u \sim U(0, 1)$ y se asigna el valor $d = 1$ si $u \leq q_x$ y $d = 0$ si $u > q_x$.

Más específicamente, el proceso de simulación se hace para cada año y comienza creando un vector, cuyo largo es la cantidad de expuestos en dicho año y que contiene la edad de cada expuesto. Se sigue el proceso, creando otro vector del mismo largo, que contiene la q_x correspondiente a la edad del expuesto. Luego se genera un vector del mismo largo que los anteriores,

con valores al azar de una distribución Uniforme entre 0 y 1. Si éste valor simulado es menor o igual al valor de q_x , se dice que la persona falleció y se le asigna valor 1, si es mayor, entonces la persona sigue viva y se le asigna valor 0. Luego, se cuenta la cantidad de fallecidos simulados para cada edad en cada año. Éste proceso se itera N veces.

El error absoluto de estimación de este método es $1/\sqrt{N}$, donde N , es el número de pruebas. En virtud del teorema del límite central, el error de estimación decrece conforme aumenta el número de pruebas. Iterando el proceso de simulación en un entorno de $N = 1000$, el error absoluto de estimación es $1/\sqrt{1000} \approx 0,032$. Considerándose adecuado dicho error, se realizan $N = 1000$ pruebas.

Por el teorema del límite central, que dice:

Sea X_1, X_2, \dots, X_N un conjunto de variables aleatorias independientes e idénticamente distribuidas de una distribución con media μ y varianza $\sigma^2 \neq 0$. Entonces, si N es suficientemente grande, la variable aleatoria

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$$

tiene aproximadamente una distribución Normal con $\mu_{\bar{X}} = \mu$ y $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{N}$,

luego de tener las 1000 simulaciones de los fallecimientos para cada tramo de edad por año, se calcula el promedio de dichos fallecimientos obteniendo el valor esperado de fallecimientos para dicho tramo de edad en dicho año.

Comparando los resultados de las muertes simuladas con los fallecidos en el BROU (cuadro 3.8), se observa que se está sobrestimando los fallecimientos realmente ocurridos. Ésto se puede ver en la figura 3.6, donde se presentan las

Capítulo 3. Resultados

Edades	2007	2008	2009	2010	2011	2012	2013	2007	2008	2009	2010	2011	2012	2013
20-24	1	1	0	0	0	0	0	0	0	0	0	0	0	0
25-29	0	1	1	1	1	1	1	0	0	0	0	1	0	0
30-34	0	0	0	0	1	1	1	0	0	0	0	0	0	0
35-39	1	0	0	0	0	0	0	0	0	0	0	0	0	0
40-44	9	7	5	3	2	1	1	0	0	1	1	0	0	0
45-49	16	19	21	21	19	16	13	1	2	1	3	0	2	2
50-54	18	22	23	25	27	28	32	1	3	0	2	2	2	4
55-59	14	15	19	23	27	32	36	3	4	3	2	1	6	3
60-64	51	43	32	26	24	23	26	10	11	7	4	3	6	2
65-69	95	98	102	101	90	79	66	4	7	8	13	12	11	9
70-74	77	93	108	121	130	143	151	12	11	11	6	11	11	10
75-79	22	37	56	75	110	125	145	1	2	4	11	4	14	18
80 y más	2	5	8	26	44	96	164	0	0	0	0	1	3	4

Cuadro 3.8: Promedio de los fallecimientos simulados y fallecimientos observados de la muestra del BROU

tasas centrales de mortalidad simuladas¹ y las tasas centrales de mortalidad observadas del BROU. El crecimiento de la tasa central de mortalidad del BROU a lo largo de las edades, es muy suave en todos los años, siendo tasas centrales de mortalidad muy bajas, que se diferencian mucho de las obtenidas con la tasa de mortalidad del BCU.

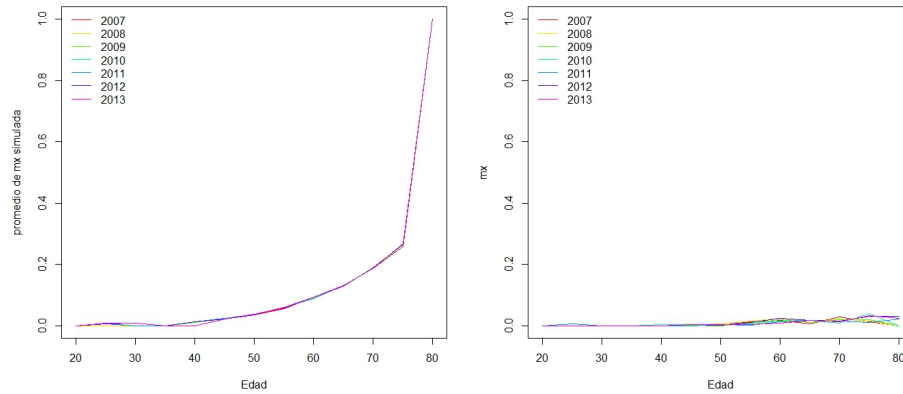


Figura 3.6: Tasas centrales de mortalidad m_x , simuladas y observadas respectivamente, para los años 2007 a 2013

Se concluye que no sería conveniente utilizar la tasa de mortalidad del BCU para estimar los fallecimientos de los funcionarios del BROU, ya que

¹Los valores de las tasas centrales de mortalidad simuladas se pueden ver en el cuadro A.8

se estarían sobrestimando los mismos.

3.2. Descripción y Proyección de la mortalidad para los funcionarios del BROU. Período 1995-2018

En esta sección, se presentan los resultados obtenidos al implementar el modelo de Lee-Carter para el estudio de la mortalidad en el tiempo y proyección de la misma.

Como se presentó en la sección 1.2.1, ecuación (1.29), el modelo de Lee-Carter a estimar es:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$$

Para comenzar con el estudio y poder utilizar las librerías nombradas al comienzo del capítulo, se debe crear un objeto que contenga los datos demográficos para el análisis. La clase de dicho objeto se denomina *demogdata* y contiene las tasas centrales de mortalidad m_x , los expuestos al riesgo de morir para las distintas edades a lo largo de los años 1995 a 2013, las edades y los años de estudio, especificando que los datos contenidos en el objeto son referentes a la mortalidad de los hombres.

Del objeto *demogdata*, se pueden generar los gráficos que se presentan en las figuras 3.7 y 3.8. En la primer figura, se muestran los gráficos de la tasa central de mortalidad observada de los funcionarios del BROU m_x por edad para los años 1995 a 2013 y el logaritmo de dichas tasas, las cuales se

pretenden estimar, por edad y por año. El último gráfico de la figura 3.7 que muestra el logaritmo de m_x por año para las distintas edades, se observa que sus curvas no son continuas, esto es debido a que, en muchos de los años, los valores de las tasas m_x valen cero. En la figura 3.8, se muestran los gráficos de los fallecimientos por edad a lo largo del período 1995 a 2013 y el logaritmo de los mismos. Como es de esperar, a medida que aumenta la edad se observan mayor cantidad de fallecimientos.

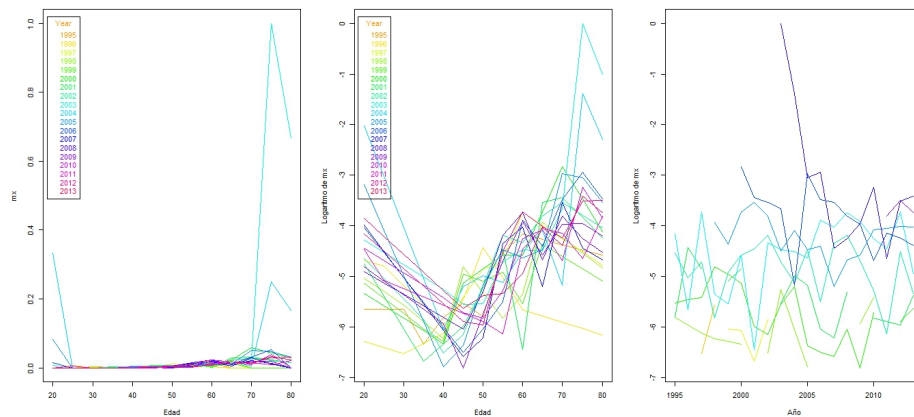


Figura 3.7: m_x y $\log(m_x)$ por edad para los años 1995 a 2013 y $\log(m_x)$ por año para las distintas edades

El modelo de Lee-Carter no se puede estimar, si a lo largo de todos los años la tasa central de mortalidad observada m_x vale 0, cosa que ocurre para el tramo de edades más jóvenes, de 20 a 24 años de edad. Por tal motivo, se saca a este grupo de edades para la estimación del modelo.

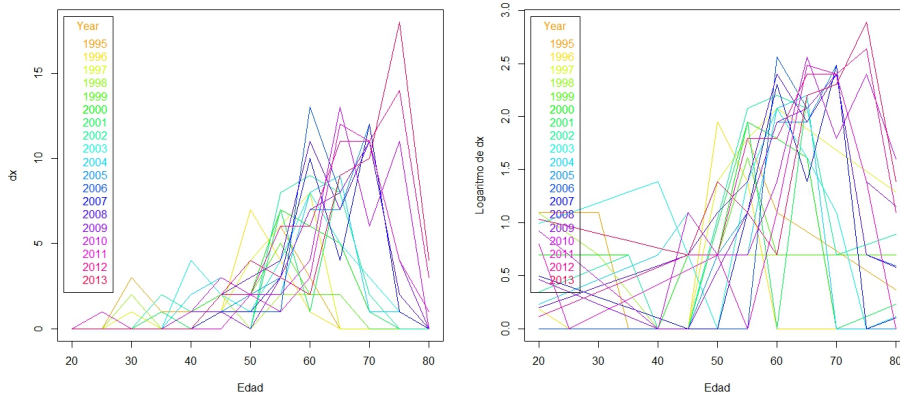


Figura 3.8: d_x y $\log(d_x)$ por edad para los años 1995 a 2013

3.2.1. Estimación del modelo mediante SVD y proyección

Estimar el modelo mediante SVD, implica suponer que sus errores son independientes y que se distribuyen Normal con media cero y varianza constante. Por lo que al obtener el modelo estimado, se deben testear sus residuos para ver si cumplen con estas hipótesis. Los test utilizados son: de gráficos para la visualización de su media y su varianza, test de Jarque-Bera para la normalidad y test de Ljung-Box para la incorrelación.

Se comienza estimando el modelo, mediante la utilización de la función *lca* que se encuentra en la librería *demography*, el modelo incluye todos los grupo de edades excluyendo al grupo de los más jóvenes (20-24 años). Se obtiene un modelo que se ajusta a los datos en 71,7%, pero al analizar los residuos, para ver si cumplen las hipótesis del modelo, los mismos muestran heterocedasticidad en el grupo de edades 75-79 años, mientras que los grupos de edades de 30-34 y el grupo de edad más avanzada (80 y más años), no

cumplen con los supuestos de normalidad para la serie de tiempo 1995-2013². Por tal motivo se vuelve a estimar el modelo para los grupos de edades 35-39 a 75-79 años y se contrastan los residuos, encontrándose correlación entre los mismos para el grupo de edades 35-39 años (ver cuadro A.10). Finalmente, se estima el modelo para los grupos de edades 40-44 hasta 75-79 años, obteniéndose un ajuste del 72,6 %, el cual cumple con los supuestos de normalidad e incorrelación con una confianza del 95 % (observar cuadro A.11) donde todos los p-valores obtenidos son menores a 0.05, también se cumple el supuesto de esperanza de los residuos igual a 0, aunque como ya se había observado, existe heterocedasticidad para el grupo de edades 75-79 años, por los que los resultados obtenidos para éste último grupo no serán del todo fiables. De la gráfica de los box plots, se deduce que los valores que generan la heterocedasticidad se encuentran en los años 1997, 2000 y 2002 (ver figura 3.9).

Parámetros del modelo

Los parámetros del modelo que se muestran en la figura 3.10 y los valores de los mismos, se pueden consultar en los cuadros A.12 y A.13.

Los valores de a_x refieren a la estructura de la mortalidad y representan cómo se comporta ésta a través de las edades. Se observa que la mortalidad crece a medida que avanzan las edades, lo que es lógico.

En cuanto a los valores de b_x , éstos describen el cambio en la mortalidad en el intervalo de edades x , frente a un cambio en k_t . Cuando b_x es grande

²La gráfica de los residuos, así como los test aplicados pueden consultarse en la figura A.3 y cuadro A.9

3.2. Descripción y Proyección de la mortalidad

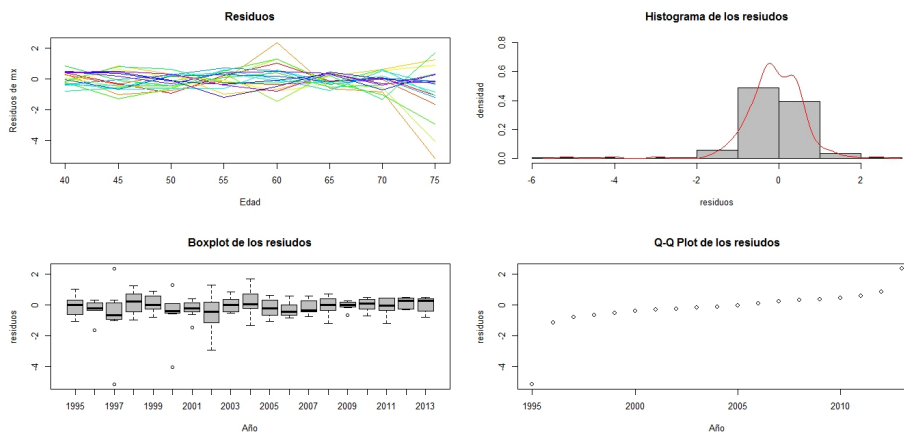


Figura 3.9: Residuos del modelo de Lee-Carter estimado para los grupos de edades 40-44 a 75-79 años durante los años 1995 a 2013, mediante la librería *demography*

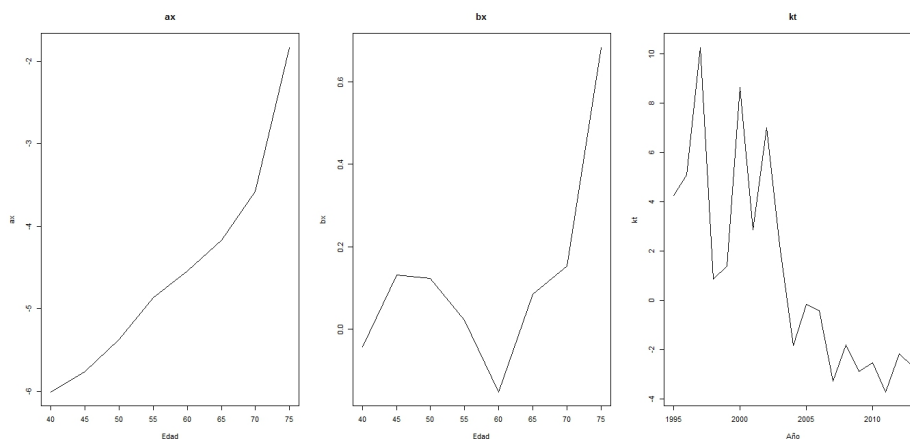


Figura 3.10: Estimaciones de a_x , b_x y k_t mediante SVD

para cierto intervalo, expresa que la tasa de mortalidad en dicho intervalo varía sustancialmente cuando la tendencia de la mortalidad k_t cambia, esto sucede en el último tramo de edades de 75-79 años. Cuando b_x es pequeño, significa que las tasas de mortalidad para esa edad varían levemente cuando la tendencia de la mortalidad cambia, como se observa en el tramo de edades de 60-64 años.

En cuanto a los valores de k_t , que describen la tendencia de la mortalidad a lo largo del tiempo, se observa que la misma sigue una tendencia decreciente, que se ha vuelto más parsimoniosa en los últimos años, fruto de la calidad de los datos, ya que las series de los últimos años, cuentan con mayor cantidad de datos para todos los grupos de edades. Esta tendencia decreciente explica el aumento de la esperanza de vida.

Proyección de la tendencia de la mortalidad para los años 2014 a 2018

Usando la función *forecast*, se proyecta la tendencia de la mortalidad k_t para los siguientes 5 años con un nivel de confianza del 90 %. Los valores estimados de la tendencia y su intervalo de confianza, se pueden consultar en el cuadro A.14.

La función *forecast* (contenida dentro de la librería con igual nombre), es una función genérica para la proyección de series temporales o modelos de series de tiempo. La función evoca métodos particulares que dependen de la clase del primer argumento. En particular para el modelo de Lee Carter, donde la clase del primer argumento es un glm (modelo lineal generalizado), genera proyecciones ARIMA(0,1,0).

De la figura 3.11, se observa una tendencia decreciente de la mortalidad k_t para los siguientes años. Observando la proyección del logaritmo de la tasa de mortalidad m_x por edades para los distintos años, se ve un decrecimiento de m_x hasta el grupo de edades de 45-49 años y gran crecimiento para las edades siguientes hasta los 60-64 años, luego se observa la inestabilidad que presentan los años siguientes, sobre todo para el tramo de edades 75-79 años,

3.2. Descripción y Proyección de la mortalidad

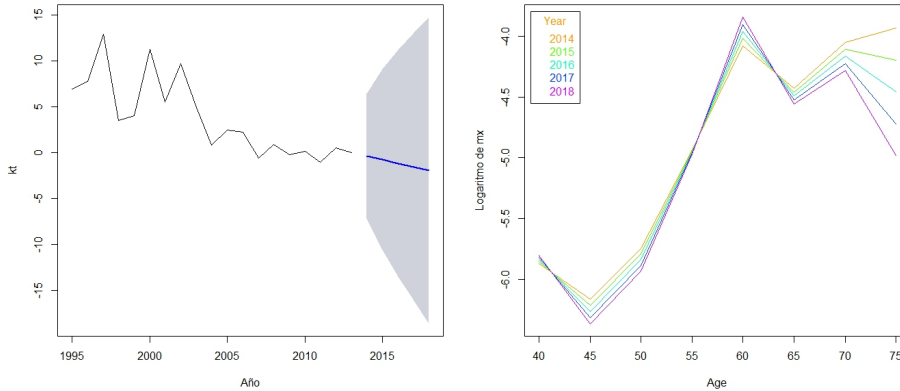


Figura 3.11: Proyección de k_t y $\log(m_x)$ por edad para los años 2014 a 2018

donde la variabilidad de la proyección para los distintos años es muy grande. Esto es debido a que los parámetros estimados para estas edades, no lograron ajustarse bien, lo que se observa en el gráfico de residuos (figura 3.9).

Observando ahora las tasas de mortalidad m_x , proyectadas en todos los tramos de edades para los años 2014 a 2018 (figura 3.12, los valores se pueden consultar en el cuadro A.15), las mismas se muestran estables para los primeros grupos de edades hasta el grupo de 55-59 años, con leve aumento en la tasa de mortalidad para el grupo de edades 40-44 años y un fuerte aumento para el grupo de edades 60-64 años. Para los grupos de edades 45-49 a 55-59 y 65-69 años, se observa disminución de las tasas de muerte para los próximos años. Observando los últimos tramos de edades, a partir del grupo de 70-74 años, los mismos se muestran con grandes fluctuaciones, lo cual proyecta una disminución de las tasas de mortalidad para estos grupos. Para el último grupo de edades, 75-79 años, la caída es muy fuerte debido a la variabilidad que presenta dicho grupo, no siendo fiable su estimación.

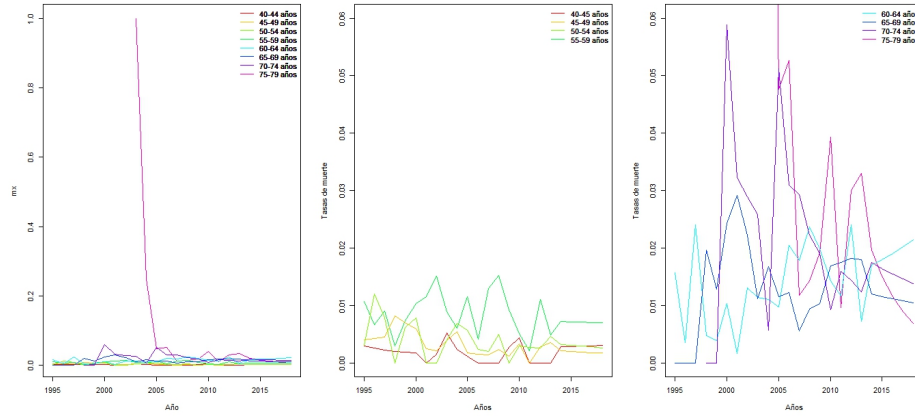


Figura 3.12: Tasas de mortalidad proyectadas para los años 2014 a 2018

3.2.2. Estimación del modelo mediante MV y proyección

Este modelo presenta la ventaja que admite heterocedasticidad entre sus residuos. Para la obtención del mismo, se utiliza la función *lca.rh*, que se encuentra en la librería *ilc*.

Parámetros del modelo

Se comienza estimando un modelo para los mismos grupos de edades con los que se estimó el modelo mediante SVD, para poder comparar los resultados.³

Se obtiene un modelo que explica 93,05% de su variabilidad.

De la figura 3.13, se observa que los parámetros a_x y b_x se comportan muy similares a los obtenidos mediante la función *lca*. En cambio, el parámetro k_t , que muestra la tendencia de la mortalidad a lo largo de tiempo, ya no se

³Los parámetros del modelo estimado, gráficos de residuos y proyección pueden consultarse en los cuadros A.16, A.17, A.4, A.18 y A.19

3.2. Descripción y Proyección de la mortalidad

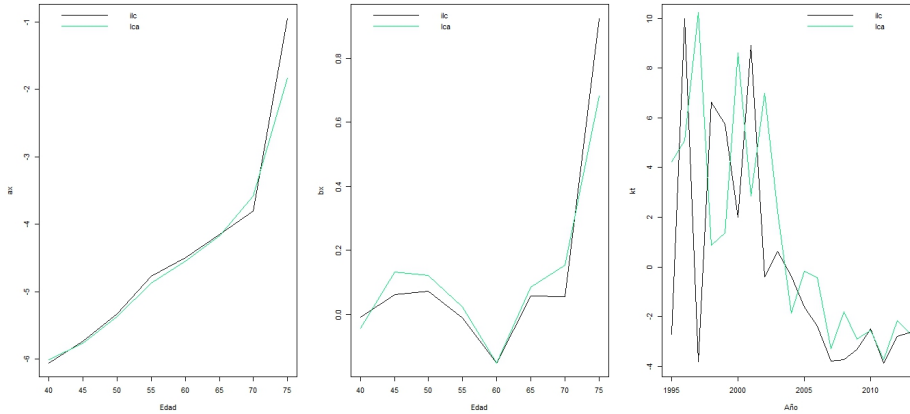


Figura 3.13: Estimaciones a_x , b_x y k_t mediante la librería *ilc*

observa decreciente como en obtenido con la función *lca*, sino que se mantiene constante en promedio, pero con fluctuaciones muy grandes en los primeros años de la serie.

Como se dijo al comienzo, este modelo estimado por MV, permite heterocedasticidad de los residuos y por tal motivo se estima el mismo para todos los tramos de edades que quedaron afuera del análisis en el modelo anterior. Obteniendo una variabilidad explicada del 95,48% (mejor ajuste que el modelo anterior) y una menor devianza.

De los valores de a_x , que refieren a la estructura de la mortalidad y representan cómo se comporta ésta a través de las edades, se observa que dicho parámetro captura el patrón de comportamiento que presenta la mortalidad por edades de los funcionarios del BROU (figura 3.7). Observando los grupos de edades que no se habían descrito hasta ahora, vemos una disminución de la mortalidad para los primeros grupos de edades hasta los 40 años. La caída que se observa para el último grupo de edades, no sería el comportamiento esperado, dado que la mortalidad tiende a ser creciente a medida que la edad

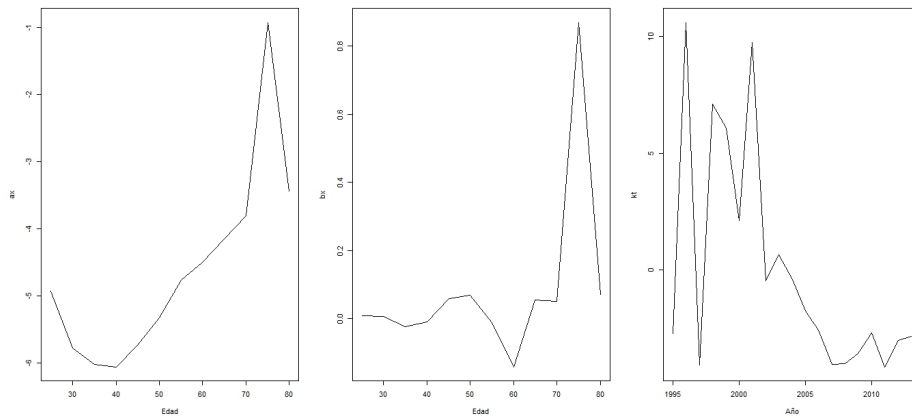


Figura 3.14: Estimaciones a_x , b_x y k_t para todos los grupos de edades mediante MV

avanza.

En cuanto a los valores de b_x , que representan qué tanto afectan los cambios de la tendencia de la mortalidad en la mortalidad misma, al igual que los parámetros estimados por SVD, se observa que los cambios más grandes de la mortalidad cuando varía la tendencia, se dan en el intervalo de edades de 75-79 años y que el grupo de edades de 60-64 años es el que menos cambia la mortalidad ante cambios en la tendencia k_t . Los grupos de edades más jóvenes y el último tramo de edades, 80 años y más, no se ven afectados en la mortalidad cuando cambia la tendencia.

Comparando el comportamiento de los parámetros a_x y b_x de este estudio, con los obtenidos por otros investigadores mediante el mismo método para poblaciones de países enteros⁴, el comportamiento de a_x siempre es creciente, sólo cae para los primeros años de edad (0 a 15 años) los cuales no forman parte de este estudio. Comparando el parámetro b_x de éste informe, con el

⁴estos estudios se pueden consultar en la Bibliografía

3.2. Descripción y Proyección de la mortalidad

de (Belliard and Williams, 2013) para la población Argentina, se observa que el grupo de edades donde la mortalidad varía con más fuerza ante cambios del nivel general de mortalidad, es el grupo 40-44 años, también varía positivamente la mortalidad en el grupo de edades 70-74 años, pero no con tanta fuerza. La tendencia de la mortalidad k_t siempre es decreciente.

Proyección de la tendencia de la mortalidad para los años 2014 a 2018

Se observa de la figura 3.15, una tendencia de la mortalidad k_t constante para los siguientes cinco años proyectados, 2014 a 2018, con una confianza del 90 %. Para los logaritmos de las tasas de mortalidad m_x , se observa una disminución hasta el grupo de edades 40-44 años, donde comienza a crecer con caídas en los grupos de edades 65-69 años y para el último grupo, 80 años y más.⁵

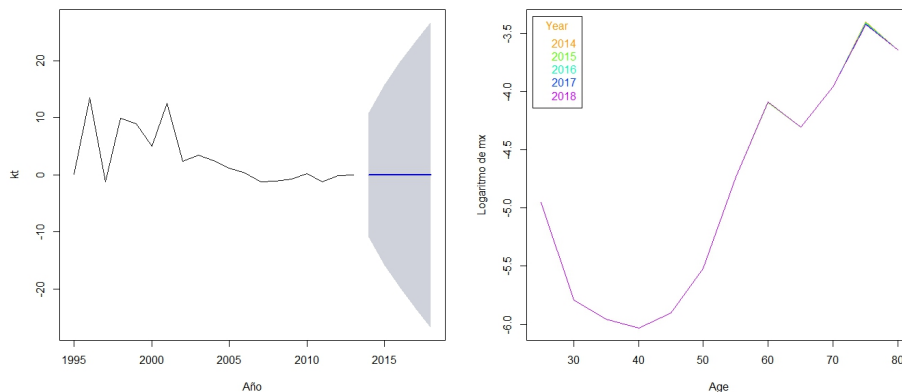


Figura 3.15: Proyección de k_t y $\log(m_x)$ por edad para los años 2014 a 2018

⁵Los resultados de las proyecciones de k_t y su intervalo de confianza, así con el valor de las tasas m_x se pueden consultar cuadros A.22 y A.23 respectivamente.

Capítulo 3. Resultados

De la figura 3.16, se observan tasas de mortalidad m_x constantes para los próximos cinco años, a excepción del grupo de edades 75-79 años, donde se espera una leve caída de la tasa.

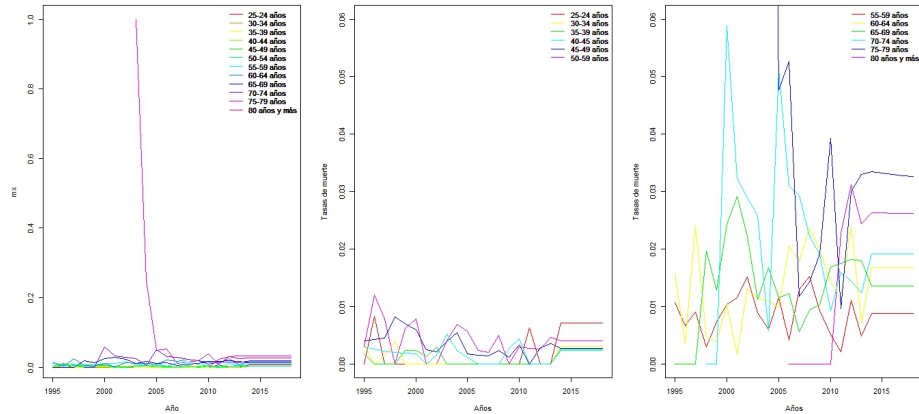


Figura 3.16: Tasas de mortalidad proyectadas para los años 2014 a 2018 de todos los grupos de edades

Comparación de resultados de ambos métodos

A continuación se presentan los valores de los modelos estimados por los métodos SVD y MV para los grupos de edades 40-44 a 75-79 años.

Edades	ax.lca	bx.lca	ax.ilc	bx.ilc
40-44	-6,0062	-0,0437	-6,0570	-0,0099
45-49	-5,7613	0,1315	-5,7359	0,0625
50-54	-5,3705	0,1224	-5,3254	0,0731
55-59	-4,8666	0,0222	-4,7655	-0,0121
60-64	-4,5434	-0,1521	-4,4937	-0,1519
65-69	-4,1674	0,0843	-4,1484	0,0581
70-74	-3,5785	0,1526	-3,8067	0,0554
75-79	-1,8325	0,6829	-0,9462	0,9247

Cuadro 3.9: Estimaciones a_x y b_x de los modelos de Lee-Carter con librerías *lca* e *ilc* para los grupos de edades 40-44 años a 75-79 años

Se observa del cuadro 3.9, que los valores estimados por ambos métodos son bastante similares, difiriendo en mayor medida en el último grupo de

3.2. Descripción y Proyección de la mortalidad

Año	kt.lca	kt.ilc
1995	4,2443697	-2,7109515
1996	5,085095	10,0077256
1997	10,2558827	-3,8145095
1998	0,8585714	6,6408311
1999	1,3726559	5,7434333
2000	8,6255637	1,9983772
2001	2,8546784	8,9180306
2002	7,0018136	-0,3984105
2003	2,2303087	0,6296841
2004	-1,8488875	-0,3611682
2005	-0,1883998	-1,6058123
2006	-0,4276358	-2,3845605
2007	-3,2970649	-3,7776793
2008	-1,8152536	-3,7477539
2009	-2,9123175	-3,3323571
2010	-2,5387839	-2,4920978
2011	-3,7263861	-3,8667571
2012	-2,1722307	-2,8019436
2013	-2,6883401	-2,6440805

Cuadro 3.10: Estimaciones k_t de los modelos de Lee-Carter con librerías *lca* e *ilc* para los grupos de edades 40-44 años a 75-79 años

edades. En cambio, los valores estimados del parámetro k_t , son similares sólo para los últimos 5 años, difiriendo en gran medida para los años anteriores, sobre todo en los primeros años.

Si se comparan los valores proyectados de la tendencia de la mortalidad k_t del modelo estimado por SVD y de ambos modelos estimados por MV: con grupos de edades de 40-44 años a 75-79 años y el modelo con los grupos de edades 25-29 años al grupo 80 años y más (los cuadros son A.14 y A.18, A.22 respectivamente), llamando a los mismos: modelo 1, modelo 2 y modelo 3 respectivamente, se observa del modelo 1 que al igual que se observó en la gráfica 3.11, sus valores son decrecientes. Para los modelos 2 y 3, los valores que se observaron gráficamente en 3.15, son constantes, de los cuadros se observa que son levemente crecientes. Los intervalos de confianza son muy amplios, siendo los más pequeños los del modelo 1, seguido del modelo 2. Al ser tan amplios los intervalos, los valores k_t proyectados, están contenidos en

cualquiera de los intervalos de los distintos modelos.

Capítulo 4

Conclusiones

A continuación se presentan las conclusiones a las que se llega luego del análisis y modelización de los datos. Comenzando con las conclusiones a las que se llega al construir la tabla de mortalidad específica para los funcionarios del BROU. Luego, se presentan las referentes a la proyección de la mortalidad mediante el modelo de Lee-Carter. Por último, se especifican las limitaciones encontradas a la hora de implementar las diferentes metodologías y se realizan algunas sugerencias para el futuro.

Conclusiones referentes a la tabla de mortalidad

Se considera apropiado el uso de la tabla de mortalidad específica para los funcionarios del BROU con la tasa de mortalidad, q_x , suavizada. Ya que no es lógico utilizar una q_x con caídas, la misma tiene que ser siempre creciente a medida que avanzan las edades. La utilización de esta tasa de mortalidad suavizada, permitirá encontrar mejores estimaciones para el resto de las funciones biométricas de la tabla.

Realizada la simulación mediante el método Monte Carlo, utilizando la tasa de mortalidad que publica el BCU para el cálculo de las reservas técnicas de las empresas de seguros y dada la muestra observada de los fallecimientos de los funcionarios del BROU, donde se observa una muy baja mortalidad, se considera muy importante que el banco cuente con una tasa de mortalidad propia. Ya que, si se utiliza la q_x del BCU, los fallecimientos de los funcionarios se estarán sobreestimando y de modo inverso (como muestra el cuadro comparativo 4.1), se estará subestimando el tiempo esperado de vida de los funcionarios. Con esto se concluye que la esperanza de vida de los funcionarios del BROU es mayor a la de la población uruguaya en general, tomando como referencia los datos del BCU. El cuadro comparativo 4.1, muestra que la esperanza de vida calculada para los funcionarios del BROU, es mayor en todos los tramos de edades.

Edad	$e_{x,BCU}$	$e_{x,BROU}$
20	53,25	59.17
25	48,56	54.17
30	43,86	49.17
35	39,15	44.17
40	34,49	39.17
45	29,90	34.17
50	25,54	29.74
55	21,42	25.38
60	17,56	20.94
65	14,12	16.62
70	10,87	12.94
75	7,79	8.61
80	4,71	4.71

Cuadro 4.1: Esperanza de vida e_x según BCU y para el BROU respectivamente

Dado que las tablas de mortalidad de período, constituyen un modelo teórico a partir de probabilidades de muerte observadas en la población real objeto de estudio en un momento dado y que representan lo que le habría

sucedido a una cohorte hipotética, que hubiera experimentado a lo largo de su vida las condiciones de mortalidad vigentes durante un periodo de tiempo en particular, se recomienda la actualización de la tabla de forma periódica.

Conclusiones respecto a la proyección utilizando el método de Lee-Carter

De los modelos obtenidos mediante la aplicación del método de Lee-Carter, se considera que el mejor modelo es el que se obtiene de la estimación de los parámetros mediante máxima verosimilitud y que se implementa en R con la librería *ilc*, dado que este modelo contempla la heterocedasticidad de los residuos observada en el grupo de edades 75-79 años y que permite incluir en el análisis a todos los grupo de edades, con un mejor ajuste a los datos y una menor devianza.

Según los autores del método y varios investigadores que lo han testeado (los cuales se encuentran en la bibliografía), se sugiere que las series de tiempo a utilizar no sean menores a los 20 años. En este trabajo, se contó con una serie de 18 años, 1995 a 2013, pero dado que se considera que los datos de los primeros años no son fiables y que los intervalos de confianza son muy amplios, no se puede tener mucha certeza que las estimaciones sean las correctas. De todos modos, permiten observar y tener una idea de cómo ha sido el comportamiento de la mortalidad a lo largo de los años.

Limitaciones

Las limitaciones que se tuvieron en la investigación fueron referentes a la cantidad y calidad de los datos.

- El bajo número de datos de fallecimientos de mujeres, únicamente 47 datos, hizo que hubiera que dejar fuera del análisis a la población femenina.
- Dado que durante el período de análisis 1995-2013, no hubieron concursos ni ingresos de personal joven hasta el año 2007, la cantidad de funcionarios menores a 36 años es muy pequeña o es 0. El mismo problema se tuvo con las edades avanzadas, donde la falta de datos es debido a la no digitalización de los mismos. Además, la muy baja cantidad de defunciones registradas en todas las edades llevó a que hubiera que agrupar los datos en intervalos de edad de 5 años (20-24, 25-29, ...), no pudiéndose realizar el análisis para edades simples.
- La calidad de los datos ha sido la mayor limitación para el análisis de los mismos. Se puede observar en el cuadro 3.1, que en los primeros años, desde 1995 hasta el 1997, no se cuenta con datos de funcionarios mayores a 70 años, los cuales se comienzan a contabilizar gradualmente a partir del año 1998. A partir del año 2003, se empiezan a contabilizar funcionarios mayores a 75 años y recién a partir del 2008 se cuenta con datos de funcionarios mayores a 85 años. Esto estaría reflejando que en el año 1995, por ejemplo, los funcionarios bancarios no superaban los 69 años de edad. Parece más lógico pensar que hay muchos funcionarios

que están quedando fuera del estudio. Este problema genera que se esté asumiendo una realidad que no es la verdadera y por lo tanto las conclusiones a las que se llegan puedan estar muy erradas.

Sugerencias

Debido a que las mujeres quedaron fuera del análisis por la baja cantidad de datos disponibles, se sugiere para el cálculo de la esperanza de vida de las mujeres, tomar la brecha que existe entre hombres y mujeres de la población uruguaya y en base a ella, calcular, con la esperanza de vida específica de los funcionarios hombres del BROU, la esperanza de vida de las funcionarias mujeres.

Dadas las limitaciones planteadas y la importancia que tienen los sistemas de información hoy en día para realizar análisis y tomar las mejores decisiones, se sugiere que se le dé mayor importancia a la recolección y calidad de los datos.

Para futuros estudios respecto a la proyección de la mortalidad y dado que el modelo de Lee-Carter tiene mucha aceptación entre los investigadores, se sugiere volver a utilizar el mismo, y probar la librería *gnm* de R para modelos no lineales generalizados. Se sugiere la lectura del estudio de Ochoa Molina, 2015, que compara las estimaciones del modelo de Lee-Carter mediante el uso de las librerías *demography*, *ilc* y *gnm*. También se sugiere: (Renshaw and Haberman, 2003), (Renshaw and Haberman, 2006) y (D'Amato et al., 2011) que refieren a técnicas de suavizado del modelo de Lee-Carter.

Bibliografía

Aguilar Fernandez, E. (2013). Estimación y proyección de la mortalidad para Costa Rica con la aplicación del método Lee-Carter con dos variantes. *Centro Centroamericano de Población*.

Alho, J. (2005). *Statistical demography and forecasting*. Springer, New York.

Álvarez, R. y Massa, F. (2015). Taller montecarlo. Notas de curso, Licenciatura de Estadística Facultad de Ciencias Económicas y Administración, Universidad de la República.

Andreozzi, L. y Blaconá, M. (2011). Estimación y pronóstico de las tasas de mortalidad y la esperanza de vida en la república argentina. In *Decimosextas Jornadas Investigaciones en la Facultad de Ciencias Económicas y Estadística*.

Belliard, M. y Williams, I. (2013). Proyección estocástica de la mortalidad. una aplicación de lee-carter en la argentina. *Revista Latinoamericana de Población*.

Butt, Z., Haberman, S., y Shang, H. L. (2014). *ilc: Lee-Carter Mortality Models using Iterative Fitting Algorithms*. R package version 1.0.

Bibliografía

- Camacho, L. (2009). Estimación de las tasas de mortalidad futuras para su aplicación en las proyecciones financieras del régimen previsional. *BPS*.
- Camacho, L. (2010). Análisis del equilibrio financiero individual de un sistema de prestación definida computando mejoras futuras en las tasas de mortalidad. *BPS*.
- Caselli, G., Vallin, J., y Wunsch, G. (2006). *Demography: Analysis and Synthesis*. Elsevier.
- Cox, P. R. (2008). *Demography*. Cambridge University Press.
- Dassattia, C. y Natalia Mariño, N. (2014). Maduración del sistema previsional: proyecciones y agenda. *SSF-BCU*.
- Débon Aucejo, A., Montes Suay, F., and Sala Garrido, R. (2008). Tablas de mortalidad dinámicas para España. una aplicación a la hipoteca inversa. Reporte técnico, Fundación ICO.
- D'Amato, V., Piscopo, G., and Russolillo, M. (2011). The mortality of the Italian population: Smoothing techniques on the Lee-Carter model. *The Annals of Applied Statistics*, 5(2A):705–724.
- Fox, J. y Weisberg, S. (2011). *An R Companion to Applied Regression*. Sage.
- García Guerrero, V. M. y Ordorica Mellado, M. (2012). Proyección estocástica de la mortalidad mexicana por medio del método de Lee-Carter. *Estudios Demográficos y Urbanos*.
- Giroi, F. y King, G. (2007). Understanding the Lee-Carter mortality forecasting method. Reporte técnico.

-
- Grolemund, G. y Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*.
- Hinde, A. (2014). *Demographic Methods*. Routledge.
- Hyndman, R. y Khandakar, Y. (2008). *Automatic time series forecasting: the forecast package for R* *Journal of Statistical Software*.
- INDEC, editor (2013). *Estimaciones y Proyecciones de población 2010-2040. Total del país*. INDEC.
- INE, editor (2004). *Chile. Tablas Abreviadas de Mortalidad por Sexo. Total País y Regiones. 2001-2002*.
- INE, editor (2015). *Tablas de mortalidad Metodología. INE España*.
- Keyfitz, N. (2005). *Applied mathematical demography*. Springer, New York, NY.
- Kintner, H. J. (2004). *The methods and materials of demography*, chapter Chapter 13 - The Life Table, pages 301–324. Elsevier Academic Press.
- Lazo, A. M. (2010). Estimación de las tasas de mortalidad específicas para los jubilados por vejez e invalidez del régimen previsional contributivo uruguayo. *BPS*.
- Lee, R. D. y Carter, L. R. (1992). Modeling and forecasting u.s. mortality. *Journal of American Statistical Association*.
- Ochoa Molina, C. A. (2015). El modelo lee-carter para estimar y pronosticar

Bibliografía

- mortalidad: Una aplicación para Colombia. Master's thesis, Universidad Nacional de Colombia.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Renshaw, A. y Haberman, S. (2003). Lee–Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33(2):255–272.
- Renshaw, A. y Haberman, S. (2006). A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38(3):556–570.
- Trapletti, A. y Hornik, K. (2015). *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-34.
- Wilmoth, J. R. (1993). Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. Reporte técnico, Department of Demography University of California.
- with contributions from Heather Booth, R. J. H., Tickle, L., y Maindonald, J. (2014). *demography: Forecasting mortality, fertility, migration and population data*. R package version 1.18.

Lista de Abreviaturas

AFAP Administradora de Fondos de Ahorro Previsional

ARIMA Autoregressive Integrated Moving Average (Autorregresivo Integrado de Media Móvil)

BCU Banco Central del Uruguay

BPS Banco de Previsión Social

BROU Banco de la República Oriental del Uruguay

CJPB Caja de Jubilaciones y Pensiones Bancarias

CJPPU Caja de Jubilaciones y Pensiones de Profesionales Universitarios

CN Caja Notarial

DNASSP Dirección Nacional de Asistencia y Seguridad Social Policial

INE Instituto Nacional de Estadística

MV Máxima Verosimilitud

NIC Norma Internacional de Contabilidad

Bibliografía

PIT-CNT Plenario Intersindical de Trabajadores - Convención Nacional de Trabajadores

SRPFAA Servicio de Retiro y Pensiones de las Fuerzas Armadas

SSF Super Intendencia del Sistema Financiero

SVD Singular Value Decomposition (Descomposición de Valores Singulares)

Apéndice A

Resultados

A.1. Datos BROU

Apéndice A. Resultados

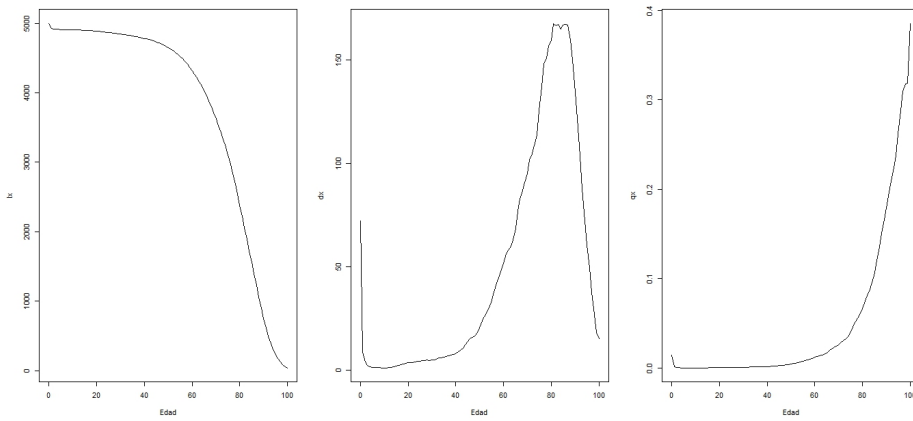
Edades	ex1	ex2	ex3	ex4	ex5
20-24	59,17	59,17	59,17	59,17	59,31
25-29	54,17	54,17	54,17	54,17	54,31
30-34	49,17	49,17	49,17	49,17	49,31
35-39	44,17	44,17	44,17	44,17	44,31
40-44	39,17	39,17	39,17	39,17	39,31
45-49	34,17	34,17	34,17	34,17	34,18
50-54	29,74	29,74	29,74	29,74	29,73
55-59	25,38	25,38	25,38	25,38	25,36
60-64	20,94	20,94	20,94	20,94	20,93
65-69	16,62	16,61	16,62	16,62	16,62
70-74	12,94	12,94	12,95	12,95	12,93
75-79	8,61	8,61	8,62	8,62	8,60
80 y más	4,71	4,71	4,73	4,73	4,71

Cuadro A.3: Esperanzas de vida e_x por grupos de edades calculadas a partir de los distintos métodos de estimación de q_x : Lineal, Exponencial, Reed-Merrel, Greville y Keyfitz

Edades	2011	2012	2013
20-24	58,76	57,44	59,17
25-29	53,76	52,44	54,17
30-34	50,39	47,44	49,17
35-39	45,39	42,44	44,17
40-44	40,39	37,44	39,17
45-49	35,39	32,44	34,17
50-54	30,39	27,86	29,74
55-59	25,78	23,19	25,38
60-64	21,03	19,36	20,94
65-69	17,14	16,53	16,62
70-74	13,48	12,87	12,94
75-79	9,39	8,64	8,61
80 y más	4,73	4,64	4,71

Cuadro A.4: Esperanzas de vida e_x por grupos de edades para los años 2011 a 2013

A.2. Datos INE

Figura A.1: Funciones l_x , d_x y q_x construidas a partir de la m_x del INE

A.3. Datos BCU

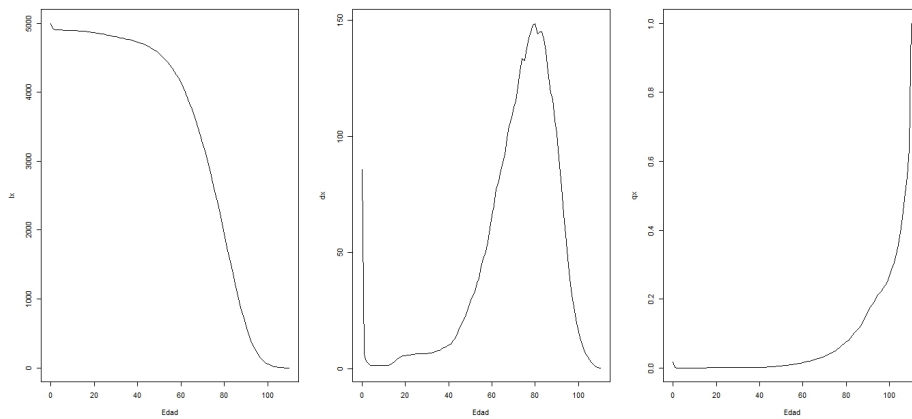


Figura A.2: Funciones l_x , d_x y q_x construidas a partir de la m_x del BCU

A.4. Tabla empírica simulada

Apéndice A. Resultados

Edades	2007	2008	2009	2010	2011	2012	2013
20-24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
25-29	0,0000	0,0000	0,0078	0,0069	0,0062	0,0071	0,0083
30-34	0,0000	0,0000	0,0000	0,0000	0,0000	0,0089	0,0072
35-39	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
40-44	0,0128	0,0129	0,0107	0,0133	0,0138	0,0128	0,0000
45-49	0,0222	0,0228	0,0233	0,0230	0,0229	0,0223	0,0216
50-54	0,0360	0,0363	0,0358	0,0371	0,0373	0,0361	0,0369
55-59	0,0560	0,0570	0,0588	0,0590	0,0586	0,0589	0,0581
60-64	0,0929	0,0929	0,0909	0,0925	0,0891	0,0924	0,0935
65-69	0,1316	0,1315	0,1313	0,1317	0,1312	0,1310	0,1297
70-74	0,1878	0,1866	0,1883	0,1867	0,1879	0,1877	0,1888
75-79	0,2588	0,2643	0,2679	0,2679	0,2657	0,2677	0,2661
80 y más	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Cuadro A.8: Tasa central de mortalidad simulada

A.5. Lee-Carter

Edades	p.valor.JB	p.valor.LB
25-29	0,362947	0,485306
30-34	0,000039	0,941405
35-39	0,123407	0,253067
40-44	0,643227	0,121113
45-49	0,637939	0,716033
50-54	0,498650	0,559495
55-59	0,442247	0,653043
60-64	0,704926	0,193825
65-69	0,727419	0,139674
70-74	0,519012	0,465270
75-79	0,642326	0,967369
80 y más	0,000000	0,991582

Cuadro A.9: p-valores de los tests Jarque-Bera y Ljung-Box, modelo Lee-Carter con librería *demography* para todos los grupos de edades de 25-29 a 80 años y más

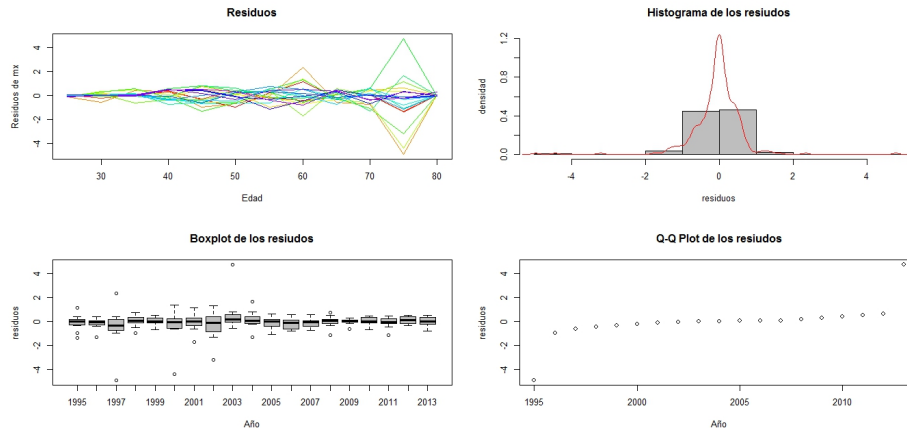


Figura A.3: Residuos del modelo de Lee-Carter, estimado mediante librería *demography* para todos los grupos de edades de 25-29 a 80 años y más

Edades	p.valor.JB	p.valor.LB
35-39	0.190251	0.045694
40-44	0.631279	0.157597
45-49	0.635008	0.471297
50-54	0.463862	0.292878
55-59	0.447760	0.762811
60-64	0.615717	0.073171
65-69	0.735811	0.088653
70-74	0.514212	0.202337
75-79	0.645202	0.736504

Cuadro A.10: p-valores de los tests Jarque-Bera y Ljung-Box, modelo Lee-Carter con librería *demography* para los grupos de edades de 35-39 a 75-79 años

Edades	p.valor.JB	p.valor.LB
40-44	0,826405	0,237833
45-49	0,685806	0,509590
50-54	0,579058	0,629665
55-59	0,472752	0,738560
60-64	0,615752	0,077422
65-69	0,624468	0,608664
70-74	0,719567	0,309001
75-79	0,103514	0,257822

Cuadro A.11: p-valores de los tests Jarque-Bera y Ljung-Box, modelo Lee-Carter con librería *demography* para los grupos de edades de 40-44 a 75-79 años

Apéndice A. Resultados

Edades	ax	bx
40-44	-6,006158	-0,04368864
45-49	-5,761261	0,13147096
50-54	-5,370549	0,12236155
55-59	-4,866617	0,02222135
60-64	-4,543372	-0,1521182
65-69	-4,16739	0,08429668
70-74	-3,578512	0,15259773
75-79	-1,832452	0,68285858

Cuadro A.12: Estimaciones de a_x y b_x del modelo de Lee-Carter con librería *demography* para los grupos de edades 40-44 a 75-79 años

Año	kt
1995	4,2443697
1996	5,085095
1997	10,2558827
1998	0,8585714
1999	1,3726559
2000	8,6255637
2001	2,8546784
2002	7,0018136
2003	2,2303087
2004	-1,8488875
2005	-0,1883998
2006	-0,4276358
2007	-3,2970649
2008	-1,8152536
2009	-2,9123175
2010	-2,5387839
2011	-3,7263861
2012	-2,1722307
2013	-2,6883401

Cuadro A.13: Tendencia de la mortalidad estimada k_t del modelo de Lee-Carter con librería *demography* para los grupos de edades 40-44 a 75-79 años

Año	kt.f.Point.Forecast	kt.f.Lo.90	kt.f.Hi.90
2014	-0,385151	-7,176347	6,406046
2015	-0,770301	-10,624005	9,083402
2016	-1,155452	-13,521751	11,210848
2017	-1,540602	-16,156006	13,074802
2018	-1,925753	-18,633520	14,782015

Cuadro A.14: Tendencia de la mortalidad k_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería *demography* para los grupos de edades 40-44 a 75-79 años e intervalos de confianza

Edad	2014	2015	2016	2017	2018
40-44	0,0028	0,0029	0,0029	0,0030	0,0030
45-49	0,0021	0,0020	0,0019	0,0018	0,0017
50-54	0,0032	0,0030	0,0029	0,0028	0,0026
55-59	0,0072	0,0071	0,0071	0,0070	0,0069
60-64	0,0170	0,0180	0,0191	0,0202	0,0215
65-69	0,0120	0,0116	0,0112	0,0108	0,0105
70-74	0,0175	0,0165	0,0155	0,0146	0,0138
75-79	0,0196	0,0151	0,0116	0,0089	0,0069

Cuadro A.15: Tasa de mortalidad m_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería *demography* para los grupos de edades 40-44 a 75-79 años

Edades	ax	bx
40-44	-6,057041	-0,009859
45-49	-5,735862	0,062531
50-54	-5,325440	0,073081
55-59	-4,765539	-0,012119
60-64	-4,493713	-0,151877
65-69	-4,148371	0,058149
70-74	-3,806705	0,055413
75-79	-0,946204	0,924681

Cuadro A.16: Estimaciones de a_x y b_x del modelo de Lee-Carter con librería *ilc* para los grupos de edades 40-44 años a 75-79 años

Año	kt
1995	-2,7109515
1996	10,0077256
1997	-3,8145095
1998	6,6408311
1999	5,7434333
2000	1,9983772
2001	8,9180306
2002	-0,3984105
2003	0,6296841
2004	-0,3611682
2005	-1,6058123
2006	-2,3845605
2007	-3,7776793
2008	-3,7477539
2009	-3,3323571
2010	-2,4920978
2011	-3,8667571
2012	-2,8019436
2013	-2,6440805

Cuadro A.17: Tendencia de la mortalidad estimada k_t del modelo de Lee-Carter con librería *ilc* para los grupos de edades 40-44 años a 75-79 años

Apéndice A. Resultados

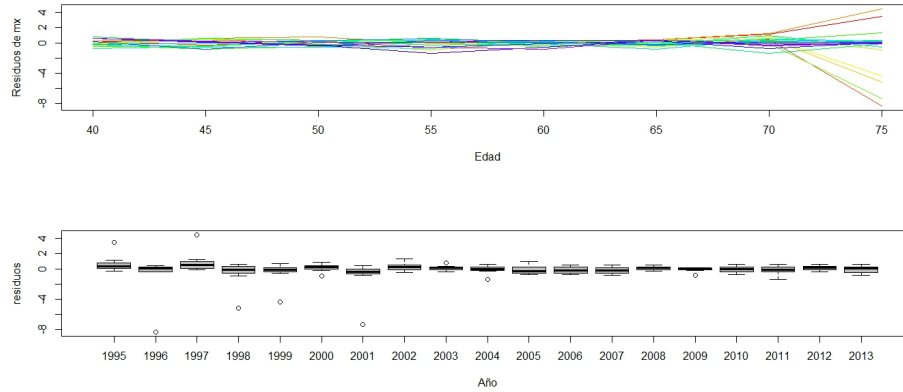


Figura A.4: Residuos del modelo Lee-Carter estimado con la librería *ilc* para los grupos de edades 40-44 años a 74-79 años

Año	kt.f.Point.Forecast	kt.f.Lo.90	kt.f.Hi.90
2014	0,003715056	-10,21592	10,22335
2015	0,007430113	-14,82077	14,83563
2016	0,011145169	-18,59810	18,62039
2017	0,014860225	-21,97892	22,00864
2018	0,018575282	-25,12386	25,16101

Cuadro A.18: Tendencia de la mortalidad k_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería *ilc* para los grupos de edades 40-44 años a 75-79 años e intervalos de confianza

Edades	2014	2015	2016	2017	2018
40-44	0,002403069	0,002402981	0,002402893	0,002402805	0,002402717
45-49	0,002736796	0,002737431	0,002738067	0,002738704	0,00273934
50-54	0,004012259	0,004013348	0,004014438	0,004015528	0,004016619
55-59	0,008795279	0,008794883	0,008794487	0,008794091	0,008793695
60-64	0,016694079	0,016684663	0,016675251	0,016665845	0,016656444
65-69	0,013542716	0,013545642	0,013548569	0,013551496	0,013554424
70-74	0,019196733	0,019200685	0,019204638	0,019208592	0,019212547
75-79	0,033786268	0,033902532	0,034019195	0,034136261	0,034253729

Cuadro A.19: Tasa de mortalidad m_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería *ilc* para los grupos de edades 40-44 años a 75-79 años

Edades	a_x	b_x
25-29	-4,926183	0,009189
30-34	-5,774407	0,005476
35-39	-6,020333	-0,023385
40-44	-6,057041	-0,009338
45-49	-5,735865	0,058013
50-54	-5,325026	0,068274
55-59	-4,765709	-0,010976
60-64	-4,494984	-0,142107
65-69	-4,148566	0,054050
70-74	-3,807511	0,051330
75-79	-0,919648	0,869871
80 y más	-3,440331	0,069604

Cuadro A.20: Estimaciones de a_x y b_x del modelo de Lee-Carter con librería *ilc* para todos los grupos de edades desde los 25-29 años a 80 y más

Año	k_t
1995	-2,717294
1996	10,594545
1997	-4,068664
1998	7,099628
1999	6,083267
2000	2,128458
2001	9,731734
2002	-0,467035
2003	0,640202
2004	-0,413681
2005	-1,740444
2006	-2,564801
2007	-4,047110
2008	-4,014147
2009	-3,572493
2010	-2,680086
2011	-4,143837
2012	-3,005421
2013	-2,842821

Cuadro A.21: Tendencia de la mortalidad estimada k_t del modelo de Lee-Carter con librería *ilc* para todos los grupos de edades desde los 25-29 años a 80 y más

Año	kt.f.Point.Forecast	kt.f.Lo.90	kt.f.Hi.90
2014	-0,006973723	-10,90416	10,89021
2015	-0,013947445	-15,82525	15,79735
2016	-0,020921168	-19,86395	19,82210
2017	-0,027894891	-23,47984	23,42405
2018	-0,034868613	-26,84423	26,77449

Cuadro A.22: Tendencia de la mortalidad k_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería *ilc* para todos los grupos de edades desde los 25-29 años a 80 y más e intervalos de confianza

Apéndice A. Resultados

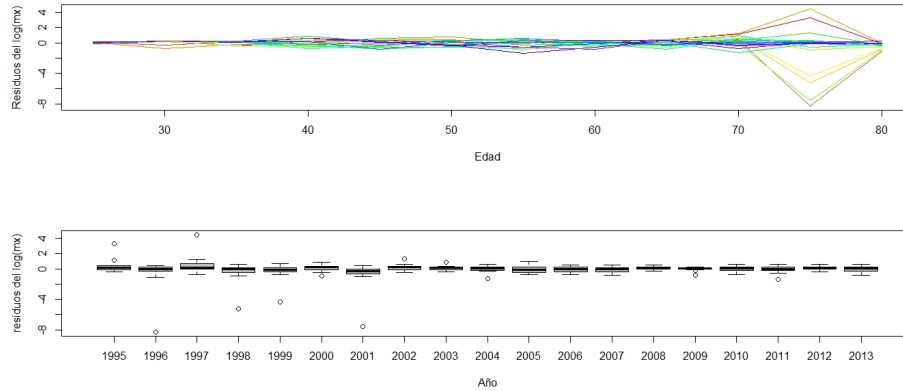


Figura A.5: Residuos del modelo Lee-Carter estimado con la librería *ilc* para todos los grupos de edades desde los 25-29 años a 80 y más

Edades	2014	2015	2016	2017	2018
25-29	0,007067	0,007066	0,007066	0,007065	0,007065
30-34	0,003058	0,003058	0,003058	0,003058	0,003057
35-39	0,002596	0,002597	0,002597	0,002598	0,002598
40-44	0,002404	0,002405	0,002405	0,002405	0,002405
45-49	0,002736	0,002735	0,002734	0,002733	0,002732
50-54	0,004007	0,004006	0,004004	0,004002	0,004000
55-59	0,008787	0,008788	0,008789	0,008789	0,008790
60-64	0,016739	0,016756	0,016772	0,016789	0,016806
65-69	0,013533	0,013528	0,013523	0,013518	0,013513
70-74	0,019182	0,019175	0,019168	0,019161	0,019154
75-79	0,033420	0,033218	0,033017	0,032817	0,032619
80 y más	0,026287	0,026274	0,026261	0,026249	0,026236

Cuadro A.23: Tasa de mortalidad m_t proyectada para los años 2014 a 2018 del modelo de Lee-Carter con librería *ilc* para todos los grupos de edades desde los 25-29 años a 80 y más

Apéndice B

Scripts en R

Se presentan, a continuación, los scripts realizados en R-project que contienen los códigos utilizados para el análisis e implementación de la metodología desarrollada en el presente trabajo.

```
#####  
##### ESTADÍSTICA DESCRIPTIVA #####  
#####  
  
library(pyramid)  
  
activos = read.csv("Datos Activos al 31.07.2014.csv", head=TRUE, sep=";")  
jubilados = read.csv("egresosJUBILADOS.csv", head=TRUE, sep=";")  
fallec = read.csv("egresosFALLEC.csv", head=TRUE, sep=";")  
sefueron = read.csv("egresosSEFUERON.csv", head=TRUE, sep=";")  
  
### ACTIVOS ###  
summary(activos)  
  
par(mfrow=c(1,2))  
  
hist(activos$EDAD, xlab="Edad", ylab="Frecuencia", main="Distribución por edades")  
  
boxplot(activos$EDAD, main="Diagrama de caja")
```

Apéndice B. Scripts en R

```
plot(activos$SEX0, ylim=c(0,2500),col=c("indianred1","lightseagreen"))

## Pirámide poblacional ##
m = min(activos$EDAD)
M = max(activos$EDAD)
EDAD = seq(m,M,5)

MUJ=activos[activos$SEX0=="FEMENINO",]
HOM=activos[activos$SEX0=="MASCULINO",]

#número de mujeres y hombres a la edad x
muj = array(0, length(EDAD))
for (i in seq(m,M,5)){
muj[which(EDAD==i)] = sum(MUJ$EDAD>=i & MUJ$EDAD<i+5)
}

hom = array(0, length(EDAD))
for (i in seq(m,M,5)){
hom[which(EDAD==i)] = sum(HOM$EDAD>=i & HOM$EDAD<i+5)
}

plot(EDAD, hom, type="l", col="lightseagreen", xlab="Edad",
ylab="# funcionarios activos",
main="Distribución de funcionarios activos por edad")
lines(EDAD, muj, type="l", col="indianred1")
legend("topleft", c("hombres", "mujeres"), col=c("lightseagreen","indianred1"),
lty=1, bty="n")

#porcentaje de mujeres y hombres a la edad x
mujeres = array(0, length(EDAD))
for (i in seq(m,M,5)){
mujeres[which(EDAD==i)] = (sum(MUJ$EDAD>=i & MUJ$EDAD<i+5)/dim(MUJ)[1])*100
}

hombres = array(0, length(EDAD))
for (i in seq(m,M,5)){
```

```

hombres[which(EDAD==i)] = (sum(HOM$EDAD>=i & HOM$EDAD<i+5)/dim(HOM)[1])*100
}

tabla_sexo=data.frame(hombres, mujeres)
dimnames(tabla_sexo)[[1]] = EDAD
tabla_sexo

pyramid(tabla_sexo, Clab="Edad", Llab="Hombres", Rlab="Mujeres",
sub="Frecuencias relativas", Cgap=0.2, Csize=0.5, Lcol="lightseagreen",
Rcol="indianred1", main="Pirámide poblacional")

### JUBILADOS ###
summary(jubilados)

par(mfrow=c(1,2))

hist(jubilados$EDAD, xlab="Edad", ylab="Frecuencia", main="")

boxplot(jubilados$EDAD, main="Diagrama de caja")

plot(jubilados$SEXO, ylim=c(0,2000),col=c("indianred1","lightseagreen"))

## Pirámide poblacional de jubilados ##
m = min(jubilados$EDAD)
M = max(jubilados$EDAD)
EDAD = seq(45,80,5)
EDAD[length(EDAD)]="más de 80"

MUJ=jubilados[jubilados$SEXO=="FEMENINO",]
HOM=jubilados[jubilados$SEXO=="MASCULINO",]

#número de mujeres y hombres a la edad x
muj = array(0, length(EDAD))
muj[length(EDAD)]=sum(MUJ$EDAD>=80)
for (i in seq(45,75,5)){
muj[which(EDAD==i)] = sum(MUJ$EDAD>=i & MUJ$EDAD<i+5)

```

Apéndice B. Scripts en R

```
}

hom = array(0, length(EDAD))
hom[length(EDAD)] = sum(HOM$EDAD >= 80)
for (i in seq(45, 75, 5)) {
  hom[which(EDAD == i)] = sum(HOM$EDAD >= i & HOM$EDAD < i + 5)
}

plot(seq(45, 80, 5), hom, type="l", col="lightseagreen", xlab="Edad",
      ylab="# jubilados", main="")
lines(seq(45, 80, 5), muj, type="l", col="indianred1")
legend("topleft", c("hombres", "mujeres"),
       col=c("lightseagreen", "indianred1"), lty=1, bty="n")

#porcentaje de mujeres y hombres a la edad x
mujeres = array(0, length(EDAD))
mujeres[length(EDAD)] = (sum(MUJ$EDAD >= 80) / dim(jubilados)[1]) * 100
for (i in seq(45, 75, 5)) {
  mujeres[which(EDAD == i)] = (sum(MUJ$EDAD >= i & MUJ$EDAD < i + 5) / dim(MUJ)[1]) * 100
}

hombres = array(0, length(EDAD))
hombres[length(EDAD)] = (sum(HOM$EDAD >= 80) / dim(jubilados)[1]) * 100
for (i in seq(45, 75, 5)) {
  hombres[which(EDAD == i)] = (sum(HOM$EDAD >= i & HOM$EDAD < i + 5) / dim(HOM)[1]) * 100
}

tabla_sexo = data.frame(hombres, mujeres)
dimnames(tabla_sexo)[[1]] = EDAD
tabla_sexo

pyramid(tabla_sexo, Clab="Edad", Llab="Hombres", Rlab="Mujeres",
         sub="Frecuencias relativas", Cgap=0.2, Csize=0.5, Lcol="lightseagreen",
         Rcol="indianred1", main="Pirámide poblacional")

### FALLECIDOS ###
summary(fallec)
```

```

par(mfrow=c(1,2))

hist(fallec$EDAD_FALLEC, xlab="Edad", ylab="Frecuencia",
main="Distribución por edades", sub="Fallecidos")

boxplot(fallec$EDAD_FALLEC, main="Diagrama de caja")

plot(fallec$SEXO, ylim=c(0,700),col=c("indianred1","lightseagreen"))

## Pirámide poblacional de fallecidos ##
m = min(fallec$EDAD_FALLEC)
M = max(fallec$EDAD_FALLEC)
EDAD = seq(20,80,5)
EDAD[length(EDAD)]="más de 80"

MUJ=fallec[fallec$SEXO=="FEMENINO",]
HOM=fallec[fallec$SEXO=="MASCULINO",]

#número de mujeres y hombres a la edad x
muj= array(0, length(EDAD))
muj[length(EDAD)]=sum(MUJ$EDAD_FALLEC>=80)
for (i in seq(20,75,5)){
muj[which(EDAD==i)] = sum(MUJ$EDAD_FALLEC>=i & MUJ$EDAD_FALLEC<i+5)
}

hom = array(0, length(EDAD))
hom[length(EDAD)]=sum(HOM$EDAD_FALLEC>=80)
for (i in seq(20,75,5)){
hom[which(EDAD==i)] = sum(HOM$EDAD_FALLEC>=i & HOM$EDAD_FALLEC<i+5)
}

plot(seq(20,80,5), hom, type="l", col="lightseagreen", xlab="Edad",
ylab="# fallecidos", main="Distribución de fallecidos")
lines(seq(20,80,5), muj, type="l", col="indianred1")
legend("topleft", c("hombres", "mujeres"), col=c("lightseagreen","indianred1"),
lty=1, bty="n")

```

Apéndice B. Scripts en R

```
#porcentaje de mujeres y hombres a la edad x
mujeres = array(0, length(EDAD))
mujeres[length(EDAD)]=(sum(MUJ$EDAD_FALLEC>=80)/dim(fallec)[1])*100
for (i in seq(20,75,5)){
mujeres[which(EDAD==i)] = (sum(MUJ$EDAD_FALLEC>=i &
MUJ$EDAD_FALLEC<i+5)/dim(MUJ)[1])*100
}

hombres = array(0, length(EDAD))
hombres[length(EDAD)]=(sum(HOM$EDAD_FALLEC>=80)/dim(fallec)[1])*100
for (i in seq(20,75,5)){
hombres[which(EDAD==i)] = (sum(HOM$EDAD_FALLEC>=i &
HOM$EDAD_FALLEC<i+5)/dim(HOM)[1])*100
}

tabla_sexo=data.frame(hombres, mujeres)
dimnames(tabla_sexo)[[1]] = EDAD
tabla_sexo

pyramid(tabla_sexo, Clab="Edad", Llab="Hombres", Rlab="Mujeres",
sub="Frecuencias relativas", Cgap=0.2, Csize=0.5, Lcol="lightseagreen",
Rcol="indianred1", main="Pirámide poblacional")

#####
## EXPUESTOS: ACTIVOS Y JUBILADOS ##

m = min(activos$EDAD)
M = max(activos$EDAD, jubilados$EDAD)
EDAD = seq(m,M,5)

MUJ1=activos[activos$SEXO=="FEMENINO",]
MUJ2=jubilados[jubilados$SEXO=="FEMENINO",]
MUJ=data.frame(EDAD=c(MUJ1$EDAD,MUJ2$EDAD))
HOM1=activos[activos$SEXO=="MASCULINO",]
HOM2=jubilados[jubilados$SEXO=="MASCULINO",]
HOM=data.frame(EDAD=c(HOM1$EDAD,HOM2$EDAD))
```

```

par(mfrow=c(1,2))

hist(c(MUJ$EDAD,HOM$EDAD), xlab="Edad", ylab="Frecuencia",
main="Distribución por edades", sub="Expuestos al riesgo")

boxplot(c(MUJ$EDAD,HOM$EDAD), main="Diagrama de caja")

#número de mujeres y hombres a la edad x
muje = array(0, length(EDAD))
for (i in seq(m,M,5)){
muj[which(EDAD==i)] = sum(MUJ$EDAD>=i & MUJ$EDAD<i+5)
}

hom = array(0, length(EDAD))
for (i in seq(m,M,5)){
hom[which(EDAD==i)] = sum(HOM$EDAD>=i & HOM$EDAD<i+5)
}

plot(EDAD, hom, type="l", col="lightseagreen", xlab="Edad", ylab="# expuestos",
main="Distribución de funcionarios activos por edad")
lines(EDAD, muj, type="l", col="indianred1")
legend("topleft", c("hombres", "mujeres"), col=c("lightseagreen","indianred1"),
lty=1, bty="n")

#porcentaje de mujeres y hombres a la edad x
mujeres = array(0, length(EDAD))
for (i in seq(m,M,5)){
mujeres[which(EDAD==i)] = (sum(MUJ$EDAD>=i & MUJ$EDAD<i+5)/dim(MUJ)[1])*100
}

hombres = array(0, length(EDAD))
for (i in seq(m,M,5)){
hombres[which(EDAD==i)] = (sum(HOM$EDAD>=i & HOM$EDAD<i+5)/dim(HOM)[1])*100
}

tabla_sexo=data.frame(hombres, mujeres)

```

Apéndice B. Scripts en R

```
dimnames(tabla_sexo)[[1]] = EDAD
tabla_sexo

pyramid(tabla_sexo, Clab="Edad", Llab="Hombres", Rlab="Mujeres",
sub="Frecuencias relativas", Cgap=0.2, Csize=0.5, Lcol="lightseagreen",
Rcol="indianred1", main="Pirámide poblacional")

#####
#EXPUESTOS HOMBRES
par(mfrow=c(1,2))
hist(HOM$EDAD,xlab="Edad", ylab="Frecuencia",ylim=c(0,900),
main="Distribución por edades de \n hombres expuestos al riesgo de muerte")
homb = array(0, length(EDAD))
for (i in seq(m,M,5)){
homb[which(EDAD==i)] = sum(HOM$EDAD>=i & HOM$EDAD<i+5)
}
lines(EDAD, homb, col="lightseagreen")
boxplot(HOM$EDAD, main="Diagrama de caja \n hombres expuestos")

#FALLECIDOS HOMBRES
HOM.F=fallec[fallec$SEXO=="MASCULINO",]
homb.f = array(0, length(EDAD))
for (i in seq(m,M,5)){
homb.f[which(EDAD==i)] = sum(HOM.F$EDAD_FALLEC>=i & HOM.F$EDAD_FALLEC<i+5)
}
hist(HOM.F$EDAD_FALLEC, xlab="Edad", ylab="Frecuencia",
main="Distribución por edades \n Fallecidos hombres")
lines(EDAD,homb.f,col="lightseagreen")
boxplot(fallec$EDAD_FALLEC, main="Diagrama de caja \n Fallecidos hombres")

#####
##### TABLAS DE MORTALIDAD #####
##### UTILIZANDO DISTINTOS METODOS DE ESTIMACION PARA qx #####
##### Y Mx DE LA POBLACION DEL BROU #####
#####

activos = read.csv("Datos Activos al 31.07.2014.csv", head=TRUE, sep=";")
jubilados = read.csv("egresosJUBILADOS.csv", head=TRUE, sep=";")
```

```

fallec = read.csv("egresosFALLEC.csv", head=TRUE, sep=";")

hactivos=activos[activos$SEXO=="MASCULINO",]
hjubilados=jubilados[jubilados$SEXO=="MASCULINO",]
hfallec=fallec[fallec$SEXO=="MASCULINO",]

library(lubridate)

hfallec$FECHAFALLECIM=as.Date(hfallec$FECHAFALLECIM,format="%d/%m/%Y")
hfallec$FECHANACIM=as.Date(hfallec$FECHANACIM,format="%d/%m/%Y")
hactivos$FECHANACIM=as.Date(hactivos$FECHANACIM,format="%d/%m/%Y")
hactivos$INGBCO=as.Date(hactivos$INGBCO,format="%d/%m/%Y")
hjubilados$FECHANACIM=as.Date(hjubilados$FECHANACIM,format="%d/%m/%Y")

#####
## EDADES SIMPLES
m=min(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
M=max(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
EDAD=seq(m,M)
AÑO=year(years(1995:2013))

#número de muertes a la edad x por año del 1995 al 2013
dx=array(0, c(length(m:M),length(1995:2013)))
for (j in 1995:2013){
  for (i in m:M){
    dx[i-m+1,j-1994]=sum(hfallec$EDAD_FALLEC==i & year(hfallec$FECHAFALLECIM)==j)
  }
  dimnames(dx)[[1]]=EDAD
  dimnames(dx)[[2]]=AÑO
}
dx

#número de vidas a la edad x por año del 1995 al 2013
Nx=array(0, c(length(m:M),length(1995:2013)))
for (j in 1995:2013){
  for (i in m:M){
    Nx[i-m+1,j-1994]=sum(j-year(hactivos$FECHANACIM)==i & year(hactivos$INGBCO)<=j)+
    sum(j-year(hjubilados$FECHANACIM)==i)+

```

Apéndice B. Scripts en R

```
sum(j-year(hfallec$FECHANACIM)==i & year(hfallec$FECHAFALLECIM)<j)
}
dimnames(Nx)[[1]]=EDAD
dimnames(Nx)[[2]]=AÑO
}
Nx

#####
## EDADES AGRUPADAS
m=min(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
M=max(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
EDAD=seq(m,80,5)
ANO=year(years(1995:2013))
EDADES=c("20-24", "25-29", "30-34", "35-39", "40-44", "45-49", "50-54", "55-59",
"60-64", "65-69", "70-74", "75-79", "80 y más")

#numero de vidas entre la edad x y la edad x+5 por año del 1995 al 2013
Nx=array(0, c(length(EDAD),length(1995:2013)))
for (j in 1995:2013){
Nx[length(EDAD),j-1994]=sum(j-year(hactivos$FECHANACIM)>=80 &
year(hactivos$INGBCO)<=j) +
sum(j-year(hjubilados$FECHANACIM)>=80)+
sum(j-year(hfallec$FECHANACIM)>=80 & year(hfallec$FECHAFALLECIM)<j)
for (i in seq(20,75,5)){
Nx[EDAD==i,j-1994]=sum(j-year(hactivos$FECHANACIM)>=i &
j-year(hactivos$FECHANACIM)<i+5 &
year(hactivos$INGBCO)<=j)+
sum(j-year(hjubilados$FECHANACIM)>=i &
j-year(hjubilados$FECHANACIM)<i+5)+
sum(j-year(hfallec$FECHANACIM)>=i & j-year(hfallec$FECHANACIM)<i+5 &
year(hfallec$FECHAFALLECIM)<j)
}
dimnames(Nx)[[1]]=EDADES
dimnames(Nx)[[2]]=ANO
}
Nx

#numero de muertes entre la edad x y la edad x+5 por año del 1995 al 2013
```

```

Dx=array(0, c(length(EDAD),length(1995:2013)))
for (j in 1995:2013){
Dx[length(EDAD),j-1994]=sum(hfallec$EDAD_FALLEC>=80 & year(hfallec$FECHAFALLECIM)==j)
for (i in seq(20,75,5)){
Dx[which(EDAD==i),j-1994]=sum(hfallec$EDAD_FALLEC>=i & hfallec$EDAD_FALLEC<i+5 &
year(hfallec$FECHAFALLECIM)==j)
}
dimnames(Dx)[[1]]=EADDES
dimnames(Dx)[[2]]=ANO
}
Dx

Mx=Dx/Nx
round(Mx,4)
dim(Mx)

##### UTILIZO SOLO LOS DATOS DE LAS MUERTES Y LAS PERSONAS VIVAS EN EL 2013
M=Mx[,19]
N=Nx[,19]
D=Dx[,19]

## MÉTODO LINEAL
qx1=array(0, length(EDAD))
for(i in 1:length(EDAD)){
qx1[i]=(2*5*M[i])/(2+5*M[i])
}

#####
## MÉTODO EXPONENCIAL SIMPLE
qx2=array(0, length(EDAD))
for(i in 1:length(EDAD)){
qx2[i]=1-exp(-5*M[i])
}

#####
## MÉTODO REED-MERREL
qx3=array(0, length(EDAD))
for(i in 1:length(EDAD)){

```

Apéndice B. Scripts en R

```
qx3[i]= 1-exp(-5*M[i]-0.008*(5^3)*(M[i]^2))
}

#####
## MÉTODO GREVILLE
qx4=array(0, length(EDAD))
for(i in 1:length(EDAD)){
qx4[i]= M[i]/(1/5+M[i]*(1/2+5/12*(M[i]-0.095)))
}

#####
## MÉTODO KEYFITZ
#Keyfitz y Caswell-2005-Applied mathematical demography-Capítulo 2:The Life Table
#Corrección para la estimación de la tasa de muerte
C=array(NaN, length(EDAD))
for (i in 2:(length(EDAD)-1)){
C[i]= (N[i-1]-N[i+1])*(M[i+1]-M[i-1])/(48*N[i])
}

#Probabilidad de vivir al menos un año más habiendo alcanzado la edad x
px5= exp(-5*(M+C))

#Probabilidad de morir durante el próximo año para una persona que ha
#alcanzado la edad x
qx5=1-px5

#####
tabla=data.frame(Nx=N,Dx=D,Mx=M,qx1,qx2,qx3,qx4,qx5)
dimnames(tabla)[[1]]=EDADES
round(tabla,6)
#El metodo lineal y el exponencial no tienen casi diferencia,
#explicado en Keyfitz y Caswell Cap.2Pg.38

plot(EDAD, M, type="l", ylab="qx", xlab="Edad",ylim=c(0,0.15),
lwd=1,main="Probabilidad de morir de los hombres")
lines(EDAD, jitter(qx1), col="firebrick1", lwd=1)
lines(EDAD, jitter(qx2), col="magenta", lwd=1)
lines(EDAD, jitter(qx3), col="mediumspringgreen", lwd=1)
```

```

lines(EDAD, jitter(qx4), col="royalblue3", lwd=1)
lines(EDAD, qx5, col="limegreen", lwd=1.7)
legend("topleft", c("Mx","Método Lineal", "Método Exponencial", "Reed-Merrel",
"Greville","Keyfitz"),
col=c("black","firebrick1", "magenta","mediumspringgreen","royalblue3","limegreen"),
lty=1, bty="n")

#### No hay casi diferencias entre los distintos metodos por lo que las Tablas
#### de vida van a dar casi igual
## con Keyfitz quedan afuera el primer y último grupo de edades y para el grupo de
## edades 40-44 da negativo lo cual es errado

#####
## TABLA DE MORTALIDAD usando tasa de muerte LINEAL ##
lx1=array(0, length(EDAD))
lx1[1]=5000
for(i in 2:length(EDAD)){
lx1[i]=lx1[i-1]-lx1[i-1]*qx1[i-1]
}

dx1=lx1*qx1

Lx1=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx1[i]=(lx1[i]+lx1[i+1])*5/2
}
Lx1[length(EDAD)]=dx1[length(EDAD)]/M[length(EDAD)]

Tx1=array(0, length(EDAD))
for(i in 1:length(EDAD)){
Tx1[i]= sum(Lx1[i:length(EDAD)])
}

ex1=Tx1/lx1

tabla1=data.frame(qx1,lx=lx1,dx=dx1,Lx=Lx1,Tx=Tx1,ex=ex1)
dimnames(tabla1)[[1]]=EADADES
round(tabla1,4)

```

Apéndice B. Scripts en R

```
par(mfrow=c(1,3))
plot(EDAD,lx1, ylab="lx", xlab="Edad", type="l")
plot(EDAD, dx1, ylab="dx", xlab="Edad", type="l")
plot(EDAD, qx1, ylab="qx", xlab="Edad", type="l")

## TABLA DE MORTALIDAD usando tasa de muerte EXPONENCIAL ##
lx2=array(0, length(EDAD))
lx2[1]=4000
for(i in 2:length(EDAD)){
  lx2[i]=lx2[i-1]-lx2[i-1]*qx2[i-1]
}

dx2=lx2*qx2

Lx2=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
  Lx2[i]=(lx2[i]+lx2[i+1])*5/2
}
Lx2[length(EDAD)]=dx2[length(EDAD)]/M[length(EDAD)]

Tx2=array(0, length(EDAD))
for(i in 1:length(EDAD)){
  Tx2[i]= sum(Lx2[i:length(EDAD)])
}

ex2=Tx2/lx2

tabla2=data.frame(qx2,lx2,dx2,Lx2,Tx2,ex2)
dimnames(tabla2)[[1]]=EADADES
round(tabla2,4)

## TABLA DE MORTALIDAD usando tasa de muerte REED-MERRELL ##
lx3=array(0, length(EDAD))
lx3[1]=5000
for(i in 2:length(EDAD)){
  lx3[i]=lx3[i-1]-lx3[i-1]*qx3[i-1]
```

```

}

dx3=lx3*qx3

Lx3=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx3[i]=(lx3[i]+lx3[i+1])*5/2
}
Lx3[length(EDAD)]=dx3[length(EDAD)]/M[length(EDAD)]

Tx3=array(0, length(EDAD))
for(i in 1:length(EDAD)){
Tx3[i]= sum(Lx3[i:length(EDAD)])
}

ex3=Tx3/lx3

tabla3=data.frame(qx3,lx3,dx3,ex3)
dimnames(tabla3)[[1]]=EADDES
round(tabla3,4)

## TABLA DE MORTALIDAD usando GREVILLE ##
lx4=array(0, length(EDAD))
lx4[1]=5000
for(i in 2:length(EDAD)){
lx4[i]=lx4[i-1]-lx4[i-1]*qx4[i-1]
}

dx4=lx4*qx4

Lx4=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx4[i]=(lx4[i]+lx4[i+1])*5/2
}
Lx4[length(EDAD)]=dx4[length(EDAD)]/M[length(EDAD)]

Tx4=array(0, length(EDAD))

```

Apéndice B. Scripts en R

```
for(i in 1:length(EDAD)){
Tx4[i]= sum(Lx4[i:length(EDAD)])
}

ex4=Tx4/lx4

tabla4=data.frame(qx4,lx4,dx4,ex4)
dimnames(tabla4)[[1]]=EADADES
round(tabla4,4)

## TABLA DE MORTALIDAD usando KEYFITZ ##
#como qx5 no tiene valores en el primer lugar del vector, ni en el último
#lo calculo por el metodo exponencial sin corrección
qx5[1]=1-exp(-5*M[1])
qx5[length(EDAD)]=1-exp(-5*M[length(EDAD)])

lx5=array(0, length(EDAD))
lx5[1]=5000
for(i in 2:length(EDAD)){
lx5[i]=lx5[i-1]-lx5[i-1]*qx5[i-1]
}

dx5=lx5*qx5

Lx5=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx5[i]=(lx5[i]+lx5[i+1])*5/2
}
Lx5[length(EDAD)]=dx5[length(EDAD)]/M[length(EDAD)]

Tx5=array(0, length(EDAD))
for(i in 1:length(EDAD)){
Tx5[i]= sum(Lx5[i:length(EDAD)])
}

ex5=Tx5/lx5
```

```

tabla5=data.frame(qx5,lx5,dx5,ex5)
dimnames(tabla5)[[1]]=EADADES
round(tabla5,4)

```

```

comp=data.frame(ex1,ex2,ex3,ex4,ex5)
dimnames(comp)[[1]]=EADADES
round(comp,4)

```

```

#como se observo en la grafica de los diferentes metodos,
#no hay diferencias en la estimacion de ex

```

```

#####

```

```

###MATRIZ DE ESPERANZA DE VIDA PARA TODOS LOS AÑOS Y EDADES

```

```

##creo una matriz que contiene las tasas de muerte por año y tramos de edades

```

```

Q=matrix(0, length(EDAD), length(ANO))

```

```

for (j in 1:length(ANO)){

```

```

M=Mx[,j]

```

```

qx1=array(0, length(EDAD))

```

```

for(i in 1:length(EDAD)){

```

```

qx1[i]=(2*5*M[i])/(2+5*M[i])

```

```

}

```

```

Q[,j]=qx1

```

```

dimnames(Q)[[1]]=EADADES

```

```

dimnames(Q)[[2]]=ANO

```

```

}

```

```

Q

```

```

##creo una matriz que contiene la esperanza de vida por año para las diferentes edades

```

```

E=matrix(0, length(EDAD), length(ANO))

```

```

for (j in 1:length(ANO)){

```

```

M=Mx[,j]

```

```

qx=Q[,j]

```

```

lx=array(0, length(EDAD))

```

```

lx[1]=5000

```

```

for(i in 2:length(EDAD)){

```

Apéndice B. Scripts en R

```
lx[i]=lx[i-1]-lx[i-1]*qx[i-1]
}

dx=lx*qx

Lx=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx[i]=(lx[i]+lx[i+1])*5/2
}
Lx[length(EDAD)]=dx[length(EDAD)]/M[length(EDAD)]

Tx=array(0, length(EDAD))
for(i in 1:length(EDAD)){
Tx[i]= sum(Lx[i:length(EDAD)])
}

ex=Tx/lx

E[,j]=ex
dimnames(E)[[1]]=EADDES
dimnames(E)[[2]]=ANO
}
E

#los NaN es porque para calcular Lx en el último tramos de edades es
#Lx=dx/Mx y dx y Mx valen 0
#el año 2014 no está completo por lo que no se toma en cuenta

par(mfrow=c(1,1))
plot(ANO[17:19], E[which(EDAD==60),17:19])

Ex=E[,17:19]
colo=rainbow(length(EDAD))
plot(round(ANO[17:19],0), Ex[1,], xlab="Año", ylab="ex",
ylim=c(0,65), type="l", col=colo[1])
for(i in (2:length(EDAD))){
lines(ANO[17:19], Ex[i,], col=colo[i])
}
}
```

```

legend("top",legend=EDAD, col=colo, lty=1, horiz=TRUE, cex=1, bty="n")

#####
## TABLA DE MORTALIDAD - SUAVIZANDO TASA DE MORTALIDAD - qx METODO LINEAL##
#####
M=Mx[,19]
qx2=array(0, length(EDAD))
for(i in 1:length(EDAD)){
qx2[i]=(2*5*M[i])/(2+5*M[i])
}
qx2=as.numeric(smooth(qx2))

lx2=array(0, length(EDAD))
lx2[1]=5000
for(i in 2:length(EDAD)){
lx2[i]=lx2[i-1]-lx2[i-1]*qx2[i-1]
}

dx2=lx2*qx2

Lx2=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx2[i]=(lx2[i]+lx2[i+1])*5/2
}
Lx2[length(EDAD)]=dx2[length(EDAD)]/M[length(EDAD)]

Tx2=array(0, length(EDAD))
for(i in 1:length(EDAD)){
Tx2[i]= sum(Lx2[i:length(EDAD)])
}

ex2=Tx2/lx2

tabla2=data.frame(qx.smooth=qx2,lx=lx2,dx=dx2,Lx=Lx2,Tx=Tx2,ex=ex2)
dimnames(tabla2)[[1]]=EDADES
round(tabla2,4)

```

Apéndice B. Scripts en R

```
par(mfrow=c(1,3))
plot(EDAD,lx2, ylab="lx", xlab="Edad", type="l")
plot(EDAD, dx2, ylab="dx", xlab="Edad", type="l")
plot(EDAD, qx2, ylab="qx", xlab="Edad", type="l")

#####
##### TABLA EMPIRICA #####
### Simulación Monte Carlo con qx del BCU ###
#####
m=min(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
M=max(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
EDAD=seq(m,80,5)
ANO=year(years(1995:2013))
EADDES=c("20-24", "25-29", "30-34", "35-39", "40-44", "45-49", "50-54", "55-59",
"60-64", "65-69", "70-74", "75-79", "80 y más")

#numero de vidas entre la edad x y la edad x+5 por año del 1995 al 2013
Nx=array(0, c(length(EDAD),length(1995:2013)))
for (j in 1995:2013){
Nx[length(EDAD),j-1994]=sum(j-year(hactivos$FECHANACIM)>=80 &
year(hactivos$INGBCO)<=j)+
sum(j-year(hjubilados$FECHANACIM)>=80)+
sum(j-year(hfallec$FECHANACIM)>=80 & year(hfallec$FECHAFALLECIM)<j)
for (i in seq(20,75,5)){
Nx[EDAD==i,j-1994]=sum(j-year(hactivos$FECHANACIM)>=i &
j-year(hactivos$FECHANACIM)<i+5 &
year(hactivos$INGBCO)<=j)+
sum(j-year(hjubilados$FECHANACIM)>=i & j-year(hjubilados$FECHANACIM)<i+5)+
sum(j-year(hfallec$FECHANACIM)>=i & j-year(hfallec$FECHANACIM)<i+5 &
year(hfallec$FECHAFALLECIM)<j)
}
dimnames(Nx)[[1]]=EADDES
dimnames(Nx)[[2]]=ANO
}
Nx

Nxb<-as.data.frame(Nx[,13:19])
ANO=ANO[13:19]
```

```

##uso tasa de muerte qx de bcu agrupada
bcu=read.table("qx_bcu_agrup80.csv", head=TRUE, dec=",", sep=";")
qx=bcu$qx
qx=qx[6:18] #corto la qx en las edades que tengo

##matriz que contiene las medianas de las simulaciones por edad y año
med.sim=matrix(0, length(EDAD), length(ANO))
prom.sim=matrix(0, length(EDAD), length(ANO))

for (k in 2007:2013){

##creo un vector de edades de los expuestos por a~no
#el vector varia de edad segun el a~no
vectora<-Nxb[,k-2006]
edades<-rep(EDAD,vectora)

##le asigno la probabilidad de muerte correspondiente a su grupo de edad
q=array(0, length(edades))
for (i in seq(20,80,5)){
q[which(edades==i)] = qx[which(EDAD==i)]
}

#Simulación Monte Carlo
set.seed(1234)
itera<-1000
d.sim = matrix(0, length(EDAD))
d.sim.itera<-matrix(0, length(d.sim),itera)

for(j in 1:itera){
ui=runif(length(edades))
D=array(0,length(edades))
for(i in 1:length(edades)){
if(ui[i]<=q[i]){D[i]=1}
}
for (i in m:M){
d.sim[which(EDAD==i)] = sum(D[which(edades==i)])
}
}

```

Apéndice B. Scripts en R

```
d.sim.itera[,j]<-d.sim
}

#calculo la mediana de las iteraciones
d.sim.itera.med<-array(0,length(EDAD))
for (i in 1:length(EDAD)){
d.sim.itera.med[i]<-median(d.sim.itera[i,])
}

med.sim[,k-2006]=d.sim.itera.med

#calculo promedio de los fallecidos iterados
d.sim.itera.prom=array(0,length(EDAD))
for(i in 1:length(EDAD)){
d.sim.itera.prom[i]=sum(d.sim.itera[i,])/itera
}

prom.sim[,k-2006]=d.sim.itera.prom
}
dimnames(med.sim)[[1]]=EDADES
dimnames(med.sim)[[2]]=ANO
med.sim

dimnames(prom.sim)[[1]]=EDADES
dimnames(prom.sim)[[2]]=ANO
prom.sim
round(prom.sim,0)

##tasa bruta de mortalidad de la tabla empirica
(med.sim.r=med.sim/Nxb)

par(mfrow=c(1,1))
colo=rainbow(7)
plot(EDAD, med.sim.r[,7], ylab="mx", xlab="Edad", type="l", col=colo[1])
for(i in 1:18){
lines(EDAD, med.sim.r[,i], col=colo[i])
}
legend("topleft", as.character(ANO), col=colo, lty=1, bty="n")
```

```

#####
##cuento los fallecidos del brou por a\~no
dx=array(0, c(length(EDAD),length(2007:2013)))
for (j in 2007:2013){
dx[length(EDAD),j-2006]=sum(hfallec$EDAD_FALLEC>=80 & year(hfallec$FECHAFALLECIM)==j)
for (i in seq(20,75,5)){
dx[which(EDAD==i),j-2006]=sum(hfallec$EDAD_FALLEC>=i &
hfallec$EDAD_FALLEC<i+5 & year(hfallec$FECHAFALLECIM)==j)
}
dimnames(dx)[[1]]=EADDES
dimnames(dx)[[2]]=ANO
}
dx

Mx=dx/Nxb

colob=topo.colors(7)
for(i in 1:7){
lines(EDAD, Mx[,i], col=colob[i])
}
legend("topright", as.character(ANO), col=colob, lty=1, bty="n")

#####
par(mfrow=c(1,2))
plot(EDAD, med.sim.r[,7], ylab="promedio de mx simulada", xlab="Edad", type="l",
col=colo[1], cex.main=0.9,main=" ")
for(i in 1:7){
lines(EDAD, med.sim.r[,i], col=colo[i])
}
legend("topleft", as.character(ANO), col=colo, lty=1, bty="n")

plot(EDAD, Mx[,7], ylim=c(0:1),ylab="mx", xlab="Edad", type="l",
col=colo[1], cex.main=0.9,main="")
for(i in 1:18){
lines(EDAD, Mx[,i], col=colo[i])
}
legend("topleft", as.character(ANO), col=colo, lty=1, bty="n")

```

Apéndice B. Scripts en R

```
#####
d.sim.itera.min<-array(0,length(EDADES))
d.sim.itera.max<-array(0,length(EDADES))
d.sim.itera.de<-array(0,length(EDADES))
d.sim.itera.med<-array(0,length(EDADES))
for (i in 1:length(EDADES)){
d.sim.itera.min[i]<-min(d.sim.itera[i,])
d.sim.itera.max[i]<-max(d.sim.itera[i,])
d.sim.itera.de[i]<-sd(d.sim.itera[i,])
d.sim.itera.med[i]<-median(d.sim.itera[i,])}
d.sim.itera.edad<-as.data.frame(cbind(EDADES,d.sim.itera.min,
d.sim.itera.max,d.sim.itera.med, d.sim.itera.de))
d.sim.itera.edad

plot(EDAD,d.sim.itera.edad$d.sim.itera.max,type="l", xlab="Edad",
ylab="mediana de los fallecimientos simulados")
lines(EDAD,d.sim.itera.edad$d.sim.itera.med,col="violet")

#####
##### TABLA DE MORTALIDAD USANDO mx DEL INE #####
#comparacion distintos metodos de estimacion de qx#
#####

ine = read.table("mx_ine.csv", head=TRUE, dec=",", sep=";")
M=ine$mx #el ultimo tramo de edad es "100 o más"
EDAD=c(0:100)
EDADES=EDAD
EDADES[length(EDAD)]=c("100 o más")

## MÉTODO LINEAL
qx1=array(0, length(EDAD))
for(i in 0:length(EDAD)){
qx1[i]=(2*M[i])/(2+M[i])
}

#####
```

```

## MÉTODO EXPONENCIAL SIMPLE
qx2=array(0, length(EDAD))
for(i in 0:length(EDAD)){
qx2[i]=1-exp(-M[i])
}

#####
## MÉTODO REED-MERREL
qx3=array(0, length(EDAD))
for(i in 0:length(EDAD)){
qx3[i]= 1-exp(-M[i]-0.008*(1^3)*(M[i]^2))
}

#####
## MÉTODO GREVILLE
qx4=array(0, length(EDAD))
for(i in 0:length(EDAD)){
qx4[i]= M[i]/(1/1+M[i]*(1/2+1/12*(M[i]-0.095)))
}

#####
tabla=data.frame(mx=M,qx1,qx2,qx3,qx4)
dimnames(tabla)[[1]]=EADDES
round(tabla,6)

plot(EDAD, M, type="l", ylab="qx", xlab="Edad",ylim=c(0,0.5), lwd=1,main="")
lines(EDAD, jitter(qx1,50), col="firebrick1", lwd=1.7)
lines(EDAD, jitter(qx2,50), col="magenta3", lwd=1.7)
lines(EDAD, jitter(qx3,50), col="royalblue3", lwd=1.7)
lines(EDAD, jitter(qx4,50), col="mediumspringgreen", lwd=1.7)
legend("topleft", c("Mx","Método Lineal", "Método Exponencial", "Reed-Merrel",
"Greville"), col=c("black","firebrick1", "magenta3","royalblue3",
"mediumspringgreen"), lty=1, bty="n")

#### No hay casi diferencias entre los distintos metodos por lo que
#### las Tablas de vida van a dar casi igual

#####

```

Apéndice B. Scripts en R

```
## TABLA DE MORTALIDAD usando tasa de muerte LINEAL ##
lx1=array(0, length(EDAD))
lx1[1]=5000
for(i in 2:length(EDAD)){
  lx1[i]=lx1[i-1]-lx1[i-1]*qx1[i-1]
}

dx1=lx1*qx1

Lx1=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
  Lx1[i]=(lx1[i]+lx1[i+1])*1/2
}
Lx1[length(EDAD)]=dx1[length(EDAD)]/M[length(EDAD)]

Tx1=array(0, length(EDAD))
for(i in 1:length(EDAD)){
  Tx1[i]= sum(Lx1[i:length(EDAD)])
}

ex1=Tx1/lx1

tabla1=data.frame(qx1,lx=lx1,dx=dx1,Lx=Lx1,Tx=Tx1,ex=ex1)
dimnames(tabla1)[[1]]=EDADES
round(tabla1,4)

par(mfrow=c(1,3))
plot(EDAD,lx1, ylab="lx", xlab="Edad", type="l")
plot(EDAD,dx1, ylab="dx", xlab="Edad", type="l")
plot(EDAD,qx1, ylab="qx", xlab="Edad", type="l")

#####
### ex USANDO qx DEL BCU Y METODO DE ESTIMACION DE
## Alho y Spencer - 2005 - Statistical demography and forecasting
## Capítulo 4: Waiting Times and Their Statistical Estimation
##expectativa de vida a la edad x
p=1-q
ex2=array(0, length(EDAD))
```

```

for(i in 0:length(EDAD)){
ex2[i]=0.5+sum(p[(i+1):length(EDAD)])/p[i]
}

tabla2=data.frame(qx1,lx=lx1,dx=dx1,Lx=Lx1,Tx=Tx1,ex.lineal=ex1,ex.alho=ex2)
dimnames(tabla2)[[1]]=EADDES
round(tabla2,4)

#####
#### TABLAS DE LA POBLACION URUGUYA ####
##Según tasas de muerte del BCU AÑO 2012##
#####
bcu = read.table("tabla bcu.csv", head=TRUE, dec=",", sep=";")

qm=bcu$qx_m #tasa de muerte masculina
qf=bcu$qx_f #tasa de muerte femenina

M_m=2*qm/(2-qm)
M_f=2*qf/(2-qf)

EDAD=c(0:110)

## TABLA MASCULINA ##
lx_m=array(0, length(EDAD))
lx_m[1]=100000
for(i in 2:length(EDAD)){
lx_m[i]=lx_m[i-1]-lx_m[i-1]*qm[i-1]
}

dx_m=lx_m*qm

Lx_m=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx_m[i]=(lx_m[i]+lx_m[i+1])*1/2
}
Lx_m[length(EDAD)]=dx_m[length(EDAD)]/M_m[length(EDAD)]

Tx_m=array(0, length(EDAD))

```

Apéndice B. Scripts en R

```
for(i in 1:length(EDAD)){
Tx_m[i]= sum(Lx_m[i:length(EDAD)])
}

ex_m=Tx_m/lx_m

tabla_m=data.frame(qm,lx_m,dx_m,Lx_m,Tx_m,ex_m)
dimnames(tabla_m)[[1]]=EDAD
round(tabla_m,4)

## TABLA FEMENINA
lx_f=array(0, length(EDAD))
lx_f[1]=100000
for(i in 2:length(EDAD)){
lx_f[i]=lx_f[i-1]-lx_f[i-1]*qf[i-1]
}

dx_f=lx_f*qf

Lx_f=array(0, length(EDAD))
for(i in 1:length(EDAD)-1){
Lx_f[i]=(lx_f[i]+lx_f[i+1])*1/2
}
Lx_f[length(EDAD)]=dx_f[length(EDAD)]/M_f[length(EDAD)]

Tx_f=array(0, length(EDAD))
for(i in 1:length(EDAD)){
Tx_f[i]= sum(Lx_f[i:length(EDAD)])
}

ex_f=Tx_f/lx_f

tabla_f=data.frame(qf,lx_f,dx_f,Lx_f,Tx_f,ex_f)
dimnames(tabla_f)[[1]]=EDAD
round(tabla_f,4)

#####
```

```

par(mfrow=c(1,3))
plot(EDAD, lx_m, xlab="Edad", ylab="lx",type="o", col="lightseagreen")
lines(EDAD, lx_f, col="indianred1")
legend("topright", c("hombres", "mujeres"), col=c("lightseagreen","indianred1"),
lty=1, pch=c(1,NA_integer_), bty="n")
plot(EDAD, dx_f, xlab="Edad", ylab="dx",type="l", col="indianred1")
lines(EDAD, dx_m, col="lightseagreen", type="o")
legend("topleft", c("hombres", "mujeres"), col=c("lightseagreen","indianred1"),
lty=1, pch=c(1,NA_integer_), bty="n")
plot(EDAD, qm, xlab="Edad", ylab="qx",type="o", col="lightseagreen")
lines(EDAD, qf, col="indianred1")
legend("topleft", c("hombres", "mujeres"), col=c("lightseagreen","indianred1"),
lty=1, pch=c(1,NA_integer_), bty="n")

```

```

#####
##### TABLA DE MORTALIDAD AGRUPADA #####
##### qx DEL BCU #####
#####

```

```

##uso tasa de muerte qx de bcu agrupada
bcu=read.table("qx_bcu_agrup80.csv", head=TRUE, dec=",", sep=";")
qx=bcu$qx
qx1=qx[6:18] #corto la qx en las edades que tengo

```

```

EDAD=seq(20,80,5)
EADDES=c("20-24", "25-29", "30-34", "35-39", "40-44", "45-49", "50-54",
"55-59", "60-64", "65-69", "70-74", "75-79", "80 y más")

```

```

lx1=array(0, length(EDAD))
lx1[1]=5000
for(i in 2:length(EDAD)){
lx1[i]=lx1[i-1]-lx1[i-1]*qx1[i-1]
}

```

```

dx1=lx1*qx1

```

```

Lx1=array(0, length(EDAD))

```

Apéndice B. Scripts en R

```
for(i in 1:length(EDAD)-1){
Lx1[i]=(lx1[i]+lx1[i+1])*5/2
}
#Lx1[length(EDAD)]=dx1[length(EDAD)]/M[length(EDAD)]
Lx1[length(EDAD)]=lx1[length(EDAD)]*4.71
#se supone e80=4,71 que es la esperanza de vida de
#los funcionarios del BROU y se calcula L80=180*e80

Tx1=array(0, length(EDAD))
for(i in 1:length(EDAD)){
Tx1[i]= sum(Lx1[i:length(EDAD)])
}

ex1=Tx1/lx1

tabla1=data.frame(qx1,lx=lx1,dx=dx1,Lx=Lx1,Tx=Tx1,ex=ex1)
dimnames(tabla1)[[1]]=EDADES
round(tabla1,4)

#####
##### LEE CARTER #####
#####

library(lubridate)
library(demography)
library(forecast)
library(ilc)
library(tseries)
library(car) #para grafica qq-plot

activos = read.csv("Datos Activos al 31.07.2014.csv", head=TRUE, sep=";")
jubilados = read.csv("egresosJUBILADOS.csv", head=TRUE, sep=";")
fallec = read.csv("egresosFALLEC.csv", head=TRUE, sep=";")

hactivos=activos[activos$SEXO=="MASCULINO",]
hjubilados=jubilados[jubilados$SEXO=="MASCULINO",]
hfallec=fallec[fallec$SEXO=="MASCULINO",]
```

```

hfallec$FECHAFALLECIM=as.Date(hfallec$FECHAFALLECIM,format="%d/%m/%Y")
hactivos$FECHANACIM=as.Date(hactivos$FECHANACIM,format="%d/%m/%Y")
hactivos$INGBCO=as.Date(hactivos$INGBCO,format="%d/%m/%Y")
hjubilados$FECHANACIM=as.Date(hjubilados$FECHANACIM,format="%d/%m/%Y")
hfallec$FECHANACIM=as.Date(hfallec$FECHANACIM,format="%d/%m/%Y")

m=min(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
M=max(hactivos$EDAD, hjubilados$EDAD, hfallec$EDAD_FALLEC)
EDAD=seq(m,80,5)
ANO=year(years(1995:2013))
EDADES=c("20-24","25-29","30-34","35-39","40-44","45-49","50-54","55-59",
"60-64","65-69","70-74","75-79","80 y más")

#número de muertes en el intervalo de edad (x,x+5)por año 1995 al 2013
Dx=array(0, c(length(EDAD),length(1995:2013)))
for (j in 1995:2013){
Dx[length(EDAD),j-1994]=sum(hfallec$EDAD_FALLEC>=80 &
year(hfallec$FECHAFALLECIM)==j)
for (i in seq(20,75,5)){
Dx[which(EDAD==i),j-1994]=sum(hfallec$EDAD_FALLEC>=i &
hfallec$EDAD_FALLEC<i+5 & year(hfallec$FECHAFALLECIM)==j)
}
dimnames(Dx)[[1]]=EDADES
dimnames(Dx)[[2]]=ANO
}
Dx

#número de expuestos en el intervalo de edad (x,x+5)por año 1995 al 2013
Nx=array(0, c(length(EDAD),length(1995:2013)))
for (j in 1995:2013){
Nx[length(EDAD),j-1994]=sum(j-year(hactivos$FECHANACIM)>=80 &
year(hactivos$INGBCO)<=j)+
sum(j-year(hjubilados$FECHANACIM)>=80)+
sum(j-year(hfallec$FECHANACIM)>=80 & year(hfallec$FECHAFALLECIM)<j)
for (i in seq(20,75,5)){
Nx[EDAD==i,j-1994]=sum(j-year(hactivos$FECHANACIM)>=i &
j-year(hactivos$FECHANACIM)<i+5 &

```

Apéndice B. Scripts en R

```
year(hactivos$INGBCO)<=j)+
sum(j-year(hjubilados$FECHANACIM)>=i & j-year(hjubilados$FECHANACIM)<i+5)+
sum(j-year(hfallec$FECHANACIM)>=i & j-year(hfallec$FECHANACIM)<i+5 &
year(hfallec$FECHAFALLECIM)<j)
}
dimnames(Nx)[[1]]=EADDES
dimnames(Nx)[[2]]=ANO
}
Nx

Mx=Dx/Nx
Mx

hbrou=demogdata(data=Mx,pop=Nx,age=EDAD,year=ANO,type="mortality",
label="BROU",name="male")
hbrou

par(mfrow=c(1,3))
plot_dd(hbrou,xlab="Edad", ylab="mx",transf=F, main="")
plot_dd(hbrou,xlab="Edad", ylab="Logaritmo de mx", main="")
plot(hbrou, plot.type="time", xlab="Año", ylab="Logaritmo de mx", main="")
#las lineas del último gráfico aparecen cortadas por las tasas cero o
#NaN que tiene la tabla

par(mfrow=c(1,2))
d=extract.deaths(hbrou)
plot_dd(d,xlab="Edad",ylab="dx", transf=F, main="")
plot_dd(d,xlab="Edad", ylab="Logaritmo de dx", main="")

##### LIBRERIA DEMOGRAPHY #####
## Estimación LC por SVD
lca(hbrou) #no funciona, hay muchos 0s en la tabla
lca(hbrou,interpolate=TRUE) #no funciona

#saco el grupo de edades mas jovenes para que funcione
(l=lca(hbrou,interpolate=TRUE,ages=hbrou$age[2:13], years=hbrou$year))
```

```

plot(1)

## Bondad de ajuste del modelo
res1=l$residuals$y
dimnames(res1)[[2]]=ANO
res1

#gráficos de residuos
par(mfrow=c(2,2))
plot(l$residuals, xlab="Edad", ylab="Residuos del log(mx)", main="Residuos")
hist(res1, probability=T, main='Histograma de los resiudos', xlab='residuos',
ylab='densidad', col='grey', ylim=c(0,1.2))
lines(density(res1), col='red')
boxplot(res1, main='Boxplot de los resiudos', col='grey', xlab="Año",
ylab="residuos")
qqplot(ANO,res1, main='Q-Q Plot de los resiudos', xlab="Año", ylab="residuos")

#Esperanza 0, Varianza constante?
par(mfrow=c(1,1))
plot(l$residuals, xlab="Edad",ylab="Residuos log mx")
#cumplen el supuesto de media 0 pero hay heterocedasticidad en el grupo
#de edades 75-79 años

#Test de normalidad
p_value1=array(0,12)
for(i in 1:12){
p_value1[i]=jarque.bera.test(res1[i,])$p.value
}
p_value1<0.05
#cofianza del 95%, segundo grupo de edades y último no cumplen Normalidad

#Test de incorrelación
p_value2=array(0,12)
for(i in 1:12){
p_value2[i]=Box.test(l$residuals$y[i,], lag=12, type="Ljung-Box")$p.value
}
p_value2<0.05 #cofianza del 95%, residuos incorrelacionados

```

Apéndice B. Scripts en R

```
pvc=cbind(p.valor.JB=p_value1,p.valor.LB=p_value2)
dimnames(pvc)[[1]]=EADDES[2:13]
round(pvc,6)

#####
## Modelo 2, saco los dos primeros grupos menores de edades y el último
## que no cumplian la normalidad y la incorrelación
(l2=lca(hbrou,interpolate=TRUE,ages=hbrou$age[4:12], years=hbrou$year))

## Bondad de ajuste
res2=l2$residuals$y
dimnames(res2)[[2]]=ANO
round(res2,4)

#gráficos de residuos
par(mfrow=c(2,2))
plot(l2$residuals, xlab="Edad", ylab="Residuos del log(mx)", main="Residuos")
hist(res2, probability=T, main='Histograma de los residuos', xlab='residuos',
ylab='densidad', col='grey', ylim=c(0,0.8))
lines(density(res2), col='red')
boxplot(res2, main='Boxplot de los residuos', col='grey', xlab="Año",
ylab="residuos")
qqplot(ANO,res2, main='Q-Q Plot de los residuos', xlab="Año", ylab="residuos")

#Esperanza 0, varianza constante?
plot(l2$residuals, xlab="Edad", ylab="Residuos de mx")
#cumple sup. media 0, pero hay heterocedasticidad en el último grupo
#de edades 74-79

#Test de Normalidad
p_value1=array(0,9)
for(i in 1:9){
p_value1[i]=jarque.bera.test(res2[i,])$p.value
}
p_value1<0.05 #cofianza del 95%, residuos normales
```

```

#Test de incorrelacion
p_value2=array(0,9)
for(i in 1:9){
p_value2[i]=Box.test(l2$residuals$y[i,], lag = 7, type = "Ljung-Box")$p.value
}
p_value2<0.05
#cofianza del 95%, residuos correlacionados en el primer grupo de
#edades(35-39 años)

pv=cbind(p.valor.JB=p_value1,p.valor.LB=p_value2)
dimnames(pv)[[1]]=EADDES[4:12]
round(pv,6)

#####
## Modelo 3, saco el grupo de 35-39 años que no cumple incorrelación
(l3=lca(hbrou,interpolate=TRUE,ages=hbrou$age[5:12], years=hbrou$year))

## Bondad de ajuste
res3=l3$residuals$y
dimnames(res3)[[2]]=ANO
round(res3,4)

#gráficos de residuos
par(mfrow=c(2,2))
plot(l3$residuals, xlab="Edad", ylab="Residuos del log(mx)", main="Residuos")
hist(res3, probability=T, main='Histograma de los resiudos', xlab='residuos',
ylab='densidad', col='grey', ylim=c(0,0.8))
lines(density(res3), col='red')
boxplot(res3, main='Boxplot de los resiudos', col='grey', xlab="Año",
ylab="residuos")
qqplot(ANO,res3, main='Q-Q Plot de los resiudos', xlab="Año", ylab="residuos")
#del histograma de residuos parece que los mismos se distribuyen normales

#Esperanza 0, varianza constante?
plot(l3$residuals, xlab="Edad", ylab="Residuos de mx")
#cumple sup. media 0, pero hay heterocedasticidad en el último grupo
#de edades 74-79

```

Apéndice B. Scripts en R

```
#Test de Normalidad
p_value1=array(0,8)
for(i in 1:8){
p_value1[i]=jarque.bera.test(res3[i,])$p.value
}
p_value1<0.05 #cofianza del 95%, residuos normales

#Test de incorrelacion
p_value2=array(0,8)
for(i in 1:8){
p_value2[i]=Box.test(13$residuals$y[i,], lag=7, type="Ljung-Box")$p.value
}
p_value2<0.05 #cofianza del 95%, residuos incorrelacionados

pv=cbind(p.valor.JB=p_value1,p.valor.LB=p_value2)
dimnames(pv)[[1]]=EDADES[5:12]
round(pv,6)

#Parámetros del modelo
mod3=data.frame(ax=13$ax,bx=13$bx)
mod3

mk3=data.frame(kt=13$kt)
dimnames(mk3)[[1]]=ANO
mk3

par(mfrow=c(1,3))
plot(seq(40,75,5),13$ax,xlab="Edad", ylab="ax", type="l",main="ax")
plot(seq(40,75,5),13$bx,xlab="Edad", ylab="bx", type="l",main="bx")
plot(13$kt,xlab="Año", ylab="kt", type="l",main="kt")

## Proyección tasas de mortalidad
f.hbrou=forecast(13,h=5,level=90)

(fk3=data.frame(kt.f=f.hbrou$kt.f))

par(mfrow=c(1,2))
```

```

plot(f.hbrou$kt.f, xlab="Año", ylab="kt", main="")
plot_dd(f.hbrou, xlab="Edad", ylab="Logaritmo de mx", main="")

edad=seq(40,75,5)
anos=c(1995:2018)
hbrou2=demogdata(data=Mx[5:12,],pop=Nx[5:12,],age=edad,year=c(1995:2013),
type="mortality",label="BROU",name="male")
tasas=cbind(hbrou2$rate$male,f.hbrou$rate$male)
dimnames(tasas)[[2]]=anos
dimnames(tasas)[[1]]=edad

par(mfrow=c(1,3))
colores=rainbow(8)
plot(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas[which(edad==75),],
col=colores[which(edad==75)], xlab="Año", ylab="mx", type="l", main= "")
for(i in seq(40,70,5)){
lines(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas[which(edad==i),],
col=colores[which(edad==i)])
legend("topright", c("40-44 años","45-49 años","50-54 años","55-59 años",
"60-64 años", "65-69 años","70-74 años", "75-79 años"), col=c(colores),
lty=1, bty="n")
}
colores=rainbow(8)
plot(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas[which(edad==55),],
col=colores[which(edad==55)],xlab="Años",
ylab="Tasas de muerte", ylim=c(0,0.06),type="l", main= "")
for(i in seq(40,50,5)){
lines(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas[which(edad==i),],
col=colores[which(edad==i)])
legend("topright", c("40-45 años","45-49 años","50-54 años", "55-59 años"),
col=colores[1:4], bty="n", lty=1)
}
plot(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas[which(edad==75),],
col=colores[which(edad==75)],xlab="Años",
ylab="Tasas de muerte", ylim=c(0,0.06),type="l", main= "")
for(i in seq(60,70,5)){
lines(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas[which(edad==i),],

```

Apéndice B. Scripts en R

```
col=colores[which(edad==i)]
legend("topright", c("60-64 años","65-69 años","70-74 años","75-79 años"),
col=colores[5:8], bty="n", lty=1)
}

##### LIBRERIA ILC #####
## Estimación LC con errores Poisson
m4 = lca.rh(hbrou2, mod='lc', interpolate=TRUE,verbose = FALSE)

plot(m4)

## Bondad de ajuste
# R^2
SCT4=sum((m4$y$m4$ax)^2)
SCR4=sum(m4$residuals$y^2)
1-SCR4/SCT4

m4$mdev

res4=m4$residuals$y
dimnames(res4)[[2]]=ANO
round(res4,4)

par(mfrow=c(1,1))
hist(res4, probability=T, main='Histograma de los residuos', xlab='residuos',
ylab='densidad', col='grey', ylim=c(0,1))
lines(density(res4), col='red')

par(mfrow=c(2,1))
plot(m4$residuals, xlab="Edad", ylab="Residuos del log(mx)", main="Residuos")
boxplot(res4, main='Boxplot de los residuos', col='grey', xlab="Año",
ylab="residuos")

## Parámetros del modelo
mod4=data.frame(ax=m4$ax,bx=m4$bx)
mod4
```

```

mk4=data.frame(kt=m4$kt)
dimnames(mk4)[[1]]=ANO
mk4

par(mfrow=c(1,3))
plot(seq(40,75,5),m4$ax, main="", xlab="Edad", ylab="ax", type="l")
lines(seq(40,75,5),l3$ax, col="springgreen2")
legend("topleft", legend=c("ilc","lca"), col=c("black","springgreen2"),
lty=1, bty="n")
plot(seq(40,75,5),m4$bx, main="", xlab="Edad", ylab="bx", type="l")
lines(seq(40,75,5),l3$bx, col="springgreen2")
legend("topleft", legend=c("ilc","lca"), col=c("black","springgreen2"),
lty=1, bty="n")
plot(m4$kt, main="", xlab="Año", ylab="kt", type="l")
lines(l3$kt, col="springgreen2")
legend("topright", legend=c("ilc","lca"), col=c("black","springgreen2"),
lty=1, bty="n")

## Proyección logaritmo de las tasas de mortalidad
f4=forecast(m4, h=5, jump='fit', level=90, shift=F)

(fk4=data.frame(kt.f=f4$kt.f))

par(mfrow=c(1,2))
plot(f4$kt.f, xlab="Año", ylab="kt", main="")
plot_dd(f4, xlab="Edad", ylab="Logaritmo de mx",main="")

tasas4=cbind(hbrou2$rate$male,f4$rate$male)
dimnames(tasas4)[[2]]=anos
dimnames(tasas4)[[1]]=edad

par(mfrow=c(1,3))
colores=rainbow(8)
plot(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas4[which(edad==75),],
col=colores[which(edad==75)],
xlab="Año", ylab="mx", type="l", main="Tasas de muerte de los hombres \n 2003-2018")
for(i in seq(40,70,5)){

```

Apéndice B. Scripts en R

```
lines(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas4[which(edad==i)],
col=colores[which(edad==i)])
legend("topright", c("40-44 años","45-49 años","50-54 años","55-59 años",
"60-64 años","65-69 años", "70-74 años", "75-79 años"), col=c(colores),l
ty=1, bty="n")
}
colores=rainbow(8)
plot(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas4[which(edad==55)],
col=colores[which(edad==55)],
xlab="Años", ylab="Tasas de muerte", ylim=c(0,0.06),type="l", main= "")
for(i in seq(40,50,5)){
lines(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas4[which(edad==i)],
col=colores[which(edad==i)])
legend("topright", c("40-45 años","45-49 años","50-54 años", "55-59 años"),
col=colores[1:4], bty="n", lty=1)
}
plot(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas4[which(edad==75)],
col=colores[which(edad==75)],xlab="Años", ylab="Tasas de muerte",
ylim=c(0,0.06),type="l", main= "")
for(i in seq(60,70,5)){
lines(seq(min(hbrou2$year),max(hbrou2$year)+5), tasas4[which(edad==i)],
col=colores[which(edad==i)])
legend("topright", c("60-64 años","65-69 años","70-74 años","75-79 años"),
col=colores[5:8], bty="n", lty=1)
}

##### Ya que el modelo permite heterocedasticidad calculo el
##### modelo para todas las edades
hbrou3=demogdata(data=Mx[2:13,],pop=Nx[2:13,],age=EDAD[2:13],
year=c(1995:2013), type="mortality",label="BROU",name="male")

m5 = lca.rh(hbrou3, mod='lc', interpolate=TRUE,verbose = FALSE)

## Bondad de ajuste
# R^2
SCT5=sum((m5$y$y-m5$ax)^2)
SCR5=sum(m5$residuals$y^2)
```

1-SCR5/SCT5

m5\$mdev

```
res5=m5$residuals$y
dimnames(res5)[[2]]=ANO
res5
```

```
par(mfrow=c(1,1))
hist(res5, probability=T, main='Histograma de los residuos', xlab='residuos',
ylab='densidad', col='grey', ylim=c(0,1))
lines(density(res4), col='red')
```

```
par(mfrow=c(2,1))
plot(m5$residuals, xlab="Edad", ylab="Residuos del log(mx)", main="")
boxplot(res5, main='', col='grey', xlab="Año", ylab="residuos del log(mx)")
```

```
## Parámetros del modelo
mod5=data.frame(ax=m5$ax,bx=m5$bx)
mod5
```

```
mk5=data.frame(kt=m5$kt)
dimnames(mk5)[[1]]=ANO
mk5
```

```
par(mfrow=c(1,3))
plot(seq(25,80,5),m5$ax, main="", xlab="Edad", ylab="ax", type="l")
plot(seq(25,80,5),m5$bx, main="", xlab="Edad", ylab="bx", type="l")
plot(m5$kt, main="", xlab="Año", ylab="kt", type="l")
```

```
## Proyección logaritmo de las tasas de mortalidad
f5=forecast(m5, h=5, jump='fit', level=90, shift=F)
```

```
(fk5=data.frame(kt.f=f5$kt.f))
```

```
par(mfrow=c(1,2))
```

Apéndice B. Scripts en R

```
plot(f5$kt.f, xlab="Año", ylab="kt", main="")
plot_dd(f5, xlab="Edad", ylab="Logaritmo de mx",main="")

edad=seq(25,80,5)
tasas5=cbind(hbrou3$rate$male,f5$rate$male)
dimnames(tasas5)[[2]]=anos
dimnames(tasas5)[[1]]=edad

par(mfrow=c(1,3))
colores=rainbow(12)
plot(seq(min(hbrou3$year),max(hbrou3$year)+5), tasas5[which(edad==75),],
col=colores[which(edad==75)], xlab="Año", ylab="mx", type="l", main= "")
lines(seq(min(hbrou3$year),max(hbrou2$year)+5),tasas5[which(edad==80),],
col=colores[which(edad==80)])
for(i in seq(25,70,5)){
lines(seq(min(hbrou3$year),max(hbrou3$year)+5), tasas5[which(edad==i)],
col=colores[which(edad==i)])
legend("topright", c("25-24 años","30-34 años", "35-39 años","40-44 años",
"45-49 años", "50-54 años", "55-59 años","60-64 años","65-69 años",
"70-74 años", "75-79 años", "80 años y más"),
col=c(colores, "pink","black"),lty=1, bty="n")
}
col=rainbow(6)
colores=c(col,col)
plot(seq(min(hbrou3$year),max(hbrou3$year)+5), tasas5[which(edad==50),],
col=colores[6],xlab="Años", ylab="Tasas de muerte", ylim=c(0,0.06),type="l", main= "")
for(i in seq(25,45,5)){
lines(seq(min(hbrou3$year),max(hbrou3$year)+5), tasas5[which(edad==i)],
col=colores[which(edad==i)])
legend("topright", c("25-24 años","30-34 años", "35-39 años","40-45 años",
"45-49 años","50-59 años"), col=colores[1:6], bty="n", lty=1)
}
plot(seq(min(hbrou3$year),max(hbrou3$year)+5), tasas5[which(edad==80),],
col=colores[12],xlab="Años", ylab="Tasas de muerte", ylim=c(0,0.06),type="l", main= "")
for(i in seq(55,75,5)){
lines(seq(min(hbrou3$year),max(hbrou3$year)+5), tasas5[which(edad==i)],
col=colores[which(edad==i)])
```

```
legend("topright", c("55-59 años","60-64 años","65-69 años","70-74 años",  
"75-79 años", "80 años y más"), col=colores[7:12], bty="n", lty=1)  
}
```