

DEPARTAMENTO DE ESTADÍSTICA E INV. OPERATIVA

GRADUACIÓN DE TABLAS DE MORTALIDAD.
APLICACIONES ACTUARIALES

ANA DEBON AUCEJO

UNIVERSITAT DE VALENCIA
Servei de Publicacions
2003

Aquesta Tesi Doctoral va ser presentada a València el dia 23 de Setembre de 2003 davant un tribunal format per:

- D. Manuel Mocholí Arce
- D. Guillermo Ayala
- D. Rafael Romero Vilafranca
- D. José Domingo Bermúdez
- D^a. Flor Marí Guerrero Casas

Va ser dirigida per:

D. Francisco Montes Suay

D. Ramón Sala Garrido

©Copyright: Servei de Publicacions
Ana Debon Aucejo

Depòsit legal:

I.S.B.N.:84-370-5893-7

Edita: Universitat de València
Servei de Publicacions
C/ Artes Gráficas, 13 bajo
46010 València
Spain
Telèfon: 963864115



Graduación de tablas de mortalidad. Aplicaciones actuariales

Tesis Doctoral

Presentada por: Ana Debón Aucejo

Dirigida por: Francisco Montes Suay
Ramón Sala Garrido

Departament d'Estadística i Investigació Operativa
Universitat de València

Junio 2003

Francisco Montes Suay, Catedrático de Estadística e Investigación Operativa del Departament d'Estadística i Investigació Operativa de la Universitat de València,
Ramón Sala Garrido, Profesor Titular de Universidad de Economía Financiera y Contabilidad del Departament de Matemática Económico-Empresarial de la Universitat de València,

CERTIFICAN: que la presente memoria, *Graduación de tablas de mortalidad. Aplicaciones actuariales*, ha sido realizada bajo su dirección en los Departament d'Estadística i Investigació Operativa y Departament de Matemática Económico-Empresarial por Doña Ana Debón Aucejo y constituye su Tesis para optar al Grado de Doctor en Matemáticas.

Y para que conste, en cumplimiento de la legislación vigente, la presenta ante la Facultat de Matemàtiques de la Universitat de València, a 4 de Junio de 2003.

Los Directores

Francisco Montes Suay

Ramón Sala Garrido

Agradecimientos

Quiero mostrar mi más sincera gratitud a mis tutores Francisco Montes Suay y Ramón Sala Garrido sin los cuales esta investigación no habría sido posible.

Mi recorrido laboral ha sido cambiante, gracias a mi primer trabajo en el Departamento de Economía Aplicada inicié mi andadura universitaria con Belén García “mi compañera del alma”, pero no puedo pasar por alto mi estancia en Florida Centre de Formació, “La Florida” allí empezó mi sueño y allí conocí a muchas personas determinantes para mi, mi agradecimiento a todos mis compañeros del Departamento de Matemáticas y Estadística. Tampoco puedo dejar de mencionar la ayuda y motivación que ha supuesto para mi el departamento donde actualmente estoy, Departamento de Estadística e Investigación Operativa Aplicadas y Calidad de la Universidad Politécnica de Valencia.

Y por último, gracias a los que me quieren y quiero, a mis padres y hermanos. Y en especial a ti Paco Puig por tu infinito apoyo.

Índice general

Introducción	1
1. Análisis de supervivencia y tablas de mortalidad	3
1.1. Introducción	3
1.2. Principales funciones de la distribución del tiempo de supervivencia	3
1.3. Estimación de las curvas de supervivencia	5
1.3.1. Distribuciones analíticas del tiempo de supervivencia	5
1.3.2. Tablas de mortalidad	7
1.4. Contenido y clasificación de las tablas de mortalidad	9
1.4.1. Presentación de una tabla de mortalidad	10
1.4.2. Tablas Estáticas y Dinámicas	11
2. Graduación de datos de mortalidad: métodos paramétricos	13
2.1. Introducción	13
2.2. Descripción del problema	14
2.2.1. Distribución de D_x según q_x	16
2.2.2. Distribución de D_x según μ_x	17
2.2.3. Distribución de E_x	18
2.3. Métodos paramétricos	18
2.3.1. El método de máxima verosimilitud	19
2.3.2. El método de mínima chi-cuadrado	20
2.3.3. El método de mínimos cuadrados	21
2.4. Metodología del CMI Bureau	22
2.5. Modelos lineales generalizados	24
2.5.1. Introducción a los modelos lineales generalizados	24
2.5.2. Modelos lineales generalizados y graduación	26
2.5.3. Modelos no lineales generalizados y graduación	30
2.6. Modelo Heligman y Pollard	30
2.7. Contrastes para la bondad del ajuste	34
3. Graduación de datos de mortalidad: métodos no paramétricos	37
3.1. Introducción	37
3.2. Técnicas de smoothing	38
3.2.1. Selección del parámetro de smoothing bandwidth o span	41

3.2.2. Kernel smoother	43
3.3. La Fórmula de Whittaker-Henderson	49
3.4. Graduación polinomial local	51
3.5. Modelos aditivos generalizados (GAM)	52
4. Modelos dinámicos	53
4.1. Introducción	53
4.2. Métodos estructurales	55
4.2.1. Modelo de Heligman y Pollard	55
4.2.2. El método logit	58
4.2.3. Método de Lee-Carter	59
4.3. Métodos no estructurales	66
4.3.1. El método logarítmico	66
4.3.2. Método del CMI BUREAU y modelos lineales generalizados	68
4.4. Métodos no paramétricos	72
4.4.1. El modelo	72
4.4.2. Estimación constante local de α	72
4.4.3. El submodelo multiplicativo y el método de estimación de integración marginal	73
4.5. Métodos basados en factores de reducción de la mortalidad	74
4.5.1. Caracterización del factor de reducción para q_x	75
4.5.2. Caracterización del factor de reducción para μ_x	75
4.5.3. Modelización de los factores de reducción	76
5. Ajustes paramétricos de tablas de mortalidad de la Comunidad Valenciana	85
5.1. Introducción	85
5.2. Estimación de las q_x	86
5.3. Modelos lineales generalizados para q_x	88
5.4. Modelos para μ_x	89
5.4.1. Modelos lineales generalizados	89
5.4.2. Modelos no lineales generalizados	97
5.4.3. Resultados	97
5.5. Modelo de Heligman y Pollard	104
5.6. Comparación de los diferentes métodos	106
6. Ajustes no paramétricos de tablas de mortalidad de la Comunidad Valenciana	111
6.1. Introducción	111
6.2. Graduación mediante kernel	111
6.3. Graduación mediante splines	120
6.4. Graduación polinomial local	125
6.5. Modelos aditivos generalizados	133
6.5.1. GAM con splines	133

6.5.2. GAM con loess	136
6.6. Comparación de los diferentes métodos	139
7. Tablas de mortalidad dinámicas para la Comunidad Valenciana	147
7.1. Introducción	147
7.2. Evolución de la mortalidad a lo largo del tiempo	147
7.3. Métodos estructurales	154
7.3.1. Modelo de Heligman y Pollard	154
7.3.2. Método de Lee-Carter	161
7.4. Métodos no estructurales	173
7.4.1. Modelos lineales generalizados con respecto a la edad y el tiempo . .	173
7.5. Comparación de los modelos	182
7.6. Métodos basados en factores de reducción de la mortalidad	182
7.6.1. Modelos lineales generalizados	185
8. Uso de las tablas de mortalidad en aplicaciones actuariales	191
8.1. Introducción	191
8.2. Rentas y seguros individuales	192
8.2.1. Determinación del valor de las rentas vitalicias	192
8.2.2. Seguros o capitales pagaderos por fallecimiento	196
8.2.3. Rentas fraccionadas	197
8.2.4. Análisis de sensibilidad y estabilidad	198
8.2.5. Aplicación	203
8.3. Planes de pensiones	213
8.3.1. Definición y conceptos básicos	213
8.3.2. Plan de pensiones de prestación definida	215
8.3.3. Plan de pensiones de aportación definida	216
8.3.4. Evaluación del coste de las pensiones mediante el sistema de capi- talización colectiva	218
8.3.5. Aplicación a la Universitat de València	219
8.3.6. Análisis de estabilidad y sensibilidad de los fondos de pensiones . . .	224
8.3.7. Aplicación	225
8.4. Conclusiones	230
Bibliografía	232

Índice de figuras

2.1. Diagrama de Lexis	14
2.2. Descomposición de la Ley de Heligman y Pollard	32
5.1. Estimaciones brutas de las probabilidades q_x	88
5.2. Comparación Modelos $LGM(0, s)$ para los hombres	93
5.3. Comparación Modelos $LGM(0, s)$ para las mujeres	93
5.4. $LGM(0, 11)$ para los hombres	94
5.5. $LGM(0, 11)$ para las mujeres	94
5.6. Comparación Modelos $GM(0, s)$ para los hombres	101
5.7. Comparación Modelos $GM(0, s)$ para las mujeres	101
5.8. Comparación de los q_x correspondientes a los Modelos $GM(0, s)$ para hombres	102
5.9. Comparación de los q_x correspondientes a los Modelos $GM(0, s)$ para mujeres	102
5.10. $GM(0, 11)$ para los hombres	103
5.11. $GM(0, 11)$ para las mujeres	103
5.12. Comparación Modelos Heligman y Pollard para los hombres	105
5.13. Comparación Modelos Heligman y Pollard para las mujeres	105
5.14. Autocorrelaciones del modelo $LGM(0, 11)$ para los hombres	108
5.15. Autocorrelaciones del modelo $LGM(0, 11)$ para las mujeres	108
5.16. Autocorrelaciones de las q_x correspondientes al modelo $GM(0, 11)$ para los hombres	109
5.17. Autocorrelaciones de las q_x correspondientes al modelo $GM(0, 11)$ para las mujeres	109
5.18. Autocorrelaciones del modelo H-P para los hombres	110
5.19. Autocorrelaciones del modelo H-P para las mujeres	110
6.1. Comparación estimación kernel para los hombres	112
6.2. Comparación estimación kernel para las mujeres	112
6.3. Residuos frente a estimación Nayadara-Watson para los hombres	114
6.4. Residuos frente a estimación Copas-Haberman para los hombres	114
6.5. Residuos frente a estimación Nayadara-Watson para las mujeres	115
6.6. Residuos frente a estimación Copas-Haberman para las mujeres	115
6.7. Nayadara-Watson con $b = 3$ para los hombres	117
6.8. Nayadara-Watson con $b = 3$ para las mujeres	117
6.9. Residuos frente a estimación del $logit(q_x)$ para los hombres	119

6.10. Residuos frente a estimación del $\text{logit}(q_x)$ para las mujeres	119
6.11. Smoothing mediante splines para los hombres	121
6.12. Smoothing mediante splines para las mujeres	121
6.13. Residuos frente a estimación de las transformaciones de q_x para los hombres	122
6.14. Residuos frente a estimación de las transformaciones de q_x para las mujeres	122
6.15. Transformación log para los hombres	124
6.16. Transformación log para las mujeres	124
6.17. Comparación de la regresión local polinomial para los hombres	126
6.18. Comparación de la regresión local polinomial para las mujeres	126
6.19. Residuos frente a estimación del $\text{logit}(q_x)$ para los hombres	130
6.20. Residuos frente a estimación del $\text{logit}(q_x)$ para las mujeres	130
6.21. Comparación regresión polinomial local sobre la transformación logit para los hombres	131
6.22. Comparación regresión polinomial local sobre la transformación logit para las mujeres	131
6.23. Comparación loess lineal y cuadrática con span elegido por validación cru- zada para los hombres	132
6.24. Comparación loess lineal y cuadrática con span elegido por validación cru- zada para las mujeres	132
6.25. Comparación de los GAM con splines para los hombres	137
6.26. Comparación de los GAM con splines para las mujeres	137
6.27. Comparación de los GAM con loess para los hombres	141
6.28. Comparación de los GAM con loess para las mujeres	141
6.29. Autocorrelaciones del modelo kernel para los hombres	143
6.30. Autocorrelaciones del modelo kernel para las mujeres	143
6.31. Autocorrelaciones del modelo splines para los hombres	144
6.32. Autocorrelaciones del modelo splines para las mujeres	144
6.33. Autocorrelaciones del modelo loess lineal para los hombres	145
6.34. Autocorrelaciones del modelo loess cuadrático para las mujeres	145
6.35. Autocorrelaciones del modelo GAM para los hombres	146
6.36. Autocorrelaciones del modelo GAM para las mujeres	146
7.1. Comparación de las probabilidades de muerte brutas	150
7.2. Comparación de estimaciones brutas de q_x para los hombres	152
7.3. Comparación de estimaciones brutas de q_x para las mujeres	152
7.4. Comparación de estimaciones graduadas de q_x para los hombres	153
7.5. Comparación de estimaciones graduadas de q_x para las mujeres	153
7.6. Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Heligman y Pollard para los hombres	159
7.7. Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Heligman y Pollard para los hombres	159
7.8. Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Heligman y Pollard para los hombres	159

7.9. Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Heligman y Pollard para las mujeres	160
7.10. Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Heligman y Pollard para las mujeres	160
7.11. Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Heligman y Pollard para las mujeres	160
7.12. Comparación de las estimaciones del parámetro a_x	164
7.13. Estimación del parámetro b_x del modelo de Lee-carter para los hombres	165
7.14. Estimación del parámetro b_x del modelo de Lee-carter para las mujeres	165
7.15. Comparación de las estimaciones del parámetro k_t	166
7.16. Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Lee-Carter para los hombres	169
7.17. Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Lee-Carter para los hombres	169
7.18. Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Lee-Carter para los hombres	169
7.19. Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Lee-Carter para las mujeres	170
7.20. Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Lee-Carter para las mujeres	170
7.21. Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Lee-Carter para las mujeres	170
7.22. Evolución de k_t para los hombres	171
7.23. Evolución de k_t para las mujeres	172
7.24. Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo lineal generalizado para los hombres	180
7.25. Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo lineal generalizado para los hombres	180
7.26. Evolución de $\log(q_x)$ para edades de 60 a 95 del modelo lineal generalizado para los hombres	180
7.27. Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo lineal generalizado para las mujeres	181
7.28. Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo lineal generalizado para las mujeres	181
7.29. Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo lineal generalizado para las mujeres	181
7.30. Bondad del ajuste $LGM(0, 10)$ para los hombres	183
7.31. Bondad del ajuste $LGM(0, 10)$ para las mujeres	183
7.32. Modelos $LGM(0, 10)$ para ambos sexos	185
7.33. Factores de mejora de la mortalidad de edades de 0 a 40 para los hombres	189
7.34. Factores de mejora de la mortalidad de edades de 45 a 85 para los hombres	189
7.35. Factores de mejora de la mortalidad de edades de 0 a 40 para las mujeres	190
7.36. Factores de mejora de la mortalidad de edades de 45 a 85 para las mujeres	190

8.1. Variación de D_x por cambio de interés del 3 al 3.2%	199
8.2. Variación del valor de la renta actual ante cambios en el tipo de interés con tabla estática para los hombres	203
8.3. Variación del valor de la renta actual ante cambios en el tipo de interés con tabla estática para las mujeres	204
8.4. Variación del valor de la renta actual ante cambios en el tipo de interés con tabla dinámica para los hombres	204
8.5. Variación del valor de la renta actual ante cambios en el tipo de interés con tabla dinámica para las mujeres	205
8.6. Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla estática para los hombres	205
8.7. Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla estática para las mujeres	206
8.8. Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla dinámica para los hombres	206
8.9. Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla dinámica para las mujeres	207
8.10. Variación de la renta cuando varían la probabilidades de forma constante para los hombres	207
8.11. Variación de la renta cuando varían la probabilidades de forma constante para las mujeres	208
8.12. Variación de la prima de seguro cuando varían la probabilidades de forma constante para los hombres	208
8.13. Variación de la prima de seguro cuando varían la probabilidades de forma constante para las mujeres	209
8.14. Variación de la renta cuando varían la probabilidades sumándoles una constante para los hombres	209
8.15. Variación de la renta cuando varían la probabilidades sumándoles una constante para las mujeres	210
8.16. Variación de la prima del seguro cuando varían la probabilidades sumándoles una constante para los hombres	210
8.17. Variación de la prima del seguro cuando varían la probabilidades sumándoles una constante para las mujeres	211
8.18. Variación de la renta calculada con tabla dinámica respecto a estática para los hombres	211
8.19. Variación de la renta calculada con tabla dinámica respecto a estática para las mujeres	212
8.20. Variación de la prima de seguro calculada con tabla dinámica respecto a estática para los hombres	212
8.21. Variación de la prima de seguro calculada con tabla dinámica respecto a estática para las mujeres	213
8.22. Distribución por edad del PAS	221
8.23. Distribución por edad del PDI	222
8.24. Aportaciones para Catedráticos de Universidad	226

8.25. Aportaciones para Catedráticos de Escuela	226
8.26. Aportaciones para Titulares de Universidad	227
8.27. Aportaciones para Titulares de Escuela Universitaria	227
8.28. Aportaciones para el grupo A	228
8.29. Aportaciones para el grupo B	228
8.30. Aportaciones para el grupo C	229
8.31. Aportaciones para el grupo D	229
8.32. Aportaciones para el grupo E	230

Índice de tablas

5.1. Medidas de bondad de ajuste para $LGM(0, s)$ con diferentes grados	90
5.2. Comparación de los modelos $LGM(0, s)$ con diferentes grados	91
5.3. Coeficientes del modelo $LGM(0, 11)$	92
5.4. Medidas de bondad de ajuste para $GM(0, s)$ con diferentes grados. Distribución Poisson	95
5.5. Comparación de los modelos $GM(0, s)$ con diferentes grados. Distribución Poisson	96
5.6. Medidas de bondad de ajuste para $GM(0, s)$ con diferentes grados. Distribución Gamma	98
5.7. Medidas de bondad de ajuste para $GM(r, 2)$ con diferentes grados	99
5.8. Coeficientes del modelo $GM(0, 11)$. Distribución Poisson	100
5.9. Comparación de los resultados según los diferentes criterios	106
5.10. Coeficientes de los modelos Heligman y Pollard	107
5.11. Resultados de los contrastes no paramétricos sobre los modelos paramétricos	107
6.1. Medidas de bondad de ajuste para suavización kernel sobre las probabilidades brutas	113
6.2. Contrastes no paramétricos para el estimador Nayadara-Watson sobre las probabilidades brutas	116
6.3. Medidas de bondad de ajuste para kernel sobre las transformadas de las probabilidades brutas	118
6.4. Medidas de bondad de ajuste para kernel sobre las transformadas logit de las probabilidades brutas para el bandwidth obtenido mediante validación cruzada	120
6.5. Medidas de bondad de ajuste para el ajuste mediante splines sobre las transformadas de las probabilidades brutas	123
6.6. Medidas de bondad de ajuste para loess sobre las probabilidades brutas	127
6.7. Medidas de bondad de ajuste para loess sobre las transformadas de las probabilidades brutas correspondientes a los hombres	128
6.8. Medidas de bondad de ajuste para loess sobre las transformadas de las probabilidades brutas correspondientes a las mujeres	129
6.9. Bandwidth y medidas de bondad de ajuste para loess sobre las transformadas logit de las probabilidades brutas	133

6.10. Medidas de bondad de ajuste para los modelos GAM con splines para hombres	134
6.11. Medidas de bondad de ajuste para los modelos GAM con splines para mujeres	134
6.12. Comparación de los modelos GAM con diferentes grados de libertad	135
6.13. Contrastes no paramétricos del modelo GAM con splines para los hombres .	136
6.14. Contrastes no paramétricos del modelo GAM con splines para las mujeres .	136
6.15. Medidas de bondad de ajuste para GAM con loess para cada sexo	138
6.16. Comparación de los modelos GAM con diferentes span	140
6.17. Resultados de los contrastes no paramétricos sobre el kernel y splines	142
7.1. Contrastes clásicos para la comparación de experiencias de mortalidad para los hombres	150
7.2. Contrastes clásicos para la comparación de experiencias de mortalidad para las mujeres	151
7.3. Coeficientes de la ley de Heligman y Pollard para los hombres	155
7.4. Coeficientes de la ley de Heligman y Pollard para las mujeres	156
7.5. Contrastes no paramétricos del modelo de Heligman y Pollard para los hombres	157
7.6. Contrastes no paramétricos del modelo de Heligman y Pollard para las mujeres	158
7.7. Modelos obtenidos para los parámetros del modelo de Heligman y Pollard .	161
7.8. Estimaciones de los parámetros del modelo de Lee-Carter	162
7.8. Estimaciones de los parámetros del modelo de Lee-Carter	163
7.8. Estimaciones de los parámetros del modelo de Lee-Carter	164
7.9. Estimaciones del parámetro k_t del modelo de Lee-Carter	166
7.10. Contrastes no paramétricos del modelo de Lee-Carter para los hombres . .	167
7.11. Contrastes no paramétricos del modelo de Lee-Carter para las mujeres . . .	168
7.12. Predicciones del índice de mortalidad k_t del modelo de Lee-Carter para los hombres	171
7.13. Predicciones del índice de mortalidad k_t del modelo de Lee-Carter para las mujeres	172
7.14. Deviance para los modelos lineales generalizados dinámicos	174
7.15. Diferencia de Deviance cuando introducimos interacciones en los modelos generalizados dinámicos	175
7.16. Coeficientes del modelo lineal generalizado para los hombres	176
7.17. Coeficientes del modelo lineal generalizado para las mujeres	177
7.18. Contrastes no paramétricos del modelo lineal generalizado para los hombres	178
7.19. Contrastes no paramétricos del modelo lineal generalizado para las mujeres	179
7.20. Comparación de los modelos $LGM(0, s)$ con diferentes grados para el periodo base 1990-1992	184
7.21. Estimaciones de los q_{x0} y factores de reducción	185
7.21. Estimaciones de los q_{x0} y factores de reducción	186
7.21. Estimaciones de los q_{x0} y factores de reducción	187
7.21. Estimaciones de los q_{x0} y factores de reducción	188
8.1. Medidas descriptivas de la edad según categorías	220

8.2. Tabla de salarios por categoría profesional 223

Introducción

El profesional del seguro de vida ha de ser capaz de asignar primas suficientes para cubrir las cantidades que habrá de pagar la compañía en el caso de muerte del asegurado. En consecuencia, la predicción adecuada de las probabilidades de muerte constituye uno de los ejes centrales de la reducción del riesgo que se asume.

Por ello el objetivo de esta tesis es una presentación exhaustiva de los métodos disponibles para ajustar tablas de mortalidad. Además pretendemos desarrollar una metodología que permita la elaboración de tablas específicas para cada uno de los colectivos asegurados a partir de datos observados provenientes de los mismos. En nuestra pretensión de aportar mejoras a la herramienta actuarial utilizada en la práctica real, nos parece imprescindible ocuparnos de todo lo relativo a las tablas de mortalidad dinámicas, puesto que la estacionariedad que el ajuste clásico exige como requisito previo es difícil de admitir cuando se estudia la mortalidad en un largo periodo de tiempo.

Un trabajo de estas características quedaría incompleto si cuanto se ha desarrollado a nivel teórico no se aplicara a datos reales que permitan valorar y comparar la bondad de los distintos métodos, por ello la última parte de la tesis esta dedicada a analizar los datos de mortalidad de la Comunidad Valenciana construyendo modelos que abarcan todo el rango de edades del colectivo y no solamente un rango parcial de las mismas. Con la tablas de mortalidad estimadas hemos elaborado un plan de pensiones para el profesorado funcionario de la Universitat de València con el que concluimos nuestro trabajo.

Con esta serie de objetivos, hemos organizado nuestro trabajo de la siguiente forma:

Capítulo 1.- Ofrecemos en él una visión global de los conceptos fundamentales en la definición de tablas de mortalidad, así como la clasificación de las mismas en estáticas, o de momento, y dinámicas.

Capítulo 2.- Está dedicado a la graduación de datos de mortalidad mediante métodos paramétricos, lo que supone ajustar una función paramétrica a las estimaciones de las probabilidades proporcionadas directamente por los datos. Se analizan en el capítulo las diferentes propuestas y los diversos contrastes para su comparación.

Capítulo 3.- Se ocupa de la graduación de datos de mortalidad mediante métodos no paramétricos, que suponen suavizar las estimaciones brutas de las probabilidades de muerte mediante técnicas de smoothing, splines, loess y GAM (modelos aditivos generalizados).

Capítulo 4.- Analizamos la dinamicidad del fenómeno de mortalidad a lo largo del tiempo del calendario introduciendo los modelos para la obtención de tablas dinámicas de mortalidad.

Capítulo 5.- La aplicación práctica de los modelos aportados en los anteriores capítulos comienza en este capítulo con la aplicación de las técnicas paramétricas del Capítulo 2 a datos de mortalidad de la Comunidad Valenciana correspondientes a los años 1999, 2000 y 2001, comparando los resultados obtenidos con cada una de ellas con el fin de elegir la que mejor ajuste proporciona.

Capítulo 6.- Con la misma estructura que el capítulo anterior, en este capítulo aplicamos las técnicas no paramétricas del Capítulo 3 a los datos de mortalidad de la Comunidad Valenciana.

Capítulo 7.- Para periodos de tiempo mayor, las tablas de mortalidad estáticas no ofrecen soluciones lo suficientemente buenas, lo que ha llevado al desarrollo de técnicas dinámicas que contemplen la evolución de la mortalidad en función del tiempo del calendario. Este capítulo se ocupa de la obtención de tablas de mortalidad dinámicas para datos de la Comunidad Valenciana correspondientes al periodo 1980-2000 y de la comparación de los resultados.

Capítulo 8.- Las tablas estáticas y dinámicas ajustadas a los datos de la Comunidad Valenciana son utilizadas en este capítulo para la obtención de rentas y seguros. En la segunda parte las mejores tablas se han utilizado para elaborar un plan de pensiones hipotético para el profesorado funcionario de la Universitat de Valencia. Se pretende analizar cómo influye el tipo de tabla, estática o dinámica, en valores actuariales de vida.

Capítulo 1

Análisis de supervivencia y tablas de mortalidad

1.1. Introducción

Como dice Betzuen (1989), antes de establecer un plan de prestaciones de un colectivo y la forma de garantizarlas es necesario realizar un estudio adecuado del colectivo desde el punto de vista demográfico. Se pretende pues investigar sobre conceptos y funciones estadísticas y actuariales que nos permitan representar de forma adecuada la evolución del número de miembros del colectivo y plasmar estos resultados en forma de herramienta capaz de pronosticar la proyección futura del colectivo.

Nuestro objetivo en el presente capítulo va a ser estudiar los modelos de probabilidad que intentan analizar el tiempo transcurrido hasta la valoración de un determinado evento. Más concretamente, vamos a analizar tiempos de supervivencia, es decir, la duración hasta que se produce la muerte. El análisis de este tipo de series de tiempo se conoce con el nombre genérico de *Análisis de Supervivencia*.

Para el estudio de la duración de un individuo dentro de una empresa, habríamos de observar un individuo hasta que fallezca o se jubile. Aunque sabemos que hay individuos que abandonan la empresa y dejamos de observarlos y por tanto representan un problema para el estudio, pues son observaciones incompletas. En este capítulo se desarrollan una serie de conceptos para describir y utilizar la distribuciones de las variables aleatorias tiempo de supervivencia T y edad de fallecimiento ξ asociada. Asimismo veremos como la distribución de la variable aleatoria ξ , se puede representar por medio de una *Tabla de supervivencia (mortalidad)*.

1.2. Principales funciones de la distribución del tiempo de supervivencia

Denotamos por x la edad de un individuo, que puede tomar cualquier valor entre cero y el límite superior de supervivencia.

Vamos a considerar una persona de edad x años y la denotaremos (x) . Denotaremos por T su tiempo futuro de supervivencia, o más explícitamente $T(x)$, y asociaremos la variable aleatoria ξ a la edad de fallecimiento, es decir, $T = \xi - x$. El tiempo futuro de supervivencia es una variable aleatoria con función de distribución de probabilidad

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

La función $G(t)$ representa la probabilidad que tiene una persona de morir dentro de t años. Supongamos que G es conocida y continua con densidad, $g(t) = G'(t)$. A partir de ella podemos definir la función de supervivencia

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

donde $G(0) = 0$, lo cual implica que $s(0) = 1$.

La función $s(t)$ se denomina *función de supervivencia*, ya que para cualquier valor positivo de t , $s(t)$ es la probabilidad que tiene una persona de sobrevivir t años.

De su definición se derivan las dos propiedades siguientes:

- Es una función no creciente.
- $s(0) = 1$ y en el extremo superior ω de x , $s(\omega) = 0$.

Además es conveniente y razonable suponer (Villalón, 1997) que $s(t)$ es una función continua de t .

La función de riesgo o tasa de hazard $h(t)$ se define como la probabilidad de que un individuo, que se supone vivo en el instante t , no sobreviva al siguiente intervalo de tiempo lo suficientemente pequeño

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T < t + \Delta t | T \geq t)}{\Delta t} \\ &= \frac{P(t < T < t + dt)}{P(T > t)} \\ &= \frac{g(t)}{s(t)}. \end{aligned}$$

De lo que podemos deducir

$$h(t) = \frac{-d \ln(s(t))}{dt}.$$

Probabilidades y valores esperados de interés pueden ser expresados en términos de las funciones g y G . Sin embargo, la comunidad internacional de actuarios usa una notación propia que mostramos a continuación, relacionándola con ambas funciones tal y como lo hace Gerber (1997). Así,

$${}_t q_x = G(t) = 1 - s(t)$$

es la probabilidad de que un individuo de edad x muera en t años. De igual forma

$${}_t p_x = 1 - G(t) = s(t), \tag{1.1}$$

denota la probabilidad de que un individuo de edad x sobreviva al menos t años. Otra notación utilizada habitualmente es

$${}_{s|t} q_x = P(s < T < s + t) = G(s + t) - G(s) = {}_{s+t} q_x - {}_s q_x,$$

que denota la probabilidad de que un individuo de edad x sobreviva s años y después muera dentro de los t años siguientes. De igual forma se usan frecuentemente

$${}_{s+t} p_x = 1 - G(s + t) = (1 - G(s)) \frac{1 - G(s + t)}{1 - G(s)} = {}_s p_x \cdot {}_t p_{x+s}$$

y

$${}_{s|t} q_x = G(s + t) - G(s) = (1 - G(s)) \frac{G(s + t) - G(s)}{1 - G(s)} = {}_s p_x \cdot {}_t q_{x+s}.$$

Si $t = 1$, el índice t se omite en los símbolos, por ejemplo q_x es la probabilidad de morir dentro de un año. La fuerza de mortalidad de (x) a la edad $x + t$ esta definida por

$$\mu_{x+t} = \frac{g(t)}{1 - G(t)} = -\frac{d}{dt} \ln(1 - G(t)).$$

Por tanto

$$\mu_{x+t} = -\frac{d}{dt} \ln({}_t p_x),$$

de forma que

$${}_t p_x = \exp\left(-\int_0^t \mu_{x+s} ds\right).$$

1.3. Estimación de las curvas de supervivencia

Para la estimación de las curvas de supervivencia tenemos fundamentalmente dos enfoques, que como veremos dan lugar a su vez a modelos específicos. El primer enfoque es postular una distribución de probabilidad sobre la variable tiempo futuro de supervivencia T . El segundo, que tiene mucha más relación con el mundo actuarial, supone la construcción de una tabla de mortalidad.

1.3.1. Distribuciones analíticas del tiempo de supervivencia

La modelización de T , mediante la explicitación de su función de distribución G , tiene la ventaja de permitir su estimación mediante un reducido número de parámetros. Ventaja nada desdeñable cuando se dispone de pocos datos. A lo largo del tiempo diversos autores han propuesto modelos para el comportamiento probabilístico de T . Algunos de ellos los exponemos a continuación (Gerber, 1997).

- De Moivre (1724) postula la existencia de una edad ω máxima y supone que T esta distribuida uniformemente entre las edades 0 y $\omega - x$, de forma que

$$g(t) = \frac{1}{\omega - x}, \quad 0 < t < \omega - x,$$
$$\mu_{x+t} = \frac{1}{\omega - x - t}, \quad 0 < t < \omega - x.$$

- Gompertz (1825) supone que la fuerza de mortalidad crece exponencialmente

$$\mu_{x+t} = Bc^{x+t}, \quad t > 0,$$

lo que expresa mejor el comportamiento de T y además no necesita la hipótesis de la edad máxima ω .

- Makeham (1860) añade una componente constante $A > 0$ al crecimiento exponencial, con lo que postula la ley siguiente

$$\mu_{x+t} = A + Bc^{x+t}, \quad t > 0.$$

La probabilidad de supervivencia en este modelo es

$${}_t p_x = \exp\left(-At - \frac{B}{\ln c} c^x (c^t - 1)\right).$$

- Weibull (1939) sugiere que la fuerza de mortalidad crece como una potencia de t en lugar de hacerlo exponencialmente

$$\mu_{x+t} = k(x+t)^n,$$

con parámetros fijos $k > 0$ y $n > 0$. La probabilidad de supervivencia se convierte entonces

$${}_t p_x = \exp\left(-\frac{k}{n+1} ((x+t)^{n+1} - x^{n+1})\right).$$

Otros autores, (Benjamin y Pollard, 1992), que teniendo en cuenta que una sola ley no pueden representar toda la experiencia de mortalidad, proponen modelos matemáticos que relacionan la fuerza de mortalidad y la edad en cada rango de edad.

- Thiele (1972) propone,

$$\mu_x = a_1 \exp(-b_1 x) + a_2 \exp\left(-\frac{1}{2} b_2 (x - c)^2\right) + a_3 \exp(b_3 x),$$

donde el primer término representa la mortalidad infantil, el último, que es una curva Gompertz, corresponde a la mortalidad para edades avanzadas y el central es una curva normal.

- Perks (1825) introduce una nueva familia de curvas cuya expresión general es,

$$\mu_x = \frac{A + Bc^x}{Kc^{-x} + 1 + Dc^x}.$$

- Estas leyes son sólo aplicables a las edades adultas, y muchas fallan al representar la joroba de los accidentes en las edades adultas. Heligman y Pollard (1980) obtuvieron resultados prometedores para la total amplitud del intervalo de vida con la ley

$$q_x/p_x = A^{(x+B)^c} + D \exp(-E(\log x - \log F)^2) + GH^x.$$

El número de parámetros puede parecer excesivo, sin embargo tienen una interpretación real. A , por ejemplo, es q_1 . C mide el ratio con el que los niños se adaptan al entorno. G indica el nivel de mortalidad de las edades elevadas, mientras H mide el incremento de esa mortalidad. D representa la intensidad de la joroba de los accidentes, mientras F indica la situación de la joroba y E su velocidad.

1.3.2. Tablas de mortalidad

Sea $K = K(x)$ una variable aleatoria discreta asociada al tiempo de vida futuro, que representa el número entero de años futuros vividos. Es decir, $K = \lfloor T \rfloor$.

Su distribución de probabilidad viene dada por

$$P(K = k) = P(k \leq T < k + 1) = {}_k p_x q_{x+k}, \quad k = 0, 1, 2, \dots$$

El valor esperado de K es la esperanza de vida abreviada y se obtiene,

$$e_x = \sum_{k=1}^{\infty} k P(K = k) = \sum_{k=1}^{\infty} k {}_k p_x q_{x+k}$$

o

$$e_x = \sum_{k=1}^{\infty} P(K \geq k) = \sum_{k=1}^{\infty} {}_k p_x$$

Sea S la fracción del año de su muerte durante la cual (x) vive, por tanto $T = K + S$. S es una variable aleatoria que tiene distribución continua entre 0 y 1. Aproximando su valor esperado por $1/2$, tenemos que

$$\dot{e}_x = E[T] \approx e_x + \frac{1}{2}.$$

La distribución de probabilidad del tiempo de vida futuro puede ser construida a partir de lo que denominamos una tabla de mortalidad. Una tabla de mortalidad es esencialmente una tabla que muestra la probabilidades de morir en un año, q_x , que definen completamente la distribución de K .

La distribución de T puede obtenerse a partir de una tabla de mortalidad mediante interpolación para lo cual son necesarias hipótesis sobre los patrones que siguen las probabilidades ${}_u q_x$, o la fuerza de mortalidad, μ_{x+u} , para edades intermedias $x + u$ (con x entero y $0 < u < 1$). Veamos alguna de ellas (Gerber, 1997).

a. Linealidad de ${}_uq_x$

Si suponemos que ${}_uq_x$ es una función lineal de u , la interpolación entre $u = 0$ y $u = 1$ conduce a

$${}_uq_x = uq_x,$$

luego

$${}_up_x = 1 - uq_x,$$

y

$$\mu_{x+u} = \frac{q_x}{1 - uq_x}.$$

b. μ_{x+u} constante

Vamos a denotar el valor constante de μ_{x+u} entre $0 < u < 1$ como $\mu_{x+\frac{1}{2}}$. Entonces

$$\mu_{x+\frac{1}{2}} = -\ln p_x,$$

de lo que se sigue,

$${}_up_x = e^{u\mu_{x+\frac{1}{2}}} = (p_x)^u.$$

Se deduce de aquí que la distribución de S , dado $K = k$, es una distribución exponencial truncada que depende de k ,

$$P(S \leq u | K = k) = \frac{1 - p_{x+k}^u}{1 - p_{x+k}}.$$

Las variables S y K no son independientes en este caso.

c. Linealidad de ${}_{1-u}q_{x+u}$

Esta hipótesis se conoce como la hipótesis de Balducci. A semejanza de lo que ocurrió en a)

$${}_{1-u}q_{x+u} = (1 - u)q_x.$$

De forma que,

$${}_up_x = \frac{p_x}{1 - uq_{x+u}} = \frac{1 - q_x}{1 - (1 - u)q_x}$$

y

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

Finalmente,

$$P(S \leq u | K = k) = \frac{u}{1 - (1 - u)q_{x+k}},$$

muestra que tampoco ahora las variables aleatorias S y K son independientes.

Observemos que en los tres supuestos la fuerza de mortalidad es discontinua en los valores enteros, pero lo más llamativo y poco creíble es que bajo la hipótesis de Balducci la fuerza de mortalidad decrece entre dos enteros consecutivos.

Para $q_{x+k} \rightarrow 0$, es decir para probabilidades pequeñas de muerte, en las hipótesis b) o c) S se distribuye aproximadamente uniforme e independiente de K .

Siguiendo a Betzuen (1995), que trata la problemática de los tantos interanuales, podemos resumir las hipótesis anteriores en el siguiente cuadro, en el que $t \in [0, 1]$ y l_x es el número de supervivientes con edad x .

hipótesis a)	hipótesis b)	hipótesis c)
${}_tq_x = a + bt$	$\mu_{x+t} = \mu$	${}_{1-t}q_x = a + bt$
${}_tq_x = tq_x$	${}_tq_x = 1 - \exp(-\mu t)$	${}_{1-t}q_{x+t} = (1-t)q_x$
$l_{x+t} = l_x - tq_x$	$l_{x+t} = l_x \exp(-\mu t)$	$l_{x+t} = \frac{l_x l_{x+t}}{l_{x+1} + td_x}$
${}_{1-h}q_{x+t} = \frac{(1-h)q_x}{1-tq_x}$		${}_{1-h}q_{x+t} = \frac{(1-h)q_x}{1-(h-t)q_x}$

Como consecuencia de los modelos planteados, en la hipótesis a) la función de supervivencia decrece linealmente dentro del intervalo anual, en la b) el tanto instantáneo de fallecimiento es constante y en la c) dicho tanto es decreciente. Además siguiendo con el mismo artículo de Betzuen y con el fin de poder clasificar a los individuos según su edad, es importante resaltar que la asignación de la edad de un individuo incide en los resultados del estudio. En general, existen diferentes criterios para ello, que son:

Edad actuarial Se trata de un método muy popular entre los actuarios y consiste en atribuir como edad de fallecimiento la edad entera más próxima al cumpleaños. Se asignaría edad x a todos los que tuvieran una edad comprendida en el intervalo $[x - 1/2, x + 1/2]$.

Edad entera alcanzada Consiste en atribuir la edad como número de años enteros vividos, es decir, como normalmente consideramos la edad. Se asignaría edad x a todos los que tuvieran una edad comprendida en el intervalo $[x, x + 1[$.

Existen también otras reglas pero que se utilizan poco en la práctica.

1.4. Contenido y clasificación de las tablas de mortalidad

En un principio, los contratos en los seguros solían ser a corto plazo. El problema de los contratos de larga duración y el establecimiento de las reservas apropiadas, exigieron establecer una teoría de la probabilidad, una estadística de la mortalidad y un instrumento matemático adecuado.

Los orígenes de estas tablas los podemos encontrar referenciados en Nieto y Vegas (1993). En su libro podemos leer que Halley (1693), elabora el primer trabajo conocido de tablas de mortalidad completas a partir de la hipótesis de estacionariedad de la población, que en la siguiente sección se explica con detalle. Posteriormente, Nicolás Titens, Jorge Barret y F. Bayly idearon los símbolos de conmutación que permitieron una agilización definitiva en el cálculo de las operaciones de seguro.

1.4.1. Presentación de una tabla de mortalidad

Según Palacios (1996) una tabla de mortalidad es una serie temporal que indica la reducción paulatina de un grupo inicial de individuos debido a los fallecimientos. Así pues, lo que realmente contiene son los que sobreviven.

La tabla de mortalidad es una abstracción matemática que representa un modelo de comportamiento de la evolución y constante decrecimiento de un colectivo, construida a partir de las observaciones de un colectivo real.

Su estructura básica, como nos describe Villalón (1994), esta constituida por las siguientes columnas,

x	l_x	d_x	q_x	p_x
-----	-------	-------	-------	-------

\mathbf{x} : La edad del individuo, $0 \leq x \leq \omega$, donde ω es la edad límite.

\mathbf{l}_x : Número de los que sobreviven a la edad x

\mathbf{d}_x : Número de los que fallecen entre x y $x + 1$,

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

\mathbf{q}_x : Tanto anual de fallecimiento a la edad x ,

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

\mathbf{p}_x : Tanto anual de supervivencia a la edad x

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

Las tablas también contienen los símbolos de conmutación D_x , N_x , S_x , C_x , M_x y R_x . Dichos símbolos son relaciones que facilitan enormemente los cálculos de primas, reservas y demás elementos correspondientes a las operaciones de seguros. Están calculados a un determinado tipo de interés, denominado tipo de interés técnico i , a partir del cual obtenemos v^x llamado factor de actualización o factor de descuento compuesto. Éste tiene la misión de actualizar un capital futuro, despojándolo de sus intereses y convertirlo en el capital inicial.

$$v^x = \frac{1}{(1+i)^n}$$

Las expresiones para el cálculo de estos símbolos son las siguientes:

- $\mathbf{D}_x = l_x v^x$
- $\mathbf{N}_x = D_x + D_{x+1} + \dots + D_\omega$
- $\mathbf{S}_x = N_x + N_{x+1} + \dots + N_\omega$
- $\mathbf{C}_x = d_x v^{x+1}$
- $\mathbf{M}_x = C_x + C_{x+1} + \dots + C_\omega$
- $\mathbf{R}_x = M_x + M_{x+1} + \dots + M_\omega$

Así pues de las tablas de mortalidad para determinados colectivos, pueden obtenerse las probabilidades de eliminación de diversas situaciones entre dos edades consecutivas, expresadas en años enteros. Estas probabilidades pueden también discriminarse por sexos.

1.4.2. Tablas Estáticas y Dinámicas

El fenómeno de la supervivencia viene caracterizado porque sus sucesos hacen referencia al hecho de que un individuo cualquiera perteneciente a un grupo específico, alcance y supere una edad concreta.

Al intentar modelizar el fenómeno de la supervivencia, aparece irremediamente como parámetro fundamental la edad, a la que denominaremos tiempo biológico para diferenciarla del tiempo cronológico que es el tiempo físico o del calendario (el carácter relativo de la dimensión temporal se ve con claridad al comparar un individuo de 40 años en 1900 y otro con la misma edad en 1995).

Los postulados recogidos en Vegas (1982), que constituyen la base fundamental de las deducciones que han de conducirnos a la construcción de una tabla de mortalidad, son:

Principio de Homogeneidad. Todos los individuos del grupo son equivalentes en lo que se refiere a mortalidad, en el sentido de que tienen la misma función de distribución de probabilidad para la variable edad de muerte ξ . El grupo es homogéneo.

Principio de Independencia. Los individuos que integran el grupo en cuestión se definen con variables estocásticamente independientes. Esto equivale a decir que el suceso de que un cierto individuo sobreviva o no a una determinada edad, tiene una probabilidad que no depende de la supervivencia de cualquier otro individuo del grupo.

Principio de Estacionariedad. La probabilidad de un individuo de no sobrevivir a una edad concreta es independiente del año de su cálculo.

Con todas estas hipótesis la probabilidad de que n individuos no sobrevivan a las edades x_1, x_2, \dots, x_n respectivamente viene dada por

$$\begin{aligned} P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n) &= G_{x_1}(x_1 - \xi_1) \cdot G_{x_2}(x_2 - \xi_1) \dots \cdot G_{x_n}(x_n - \xi_n) \\ &= {}_{x_1 - \xi_1}q_{x_1} \cdot {}_{x_2 - \xi_2}q_{x_2} \dots \cdot {}_{x_n - \xi_n}q_{x_n} \end{aligned}$$

Es evidente que si el estudio del fenómeno de la supervivencia se refiere solo al tiempo biológico es porque se admite, implícitamente, la hipótesis de estacionariedad del fenómeno. En consecuencia, si todas las consideraciones y formulaciones que se hacen vienen referidas al tiempo biológico o edad, con exclusión de toda referencia al tiempo cronológico, construimos una *tabla de mortalidad estática o de momento*.

Pero un estudio completo debería abarcar ambos conceptos temporales, puesto que, en su formulación más general la expresión matemática del fenómeno de la supervivencia depende de dos parámetros que son el tiempo físico (cronológico) y el tiempo biológico (edad). Con semejante planteamiento se obtienen las *tablas de mortalidad dinámicas*.

Como reflexión a la comparación teórica de las tablas estáticas y dinámicas, hemos de añadir que las primeras nacen con una fecha de caducidad implícita, puesto que la mortalidad desciende y la esperanza de vida aumenta con el paso de los años, de forma que necesitaríamos pedirle al asegurado una dotación adicional cuando pasaran un número determinado de años, mientras que con las segundas las posibles modificaciones serían menores.

En la tablas PERM/F 2000 recientemente elaboradas en España, se construye una tabla para cada generación, entendiendo por generación individuos nacidos todos en el mismo año. La tabla base que se toma para la formación de todas éstas es la llamada tabla dinámica, con los dos parámetros fundamentales: la edad y el tiempo cronológico.

Capítulo 2

Graduación de datos de mortalidad: métodos paramétricos

2.1. Introducción

Históricamente la Ciencia Actuarial ha trabajado con los datos de mortalidad de una población. El primer paso y quizá una de las partes fundamentales en las que interviene la estadística es la graduación de los mismos.

Definimos la graduación (Haberman y Renshaw, 1996) como los principios y métodos por los que un conjunto de probabilidades observadas (o brutas) se ajustan para proporcionar una base suavizada que nos permitirá hacer inferencias y además cálculos prácticos de primas, reservas, etc.... El objetivo de este capítulo es una de las principales aplicaciones de la graduación, la construcción de modelos de supervivencia que presentaremos finalmente en forma de tabla de mortalidad.

La graduación es necesaria y tiene una naturaleza eminentemente estadística de estimación. London (1985) explica la razón de por qué, obtenida nuestra secuencia de estimaciones iniciales, hemos de cambiarla y por tanto graduarla. Esto es debido a que para cada período concreto, dados los correspondientes datos, podemos obtener la secuencia de estimaciones iniciales correspondientes a las edades, estimación inicial que posee cambios a veces incluso bruscos, pero es una particular concreción de la evolución de la mortalidad en la que debería cumplirse que la diferencia entre las probabilidades de muerte de dos edades consecutivas no fuera excesivamente alta, lo que exige el ajuste de una función que cumpla con dicha condición.

Llegados a este punto, graduación y smoothing pueden considerarse sinónimos, pero el objetivo de aquella no es exactamente obtener una curva suave, sino estimar más adecuadamente las verdaderas probabilidades de muerte. Los métodos de graduación sugeridos en la literatura y usados en la práctica se pueden clasificar en dos tipos fundamentales, paramétricos y no paramétricos, según ajusten los datos a una función matemática o por el contrario sólo se realice un suavizado.

El desarrollo del presente capítulo se expone a continuación. En primer lugar describiremos las distribuciones de las principales medidas de mortalidad, q_x y μ_x , ya definidas

en el capítulo anterior.

En la sección 3 definimos graduación paramétrica, los pasos a seguir para su obtención y las opciones revisadas en cuanto a criterios de optimización se refiere.

Posteriormente en la sección 4 mostramos la metodología desarrollada por el Continuous Mortality Investigation (CMI) Bureau, cuya descripción se encuentra pormenorizada en el artículo de Forfar, McCutcheon y Wilkie (1988).

La extensión de estos conceptos a modelos lineales generalizados usando el esquema de Haberman y Renshaw (1996) y Verrall (1996) se expone en la sección 5.

En la sección 6 describimos los pasos a seguir en el proceso de graduación paramétrica con el modelo de Heligman y Pollard.

Con todo ello podemos obtener modelos del comportamiento de las medidas de la mortalidad frente a la edad, con diferentes metodologías que podemos discutir y comparar. Finalmente en la sección 7 recogemos los principales estadísticos y tests para evaluar los diferentes procedimientos.

2.2. Descripción del problema

Una vez fijado el periodo o duración del estudio considerado, para estimar los q_x debemos observar los individuos desde la edad x hasta la $x + 1$. Si todas la observaciones fueran completas, entendiendo por tal que han sido observadas desde la edad x hasta la edad $x + 1$, nuestro análisis sería sencillo, pero desgraciadamente no ocurre así. El diagrama de Lexis, que mostramos a continuación, nos ayudará a ilustrar las diferentes situaciones que pueden presentarse.

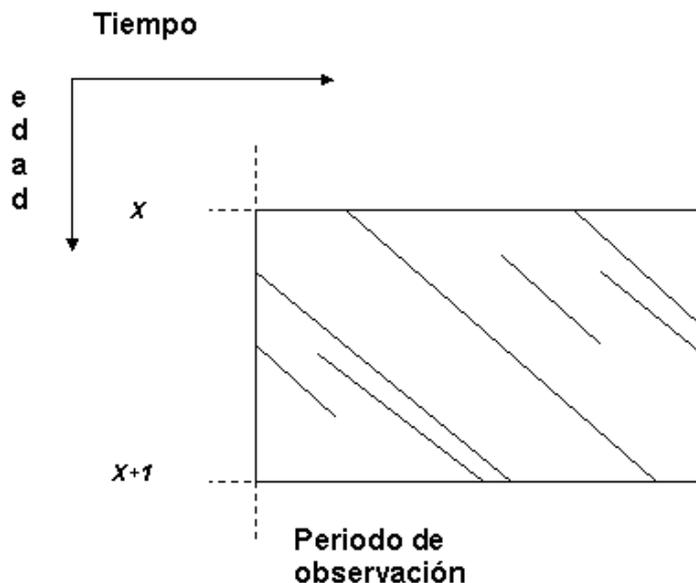


Figura 2.1: Diagrama de Lexis

El eje horizontal representa el tiempo de observación y el vertical las edades. Cada individuo queda representado por una línea diagonal que comienza cuando el individuo entra en observación y finaliza cuando muere o deja de ser observado. Sea $x + t_i$ la edad en la que entra en el estudio y $x + s_i$ edad a la que sale con $0 \leq t_i \leq s_i \leq 1$. Por tanto, la duración de cada uno es $s_i - t_i$ y la suma $E_x = (s_1 - t_1) + (s_2 - t_2) + \dots + (s_n - t_n)$ es la duración total de los n individuos.

Sea D_x el número observado de muertes, y sea I el conjunto de subíndices correspondientes a los individuos que acaban falleciendo. Para estimar las probabilidades de muerte para cada una de las edades a partir de los datos tenemos varias opciones.

1. Un primer método consiste en igualar las muertes esperadas a las observadas. El número esperado de muertes es

$$\sum_{i=1}^n {}_{1-t_i}q_{x+t_i} - \sum_{i \notin I} {}_{1-s_i}q_{x+s_i}, \quad (2.1)$$

pues sumando todas las ${}_{1-t_i}q_{x+t_i}$, probabilidad de que un individuo de edad $x + t_i$ no alcance la edad $x + 1$, tendríamos la duración si observáramos todas las muertes. A este total se le resta el tiempo que hemos contado de más para los que no observamos su muerte, es decir la suma de las probabilidades de que no alcancen $x + 1$ para los que cumplen $x + s_i$ (${}_{1-s_i}q_{x+s_i}$).

Utilizando la hipótesis de Balducci (capítulo 1, sección 3.2) ${}_{1-u}q_{x+u} = (1 - u)q_x$, y teniendo en cuenta que E_x puede expresarse de la forma

$$\begin{aligned} E_x &= (1 - t_1) - (1 - s_1) - (1 - t_2) - (1 - s_2) - \dots - (1 - t_n) - (1 - s_n) \\ &= \sum_{i=1}^n (1 - t_i) - \sum_{i=1}^n (1 - s_i), \end{aligned} \quad (2.2)$$

la expresión (2.1) se convierte en

$$\begin{aligned} &\sum_{i=1}^n (1 - t_i)q_x - \sum_{i \notin I} (1 - s_i)q_x = \\ &= q_x \left(\sum_{i=1}^n (1 - t_i) - \sum_{i \notin I} (1 - s_i) \right) \\ &= q_x \left(\sum_{i=1}^n (1 - t_i) + \sum_{i \in I} (1 - s_i) - \sum_{i \in I} (1 - s_i) - \sum_{i \notin I} (1 - s_i) \right) \\ &= E_x + \sum_{i \in I} (1 - s_i)q_x. \end{aligned}$$

Procedemos entonces a igualar

$$D_x = E_x q_x + \sum_{i \in I} (1 - s_i) q_x,$$

obteniendo el estimador clásico de q_x

$$\hat{q}_x = \frac{D_x}{E_x + \sum_{i \in I} (1 - s_i)}. \quad (2.3)$$

Al denominador de esta expresión es al que habitualmente se denomina *expuestos inicialmente al riesgo*, que suponiendo la hipótesis de que las muertes ocurren a edad $x + \frac{1}{2}$, aproximadamente en la media de x a $x + 1$, puede aproximarse por

$$E_x + \frac{1}{2} D_x. \quad (2.4)$$

Al valor E_x se le denomina *expuestos centrales al riesgo*.

El estimador (2.3) funciona bien con un número elevado de datos pero no con pocos datos. Uno de los problemas es que el denominador puede exceder al numerador, proporcionando una estimación absurda e inservible.

2. Un método alternativo consiste en partir de la hipótesis de que la fuerza de la mortalidad es constante $\mu_{x+u} = \mu_{x+\frac{1}{2}}$ para $0 < u < 1$.

El número de muertes esperadas sería $E_x \mu_{x+\frac{1}{2}}$, igualando de nuevo a las observadas obtenemos un estimador para la fuerza de mortalidad

$$\hat{\mu}_{x+\frac{1}{2}} = \frac{D_x}{E_x},$$

a partir del cual podemos obtener una estimación para la probabilidad q_x

$$\hat{q}_x = 1 - \exp(\mu_{x+\frac{1}{2}}) = 1 - \exp\left(\frac{D_x}{E_x}\right).$$

En realidad tanto D_x como E_x son variables aleatorias. Sin embargo suele ser conveniente tratar a E_x como una cantidad no aleatoria, aproximando su valor a partir de los datos observados de población y muertes por edades, y a partir del valor fijado para E_x obtener la distribución de D_x .

2.2.1. Distribución de D_x según q_x

Supongamos que E_x personas entran en observación a edad exacta x y continúan en observación hasta que sobreviven a una edad exacta $x + 1$ o mueren antes. En este caso denotamos por E_x los expuestos iniciales al riesgo, que es una estimación de las

observaciones que inician el estudio y determina por tanto el número de individuos en estudio. Supongamos también que la probabilidad de muerte dentro del año para cada uno de ellos es q_x , y que la muerte o supervivencia de uno es independiente de la muerte o supervivencia de los otros. Si llamamos D_x a la variable aleatoria que representa el número de muertes que ocurren en el año, D_x se distribuye como una binomial de parámetros E_x y q_x .

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

$$\hat{q}_x = \frac{D_x}{E_x}.$$

Si D_x toma un valor concreto d_x , número observado de muertes, podemos calcular $\hat{q}_x = \frac{d_x}{E_x}$ estimaciones a las que denominaremos probabilidades brutas.

Cuando el número de muertes observadas, es suficientemente elevado respecto del valor estimado de E_x , la aproximación Normal es generalmente satisfactoria. En estas condiciones un intervalo de confianza $100(1 - 2\alpha)$ para q_x viene dado por

$$\left[\hat{q}_x - z_\alpha \left(\frac{\hat{q}_x(1 - \hat{q}_x)}{E_x} \right)^{1/2}, \hat{q}_x + z_\alpha \left(\frac{\hat{q}_x(1 - \hat{q}_x)}{E_x} \right)^{1/2} \right].$$

Cuando d_x es muy pequeño o muy grande respecto de E_x esta aproximación no es buena y deberíamos obtener el intervalo exacto, $q_x^{inf} \leq q_x \leq q_x^{sup}$ con

$$q_x^{inf} \quad \text{tal que} \quad \sum_{d=d_x}^{E_x} \frac{E_x!}{d!(E_x - d)!} (q_x^{inf})^d (1 - q_x^{inf})^{E_x - d} = \alpha$$

$$q_x^{sup} \quad \text{tal que} \quad \sum_{d=0}^{d_x} \frac{E_x!}{d!(E_x - d)!} (q_x^{sup})^d (1 - q_x^{sup})^{E_x - d} = \alpha.$$

2.2.2. Distribución de D_x según μ_x

Supongamos que E_x personas entran en observación bajo la hipótesis de que la fuerza de mortalidad es constante $\mu_{x+\frac{1}{2}}$ durante el periodo de observación y que la muerte o supervivencia de cada uno es independiente. En este caso E_x representa los expuestos centrales al riesgo, que pueden modificarse a lo largo de la duración del estudio, con lo que el número de individuos en estudio no está determinado y por tanto D_x , número de muertes que ocurren en el periodo de observación, tendrá una distribución Poisson con media y varianza igual a $E_x \mu_{x+\frac{1}{2}}$. De donde obtenemos que el estimador máximo-verosímil de $\mu_{x+\frac{1}{2}}$ es

$$\hat{\mu}_{x+\frac{1}{2}} = \frac{D_x}{E_x},$$

cuyas propiedades de insesgaredad y consistencia son ampliamente conocidas.

Si el número de muertes es razonablemente grande, mayor de 5, entonces D_x tiene una distribución aproximadamente Normal de media y varianza $\mu_{x+\frac{1}{2}}E_x$ y podemos obtener un intervalo de confianza $100(1 - 2\alpha)$ para $\mu_{x+\frac{1}{2}}$ cuyos extremos son

$$\left[\hat{\mu}_{x+\frac{1}{2}} - z_\alpha \left(\frac{\hat{\mu}_{x+\frac{1}{2}}}{E_x} \right)^{1/2}, \hat{\mu}_{x+\frac{1}{2}} + z_\alpha \left(\frac{\hat{\mu}_{x+\frac{1}{2}}}{E_x} \right)^{1/2} \right].$$

De forma análoga al caso anterior, cuando la aproximación no es buena se puede obtener el intervalo exacto $\mu_{x+\frac{1}{2}}^{inf} \leq \mu \leq \mu_{x+\frac{1}{2}}^{sup}$ con

$$\mu_{x+\frac{1}{2}}^{inf} \quad \text{tal que} \quad \sum_{d=d_x}^{\infty} \exp\left(-E_x \mu_{x+\frac{1}{2}}^{inf}\right) \frac{\left(E_x \mu_{x+\frac{1}{2}}^{inf}\right)^d}{d!} = \alpha$$

$$\mu_{x+\frac{1}{2}}^{sup} \quad \text{tal que} \quad \sum_{d=0}^{d_x} \exp\left(-E_x \mu_{x+\frac{1}{2}}^{sup}\right) \frac{\left(E_x \mu_{x+\frac{1}{2}}^{sup}\right)^d}{d!} = \alpha.$$

2.2.3. Distribución de E_x

Sin embargo es también posible tratar a D_x como una cantidad no aleatoria fijando su valor a partir del número de muertes observadas d_x , y entonces asumir que E_x sigue una distribución gamma con parámetros $\alpha = D_x$ y $\beta = \mu_{x+\frac{1}{2}}$. Esta aproximación es compatible con la expresión de la función de verosimilitud

$$\left(\mu_{x+\frac{1}{2}}\right) \exp\left(-\mu_{x+\frac{1}{2}} E_x\right), \quad (2.5)$$

que habitualmente se identifica con la distribución de las D_x como variables Poisson cuando E_x se trata como una cantidad no aleatoria. Esta interpretación la plantea Gerber (1997) y es utilizada por Renshaw, Haberman y Hatzopoulos (1997) para introducir un modelo lineal generalizado que explicamos más adelante.

2.3. Métodos paramétricos

La representación de los datos de mortalidad via modelos paramétricos ha atraído la atención de actuarios, demógrafos y estadísticos a lo largo de todo el siglo pasado.

Los métodos paramétricos se caracterizan porque se basan en la hipótesis de que la medida de mortalidad elegida, q_x o μ_x , es función de x la edad, $f_\alpha(x)$, con $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ parámetros a determinar.

Estos métodos permiten obtener la graduación, determinando la función matemática supuesta, una vez calculadas las estimaciones brutas a partir de los datos de población y mortalidad.

Así pues, la técnica descrita aquí no es sólo una graduación con aplicaciones actuariales, sino técnicas de regresión ampliamente descritas y utilizadas en la literatura estadística

que ahora aplicamos con fines concretos de tratamiento de datos de mortalidad. Este método tiene muchas ventajas sobre otras formas de describir los datos, particularmente porque facilita las comparaciones a lo largo del tiempo y del espacio, por ejemplo, entre cohortes, periodos y regiones. En cuanto a una revisión de modelos demográficos, Congdon (1993) describe en detalle las ventajas de los métodos paramétricos: suavidad, parsimonia, interpolación, comparaciones, tendencias, predicciones y manipulación analítica.

Al aplicar estos procedimientos el objetivo es conseguir el mejor ajuste posible con el mínimo número de parámetros. Es por tanto necesario alcanzar un equilibrio entre el número de parámetros y la bondad del ajuste. Muchos estudios de graduación demográfica han enfatizado la bondad del ajuste sin considerar la estabilidad estadística de los parámetros en términos de la regresión no lineal asociada (es decir, en términos de los errores estándar y correlaciones entre los parámetros, especificación de los errores y una apropiada forma de la función de pérdida, validación del modelo, etc.). Asociado a este enfoque suele estar la sobreparametrización del modelo, que puede ponerse de manifiesto cuando se observan errores estándar demasiado grandes para dar la significación estadística del parámetro, altas correlaciones entre los parámetros y fallos de la convergencia en las rutinas iterativas de ajuste no lineal. Además la sobreparametrización tiene implicaciones prácticas sobre el uso de la graduación, por ejemplo en la comparación de series temporales de los parámetros a lo largo del tiempo y en la predicción para futuros años, en que los parámetros que se obtienen para diferentes años pueden mostrar fluctuaciones erráticas irregulares que dificulten la predicción. Hay una fuerte relación entre la sobreparametrización y la inestabilidad de los parámetros a lo largo del tiempo, por lo que existen razones estadísticas para preferir funciones parsimoniosas (con pocos parámetros) aun a costa de ligeras pérdidas de bondad de ajuste.

La forma de las funciones matemáticas que se ajustan a los datos son diversas y vienen sugeridas fundamentalmente por el perfil que presentan las estimaciones en los estudios realizados por diferentes autores clásicos como Gompertz, Makeham y otros que citamos en el primer capítulo. Las propuestas de estos autores dieron buenos resultados para datos de finales del siglo XIX y principios del XX. Más tarde cambió el patrón básico, aumentó la mortalidad en edades jóvenes y una relativa joroba de mortalidad en edades medias, de forma que era difícil obtener una buena graduación con la fórmula de Makeham y esto dió lugar a nuevos modelos como Heligman y Pollard.

Para estimar el valor de los parámetros incluidos en $f_\alpha(x)$ podemos considerar varios criterios de optimización, que con el signo correspondiente minimizan lo que se denomina función de pérdida. A continuación nos ocupamos de los más utilizados.

2.3.1. El método de máxima verosimilitud

Este criterio pretende encontrar los valores de los parámetros que maximizan la función de verosimilitud de los valores observados de defunciones d_x , para edades $n_1 \leq x \leq n_2$, sera pues diferente si planteamos la graduación de q_x o de μ_x . Consideremos la distribución de la variable aleatoria número de muertos D_x , que en el caso de la graduación de q_x según la sección 2.2.1 es $D_x \sim Bi(E_x, q_x)$ donde E_x estima los expuestos iniciales al riesgo, mientras que en el caso de μ_x consideramos $D_x \sim Po(\mu_{x+\frac{1}{2}} E_x)$ siendo ahora E_x estimación de los

expuestos centrales al riesgo (sección 2.2.2).

▪ **Graduación de q_x**

La función de verosimilitud (sin especificar las constantes) dados los datos es

$$L(\alpha) \propto \prod_{x=n_1}^{n_2} (q_x(\alpha))^{d_x} (1 - q_x(\alpha))^{E_x - d_x},$$

siendo equivalente maximizarla a ella o a su logaritmo,

$$L_1(\alpha) \propto \sum_{x=n_1}^{n_2} (d_x \log(q_x(\alpha)) + (E_x - d_x) \log(1 - q_x(\alpha))).$$

▪ **Graduación de μ_x**

Conocidos los datos, la función de verosimilitud es

$$L(\alpha) = \prod_{x=n_1}^{n_2} \exp(-E_x \mu_{x+\frac{1}{2}}(\alpha)) \frac{(E_x \mu_{x+\frac{1}{2}}(\alpha))^{d_x}}{d_x!},$$

y su logaritmo

$$\log L(\alpha) = \sum_{x=n_1}^{n_2} (-E_x \mu_{x+\frac{1}{2}}(\alpha) + d_x \log E_x \mu_{x+\frac{1}{2}}(\alpha) - \log(d_x!)).$$

Como d_x y E_x no dependen de α maximizar $\log L(\alpha)$ equivale a maximizar

$$L_1(\alpha) = \sum_{x=n_1}^{n_2} (-E_x \mu_{x+\frac{1}{2}}(\alpha) + d_x \log(\mu_{x+\frac{1}{2}}(\alpha))).$$

2.3.2. El método de mínima chi-cuadrado

El estadístico χ^2 que definimos a continuación mide la discrepancia entre las muertes observadas y las esperadas bajo la hipótesis de que el modelo planteado en la sección anterior, para cada una de las medidas de mortalidad, es correcto.

Así definido, es una medida de la diferencia global entre los datos observados y el modelo propuesto por $f_\alpha(x)$ y por tanto debemos encontrar los valores de los parámetros que minimicen esta diferencia.

▪ **Graduación de q_x**

El número de muertos a edad x tiene media $E_x q_x(\alpha)$ y varianza $E_x q_x(\alpha)(1 - q_x(\alpha))$, con lo que podemos definir la desviación relativa en cada edad x

$$z_x = \frac{d_x - E_x q_x(\alpha)}{(E_x q_x(\alpha)(1 - q_x(\alpha)))^{1/2}},$$

y

$$\chi^2 = \sum_{x=n_1}^{n_2} z_x^2 = \sum_{x=n_1}^{n_2} \frac{(d_x - E_x q_x(\alpha))^2}{E_x q_x(\alpha)(1 - q_x(\alpha))}.$$

■ **Graduación de μ_x**

En el caso de μ_x podemos calcular las desviaciones relativas de la siguiente forma

$$z_x = \frac{d_x - E_x \mu_{x+1/2}}{E_x \mu_{x+1/2}},$$

y de acuerdo con esto definir la siguiente medida de discrepancia entre las muertes observadas y las que proporcionaría el modelo

$$\chi^2 = \sum_{x=n_1}^{n_2} z_x^2 = \sum_{x=n_1}^{n_2} \frac{(d_x - E_x \mu_{x+1/2}(\alpha))^2}{E_x \mu_{x+1/2}(\alpha)}.$$

Definidas así, las z_x siguen una distribución $N(0, 1)$ y son independientes por la hipótesis asumida, de forma que χ^2 sigue una distribución ji-cuadrado cuyos grados de libertad vienen determinados por el número de edades que graduemos, $n_2 - n_1$, menos el número de parámetros de la función $f_\alpha(x)$ y menos el número de restricciones para los parámetros si hubiera condiciones que éstos debieran de cumplir.

2.3.3. El método de mínimos cuadrados

Otro criterio muy utilizado para obtener los valores de los parámetros es el conocido método de los mínimos cuadrados, que se basa, como indica su nombre, en minimizar las distancias entre los valores observados y la fórmula matemática elegida para la graduación, es decir,

$$\sum_{x=n_1}^{n_2} \left(\frac{d_x}{E_x} - f_\alpha(x) \right)^2.$$

En la mayoría de ocasiones la varianza varía con la edad y por tanto es deseable que la curva se acerque más a las edades cuya varianza es menor, para ello se utilizan los mínimos cuadrados ponderados

$$\sum_{x=n_1}^{n_2} \omega_x \left(\frac{d_x}{E_x} - f_\alpha(x) \right)^2,$$

donde las ponderaciones ω_x suelen venir dadas por un valor inversamente proporcional a la varianza.

Por ejemplo en el caso de la graduación de q_x su varianza es $\frac{q_x(1 - q_x)}{E_x}$ que como q_x suele ser relativamente pequeño se aproxima por $\frac{q_x}{E_x}$ podemos encontrarnos con ponderaciones tales como: $\frac{E_x}{q_x}$ o $\frac{1}{q_x^2}$.

2.4. Metodología del CMI Bureau

El Continuous Mortality Investigation (CMI) Bureau del Instituto y Facultad de Actuarios de Londres, tiene su origen en 1924 cuando la recolección continua de datos de mortalidad comienza. Tiene la responsabilidad de la construcción de tablas de mortalidad estándar para su uso en la industria de seguros de Gran Bretaña. Forfar, McCutcheon y Wilkie (1988) han dado una descripción comprensible de la metodología usada normalmente por el CMI para producir tales tablas y nosotros la resumimos a continuación, pues generaliza modelos clásicos como los de Gompertz y Makeham, además de haber sido ampliamente desarrollada por diversos autores de nuestro interés, que como expondremos en un capítulo posterior, han adaptado esta metodología a modelos lineales generalizados y la han hecho servir de base para desarrollo de modelos dinámicos.

Dado un conjunto de edades consecutivas consideramos, d_x , secuencia del número de muertes y, E_x , secuencia de expuestos al riesgo, ambos a edad x , aunque en la práctica las muertes no se producen en una edad exacta x sino que se producen a una edad $x + b$. La determinación de la edad de muerte puede hacerse siguiendo distintos criterios, lo que da lugar a otras tantas formas de clasificar muertes. El criterio seguido por el CMI Bureau es el de asignar la muerte a la edad más cercana al cumpleaños en el momento en el que se produjo. Este criterio de clasificación equivale a tomar $b = \frac{-1}{2}$ lo que supone que estamos estimando $q_{x+\frac{1}{2}}$ y μ_x .

Los expuestos inicialmente al riesgo se calculan añadiendo un periodo extra a los centrales, un medio de las muertes, a través de la aproximación usual explicada según la expresión (2.4).

Suponemos que no hay duplicidades, es decir, que la investigación recoge individuos o pólizas independientes de forma que a un solo individuo le corresponde una póliza y por tanto todas las muertes son independientes. Si las hubiera las distribuciones Poisson y binomial normalmente utilizadas tendrían un parámetro de sobredispersión que afecta a su varianza.

Para obtener la graduación el CMI Bureau utiliza las funciones llamadas *Gompertz-Makeham de tipo* (r, s) . Son funciones con $r + s$ parámetros de la forma

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

con la convención de que si $r = 0$ sólo tendríamos parte exponencial y si $s = 0$ tendríamos sólo el término polinómico.

Derivadas de las anteriores podemos considerar los *Logit Gompertz-Makeham de tipo* (r, s) cuya la expresión es

$$LGM_{\alpha}^{r,s}(x) = \frac{GM_{\alpha}^{r,s}(x)}{1 + GM_{\alpha}^{r,s}(x)}.$$

De acuerdo con las circunstancias, el rango de valores puede tener mucha relevancia en la elección de la función adecuada. En el caso de μ_x el rango de valores posibles es de 0 a ∞ ,

mientras que q_x toma valores entre 0 y 1, de forma que para la primera en principio podría ser más adecuada la función GM y para la segunda LGM , aunque esto es orientativo y no una regla absoluta.

Además utilizan para la expresión de las funciones $GM(r, s)$ una base de polinomios ortogonales, como por ejemplo los polinomios de Chebycheff $C_i(x)$ definidos las ecuaciones

$$C_0(x) = 1,$$

$$C_1(x) = x,$$

y la relación recurrente

$$C_{n+1}(x) = 2xC_n(x) - C_{n-1}(x), \quad n \geq 1,$$

que son ortogonales en el intervalo de $[-1, 1]$ con la función peso $W(x) = (1 - x^2)^{-\frac{1}{2}}$. En relación al mismo intervalo pero cuando la función peso es $W(x) = 1$, obtenemos los polinomios de Legendre $\{L_i(x)\}$, definidos con los valores iniciales

$$L_0(x) = 1,$$

$$L_1(x) = x,$$

y la relación recurrente

$$(n + 1)L_{n+1}(x) = (2n + 1)xL_n(x) - nL_{n-1}(x), \quad n \geq 1,$$

La idea de utilizar una base ortogonal es porque a la hora de encontrar el polinomio de grado n , $P(x) = \sum_{i=0}^n a_i x^i$, que mejor ajusta a los datos según el criterio de mínimos cuadrados los coeficientes en la base natural $\{1, x, x^2, \dots\}$, son diferentes si consideramos el de grado $n + 1$. Pero si su expresión es en una base ortogonal, $P(x) = \sum_{i=0}^n \alpha_i C_i(x)$, solo hemos de añadir un nuevo coeficiente α_{n+1} . La utilización de polinomios ortogonales requiere reescalar la edad de $[-1, 1]$, para ello definimos la transformación $t = \frac{x - u}{v}$, donde $u = \frac{\max(x) + \min(x)}{2}$ y $v = \frac{\max(x) - \min(x)}{2}$.

Para estimar el valor de los parámetros incluidos en $f_\alpha(x)$ consideran los dos criterios de optimización

1. Maximizar la función de verosimilitud.
2. Minimizar χ^2 .

Es interesante hacer notar que en la práctica la alternativa de maximizar la verosimilitud y minimizar la χ^2 generalmente produce graduaciones muy similares. De hecho se demuestra que maximizar $L_1(\alpha)$ es equivalente a maximizar $-\frac{1}{2}\chi^2(\alpha)$, lo cual es igual a minimizar χ^2 .

2.5. Modelos lineales generalizados

Es posible reformular la metodología del CMI y extenderla usando los esquemas de modelos lineales y no lineales generalizados. Si hemos de estimar los parámetros de las funciones *GM* que mejor ajustan los valores brutos, conocidas las correspondientes distribuciones, no podemos recurrir a modelos lineales clásicos por motivos obvios de no linealidad de las funciones presentadas.

Vamos a introducir los principales conceptos de los Modelos lineales generalizados (McCullagh y Nelder, 1989), para después relacionarlos con la graduación actuarial como han hecho Haberman y Renshaw (1996) y Verrall (1996).

2.5.1. Introducción a los modelos lineales generalizados

Los modelos lineales generalizados (GLM) son una extensión de los modelos lineales para distribuciones de la variable respuesta no normales y transformaciones no lineales.

En general, consideremos un vector de observaciones y con n componentes, que son realizaciones de una variable aleatoria Y , independientes e idénticamente distribuidas y con media m .

Un modelo de regresión cualquiera constituye una especificación para la media m de la variable en términos de un número pequeño de parámetros desconocidos $\beta_0, \beta_1, \dots, \beta_p$. En el caso particular de modelos lineales se desea encontrar una función lineal tal que $E(Y|X) = m = \beta_0 + \sum_{i=1}^p \beta_i x_i$, y para ello se supone la varianza de Y constante, $var(Y) = \sigma^2$.

A diferencia de éstos, un modelo lineal generalizado (GLM) proporciona un método de estimación de una función de la media de la variable respuesta como una combinación lineal de un conjunto de variables predictoras, es decir

$$l(E(Y|X)) = l(m) = \beta_0 + \sum_{i=1}^p \beta_i x_i = \eta(x).$$

La función de la media de la respuesta, $l(m)$, se llama función *link*, y se supone igual a una función lineal de los predictores, $\eta(x)$, que se llama *predictor lineal*.

Un modelo lineal generalizado puede ser descrito por las siguientes hipótesis:

- Hay una variable respuesta, Y , observada independientemente a valores fijos de variables explicativas X_1, \dots, X_p .
- Las variables explicativas sólo pueden influir en la distribución de Y a través de una función lineal llamada predictor lineal $\eta = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$.
- La distribución de cada componente y_i de Y tiene densidad de la forma

$$f(y_i; \theta_i, \varphi) = \exp [A_i \{y_i \theta_i - \gamma(\theta_i)\} / \varphi + \tau(y_i, \varphi / A_i)]$$

donde φ es un parámetro de escala (posiblemente conocido), A_i es conocido, y el parámetro θ_i depende del predictor lineal.

- La media, m , es una función invertible del predictor lineal,

$$m = g(\eta), \quad \eta = g^{-1}(m) = l(m).$$

La función inversa, $l(\cdot)$, es la que denominada link.

Señalemos que θ_i es también una función invertible de m_i , de hecho $\theta_i = (\gamma')^{-1}(m_i)$ como se muestra después.

Los GLM proporcionan un tratamiento estadístico unificado para muchas importantes clases de modelos. Consideremos algunos ejemplos que luego utilizaremos.

Binomial. Para una distribución Binomial con un número fijo de pruebas a_i y parámetro p_i podemos tomar las respuestas $y = s_i/a_i$ donde s_i es el número de éxitos. La densidad puede expresarse de la forma

$$\log f(y_i) = a_i \left[y_i \log \left(\frac{p_i}{1-p_i} \right) + \log(1-p_i) \right] + \log \binom{a_i}{a_i y_i},$$

con lo que $A_i = a_i$, $\varphi = 1$, θ_i será la transformación logit de p_i $\log(p_i/(1-p_i))$ y $\gamma(\theta_i) = -\log(1-p_i) = \log(1+e^{\theta_i})$.

Poisson. Para una distribución Poisson con media m tenemos

$$\log f(y_i) = y_i \log m_i - m_i - \log(y_i!),$$

por tanto $\theta = \log m_i$ y $\gamma(\theta_i) = m_i = e^{\theta_i}$.

Gamma. Para una distribución Gamma con media m_i y varianza m_i^2/ν tenemos

$$\log f(y_i) = \frac{-y_i/m_i + \log(1/m_i)}{1/\nu_i} + \nu_i \log y_i + \nu_i \log \nu_i - \log \Gamma(\nu_i),$$

por tanto $\theta_i = -1/m_i$ y $\gamma(\theta_i) = -\log(-\theta_i)$.

La estimación en el caso de los modelos lineales generalizados, se realiza maximizando la cuasi-log-verosimilitud, que para los miembros de la familia exponencial es sinónimo de log-verosimilitud.

Si $f(y; \theta)$ es la función de densidad de probabilidad para la observación y dado el parámetro θ , entonces la log-verosimilitud basada en el conjunto de observaciones, y expresada como una función del parámetro valor medio, $m = E(Y)$, es la suma de la contribuciones individuales y por tanto

$$l(m; y) = \sum_{i=1}^n \log f(y_i; \theta_i).$$

Una medida de bondad de ajuste entre los datos observados y los valores ajustados generados por el modelo es la *Deviance*, cuyo valor se obtiene a partir de la expresión

$$D(y; \hat{m}) = 2l(y; y) - 2l(\hat{m}; y),$$

donde $l(y; y)$ es la log-verosimilitud cuando los valores estimados coinciden con los observados y $l(\hat{m}; y)$ la de los valores estimados \hat{m} .

La Deviance mide la discrepancia entre la log-verosimilitud de los datos y la del modelo utilizado, en consecuencia y puesto que $l(y; y)$ no depende del modelo utilizado, maximizar la verosimilitud equivale a minimizar $D(y; \hat{m})$. La distribución asintótica de $D(y; \hat{m})$ es una χ_{n-p}^2 , donde p es el número de parámetros ajustados bajo H_0 , que es la hipótesis del modelo.

La Deviance más que como medida absoluta de bondad de ajuste puede utilizarse para comparar dos modelos. Por ejemplo, si queremos contrastar si añadir una nueva covariable mejora el modelo, si H_0 es el modelo a validar y H_A la extensión del modelo conteniendo una covariable adicional y m_0 y m_A los correspondientes valores ajustados, la diferencia de Deviance,

$$\begin{aligned} D(y; \hat{m}_0) - D(y; \hat{m}_A) &= 2l(y; y) - 2l(\hat{m}_0; y) - (2l(y; y) - 2l(\hat{m}_A; y)) \\ &= 2l(\hat{m}_A; y) - 2l(\hat{m}_0; y), \end{aligned}$$

corresponde al estadístico ratio de verosimilitud para contrastar H_0 versus H_A y se distribuye aproximadamente χ_1^2 .

Pero el problema de la modelización requiere un equilibrio entre la mejora del modelo y el no deseable incremento de la complejidad de éste al ir añadiendo cada nuevo parámetro. Para esto utilizaremos el valor del estadístico C_p de Mallows descrito en Venables y Ripley (1994), pag.219. Dicho estadístico está definido como

$$C_p = RSS + 2\sigma^2 p,$$

donde RSS es la suma de cuadrados residuales y p el número de parámetros del modelo. Así definido, puede ser interpretado como una suma de cuadrados residuales penalizada por la complejidad del modelo. Una disminución del valor de C_p implica una mejora del modelo.

2.5.2. Modelos lineales generalizados y graduación

En esta sección vamos a desarrollar la teoría de los modelos lineales generalizados en el contexto de la graduación, primero de los q_x y después de los μ_x . La experiencia en graduación usando modelos lineales generalizados ha sido recopilada en la literatura actuarial por Renshaw (1991), Renshaw y Hatzopoulos (1996), Haberman y Renshaw (1996). Haberman y Renshaw (1996) ofrecen una descripción de todas las aplicaciones de los modelos lineales generalizados en la Ciencia Actuarial. Renshaw (1991) se refiere en concreto a la graduación y Verrall (1996) se ocupa de su implementación en S-plus.

Vamos a explicar detenidamente la aplicación de la metodología de GLM a la graduación de q_x cuando las funciones $LGM(r, s)$ carecen de la parte polinómica y considerando la transformación $\text{logit}(q_x)$, y también a la graduación de μ_x con la transformación $\log(\mu_{x+\frac{1}{2}})$ y las funciones $GM(r, s)$ con $r = 0$. Nos ocuparemos también de algunos casos más generales.

Modelos para q_x

Consideramos la graduación de la probabilidad de muerte a edad x , q_x , basada en la variable aleatoria muertes a edad x , D_x , sus valores observados, d_x , y la estimación de los expuestos iniciales al riesgo, E_x . Respecto al modelo de distribución del número de muertes, supondremos que $D_x \sim Bi(E_x, q_x)$, y que las muertes son independientes unas de otras.

Ya hemos comentado en la sección 2.5.1 que la distribución binomial es un miembro de la familia exponencial, como se requiere en McCullagh y Nelder (1989). La expresión de su función de verosimilitud que nos servirá para desarrollos posteriores es

$$\begin{aligned}
 L(q) &= \prod_{x=n_1}^{n_2} \binom{E_x}{d_x} q_x^{d_x} (1 - q_x)^{E_x - d_x} \\
 &\propto \prod_{x=n_1}^{n_2} q_x^{d_x} (1 - q_x)^{E_x - d_x} \\
 &= \exp \left(\sum_{x=n_1}^{n_2} (d_x \log(q_x) + (E_x - d_x) \log(1 - q_x)) \right) \\
 &= \exp \left(\sum_{x=n_1}^{n_2} \left(d_x \log \left(\frac{q_x}{1 - q_x} \right) + E_x \log(1 - q_x) \right) \right).
 \end{aligned}$$

La transformación de $\log \left(\frac{q_x}{1 - q_x} \right) = \text{logit}(q_x)$ es el link canónico en la familia binomial y es el que vamos a describir en profundidad. No obstante, Renshaw (1991) extiende esta metodología a otras funciones link e incluso a modelos lineales no generalizados que después detallaremos.

Comenzando pues con el link logit, las funciones que vamos a ajustar son

$$q_x = LGM(r, s) = \frac{GM(r, s)}{1 + GM(r, s)},$$

de donde fácilmente se obtiene

$$\frac{q_x}{1 - q_x} = GM(r, s),$$

de forma que su logaritmo, si $r = 0$, corresponde a un predictor lineal.

Además tenemos que

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

Tomando logaritmos en la verosimilitud y sin considerar las constantes tenemos

$$\begin{aligned}
 \log L(q) &= \sum_{x=n_1}^{n_2} \left(d_x \log \left(\frac{q_x}{1-q_x} \right) + E_x \log(1-q_x) \right) \\
 &= \sum_{x=n_1}^{n_2} \left(d_x \log \left(\frac{q_x}{1-q_x} \right) + \frac{-1}{-1} E_x \log(1-q_x) \right) \\
 &= \sum_{x=n_1}^{n_2} \left(d_x \log \left(\frac{q_x}{1-q_x} \right) - E_x \log \left(\frac{1}{1-q_x} \right) \right) \\
 &= \sum_{x=n_1}^{n_2} (d_x \log(GM(r, s)) - E_x \log(1 + GM(r, s))).
 \end{aligned}$$

Los coeficientes de los que dependen las funciones GM se obtienen por máxima verosimilitud o, lo que es equivalente, minimizando

$$\sum_{x=n_1}^{n_2} E_x \log(1 + GM(r, s)) - d_x \log(GM(r, s)).$$

Dentro del modelo binomial, para la graduación de los q_x podemos considerar otros link como son,

- Complementario log-log.

Cuyo predictor lineal es $\eta_x = \log(-\log(1 - q_x))$ con inversa $q_x = 1 - \exp(-\exp(\eta_x))$

- Probit.

Cuyo predictor lineal es $\eta_x = \Phi^{-1}(q_x)$, donde Φ es la función de distribución de probabilidad de una normal estándar, con inversa $q_x = \Phi(\eta_x)$.

Renshaw (1991), partiendo de la fórmula ya muy familiar de Gompertz $\mu_x = Bc^x$, no lineal pero que puede linealizarse, por ejemplo en términos de q_x con

$$\eta_x = \log(-\log(1 - q_x)) = \beta_0 + \beta_1 x,$$

donde β_0 y β_1 son una reparametrización de B y C , introduce esta graduación como un GLM de un modelo de distribución Binomial, con link complementario log-log y un predictor lineal correspondiente a una recta. Lo que propone el autor es extender este modelo aumentando el grado del predictor lineal obteniendo así graduaciones con la siguiente expresión

$$q_x = 1 - \exp(-\exp(\hat{\eta}_x)),$$

donde la estimación del predictor lineal es $\hat{\eta}_x = \sum_{i=1}^{s-1} \beta_i x^i$.

El modelo binomial junto con el link probit no se ha utilizado en el contexto actuarial antes del citado artículo de Renshaw. En su trabajo el autor si lo utiliza y compara con los otros aunque en principio no represente ningún modelo paramétrico reconocible a diferencia de los otros dos link.

Modelos para μ_x

Consideramos ahora la graduación de μ_x , fuerza de mortalidad, en la que se utiliza E_x como expuestos centrales al riesgo y d_x valores observados de las muertes. El modelo usual en este caso es $D_x \sim Po(E_x \mu_{x+\frac{1}{2}})$, con muertes independientes.

De nuevo tenemos un miembro de la familia exponencial como estudiamos en la sección 2.5.1, la expresión de su función de verosimilitud es

$$\begin{aligned} L(\mu) &= \prod_{x=n_1}^{n_2} \frac{(E_x \mu_{x+\frac{1}{2}})^{d_x} \exp(-E_x \mu_{x+\frac{1}{2}})}{d_x!} \\ &\propto \exp\left(\sum_{x=n_1}^{n_2} \left(D_x \log E_x + d_x \log\left(\mu_{x+\frac{1}{2}}\right) - E_x \mu_{x+\frac{1}{2}}\right)\right). \end{aligned}$$

En este caso el link utilizado es $\log\left(\mu_{x+\frac{1}{2}}\right)$ que es el canónico de la familia Poisson. Si además la fórmula que se utiliza es $\mu_{x+\frac{1}{2}} = GM(r, s)$, cuando $r = 0$, tenemos un predictor lineal. Sustituyendo en el logaritmo de la función de verosimilitud cambiada de signo y sin considerar los términos constantes, que sólo dependen de d_x y E_x , obtenemos

$$\sum_{x=n_1}^{n_2} (E_x GM(r, s) - d_x \log GM(r, s)),$$

de la que se obtienen los parámetros minimizando.

También es posible como comentábamos en la introducción, considerar E_x como variable aleatoria (Renshaw, Haberman y Hatzopoulos, 1997) fijados los valores observados para las muertes d_x . Consideramos ahora la graduación de $1/\mu_x$, fuerza de vitalidad según terminología de Lambert (1772). El modelo es en este caso $E_x \sim Ga\left(d_x, \mu_{x+\frac{1}{2}}\right)$ e independientes para las distintas x . Podemos utilizar entonces variables respuesta E_x , con medias $\lambda_x = d_x \frac{1}{\mu_{x+\frac{1}{2}}}$, $Var(\lambda_x) = \lambda_x^2$ y pesos $\omega_x = d_x$.

Si tomamos el link log, tenemos

$$\log \lambda_x = \log a_x - \log \mu_{x+\frac{1}{2}} = \log d_x + \eta_x,$$

donde η_x es el predictor lineal.

En su artículo, Renshaw (1991) llega a la conclusión de que la graduación convencional de μ_x , basada en el ajuste de la fórmula matemática maximizando la función de verosimilitud del número de muertes como una variable aleatoria Poisson condicionada a los expuestos centrales a el riesgo, es efectivamente equivalente a la aproximación planteada por los expuestos centrales al riesgo modelados como variables aleatorias Gamma condicionadas al número de muertes. Las ventajas que presenta esta segunda alternativa, están discutidas en el artículo de Renshaw y Haberman (1997), cuando ésta se aplica a la construcción de tablas de mortalidad para individuos seleccionados.

2.5.3. Modelos no lineales generalizados y graduación

Siguiendo con el artículo de Renshaw (1991) vemos como es posible, mediante un método iterativo, aplicar una metodología similar cuando el predictor no es lineal. De nuevo recurrimos a una fórmula histórica que es la de Makeham $\eta_x = A + Bc^x$.

Dado que no es posible transformar esta forma no lineal en una lineal a menos que $A = 0$, podemos introducir una reparametrización trivial a la forma exponencial y escribir

$$\eta_x = \alpha + \beta \exp(\phi x),$$

que puede considerarse lineal si se supone el parámetro ϕ conocido.

Sea

$$g(x; \phi) = \exp(\phi x), \tag{2.6}$$

denotando el término no lineal, podemos aproximar $g(x; \phi)$ mediante

$$g(x; \phi) \simeq g(x; \phi_0) + (\phi - \phi_0) \left(\frac{\partial g}{\partial \phi} \right)_{\phi=\phi_0},$$

con lo que $\beta g(x; \phi)$ puede ser reemplazado por $\beta u + \gamma v$ con

$$u = g(x; \phi_0) \quad y \quad v = \left(\frac{\partial g}{\partial \phi} \right)_{\phi=\phi_0}$$

y donde $\gamma = \beta(\phi - \phi_0)$. De esta forma el término no lineal (2.6) lo hemos convertido en una expresión lineal que puede ser insertada en el predictor de un modelo lineal generalizado.

Partiendo pues de un valor inicial ϕ_0 , calculamos las covariables

$$u = \exp(\phi_0 x) \quad y \quad v = x \exp(\phi_0 x),$$

y los parámetros β y γ estimados a partir del ajuste del modelo como en cualquier estimador lineal. Entonces actualizamos

$$\phi_1 = \phi_0 + \frac{\hat{\gamma}}{\hat{\beta}}$$

y este proceso se repite hasta la convergencia. La convergencia no está garantizada para valores iniciales muy distantes. El autor ha encontrado que un valor inicial de $\phi_0 = 0.0005$ induce a convergencia para muchos de los conjuntos de datos típicos que ha graduado de esta forma. Considerando la graduación de μ_x , el modelo de distribución Poisson y seleccionando link identidad, este procedimiento permite ajustar las fórmulas $GM_x(r, 2)$ con $r \neq 0$.

2.6. Modelo Heligman y Pollard

Una alternativa a las anteriores funciones son las leyes de Heligman y Pollard (1980) que han sido ampliamente utilizadas por diferentes países de nuestro entorno europeo

(Inglaterra, Suecia, Alemania y España) y de otros países desarrollados (Estados Unidos de América y Australia) desde que la ONU promovió el ajuste de la mortalidad a través de la primera ley de Heligman y Pollard (Felipe y Guillén, 1999). Los autores, inspirándose en Thiele (1972), ajustan una nueva ley de mortalidad en la Australia de la posguerra, cuya expresión genérica es

$$\frac{q_x}{1 - q_x} = \sum_{i=1}^n A_i \exp(-B_i(f_i(x) - C_i)^{D_i}),$$

y los parámetros a estimar son A_i, B_i, C_i, D_i con $i = 1, 2, \dots, n$. Normalmente con $n = 3$ se obtienen buenos ajustes. Respecto a $f_i(x)$ suele utilizarse x o $\ln(x)$.

Las tres expresiones que realmente se ajustaron a la mortalidad australiana fueron

- Primera ley (versión 1) de Heligman y Pollard.

$$\frac{q_x}{1 - q_x} = A^{(x+B)^C} + D \exp(-E(\ln x - \ln F)^2) + GH^{(x-x_0)}$$

expresión que consideran no puede distinguirse de,

- Primera ley (versión 2) de Heligman y Pollard.

$$q_x = A^{(x+B)^C} + D \exp(-E(\ln x - \ln F)^2) + \frac{GH^x}{1 + GH^x}$$

- Segunda ley de Heligman y Pollard.

$$q_x = A^{(x+B)^C} + D \exp(-E(\ln x - \ln F)^2) + \frac{GH^x}{1 + KGH^x}$$

- Tercera ley de Heligman y Pollard.

$$q_x = A^{(x+B)^C} + D \exp(-E(\ln x - \ln F)^2) + \frac{GH^{x^k}}{1 + GH^{x^k}}$$

Según Felipe y Guillén la curva básica, primera ley, presenta la siguientes ventajas:

- 1) Es una función continua.
- 2) Es aplicable a todo el rango de edades $0 < x < \omega$.
- 3) La probabilidad ajustada está entre los límites apropiados $0 \leq q_x \leq 1$.
- 4) Tiene pocos parámetros si consideramos que es para todo el rango de edades.
- 5) Los parámetros tienen una interpretación demográfica o biológica.
- 6) Es suficientemente flexible para ajustar una gran variedad de modelos de mortalidad.

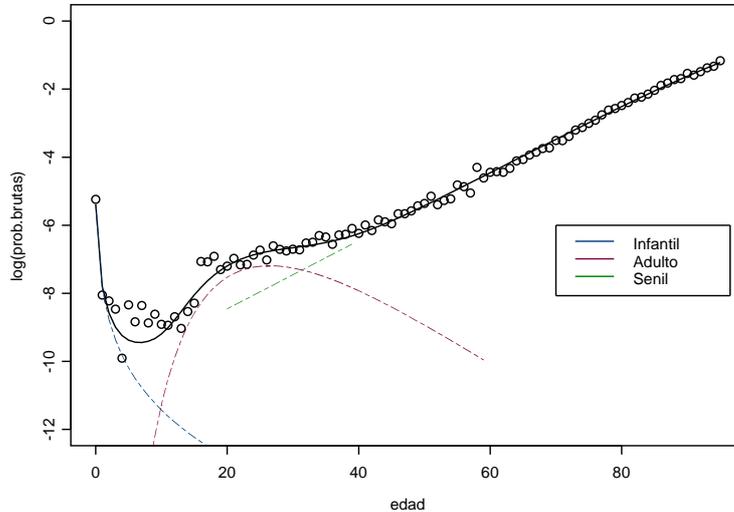


Figura 2.2: Descomposición de la Ley de Heligman y Pollard

Cada uno de los tres términos de la ecuación básica representa una componente distinta de la mortalidad, cuya representación gráfica puede verse en la Figura (2.2). La interpretación de cada uno de los tres sumandos que aparecen en la primera ley es la siguiente:

El primer sumando, una exponencial de decrecimiento rápido, representa la curva de mortalidad infantil. Refleja la caída de la mortalidad durante los primeros años de vida, conforme el niño se adapta a su nuevo ambiente.

- El parámetro A tiene un valor similar a la probabilidad q_1 y mide el nivel de mortalidad infantil.
- El parámetro B representa la mortalidad para los niños con un año de edad. Toma valores entre 0 y 1. En la práctica B está próximo a cero (entre 0.01 y 0.03) y su efecto sobre las probabilidades de fallecimiento en edades distintas de 0 es prácticamente nulo.
- El parámetro C toma valores entre 0 y 1 y está íntimamente asociado al decrecimiento del ratio de mortalidad. Cuanto mayor es C más rápidamente decrece la mortalidad conforme aumenta la edad.

El segundo sumando, una función similar a la lognormal, representa la mortalidad para edades más adultas. Esta es referenciado en la literatura demográfica y actuarial como joroba de los accidentes y recoge las muertes por accidentes más la mortalidad maternal para la población se femenina. Se trata de una mortalidad adicional, sobrepuesta a la curva natural de la mortalidad que le correspondería según la edad. Generalmente la 'joroba' se da entre los 10 y los 40 años.

- El parámetro D establece la severidad de la joroba, toma valores entre 0 y 1, y a mayor D mayor será el efecto de los accidentes sobre la mortalidad total.
- El parámetro E esta relacionado con la dispersion, toma valores entre 0 e ∞ , y a mayor E menor dispersion indicando una joroba concentrada.
- El parámetro F indica la localización de la joroba de los accidentes, toma valores entre 15 y 110, y es el máximo de la misma.

El tercer sumando es la ley de Gompertz. Refleja crecimiento de la mortalidad en edades altas y representa el deterioro del cuerpo con los años, es decir, la mortalidad senil.

- El parámetro G indica el nivel base de la mortalidad senil y toma valores entre 0 e ∞ .
- El parámetro H es la tasa de crecimiento de dicha mortalidad senil.
- La edad x_0 que aparece en la versión 1 de la primera ley es una edad cercana a ω (edad límite). Es aquella edad para la cual $q_x = 0.5$.

Hartmann (1983) indica que todos los parámetros son positivos y las expresiones solo están definidas para edades mayores de 0.

La segunda y tercera ley, especialmente para edades altas, mejoran el ajuste respecto a la primera. En la segunda ley, K es el único parámetro que puede llegar a tomar valores negativos. Heligman y Pollard (1980) consideran que la interpretación de los parámetros sigue siendo válida. No obstante, los valores de G y H para la tercera ley están muy afectados por la dirección y extensión de la curvatura; y no queda claro que tengan todavía una interpretación demográfica, aunque x^k puede entenderse como una transformación en la escala de las edades. En el trabajo de Felipe y Guillén (1999) se sigue la línea iniciada por Rue (1992) quien tras ajustar las tres leyes a la población española para datos agrupados de 1979-1982, obtuvo el mejor ajuste para la segunda ley. Estas autoras utilizan también estos modelos para la construcción de tablas dinámicas, en cuyo caso incorporan el análisis temporal de los parámetros a la segunda ley de Heligman y Pollard, siguiendo a McNown y Rogers (1992) que, a diferencia de ellas, utilizan la primera ley.

Los parámetros se obtienen por mínimos cuadrados ponderados no lineales para todo el rango de edades,

$$\sum_x \omega_x (q_x - f(x))^2,$$

donde la varianza de la observación a edad x es proporcional a ω_x^{-1} y $f(x)$ es la función a ajustar a los valores q_x observados.

La necesidad de estos pesos es debida a la desigualdad de varianzas en el modelo $q_x = F(x)$, las cuales siguiendo el modelo binomial son $var(q_x) = \frac{q_x(1-q_x)}{E_x}$, de ahí que los pesos propuestos sean $1/q_x$ o alguna potencia inversa de q_x .

Congdon (1993), plantea el problema de la sobreparametrización y cómo puede afectar ésta en la estabilidad de los parámetros a lo largo del tiempo, por esta razón decide fijar los

valores de alguno de los parámetros (por ejemplo, A y B los fija a q_0 y q_1 , respectivamente) al utilizar la segunda ley de Heligman-Pollard para ajustar sus datos.

Finalmente compara este ajuste para las edades adultas, con funciones Gompertz-Makeham concluyendo que estas últimas funcionan peor, tienen menor bondad del ajuste (un R^2 más pequeño) y las correlaciones entre sus parámetros son más altas. También realiza una breve comparación con algunas técnicas no paramétricas.

Es novedosa la aportación de Dellaportas, Smith y Stavropoulos (2001) que adoptan el análisis bayesiano y usan técnicas MCMC de simulación en la fórmula de Heligman y Pollard.

2.7. Contrastes para la bondad del ajuste

En este apartado se describen los test clásicos utilizados, que nos servirán para ver si una determinada graduación representa un ajuste razonable de los datos, permitiéndonos comparar los diferentes ajustes para poder elegir el más adecuado. El objetivo final de cada uno de los ajustes es que los valores obtenidos sean muy parecidos a la experiencia observada, y para ello hemos de utilizar medidas estadísticas de bondad de ajuste, que nos ayuden a decidir si es aceptable o no la graduación, y finalmente que nos permita elegir entre todas ellas. Deseamos pues contrastar si los resultados de la graduación constituyen una razonable representación de los datos de mortalidad correspondientes a la experiencia investigada. Los contrastes propuestos por Forfar, McCutcheon y Wilkie (1988), y que posteriormente han sido aplicados por muchos autores a sus graduaciones, son los siguientes:

Desviaciones y desviaciones relativas

A partir de la graduación obtenemos los correspondientes valores ajustados de las medidas de mortalidad, que dan lugar al cálculo de las muertes esperadas, o_x , sin más que multiplicarlas por los respectivos expuestos al riesgo. Se definen entonces las desviaciones en cada edad como

$$Dev_x = d_x - o_x,$$

y las desviaciones relativas como

$$z_x = Dev_x / Var_x,$$

donde Var_x es la varianza correspondiente a la variable aleatoria número de muertes, tratada en sección 2 del capítulo de graduaciones paramétricas, que se estima a partir del ajuste obtenido. Después de su cálculo hay que comprobar si están aleatoriamente distribuidas y si su distribución esta de acuerdo con las hipótesis inherentes al modelo utilizado, Binomial o Poisson. La mayoría de los test que vamos a describir se basan en la hipótesis de que el número de fallecidos se puede aproximar a una Normal, lo cual puede no ocurrir en edades extremas donde el número de expuestos al riesgo suele ser pequeño. La hipótesis de normalidad sólo es razonable si el número esperado de muertes es suficientemente grande, suele fijarse como mayor o igual a 5, si no es así se agrupan las edades hasta conseguirlo antes de aplicar los test.

Test de signos

El primer test descrito es el test de signos. El número de signos positivos (o negativos) denotado por N_1 está distribuido como binomial de media $1/2$ y número de pruebas igual a número de edades consideradas N , $N_1 \sim B(N, 1/2)$. Denotamos $p(pos)$ la probabilidad de que el número de signos positivos no exceda el número observado. Si el valor de $p(pos)$ es demasiado pequeño, menor de 0.025, entonces el número de desviaciones positivas es inesperadamente bajo, y si es demasiado grande, mayor de 0.975, entonces el número de desviaciones positivas es inesperadamente alto. En cualquier caso las probabilidades graduadas están demasiado desviadas hacia un lado u otro de las probabilidades observadas, y no representan adecuadamente la experiencia.

Test de Rachas

El segundo contraste no paramétrico es el test de rachas o test de cambios de signo. Si los signos de las desviaciones están aleatoriamente distribuidos, no debería haber ni pocos ni demasiados cambios de signo.

Si el número de signos positivos es N_1 y N_2 el número observado de signos negativos (con $N = N_1 + N_2$), si los signos están dispuestos aleatoriamente y N_1 y N_2 son mayores de 20, entonces la distribución del número de rachas NR es aproximadamente Normal de media

$$E(NR) = \frac{2N_1N_2}{N} + 1$$

y varianza

$$Var(NR) = \frac{2N_1N_2(2N_1N_2 - N)}{N^2(N - 1)}$$

Llamando $p(run)$ a la probabilidad de que NR sea menor o igual que el número de rachas observadas, debe ser un valor no demasiado grande ni pequeño de forma que indique la adecuada aleatoriedad en los signos.

Test de Kolmogorov-Smirnov

Otro test no paramétrico es el de Kolmogorov -Smirnov, considera la distribución del estadístico KS una transformación de la máxima desviación absoluta entre dos funciones de distribución, cuya distribución es conocida. En las graduaciones lo utilizaremos para comparar la distribución de las muertes observadas y esperadas para cada edad, calculando $p(KS)$ probabilidad de que la variable tome un valor mayor al observado. Valores muy altos de la desviación, KS , o equivalentemente, $p(KS)$ probabilidades pequeñas corresponden a graduaciones no satisfactorias. Es un test que aunque no prueba la bondad de una determinada graduación pues no se cumple la hipótesis de independencia de las dos series de datos que se comparan, sí sirve para descartar malos ajustes, pues empíricamente se ha comprobado que graduaciones satisfactorias proporcionan valores de $p(KS)$ mayores de 0.9, incluso algunas veces excediendo el 0.99.

Test de autocorrelaciones

El siguiente test se trata de medir las autocorrelaciones correlaciones entre la desviaciones relativas, obteniendo su secuencia de valores, $cov(z_x, z_{x-j})$ para $j = 1, 2, \dots, n$.

Test χ^2

El estadístico

$$\chi^2 = \sum_x z_x^2$$

si las z_x se consideran $N(0, 1)$, esta distribuido como una $\chi^2(N - m)$ los grados de libertad son $N - m$ donde N es el número de edades consideradas y m el número de parámetros de la función utilizada en la graduación. Sin embargo, en la práctica muchas graduaciones satisfactorias muestran altos valores de χ^2 , una inspección detallada de los valores individuales de los z_x podrían indicar errores en los datos que deberían ser investigados.

Capítulo 3

Graduación de datos de mortalidad: métodos no paramétricos

3.1. Introducción

Sin duda los modelos paramétricos constituyen un método eficiente cuando se tiene información del modelo subyacente a las variables y sólo resta por determinar un número finito de parámetros, sin embargo una fuente de error puede ser elegir una familia paramétrica no adecuada. En estos casos podemos utilizar los métodos no paramétricos que además de permitir graduar probabilidades brutas que no siguen una fórmula paramétrica clara, pueden utilizarse para proporcionar una prueba de diagnóstico de los modelos paramétricos o simplemente para explorar los datos.

Los métodos no paramétricos pretenden obtener unos nuevos valores a partir de los observados, en los que se haya eliminado cualquier influencia no procedente de la variable predictora y a diferencia de los paramétricos no proporcionan una función que exprese la relación entre las probabilidades de muerte y la edad. Los métodos paramétricos y no paramétricos pueden complementarse en algunos casos, en el sentido que el resultado de una regresión no paramétrica puede describir de forma más adecuada la forma funcional a ajustar.

En este capítulo se pretende revisar las alternativas que se ofrecen para realizar un suavizamiento de los datos observados sin necesidad de obtener una fórmula matemática, sino a través de unos valores que describen de forma menos confusa la tendencia de la mortalidad. El primero de estos métodos que se utilizó es medias móviles ((Hoem y Linne-
mann, 1988), (Benjamin y Pollard, 1992)) en el que no profundizamos demasiado porque ha sido claramente mejorado por otras alternativas.

En primer lugar (sección 2) haremos una descripción de las técnicas de smoothing, para después ocuparnos más detalladamente del método kernel y de como elegir el parámetro de smoothing. En la siguiente sección introducimos el criterio de Whittaker-Henderson, considerando como caso particular los splines. Finalmente nos ocupamos de la regresión

dinámica relacionada con la regresión local ponderada.

Todas estas técnicas pueden utilizarse también considerando la distribuciones binomial y Poisson ya utilizadas en la graduación paramétrica y pueden implementarse en S-plus mediante Modelos Aditivos Generalizados (GAM) tal y como sugiere Verrall (1996). Dedicaremos la sección 5 a la descripción de estos.

3.2. Técnicas de smoothing

En esta sección pretendemos definir y revisar las diferentes técnicas de smoothing para desarrollar algunas de ellas con más detalle, relacionarlas y analizar como se han aplicado en la graduación de los datos de mortalidad.

El objetivo de las técnicas de smoothing es disminuir la variabilidad y facilitar el análisis. Para ello se han de modificar los datos observados mediante procedimientos que permiten obtener una nueva serie de la que se han eliminado las excesivas y no deseadas fluctuaciones.

El procedimiento utilizado se llama smoother que en términos generales se basa en obtener un promedio local. Se diferencian unos de otros por el método utilizado para promediar y por tanto de asignar ponderaciones, pero además dentro de un tipo determinado podemos obtener diferentes resultados según sea el tamaño de la vecindad utilizada, es decir, el número de observaciones que intervienen en el promedio. Cuanto mayor es el tamaño de la vecindad menor es la varianza pero mayor será el sesgo y viceversa. Hay diversos métodos para determinar el tamaño de la vecindad o bandwidth que también puede representarse por lo que llamamos span (proporción del número de elementos de cada vecindad entre el número de elementos total), éste es un tema ampliamente tratado y de gran importancia en este tipo de técnicas. Suele utilizarse como procedimiento la validación cruzada, que explicaremos más adelante después de definir brevemente todos los tipos de smoother.

Smoother tipo "bin". Divide el intervalo de valores de la variable predictora escogiendo puntos de corte c_0, c_1, \dots, c_n y define las regiones

$$R_k = \{c_k \leq x_i \leq c_{k+1}\} \quad k = 1, 2, \dots, n - 1$$

para después promediar los valores respuesta de cada región. El smoother vendría dado por la secuencia de valores,

$$\bar{y}_k = \underset{x_i \in R_k}{\text{media}} \ y_i. \tag{3.1}$$

Las estimaciones no son buenas y describe los datos con poca suavidad. No es adecuado para la graduación.

Media móvil. Es un smoother muy utilizado por su facilidad de cálculo y no presenta el problema de los smoother tipo bin. Para cada valor x_i definimos la vecindad simétrica de tamaño k

$$N_k^S(x_i) = \{x_{\max\{1, i-k\}}, \dots, x_i, \dots, x_{\min\{n, i+k\}}\} \tag{3.2}$$

El método de medias móviles supone sustituir la observación y_i por la media de las observaciones correspondientes a su vecindad,

$$\bar{y}_k = \underset{j \in \text{ind}(N^S(x_i))}{\text{media}} y_j,$$

con $\text{ind}(N^S(x_i))$ conjunto de índices de la vecindad. Para los puntos iniciales y finales de la serie que disten menos de k unidades de los extremos, la vecindad contiene menos puntos que en el resto de forma que la estimación aparece sesgada. La ventaja que su facilidad de obtención suponía ha desaparecido con el auge de la informática y los inconvenientes que suponen su movilidad y el sesgo en los extremos han ido disminuyendo su uso.

Smoother de vecindad más cercana. Para definir la vecindad de un valor x_i se suele tomar la expresión (3.2), pero es posible, ignorando la simetría, tomar los $2k$ puntos más cercanos a x_i independientemente de a que lado se encuentren y después promediarlos. En este caso estamos utilizando la vecindad más cercana y en consecuencia un smoother de vecindad más cercana. Se aprecia la diferencia cuando los valores de la variable predictora no están equiespaciados, que no suele ser el caso de las edades en la graduación.

Regresión Local. Un método que resuelve la fluctuación de las medias móviles es el smoother de recta móvil. Se trata de ajustar una recta de mínimos cuadrados a cada vecindad de tamaño fijo. Una mejora de éste método es el uso de rectas de regresión utilizando ponderaciones que disminuyan en relación con la lejanía del punto. Cleveland propuso un smoother que se encuentra implementado en el S-plus con el comando *loess*, que utiliza la función tricúbica para asignar los pesos a los puntos de la vecindad.

Hemos visto dos tipos de vecindad una que considera los $2k$ puntos más cercanos ignorando la simetría y en la que se fija a priori el número de elementos con los que se realiza el ajuste y que da lugar a los smoother de vecindad más cercana. Otra que escoge la vecindad como un intervalo centrado en cada punto donde se aplica el smoother y que corresponde a los smoother de vecindad tipo bandwidth, en los que, a diferencia de los anteriores, lo que se fija a priori es el número de puntos de la vecindad, que en los puntos extremos no coincidirá con los puntos que intervienen en el ajuste. Los siguientes smoother son de este segundo tipo.

Kernel smoother. Se trata de tomar una media ponderada de los puntos de la vecindad de uno dado, donde los pesos vienen dados por una función kernel. El peso j -ésimo que produce la estimación en x_0 viene dado por la expresión

$$S_{0j} = c_0 d \left(\frac{|x_0 - x_j|}{b} \right).$$

La constante c_0 es la nos permite ajustar que los pesos sumen 1 si de inicio no es así, el parámetro b es el bandwidth o ancho de banda y se refiere al tamaño de la vecindad. La función $d(t)$ par y decreciente en $|t|$, es el kernel, que también suele denotarse $K(t)$, y es deseable que cumpla además las condiciones siguientes:

1.

$$\int_I K(t)dt \geq 0, \quad I \text{ un intervalo cualquiera}$$

2.

$$\int_{-\infty}^{\infty} K(t)dt = 1$$

Según las diferentes elecciones de la función kernel tenemos los diferentes tipos: Normal, Gaussiano, Triángulo, Parzen, y entre los más populares y que mejores resultados dan son los Epanechnikov.

Splines cúbicos. Un spline es una función polinomial definida a trozos para la cual el número máximo de derivadas existe. Consideremos una partición del intervalo $[a, b]$ dada por los puntos $a = x_1 < x_2 < \dots < x_n = b$, s es una función spline de grado k con nodos x_1, x_2, \dots, x_n si es un polinomio de grado k o menor en cada intervalo $[x_i, x_{i+1}]$ y s es $(k - 1)$ veces diferenciable.

Podemos definir la interpolación de una función $g(x)$, definida en $[a, b]$, con splines cúbicos de nodos x_1, x_2, \dots, x_n , como una función $s(x)$ que verifica:

1.

$$s(x) \in C^2(a, b)$$

2. En cada intervalo $[x_i, x_{i+1}]$, es un polinomio cúbico de la forma:

$$s(x) = s_i(x) = \sum_{l=0}^3 a_l^{(i)} (x_i - x)^l, \quad i = 1, \dots, n$$

3. En los nodos x_i ($i = 1, \dots, n$)

$$s(x_i) = g(x_i)$$

4. Las condiciones de borde

$$s''(a) = s''(b) = 0$$

La función de interpolación cúbica esta definida de forma unívoca con estas condiciones.

De entre todas las funciones $f(x)$ con dos derivadas continuas, los splines cúbicos tienen una propiedad conocida e interesante, son los que minimizan la suma de cuadrados penalizados

$$\sum (y_i - f(x))^2 + \lambda \int_a^b (f''(t))^2 dt,$$

donde λ es una constante fija. El primer término de la expresión mide la cercanía a los datos mediante los cuadrados de la diferencia entre observados y ajustados, como habitualmente, pero está penalizado por el término segundo que aumenta con

los cambios de curvatura de la función $f(x)$, de forma que los splines cúbicos intentan maximizar la bondad del ajuste pero además con suavidad, con pocas oscilaciones. El parámetro λ es el que determina la importancia de uno u otro término y tiene el mismo papel que el span del smoother de recta móvil o el bandwidth del kernel. Este método es conocido en términos más generales con fórmula de Whittaker-Henderson, que más tarde desarrollaremos.

3.2.1. Selección del parámetro de smoothing bandwidth o span

La elección óptima del parámetro bandwidth (o, equivalentemente, del span) es un tema ampliamente discutido en la literatura del smoothing de datos. Una técnica estadística habitual en la ciencia actuarial es primero elegir el modelo que mejor ajusta a los datos y después contrastar su suavidad. En la estadística moderna se combinan ambos pasos utilizando un método para elegir el bandwidth que intenta equilibrar varianza y sesgo. Gavin, Haberman y Verrall (1994) utilizan para ello la validación cruzada, que comparan después con el debido a Bloomfield y Haberman (1987), que ajustan una curva a los datos y después de forma separada contrastan la suavidad de los datos con tests de ajuste actuariales estándar.

Verrall (1996) nos describe brevemente el método de validación cruzada en el contexto de la graduación como aquel que minimiza

$$\frac{1}{n} \sum_{i=1}^n (\hat{q}_i - \hat{q}_i^{(-i)})^2.$$

Dado un estimador cualquiera \hat{q}_x de la verdadera probabilidad de muerte q_x , elegiremos el valor del bandwidth b , que minimiza:

$$CV(b) = \frac{1}{n} \sum_{i=1}^n (\hat{q}_i - \hat{q}_i^{(-i)})^2$$

donde $\hat{q}_i^{(-i)}$ es la estimación utilizando todos los valores brutos salvo el i -ésimo.

La validación cruzada consiste pues en estimar sucesivamente, de uno en uno, el suavizado en x_i a partir de los $n - 1$ puntos restantes, todos salvo (x_i, q_i) (por simplificar, escribimos q_i en lugar de q_{x_i}).

Teóricamente, minimizar $CV(b)$ es aproximadamente equivalente a minimizar el error medio entero (MISE) de \hat{q} definido por

$$\begin{aligned} MISE(\hat{q}) &= E(\int (\hat{q}_x - q_x)^2 dx) \\ &= \int (E(\hat{q}_x) - q_x)^2 dx + \int V(\hat{q}_x) dx \\ &= \text{Integral del cuadrado del sesgo} + \text{Integral de la varianza.} \end{aligned}$$

con lo que conseguimos un compromiso entre varianza y sesgo. Autores como Bloomfield y Haberman (1987) para elegir el bandwidth utilizan tres tests: test chi-cuadrado, el test de rachas y el test de correlaciones retardadas. El trabajo de Gavin, Haberman y Verrall (1994) compara los dos métodos para las tablas de asegurados 1987-70, datos del CMI, y

ambos producen similares elecciones, pero la validación cruzada es más intuitivo y teóricamente más sólido. En nuestro caso elegiremos el bandwidth por validación cruzada, comprobando después las buenas propiedades mediante los test.

Validación cruzada para smoothers lineales

Un smoother S es lineal si puede ser escrito como

$$\hat{q} = S\dot{q}$$

donde $S = s_{ij}$ es una matriz $n \times n$ llamada matriz smoother y \dot{q} son la probabilidades brutas. Los smoother media móvil, recta móvil, smoothing spline, kernel, loess, son todos ejemplos de smoothers lineales.

Denotamos por $\hat{q}_i^{(-i)}$ el ajuste obtenido asignando cero a la i -ésima observación e incrementando los pesos para hacer que su suma sea uno.

Formalmente,

$$\hat{q}_i^{(-i)} = \sum_{j=1, j \neq i} \frac{S_{ij}(\lambda)}{1 - S_{ii}(\lambda)} \dot{q}_j$$

de lo cual resulta

$$\hat{q}_i^{(-i)}(1 - S_{ii}(\lambda)) = \sum_{j=1, j \neq i}^n S_{ij}(\lambda) \dot{q}_j \quad (3.3)$$

y dado que

$$\hat{q}_i = \sum_{j=1, j \neq i}^n S_{ij}(\lambda) \dot{q}_j + S_{ii}(\lambda) \dot{q}_i, \quad (3.4)$$

sustituyendo la ecuación (3.3) en (3.4) tenemos

$$\begin{aligned} \hat{q}_i &= \hat{q}_i^{(-i)} - S_{ii}(\lambda) \hat{q}_i^{(-i)} + S_{ii}(\lambda) \dot{q}_i \\ &= (1 - S_{ii}(\lambda)) \hat{q}_i^{(-i)} + S_{ii}(\lambda) \dot{q}_i - \dot{q}_i + \dot{q}_i \\ &= (1 - S_{ii}(\lambda)) (\hat{q}_i^{(-i)} - \dot{q}_i) + \dot{q}_i \end{aligned}$$

de donde obtenemos

$$\dot{q}_i - \hat{q}_i^{(-i)} = \frac{\dot{q}_i - \hat{q}_i}{1 - S_{ii}(\lambda)} \quad (3.5)$$

Por lo tanto el ajuste $\hat{q}_i^{(-i)}$ puede calculado sin necesidad de eliminar el punto i -ésimo y recalculer el suavizamiento.

Con (3.5), la suma de cuadrados de validación cruzada puede ser reescrita como

$$CV(\lambda) = \frac{1}{n} \sum_{i=1}^n \left(\frac{\dot{q}_i - \hat{q}_i}{1 - S_{ii}(\lambda)} \right)^2.$$

Grados de libertad de un smoother

Conocer los grados de libertad nos servirá para poder comparar los distintos smoother y saber la cantidad de suavizado que realizan, así como la calidad de su ajuste medida en términos de algunos test. Los grados de libertad se obtienen sumando los valores propios de S_λ . Otras definiciones para los grados de libertad de un smoother lineal, propuestas por analogía con modelos de regresión lineal, son $n - \text{tr}(2S_\lambda - S_\lambda S_\lambda^T)$, $\text{tr}(S_\lambda S_\lambda^T)$ o $\text{tr}(S_\lambda)$.

Si suponemos el modelo $q_i = \hat{q}_i + r_i$, donde \hat{q}_i son las estimaciones por la función y r_i son los errores independientes e idénticamente distribuidos de media 0 y varianza σ^2 , los grados de libertad, $df = \text{tr}(S_\lambda)$, se pueden ver como una corrección para el promedio de los cuadrados residuales ($ASR = \frac{1}{n} \sum_{i=1}^n (q_i - \hat{q}_i)^2$). En particular el estadístico C_p es un corrector de ASR para hacerlo insesgado mediante la suma de la cantidad $2\text{tr}(S_\lambda)\hat{\sigma}^2/n$, donde $\hat{\sigma}^2$ es una estimación de σ^2 .

3.2.2. Kernel smoother

El poder de las modernas computadoras ha hecho que las técnicas smoothing sean mucho más accesibles y de fácil implementación (véase como ejemplo el software S-plus).

En la sección anterior hicimos una breve descripción del smoother tipo kernel, en esta sección vamos a revisar su aplicación al campo de la graduación de datos de mortalidad según Gavin, Haberman y Verrall (1993, 1994 y 1995).

En sus trabajos los autores estudian la relación entre las medias móviles y la graduación por kernel smoothing, muestran además que el método kernel proporciona una aproximación satisfactoria, y que muchos desarrollos estadísticos pueden ser incorporados dentro del área de la graduación.

Gavin, Haberman y Verrall (1993) exploran la flexibilidad de la estimación kernel como medio de graduación no paramétrica y lo relacionan con las medias móviles. Gavin, Haberman y Verrall (1995) busca en primer lugar un modelo que haga explícita la variación en los expuestos con la edad, considera además varias transformaciones de los datos, evalúan la validación cruzada como método de elección del parámetro de smoothing y también utilizan métodos de diagnóstico para validar las hipótesis. Finalmente Gavin, Haberman y Verrall (1995) discuten una función kernel para mejorar las estimaciones en los extremos, aplicando esta función a dos tablas de mortalidad.

Para un conjunto específico de probabilidades de mortalidad observadas \hat{q}_i , para un conjunto de edades $\{x_i\}_{i=1}^n$, se desea revisar estas estimaciones iniciales y producir una nuevas ya suavizadas \hat{q}_i de las que se han eliminado las fluctuaciones aleatorias. Se considera un modelo general

$$\hat{q}_i^t = \hat{q}_i^t + r_i, \quad (3.6)$$

donde t denota alguna transformación de las probabilidades, pues los datos pueden no tener varianza constante y los r_i se suponen variables aleatorias independientes e idénticamente distribuidas, con media cero y varianza constante y finita. Hipótesis que deberemos verificar con el diagnóstico del modelo, mediante gráficos y los tests no paramétricos habituales. En sus trabajos los autores se refieren solo a las probabilidades de muerte, pero los resultados podrían extenderse a tablas de múltiples decrementos o a la fuerza de mortalidad μ_x .

Gavin, Haberman y Verrall (1993) describen y comparan en su artículo dos estimadores de las q_x debidos a Copas y Haberman (1983) y Nayadara-Watson ((Nayadara, 1964), (Watson, 1964)), en los que no consideran transformación alguna. Ambos métodos kernel son discutidos en más profundidad, junto con los criterios para la elección de “bandwidth”, en Gavin, Haberman y Verrall (1994).

- **Estimador Copas-Haberman.**

Utilizado para graduar datos de mortalidad por Copas y Haberman (1983), Bloomfield y Haberman (1983) y por Ramlau-Hansen (1983), que también resalta la conexión con medias móviles ponderadas y estudia algunas de sus propiedades de optimalidad, la expresión de este estimador, obtenida a partir de la expresión de q_x según el teorema de Bayes, es

$$\hat{q}_x^{CH} = \frac{\sum_{i=1}^n d_i K_b(x - x_i)}{\sum_{i=1}^n e_i K_b(x - x_i)},$$

donde $\{e_i\}_{i=1}^n$ y $\{d_i\}_{i=1}^n$ son las sucesiones de expuestos inicialmente al riesgo y de muertes correspondientes a cada una de las edades estudiadas. Hay que resaltar el hecho de que en el numerador y en el denominador se usa la misma función kernel y el mismo bandwidth. Si el kernel es no negativo la estimación tomará valores en el intervalo de 0 a 1. Este estimador minimiza la log verosimilitud local binomial

$$\sum_{i=1}^n K_b(x - x_i) \{d_i \log(\hat{q}_i) + (e_i - d_i) \log(1 - \hat{q}_i)\},$$

donde $K_b(x) = K(x/b)$.

- **Estimador Nadaraya-Watson.**

Propuesto por Nayadara (1964) y Watson (1964), su expresión viene dada por

$$\hat{q}_x^{NW} = \frac{\sum_{i=1}^n \left(\frac{d_i}{e_i}\right) K_b(x - x_i)}{\sum_{i=1}^n K_b(x - x_i)}.$$

Este estimador minimiza los cuadrados medios locales

$$n^{-1} \sum_{i=1}^n K_b(x - x_i) (\dot{q} - \hat{q})^2.$$

Está claro que hay pérdida de información del primer estimador al segundo, puesto que éste último solo tiene en cuenta el ratio de muertes e ignora la cantidad de expuestos al riesgo en cada edad. Se podría esperar por ello que fuera mejor estimador el primero, sin embargo en cualquiera de las probabilidades brutas o iniciales hay una hipótesis implícita de que contiene la suficiente información respecto al número de muertes y expuestos al riesgo, puesto que son estadísticos suficientes.

En el ejemplo que analizan Gavin, Haberman y Verrall (1993) observan que el estimador de Nayadara-Watson tiende a infraestimar la verdadera curva de la mortalidad en las edades altas, donde la verdadera curva es cóncava, y a sobreestimar en las edades tempranas, donde la curva es convexa. Copas y Haberman (1983) llegan a la misma conclusión con el segundo estimador. En general el problema de la estimación kernel es que el sesgo se incrementa en los extremos de la tabla de mortalidad. Para evitar esto podrían transformarse los valores como consideramos en la expresión general, pero esto supone un problema para las edades que no tienen muertes en el cero (transformaciones logarítmicas, que suelen ser bastante frecuentes).

Verrall (1996) trata de soslayar el problema mediante una transformación de los datos a la que se pueda ajustar una recta, de forma que el sesgo inherente en cualquier graduación se reduce mucho a lo largo del rango ya que no hay curvatura y desaparecen los términos del sesgo a partir de la derivada segunda según su desarrollo de Taylor. En Gavin, Haberman y Verrall (1995) se analiza la conveniencia de estas transformaciones y otras soluciones. Después de exponer las ideas básicas de la graduación kernel, describen métodos para mejorar el estimador Nayadara-Watson cerca de los extremos, finalmente definen y discuten un kernel adaptativo general con bandwidth que varía con la edad. Nielsen (1992) ofrece una aproximación a la decisión teórica de reducción del sesgo via transformaciones.

Como transformaciones de los datos brutos podemos considerar las siguientes:

Transformación	Variable respuesta	Variable explicativa
logit	$\log\left(\frac{q_x}{1-q_x}\right)$	x_i
Gompertz	$\ln(-\ln(1-q_x))$	x_i
Weibull	$\ln(-\ln(1-q_x))$	$\ln(x_i)$

Muchas otras transformaciones son posibles sus referencias pueden encontrarse en Gavin, Haberman y Verrall (1995), pero su elección es subjetiva y depende de los relativos éxitos conseguidos con un determinado conjunto de datos. Ellos finalmente trabajan con la transformación logit, cuya inversa se encuentra dentro del intervalo $[0, 1]$ y refleja el hecho de que pequeños cambios cuando los ratios son pequeños son tan importantes como grandes cambios cuando los ratios son altos. La solución que se propone al problema de tener que aplicar transformaciones a los datos cuando existen edades con 0 muertes, es

agrupar las edades con bajo número de muertes que suelen ser las edades del final de la tabla y atribuirles la edad media del grupo.

Un análisis de la literatura relativa a la utilización de la técnica que nos ocupa para la graduación nos lleva a concluir que la elección de la función kernel K no tiene tanta influencia como el valor del parámetro de smoothing b . Gavin, Haberman y Verrall (1995) utiliza kernel estandarizado normal, aunque en Gavin, Haberman y Verrall (1993) se explora otro kernel que minimiza la varianza estimada por la curva.

Para el diagnóstico del modelo hemos de confirmar que las hipótesis respecto a los residuos, r_i , sean válidas. Debemos dibujar los residuos de la ecuación (3.6) frente a los estimadores de las probabilidades de muerte en la escala transformada para observar que no siguen ningún patrón claro, y por tanto se disponen de forma aleatoria.

Se obtienen además, mediante la media y varianza de una binomial, los residuos que el S-plus llama "Pearson residuals", las desviaciones estandarizadas entre las muertes observadas y las esperadas,

$$\frac{d_i - e_i \hat{q}_i}{\sqrt{e_i \hat{q}_i (1 - \hat{q}_i)}} \quad \text{para } i = 1, 2, \dots, n.$$

Esperamos que estos residuos tengan media cero y estén comprendidos entre $[-2, 2]$, la distribución puede no ser Normal pero deben cumplir la hipótesis de independencia, que puede comprobarse mediante el análisis de sus autocorrelaciones. Estos residuos ya son utilizados en los tests clásicos propuestos por Forfar, McCutcheon y Wilkie (1988).

Muchos otros gráficos y test no paramétricos pueden ser considerados como los sugeridos por Azzalati y Bowman (1993), Cleveland, Devlin y Groose (1988), Eubank y Thomas (1993).

Kernel con corrección en los extremos

Como hemos comentado anteriormente uno de los problemas del kernel smoothing es que el sesgo aumenta en los extremos. Una comparación del sesgo del estimador Nayadara-Watson en relación con el de Copas-Haberman podemos encontrarla en Gavin, Haberman y Verrall (1994). En general, el sesgo del estimador puede obtenerse de la siguiente forma siguiendo un desarrollo de Taylor,

$$E(\hat{q}_i^t) = q_i^t + \frac{\sum_{j=1}^n (x_i - x_j) K_b(x_j - x_i)}{\sum_{j=1}^n K_b(x_j - x_i)} (q_i^t)' + R, \quad (3.7)$$

donde $(q_i^t)'$ (primera derivada) denota la pendiente de la transformada de la verdadera curva en x_i y R es el resto, consistente en un término que agrupa las derivadas de orden superior. Para las edades centrales de la tabla el coeficiente de $(q_i^t)'$ resulta ser cero pues se compensan las diferencias, si están igualmente espaciadas, a derecha e izquierda del valor. Sin embargo en los extremos hay más valores en uno de los lados, de forma que el

coeficiente no es cero, y se incrementa el sesgo. Para mejorar la estimación en los extremos existen varias propuestas. Por ejemplo:

- Hall y Wehrly (1991) sugieren reflejar los datos de forma que los datos originales estén en el interior de una serie más amplia, cuyo tamaño hemos aumentado como si colocáramos un espejo en el extremo.
- Otra solución que es la desarrollada por Gavin, Haberman y Verrall (1995), está basada en la combinación lineal de dos estimadores. Consideremos dos estimadores q_i^1 y q_i^2

$$E(\hat{q}_i^1) = q_i + C_1 q_i' + R_1$$

$$E(\hat{q}_i^2) = q_i + C_2 q_i' + R_2,$$

expresados según la ecuación (3.7), donde C_1 y C_2 son los coeficientes de primer orden y R_1 y R_2 son los restos. Los coeficientes C_1 y C_2 dependen solo de la edad y no de la verdadera curva de mortalidad, por tanto podemos construir una combinación lineal de los estimadores en la que eliminemos el término de q_i' . Esto significa que el sesgo de nuestro estimador en los extremos no depende de la pendiente de la mortalidad sino de los términos de orden superior tales como la curvatura. A partir de esta idea, pero en el contexto de la estimación de la densidad, Jones (1993) sugiere redefinir la función kernel como una combinación lineal de $K(x)$ y $xK(x)$. Todo ello conduce al autor a definir una función kernel para el extremo de la parte derecha, con expresión

$$K_b^R(x) = \frac{[a_2(p) - a_1(p)x]K_b(x)}{a_0(p)a_2(p) - a_1^2(p)},$$

donde $a_l(p) = \int_{-\infty}^p u^l K_b(u)$ y $p = x/b$, que puede ser usada para reducir el sesgo cerca de la cota más alta. La variable p mide la distancia desde el punto en el que se calcula la estimación a la cota de la parte derecha, expresada en unidades de bandwidth. La variable x mide la distancia desde el punto a cada uno de las n probabilidades brutas, de nuevo en unidades de bandwidth. Si sustituimos en el kernel normal tenemos

$$K_b^R(x) = \frac{[\Phi(p) + (x-p)\phi(p)]\phi(x)}{\Phi(p)[\Phi(p) - p\phi(p)] - \phi^2(p)},$$

donde $\Phi(x) = \int \phi(y)dy$ y $\phi(y)$ son respectivamente las funciones de densidad y de distribución de una Normal tipificada. Para edades que están muy cerca del extremo izquierdo, es decir, para edades jóvenes, un ajuste de la fórmula nos conduce a

$$K_b^L(x) = \frac{\{[1 - \Phi(p)] + (p-x)\phi(p)\}\phi(x)}{[1 - \Phi(p)]\{[1 - \Phi(p)] + p\phi(p)\} - \phi^2(p)}.$$

Las funciones transformadas, K_b^L y K_b^R , se comportan como kernel Normal para edades de mitad del rango. Si $p > 2$, entonces $a_0(p) \rightarrow 1$ y $a_i(p) \rightarrow 0$, por tanto $K_b^L \approx K_b^N$ y $K_b^R \approx K_b^N$. Pero cuando las estimaciones se acercan a los extremos, los pesos cambian el perfil, convirtiéndose en asimétricos y negativos sobre algunas regiones.

Kernel con bandwidth variable

En todos los desarrollos anteriores hemos considerado el bandwidth constante a lo largo de todo el rango de edades, un valor fijo denotado b . Sin embargo resultaría una aproximación mucho más flexible si lo consideráramos variable, de forma que pudiera adaptarse al número de expuestos (tamaño de la muestra). En regiones donde el número de expuestos es muy pequeño, tales como las edades altas, un valor más alto del bandwidth conseguiría estimar las probabilidades de manera que aumentarían progresivamente de forma más suave. Así, en estas regiones calcularíamos una media local sobre un número más grande de observaciones lo cual reduciría la varianza. Esta técnica es la que se llama estimador kernel adaptativo o variable.

Algunos modelos adaptativos son:

- Considerar un bandwidth b_i diferente para cada edad x_i en la que la curva va a ser estimada. Podemos entonces medir la distancia de esa edad a cualquier edad observada x_j usando b_i . Por tanto el modelo es

$$\hat{q}_i^t = \sum_{j=1}^n s_{ij} \hat{q}_j^t, \quad \text{donde } s_{ij} = \frac{K_{b_i}(x_i - x_j)}{\sum_{j=1}^n K_{b_i}(x_i - x_j)}, \quad i = 1, \dots, n. \quad (3.8)$$

Si la edad en que se realiza la estimación no es una edad observada, podemos suavizar la densidad de probabilidad estimada

$$\hat{f}_i = \frac{e_i}{\sum_{j=1}^n e_j} \quad \text{para } i = 1, \dots, n. \quad (3.9)$$

- Otra alternativa es considerar un bandwidth b_j para cada edad observada x_j $j = 1, \dots, n$, y entonces medir la distancia a la edad x_i en la que hacemos la estimación. El resultado es el nuevo smoother

$$\hat{q}_i^t = \sum_{j=1}^n s_{ij} \hat{q}_j^t, \quad \text{donde } s_{ij} = \frac{K_{b_j}(x_i - x_j)}{\sum_{j=1}^n K_{b_j}(x_i - x_j)}. \quad (3.10)$$

El bandwidth local es simplemente el bandwidth global multiplicado por un factor local, $b_i = b l_i^s$ para $i = 1, \dots, n$. Como la variación de los expuestos entre las edades jóvenes y las mayores puede ser enorme, vamos a elegir el factor local de la manera siguiente

$$l_i^s \propto \hat{f}_i^{-s} \quad \text{para } i = 1, \dots, n; \quad 0 \leq s \leq 1, \quad (3.11)$$

donde s es el parámetro de sensibilidad. Tomando $s = 0$ reducimos ambos casos al de bandwidth fijo, mientras que si tomamos $s = 1$ puede resultar un bandwidth con mucha variación. Por convenio, vamos a tomar la inversa de $\max\{\hat{f}_i^{-s} : i = 1, \dots, n\}$ como

constante de proporcionalidad en (3.11), de forma que $0 < l_i^s \leq 1$ y además a una edad x_i , con un reducido número de expuestos, le corresponde un factor local l_i^s cuyo valor es grande. Gavin, Haberman y Verrall (1995) exploran dos ejemplos, en primero aplica (3.8) y en el segundo (3.10).

Otras alternativas, propuestas por el autor, que podríamos considerar son:

- Elegir $l_{ij}^s = (e_i/e_j)^s$, para $i, j = 1, \dots, n$. En este caso el bandwidth global se ajusta mediante el cociente de expuestos de forma que el peso correspondiente a la j -ésima probabilidad bruta para estimar el correspondiente a la i -ésima viene dado por

$$\hat{q}_i^t = \sum_{j=1}^n s_{ij} \hat{q}_j^t, \quad \text{donde } s_{ij} = \frac{K_{b_{ij}}(x_i - x_j)}{\sum_{j=1}^n K_{b_{ij}}(x_i - x_j)} \tag{3.12}$$

y $b_{ij} = b l_{ij}^s$ y $i, j = 1, \dots, n$.

- También podríamos tener en cuenta el perfil de la curva utilizando $l_i^s = (|(q_i^t)''| f_i)^{-s}$. Los verdaderos valores de los q_i no son conocidos, de forma que deberíamos utilizar la estimación inicial, es decir, sus segundas diferencias, para aproximar la curvatura sin distinguir entre positiva y negativa. Podemos usar los \hat{f}_i definidos en (3.9) como estimación de f_i .
- Una razonable alternativa propuesta por Jones (1993) es

$$\hat{q}_i^t = \frac{\sum_{j=1}^n w_j \hat{q}_j^t K_b(x_i - x_j)}{\sum_{j=1}^n w_j K_b(x_i - x_j)}$$

donde los w_j son pesos que dependen de la cantidad de expuestos.

Para la presentación de todos los modelos expuestos hemos requerido de dos parámetros s y b que hemos supuesto determinados. Pero en la práctica el criterio de elección de los parámetros propuestos, sensibilidad, s , y bandwidth global, b , debe ser anterior a la aplicación de dichos modelos. El parámetro de sensibilidad puede ser elegido por validación cruzada, no obstante cuando la variación de la cantidad de expuestos es grande el kernel adaptativo se vuelve muy sensible a ella con lo que resultan valores de s no razonables por ser demasiado elevados. Así pues, se suele elegir subjetivamente el valor de s y después utilizar la validación cruzada para obtener el valor de b .

3.3. La Fórmula de Whittaker-Henderson

Método iniciado por Whittaker (1923) y Henderson (1924) tiene como objetivo minimizar una función que auna los criterios de ajuste y suavidad.

Sea x la edad, d_x el número observado de muertes y e_x la estimación de los expuestos al riesgo iniciales o centrales, según que las cantidades a estimar sean q_x o μ_x , respectivamente, y cuyas estimaciones brutas son los cocientes d_x/e_x . Denotamos en general θ_x a la cantidad a estimar y $\hat{\theta}_x$ a su estimación bruta. Entonces, el método de Whittaker consiste en encontrar el estimador de θ_x , $\hat{\theta}_x$, que minimiza

$$W(\theta) = \sum_{x=1}^n w_x (\hat{\theta}_x - \theta_x)^2 + \alpha \sum_{x=z+1}^n (\Delta^z \theta_x)^2,$$

donde w_x son las ponderaciones, α es el parámetro de smoothing, Δ es el operador diferencia definido por

$$\Delta \theta_x = \theta_x - \theta_{x-1}$$

y Δ^z es el operador diferencia aplicado z veces. Minimizar esta función es encontrar valores que equilibren la bondad del ajuste, primer término de la función, y la suavidad o smoothing, segundo término.

Para el caso particular de $z = 2$ el criterio de Whittaker intenta encontrar unas estimaciones $f(x)$ que alcancen un equilibrio entre bondad de ajuste y suavidad. Puede explicitarse entonces de la siguiente forma

$$\sum (y - f(x))^2 + \lambda \sum (\Delta^2 f(x))^2.$$

Donde $\Delta^2 f(x)$ indica la segunda diferencia en las estimaciones, es decir, aplicar diferencias dos veces $[f(x) - f(x-1)] - [f(x-1) - f(x-2)] = f(x) - 2f(x-1) + f(x-2)$. El parámetro λ es el que da la importancia relativa entre sesgo y varianza, puede ser elegido subjetivamente o por métodos automáticos como validación cruzada. Se puede demostrar (Reinch, 1967) que la estimación $\hat{f}(x)$ es un spline cúbico, con nodos en cada valor de x .

En el artículo de Forfar, McCutcheon y Wilkie (1988) tenemos una breve introducción de la graduación por splines, que fue utilizada en las ELT14 (English life tables nº 14). Además Betzuen (1997) ha utilizado este método para la graduación de los datos de un colectivo de activos. Chan et al. (1986) muestran como esta graduación puede formularse como un problema de programación lineal en unos casos y como un problema de programación cuadrática en otros.

Posteriormente ha sido tratada en diversos trabajos, de los que destacamos dos de Verrall. El primero Verrall (1993b), describe brevemente el método y se relaciona claramente con la metodología Bayesiana, el segundo Verrall (1996), tiene una descripción muy interesante del método en relación con los Modelos Aditivos Generalizados y su implementación en S-plus. Verrall (1993a) muestra como la graduación de Whittaker puede ser expresada como un modelo lineal generalizado dinámico, estimación que puede ser hecha recursivamente mediante el filtro de Kalman. Demuestra que es equivalente a un análisis de regresión dinámica, ajuste de una recta con varianza estocástica. Por lo que es de hecho una combinación de método paramétrico y método de smoothing.

3.4. Graduación polinomial local

El smoother de regresión local ponderada es un método que ajusta localmente polinomios con bajo grado y parece ser prometedor según Verrall (1996). Este método, que se debe a Cleveland (1979) y está discutido en detalle en Fan y Gijbels (1996), se encuentra implementado en S-plus con el comando loess.

Consideremos una variable respuesta Y con distribución Normal y un predictor simple X , de los que tenemos los datos $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. Dado un valor x_0 llamamos $\hat{f}(x_0)$ al estimador que se obtiene de la siguiente forma:

- 1 Para cada valor x_0 , sea $N(x_0)$ el conjunto de los k vecinos más próximos a x_0 .
- 2 Se define $\Delta(x_0) = \max_{x_i \in N(x_0)} |x_0 - x_i|$, la distancia más grande en la vecindad de x_0 .
- 3 Se asignan pesos ω_i a cada punto de $N(x_0)$, usando la función de pesos tricúbica

$$\omega_i = T\left(\frac{|x_0 - x_i|}{\Delta(x_0)}\right)$$

donde

$$T(u) = \begin{cases} (1 - u^3)^3 & \text{para } 0 \leq u \leq 1 \\ 0 & \text{en otro caso.} \end{cases}$$

- 4 El valor, $\hat{f}(x_0)$ obtenido por el ajuste de Y sobre X por mínimos cuadrados ponderados (con pesos ω_i) en la vecindad $N(x_0)$.

La cantidad de suavizamiento viene dada por el parámetro λ y denota la proporción del número de elementos de la vecindad frente al número total de puntos, $\lambda = \frac{\#N_k^s(x_0)}{n}$.

Hay dos casos extremos interesantes, cuando $\lambda = \frac{1}{n}$ cada vecindad contiene un solo dato, el smoother corresponde a la interpolación, y cuando $\lambda = 2$, que da lugar a la recta de regresión por mínimos cuadrados ponderados. Por defecto tenemos la regresión local lineal pero en la práctica suele resultar mejor ajustar mediante regresión local cuadrática utilizando un polinomio de grado 2. Por su característica de ser ponderado, la importancia relativa de cada punto en la estimación viene dada por la distancia al punto x_0 correspondiente. Así pues se ajusta mejor a los valles y a los picos y por tanto, de los smoothers de vecindad más cercana es el de mejor rendimiento.

Este método de graduación posee varias similitudes con los modelos lineales generalizados lineales sugeridos por Verrall (1993b), y con la graduación de Whittaker, modelos todos ellos basados en una recta cuyo gradiente varía suavemente con la edad (Verrall, 1993a)

En su artículo Verrall propone la graduación por modelos lineales generalizados dinámicos, que consiste en extender la teoría de modelos lineales generalizados (Renshaw, 1991) a modelos cuyos coeficientes no sean constantes a lo largo del rango de edades sino que varíen con ella. Este método consigue una regresión que es localmente lineal pero no globalmente. Puede ser considerado como una combinación de métodos paramétricos, pues

se considera una función dependiente de parámetros, y no paramétricos, pues a su vez considera un procedimiento de smoothing que es el filtro de Kalman.

3.5. Modelos aditivos generalizados (GAM)

Muchos métodos de graduación no paramétricos descritos hasta el momento, pueden ser reformulados como Modelos Aditivos Generalizados. Los GAM se estudian en profundidad en el texto de Hastie y Tibshirani (1990). Estos modelos mejoran la formulación que ya tenían en la teoría actuarial los métodos no paramétricos. Utilizan la misma estructura que los GLM, pero los predictores lineales son funciones en las que se aplican diferentes técnicas de smoothing.

Los trabajos con planteamientos similares son Verrall (1993b) y Verrall (1996) que muestran la graduación de Whittaker, descrita en la sección 3, expresada como un modelo lineal generalizado dinámico y puede ser mejorada recursivamente por el filtro de Kalman. Además, en sus artículos relaciona la graduación de Whittaker con los métodos paramétricos, y describe como puede obtenerse como extensión de un modelo lineal particular. Otro trabajo en este sentido es Verrall (1996) que describe varios métodos paramétricos y no paramétricos dentro del esquema de implementación en S-plus, utilizando esquemas GLM ya descritos.

Los GAM aportan una mejora respecto a la práctica habitual, pues la forma usual de trabajar con la graduación de Whittaker es por aproximación de una Normal, sin embargo con el enfoque de los modelos aditivos generalizados esto no es necesario, puesto que permiten utilizar la distribución exacta. En el caso de graduar los q_x se considera la distribución binomial y el link logit, y si graduamos μ_x tendríamos la distribución Poisson y el link log, como ya ocurría con los GLM, pero sin embargo como hemos dicho, varía el predictor. En efecto, el predictor η ya no tiene porque ser una función lineal de las variables predictivas, sino que es

$$\eta = \alpha + \sum_{j=1}^p f_j(x_j) + \epsilon$$

donde f_j es una función smooth (splines, regresión local, ...). El predictor aditivo η puede aplicarse en todas las situaciones en las que los modelos lineales generalizados utilizaban el predictor lineal.

Capítulo 4

Modelos dinámicos

4.1. Introducción

Después de haber estudiado en anteriores capítulos la graduación de datos de mortalidad por diferentes métodos paramétricos y no paramétricos, centrándonos en la influencia que la edad tiene sobre ésta, tenemos ahora que analizar el efecto del tiempo del calendario, dado que para experiencias de mortalidad correspondientes a periodos diferentes las probabilidades de muerte a una determinada edad son diferentes y entonces la hipótesis de estacionariedad no se cumple.

Este hecho se ve corroborado al observar que la mortalidad muestra un gradual declive a lo largo del tiempo, con los tantos de descenso no necesariamente uniformes a través de los rangos de edad. Así pues, es importante ser capaces de medir con exactitud cambios en la mortalidad a través del tiempo, ya que los beneficios de las pólizas, por ejemplo para pensionistas y anualidades, dependen de la supervivencia. Por este motivo, si la tabla estándar utilizada para calcular tantos anuales y reservas predice unas probabilidades de mortalidad más altas que las realmente experimentadas por los asegurados, se habrá cobrado de menos a éstos, y la compañía incurrirá en pérdidas. Y no sólo eso, sino además las reservas para tal fin serán insuficientes.

Es un hecho reconocido por la literatura actuarial más reciente que la mortalidad evoluciona con el transcurrir de los años. Esto es equivalente en nuestro estudio a que la probabilidad de fallecimiento para una edad x es diferente para cada año de calendario t , es por ello que necesitamos desarrollar modelos que nos permitan estimar esta dinamicidad y realizar predicciones.

Por todas estas razones el objetivo de este capítulo será describir los métodos de graduación de la mortalidad incorporando el tiempo del calendario o cronológico. Dichos métodos son en su mayoría métodos paramétricos porque, como ya señalábamos en el Capítulo 2, la graduación paramétrica favorece la comparación a lo largo del tiempo y del espacio, estudiando la evolución de los parámetros a estimar o incorporando una nueva variable explicativa que represente su influencia en la función.

Estadísticamente se trata pues de modelizar las medidas de la mortalidad con una función $f(x, \alpha, t) + \epsilon_x$, expresión en la cuál las propuestas sobre la forma funcional f han

sido múltiples. Para cubrir nuestro objetivo vamos a recopilarlas y clasificarlas.

Una recopilación de algunas de estas técnicas podemos encontrarla en Benjamin y Soliman (1995). Posteriormente Felipe y Guillén (1999) han seguido dicha clasificación de métodos paramétricos aunque la generalizan para incluir los modelos desarrollados desde entonces. Para introducir tanto los modelos como su clasificación hemos de considerar las medidas habituales de la mortalidad pero dependientes además del tiempo cronológico. Así,

q_{xt} , es la probabilidad de que un individuo de x años de edad muera antes de cumplir $x + 1$ en el año cronológico t .

μ_{xt} , es la fuerza de mortalidad de que un individuo de x años en el año cronológico t .

m_{xt} , tanto central de mortalidad, es el ratio central de muerte o el ratio de mortalidad medio sobre el intervalo de edad x a $x + 1$ en el año cronológico t . Relaciona las muertes con la población en el centro del año. Además puede relacionarse con las medidas anteriores a través de la expresión,

$$m_{xt} = \frac{2q_{xt}}{2 - q_{xt}}.$$

Pasamos a describir la clasificación seguida en nuestro trabajo que amplía la de Felipe y Guillén (1999).

Métodos paramétricos.

Son modelos que, como ya los estudiamos en el Capítulo 2, ajustan una función dependiente de unos parámetros a las medidas de la mortalidad, pero ahora incorporan la influencia del tiempo cronológico. Para ello son dos, básicamente, los tratamientos:

1. Modelizar las medidas de la mortalidad q_{xt} , μ_{xt} o m_{xt} mediante una misma ley de mortalidad a lo largo de los años y considerar que la influencia del tiempo del calendario sólo afecta a los parámetros, de forma que el modelo se expresa mediante una función $f(x, \alpha_1(t), \alpha_2(t), \dots, \alpha_n(t))$. A estos métodos los denominaremos **métodos estructurales**.
2. Incorporar el tiempo cronológico como variable t en la función $f(x, \alpha_1, \alpha_2, \dots, \alpha_n, t)$ que modeliza la medida de mortalidad q_{xt} , μ_{xt} o m_{xt} . A estos métodos los denominaremos **no estructurales**.

Métodos no paramétricos

Modelos que ya introdujimos en el Capítulo 3, fundamentalmente describimos un método kernel que además ahora va a depender del tiempo.

Métodos basados en factores de reducción

Son modelos que después de modelizar las medidas de mortalidad incorporan la evolución de éstas a lo largo del tiempo a través de los factores de reducción. Factores que a su vez vamos a modelizar.

El contenido del capítulo está estructurado de la siguiente forma. La sección 2 se dedica a los métodos estructurales, dentro de la cuál se describe en primer lugar la ley de mortalidad de Heligman y Pollard, pero incorporando dinamicidad a sus parámetros mediante el tiempo, siguiendo el trabajo de Felipe y Guillén (1999). Después el método logit debido a Brass (1969) que se encuentra descrito en el texto de Benjamin y Soliman (1995), para terminar con el más reciente, el método de Lee-Carter, que comienza con el trabajo de Lee y Carter (1992).

En la sección 3 estudiamos la influencia del tiempo del calendario como variable según los métodos no estructurales, el primero de los cuales es el método logit que describimos según Benjamin y Soliman (1995), después los métodos lineales generalizados que incorporan dinamicidad a los ajustes del CMI Bureau, técnica recogida en los Informes 10 (1990) y 17 (1999) y finalmente el esquema de GLM (Renshaw, Haberman y Hatzopoulos, 1996) que se utiliza para conseguir no sólo un buen ajuste a los datos sino también una predicción adecuada en un futuro (Sithole, Haberman y Verrall, 2000).

La sección 4 recoge métodos no paramétricos dinámicos correspondientes a un suavizado de tipo kernel que depende de la edad x y el tiempo t .

Por último la sección 5 está dedicada a los factores de reducción de la mortalidad para obtener tablas de mortalidad proyectadas, incluyendo diferentes metodologías.

4.2. Métodos estructurales

La representación de la evolución de la mortalidad via métodos estructurales es un tema de actualidad tanto para actuarios como para demógrafos y estadísticos.

Los métodos estructurales se basan en la hipótesis de que la medida de mortalidad elegida, q_{xt} , μ_{xt} o m_{xt} , es función de la edad x , de forma que dicha función, o ley de mortalidad, es la misma a lo largo de los años y depende del tiempo a través del valor de sus parámetros, $f(\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t), x)$. Los valores de los parámetros, $\alpha_1(t), \alpha_2(t), \dots, \alpha_n(t)$, correspondientes a los años observados, se determinan ajustando la ley mediante los métodos habituales descritos en el capítulo de métodos paramétricos y, posteriormente, pueden ser modelados con la metodología Box-Jenkins puesto que cada parámetro está representado por una serie temporal.

Estos modelos de series temporales se utilizan para obtener las estimaciones de los parámetros de años futuros, que sustituidos en la función matemática supuesta, o ley de mortalidad, nos permite realizar la predicciones correspondientes a la mortalidad de cada año y edad. Pasemos a describir cada una de las propuestas.

4.2.1. Modelo de Heligman y Pollard

En el capítulo de modelos paramétricos comentamos la utilización del modelo de Heligman-Pollard para la obtención de tablas de mortalidad. La utilización que de este modelo hacen McNown y Rogers (1988), adaptado al modelo español, aparece ampliamente descrito en Felipe y Guillén (1999) y es en este trabajo en el que basaremos nuestra descripción. Es un método estructural en el sentido de que los cambios, según la variable

tiempo t , en las medidas de la mortalidad, q_{xt} , afectan a cambios en los parámetros de la ley de mortalidad ajustada a los diferentes años.

Método de ajuste

La ley de mortalidad empleada para cada uno de los años del periodo estudiado es la segunda ley de Heligman y Pollard,

$$q_{xt} = A_t^{(x+B_t)C_t} + D_t \exp(-E_t(\ln x - \ln F_t)^2) + \frac{G_t H_t^x}{1 + K_t G_t H_t^x}.$$

La descripción de las leyes de Heligman y Pollard, así como la interpretación de sus parámetros está ampliamente descrita en la sección 2.6.

Los parámetros se estiman para cada uno de los años mediante mínimos cuadrados ponderados no lineales para todo el rango de edades,

$$\sum_x \omega_x (q_x - F(x))^2,$$

donde la varianza de la observación a edad x es proporcional a ω_x^{-1} y $F(x)$ es la función a ajustar a los valores q_x observados.

La necesidad de estos pesos es debida a la desigualdad de varianzas en el modelo $q_x = F(x)$, las cuáles, siguiendo el modelo binomial, valen $var(q_x) = \frac{q_x(1 - q_x)}{E_x}$ de ahí que los pesos propuestos sean $1/q_x$ o alguna potencia inversa de q_x .

La ponderación que utilizan Felipe y Guillén (1999) es la inversa de la probabilidad de fallecimiento observada, criterio ya utilizado por Heligman y Pollard (1980). Los valores de la estimaciones de los parámetros dependen de varios condicionantes que afectan al algoritmo de programación no lineal utilizado, en su caso Gauss-Newton, como por ejemplo

- los valores iniciales de las estimaciones,
- el rango de edades a ajustar.
- el método de cálculo de la longitud de cada salto del logaritmo iterativo.

Las autoras varían los condicionantes hasta dar con una estimación adecuada, para después seleccionar las combinaciones que proporcionaban los valores de los parámetros con menor suma de cuadrados de los residuos y con mayor porcentaje de variabilidad explicada por la regresión sobre la variabilidad total (R^2). El objetivo del ajuste es encontrar unos valores de los parámetros de forma que el rango de edad sea el mayor posible, sin embargo se observa que cuando el rango aumenta la variabilidad de la curva también, siendo más difícil la estimación y pudiendo ocurrir que en algunos casos se vuelva singular la matriz hessiana. Las autoras eligen edades máximas de 88, 89, 90 pues para sus datos la edad máxima de 95 produce una matriz hessiana singular.

Predicción

Dado que el modelo estudiado es estructural, una vez estimados los valores de los parámetros de la segunda ley de Heligman y Pollard en cada uno de los años según el apartado anterior, el siguiente paso es determinar el modelo que sigue cada uno de los parámetros.

Para ello, se aplica el análisis univariante de series temporales a cada uno de los parámetros de la ley para poder obtener las predicciones de cada uno de ellos y, más tarde, obtener la predicción deseada de la probabilidad de fallecimiento. Se sigue la metodología habitual,

1. Primero se trata de identificar el modelo, para lo cual se representa la serie gráficamente, se obtienen los valores de la función de autocorrelación simple y autocorrelación parcial y sus correspondientes autocorrelogramas que muestran la significatividad de algunos de ellos.
2. Después se proponen varios modelos ARIMA. Para validar el modelo ARIMA se estiman sus parámetros con todas las componentes de la serie temporal salvo las dos últimas. Con el modelo así obtenido, se calculan predicciones para las dos últimas componentes y se comparan con sus valores iniciales, lo que nos sirve para evaluar la capacidad predictiva del modelo propuesto. La bondad del ajuste se mide mediante los siguientes estadísticos;

- El error cuadrático medio (ECM)

$$\frac{1}{T} \sum_{t=1}^T e_t^2$$

- El error absoluto medio (MAE)

$$\frac{1}{T} \sum_{t=1}^T |e_t|$$

- El error porcentual absoluto medio (EPAM o MAPER)

$$\frac{100}{T} \sum_{t=1}^T \frac{|e_t|}{X_t}$$

- El error medio (ME)

$$\frac{1}{T} \sum_{t=1}^T e_t$$

- El error porcentual medio (MPE)

$$\frac{100}{T} \sum_{t=1}^T \frac{e_t}{X_t},$$

siendo X_t la serie temporal del parámetro estudiado.

También se realizan contrastes sobre los residuos para ver su comportamiento aleatorio,

- RUNS, contraste para un valor excesivo de rangos positivos y negativos.
 - RUNM, contraste para valores excesivos por encima o por debajo de la media.
 - AUTO, contraste de Box-Pierce para una autocorrelación excesiva, basada en la suma de cuadrados de los 24 coeficientes de autocorrelación.
 - MEAN, contraste para la diferencia en media entre la primera mitad y la segunda.
 - VAR, contraste para la diferencia en varianza entre la primera mitad y la segunda.
3. El modelo que mayor bondad de ajuste y capacidad predictiva tenga será el elegido, pero la estimación final se llevará a cabo utilizando todas las observaciones.

Este tipo de análisis, con la metodología propuesta, podría hacerse teniendo en cuenta intervalos de confianza y errores en la predicción y de esta forma las posibilidades se amplían enormemente.

4.2.2. El método logit

Otra metodología a considerar es la propuesta por Brass (1969) descrita en Benjamin y Soliman (1995), que observa que el comportamiento del $\text{logit}(1 - l_{xt})$ con la edad es similar para los diferentes periodos de observación. Considera la transformación logit para cada edad x y periodo t ,

$$\text{logit}(1 - l_{xt}) = 0,5 \ln \left(\frac{1 - l_{xt}}{l_{xt}} \right),$$

donde l_{xt} es la función de supervivencia de una tabla de mortalidad, y la relaciona con la de un periodo estándar s ,

$$\text{logit}(1 - l_{xt}) = \alpha_t + \beta_t \text{logit}(1 - l_{xs}).$$

Sin embargo, como la linealidad entre las curvas logit no siempre se cumple, Brass(1971) sugiere que la linealidad puede ser obtenida empleando la diferencia entre el logit para un periodo t y el logit del periodo estándar

$$\text{logit}(l_{xt}) - \text{logit}(l_{xs}) = \alpha_t + \beta_t \text{logit}(l_{xs}).$$

El autor utiliza este sistema para obtener resultados tanto para Suiza como para Inglaterra y Gales con datos generacionales.

Benjamin y Soliman (1995) aplican este método a valores de l_{xt} para grupos de edad de 5 años estimando los parámetros mediante mínimos cuadrados. El periodo estándar se fija en el año inicial (método prospectivo) o final (método retrospectivo) del periodo de años

correspondiente a los datos observados considerados. Una vez estimados los parámetros puede calcularse una predicción para el periodo t mediante

$$l_{xt} = \frac{1}{1 + \exp(2\text{logit}(l_{xt}))}.$$

Benjamin y Soliman (1995) observan que el método logit sobreestima la mejora de la mortalidad para el rango total de edades para el grupo de 15-19 años proporciona una predicción distante de la realmente observada, mientras que es mucho más cercana para edades extremas. Felipe y Guillén (1999) añaden también el inconveniente de que es un método meramente empírico.

4.2.3. Método de Lee-Carter

Continuando con la idea de desarrollar métodos con funciones paramétricas que dependen del tiempo a través de los parámetros, tenemos el desarrollado por Lee y Carter (1992), que trata de ajustar una función lineal a los logaritmos de los tantos centrales de mortalidad observados para cada edad específica. Los parámetros de dicha función dependen de la edad x y la variable independiente es un índice de intensidad k_t (no observado) dependiente de cada periodo específico t .

Se ajusta a la matriz formada por los logaritmos de los tantos centrales de muerte dispuestos por filas según la edad, x , y columnas según el año, t , dicho modelo, utilizando el método de descomposición de valor singular (SVD) que permite obtener la estimación por mínimos cuadrados de los parámetros. Dada una matriz A , se denomina SVD al problema de encontrar la descomposición siguiente

$$A = UDV^t,$$

donde U y V son ortogonales y D diagonal. A los elementos de la matriz D se les llama valores singulares, a las columnas de U , vectores singulares por la izquierda, y a las columnas V , vectores singulares por la derecha. La descripción extensa del método puede encontrarse en muchos textos de algebra, sirva por ejemplo Lay (2001).

Posteriormente se utilizan series temporales para obtener predicciones e intervalos de confianza para las medidas de la mortalidad a edades específicas

Lee y Carter (1992) muestran la aplicación de este método con datos de los EEUU de 1933 a 1987 y obtienen predicciones desde 1990 a 2065. En su trabajo destacan como ventaja de su modelo demográfico de mortalidad que representa el nivel de mortalidad mediante un único índice, k_t , además de diferenciarlo claramente del método propuesto por McNown y Rogers (1988) y McNown y Rogers (1992), descrito con anterioridad en la sección 4.2.1, que necesita de muchos más parámetros.

El modelo y su ajuste

Sea m_{xt} el ratio central para la edad x en el año t . Lee y Carter proponen ajustar a la matriz de ratios de mortalidad la función

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

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

$$\ln(m_{xt}) = a_x + b_x k_t + \epsilon'_{xt}. \quad (4.1)$$

En ambos casos siguiéndose una apropiada elección del conjunto de constantes específicas para la edad, $\{a_x\}$ y $\{b_x\}$, y para el índice de variación de año k_t . Hay que señalar que por tratarse de una función exponencial las predicciones de tantos centrales de mortalidad no pueden tomar valores negativos, lo que por otra parte sería totalmente absurdo.

La interpretación de los parámetros es la que sigue,

1. La sucesión de valores e^{a_x} describen el perfil general a lo largo de la edad del esquema de mortalidad.
2. La evolución de b_x da idea de lo rápidamente que decrecen los ratios en respuesta a cambios en k_t $\left(\frac{d \ln m_{xt}}{dt} = b_x \frac{dk_t}{dt}\right)$.
3. Los valores k_t representan la tendencia de la mortalidad a lo largo del periodo t

Los errores ϵ_{xt} , con media cero y varianza σ_ϵ^2 , reflejan influencias históricas particulares de cada edad específica que no son explicadas por el modelo.

Para estimar el modelo para una matriz dada de ratios m_{xt} , hemos de recurrir a la solución por mínimos cuadrados de la ecuación (4.1). Pero este modelo es indeterminado puesto que dada una solución (a, b, k) son soluciones también $(a - bc, b, k + c)$ y $(a, bc, k/c)$, donde c es un escalar cualquiera. Así pues, para que la solución sea única, es necesario normalizar b_x para que sumen 1 y k_t para sumar 0, lo cual implica que los a_x son simplemente las medias a lo largo del tiempo de $\ln(m_{xt})$. En efecto,

$$\begin{aligned} \sum_t \sum_x \ln(m_{xt}) &= \sum_t \sum_x (a_x + b_x k_t) \\ &= \sum_t \sum_x a_x + \sum_t \sum_x b_x k_t \\ &= \sum_x a_x \sum_t 1 + \left(\sum_t k_t\right) \left(\sum_x b_x\right) \end{aligned} \quad (4.2)$$

Si

$$\sum_t k_t = 0 \quad \sum_x b_x \neq 0 \quad (4.3)$$

desaparece el segundo sumando de (4.2) y entonces,

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

de aquí

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

Si $a_x = \frac{\sum_t \ln(m_{xt})}{T}$ se cumpliría con la restricción, pues entonces,

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

Alternativamente el segundo sumando de (4.2) podría ser cero tomando $\sum_x b_x = 0$, no obstante se utiliza (4.3) por causas empíricas. Según Lee y Carter (1992) los $\{b_x\}$ podrían ser negativos indicando que la mortalidad a esas edades tiende a subir mientras cae para otras edades, pero esto en la práctica no les ocurre a lo largo de un periodo largo de años. Por otra parte Wilmoth (1993) expone que es preferible hacer la suma de los $\{k_t\}$ cero pues esto facilita la comparación entre poblaciones.

El modelo (4.1) no puede ser ajustado por métodos de regresión habituales, pues a su izquierda tiene la variable dependiente $\ln(m_{xt})$ pero a su derecha tiene una función cuyos parámetros acompañan a un índice k_t y no a una variable independiente con valores observables.

Recurrimos al método de descomposición de valor singular (SVD) aplicado a la matriz de los logaritmos de los tantos centrales de mortalidad menos los valores de la a_x , que es el que emplean los autores para encontrar la solución de mínimos cuadrados. Tenemos pues que resolver el sistema

$$\ln(m_{xt}) - \frac{\sum_t \ln(m_{xt})}{T} = b_x k_t. \quad (4.4)$$

Si por simplicidad llamamos $Z_{xt} = \ln(m_{xt}) - \frac{\sum_t \ln(m_{xt})}{T}$, entonces podemos estimar, respectivamente, k_t y b_x como los primeros vectores derecho e izquierdo correspondientes al principal valor de la SVD de la matriz $Z_{x,t}$, o bien, estimar k_t como $\hat{k}_t = \sum_x Z_{xt}$ para después determinar b_x a partir de la regresión (4.4), una vez sustituidos los valores \hat{k}_t .

Los tantos centrales de mortalidad estimados de esta forma generalmente no conducen al número real de muertes para el total de edades, puesto que k_t ha sido estimada minimizando los errores en los logaritmos y no en los mismos ratios de muerte. Por esto, utilizando la ecuación

$$D_t = \sum_x (N_{xt} \exp(a_x + k_t b_x)), \quad (4.5)$$

donde D_t es el número total de muertes y N_{xt} es la población de edad x en el año t , reestimamos k_t en un segundo paso, sustituyendo los a_x y b_x por las estimaciones obtenidas en el primer paso. Estos valores de k_t se encuentran mediante búsqueda iterativa y difieren de la estimaciones directas mediante SVD.

La expresión (4.5) que puede también ser empleada con el propósito inverso, es decir, para resolver analíticamente tablas de mortalidad particulares que produzcan un número

de muertes observado D_t , para una distribución de la población por edades, N_{xt} . De esta forma, resulta útil para periodos en los cuales las distribuciones de población por edades y el total de muertes son conocidas pero los tantos de mortalidad centrales no están disponibles puesto que no se conocen las muertes por edad en cada año.

Después de haber desarrollado y ajustado el modelo demográfico, vamos a centrarnos en el problema de la predicción. El último paso del método de Lee-Carter sería encontrar un modelo para los valores de los índices de mortalidad, $\{k_t\}$, utilizando metodología Box-Jenkins. En muchas de las aplicaciones, se obtiene un buen modelo para k_t mediante la expresión

$$k_t = c + k_{t-1} + u_t.$$

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

Extensiones del método de Lee-Carter

Las modificaciones del método para predecir otras medidas de mortalidad, mejorar las predicciones o modelizar subpoblaciones han sido varias. A continuación enumeramos algunas de ellas.

Una primera modificación tiene por objetivo analizar y predecir variaciones en las expectativas de muerte por sexo y ha sido desarrollada por Carter y Lee (1992). Se basa en examinar diferencias en las observaciones y en las predicciones de las esperanzas de vida y longevidad de cada sexo.

Otra modificación del método es debida a Wilmoth (1993) que a pesar de que la ejecución del método en dos etapas funciona bien, desarrolla un método de una sola etapa. Las dos fases se unen en una sola mediante la utilización del SVD ponderado para ajustar el modelo, utilizando el número de muertes como ponderación. Minimiza pues la ecuación

$$\sum_{x,t} d_{xt} (\ln(m_{xt}) - a_x - b_x k_t)^2, \quad (4.6)$$

para lo que es necesario calcular las primeras derivadas con respecto a los parámetros a_x , b_x y k_t e igualarlas a cero. Las ecuaciones resultantes pueden ser resueltas numéricamente con un procedimiento iterativo que tome como punto inicial, por ejemplo, la solución de mínimos cuadrados ordinarios o sin ponderar. También puede utilizarse métodos quasi-Newton y simplex para minimizar directamente (4.6). El método permite también la obtención de errores estándar para a_x y b_x , que en el método de dos etapas se llevaba a cabo mediante técnicas bootstrap.

Wilmoth (1993) ha desarrollado también un método de máxima verosimilitud. Considerando en la notación habitual el número de muertes D_{xt} una variable aleatoria con distribución de Poisson, $D_{xt} \sim Po(m_{xt} E_{xt})$, con E_{xt} , los expuestos al riesgo en edad x y tiempo t y sustituyendo el modelo de Lee-Carter en el logaritmo de la función de verosi-

militud, tenemos

$$\begin{aligned} & \sum_{x,t} (d_{xt}(\ln(m_{xt}) + \ln(E_{xt})) - m_{xt}E_{xt}) = \\ & = \sum_{x,t} (d_{xt}((a_x + b_x k_t) + \ln(E_{xt})) - \exp(a_x + b_x k_t)E_{xt}) \end{aligned}$$

que hay que maximizar respecto a los parámetros para obtener la solución, para lo cual utiliza de nuevo métodos quasi-Newton y simplex. La etapa siguiente es ajustar series temporales a k_t para poder predecir del periodo deseado.

En el congreso IME (2002), Renshaw y Haberman (2002) presentaron un método que, basado en Lee-Carter, considera los dos primeros vectores singulares izquierdos y derechos. El método, que denominan LC2, es como sigue

$$\ln(m_{xt}) = a_x + \sum_{j=1}^2 b_{xj} k_{tj} + \epsilon_{xt}$$

debiéndose cumplir las condiciones

$$b_{xj} k_{tj} = s_j u_j v_j$$

sujeto a,

$$\begin{aligned} \sum_t k_{tj} &= 0, \\ \sum_t b_{tj} &= 1, \quad j = 1, 2 \end{aligned}$$

y cuyos grados de libertad son $\nu = (N - 2)(T - 2) - N$. Para asegurar que el número total de muertes observado coincide con el esperado para cada t se reproduce la estrategia del método Lee-Carter inicial. Finalmente los autores modelan $\{\hat{k}_{tj} : j = 1, 2\}$ como procesos ARIMA separados.

El propio Lee (2000), haciendo una referencia a un trabajo anterior de Carter y Lee (1992), presenta una modificación del método original que permite analizar las subpoblaciones masculina y femenina o cualesquiera dos poblaciones. Propone,

1. Modelizar separadamente cada una de las subpoblaciones. La relación entre ambas puede analizarse a partir de las dos series de k_t .
2. Aproximar un sólo k_t , el cual dirige los cambios en todos los tantos centrales de mortalidad de cada edad para ambas subpoblaciones. Este procedimiento conduce a una sola k_t y dos series de a_x y b_x , una para hombres y otra para mujeres. La reestimación de k_t conduce a un parámetro, variando en el tiempo, consistente con el total de muertes.

En este trabajo el autor describe el método básico de Lee-Carter y discute sus predicciones, revisa sus extensiones, aplicaciones, y las mejoras metodológicas que han sido hechas en los recientes años. También describe brevemente como ha sido utilizada como componente de proyecciones de una población estocástica más general de la finanzas del sistema de la Seguridad Social de los Estados Unidos.

Ventajas y desventajas del método de Lee-Carter

Nos ocuparemos ahora de revisar las ventajas e inconvenientes que los autores han encontrado a su método comparándolo con otras alternativas.

En primer lugar resaltan que este procedimiento puede ser utilizado aún cuando haya pocos datos disponibles, como por ejemplo en el tercer mundo. Es posible generar una familia entera de tablas de mortalidad a partir de dos tablas de mortalidad observadas en los periodos que denotamos 0 y 1, para ello

1. Tomamos $k_0 = 0$ para una de las tablas y $k_1 = 1$ para la otra.
2. Sean los $a_x = \ln(m_{x,0})$ y los $b_x = \ln(m_{x,0}) - \ln(m_{x,1})$.
3. Para obtener predicciones daríamos valores a k_t . Para k_t entre 0 y 1, el modelo interpolaría geoméricamente entre las dos tablas; para k_t menor de cero y k_t mayor de 1, el modelo extrapola a partir de las dos tablas.

Aunque en la mayoría de casos este procedimiento tan sencillo es ineficiente y es mucho mejor estimar a_x y b_x con k_t minimizando las desviaciones al cuadrado a partir de una matriz de ratios específicos para cada edad.

Seguidamente enumeramos las ventajas del método frente a la estimación independiente de cada tanto central de mortalidad,

1. Si cada tanto central de mortalidad fuera predicho independientemente, entonces necesitaríamos calcular $N(N - 1)/2$ covarianzas de error diferentes, donde N es el número de grupos de edad. Estas covarianzas son necesarias para encontrar cotas de confianza para la esperanza de vida. Sin embargo, el modelo de Lee-Carter explota el alto grado de correlación entre años a lo largo de las edades de forma que todas las edades dependen de un mismo parámetro que varía con el año. Sus varianzas y covarianzas se obtienen del modelo de media móvil autoregresiva (ARIMA) ajustado para k_t .
2. Si cada tanto central de mortalidad se modela individualmente por un proceso ARIMA de diferente orden, se requieren de muchos más parámetros.
3. Con predicciones individuales, los tantos centrales de mortalidad en el futuro lejano pueden combinarse para formar perfiles de edad muy poco plausibles. En cambio, utilizando el método de Lee-Carter los tantos centrales de mortalidad individuales están siempre sujetos a pertenecer a un sistema de tablas de mortalidad que se ajusta a los datos históricos.

4. Esta cohesión se obtiene al predecir un sólo parámetro k_t , el cual por si mismo es una especie de compromiso entre las tendencias de todos los tantos individuales correspondientes a cada edad.

Comparado con otros métodos como el propuesto por McNown y Rogers (1989, 1992) descrito en la sección 2.1, el método de Lee-Carter mejora algunos de sus aspectos desfavorables que son:

1. Mientras que el número parámetros necesarios para expresar una sola distribución de la edad es relativamente pequeño, el número que es necesario para predecir es más bien grande, 3 en McNown y Rogers (1992) y 8 en McNown y Rogers (1989), al contrario que en Lee-Carter.
2. El elevado número de parámetros dificulta la obtención de intervalos de confianza para predecir, por esta razón en ningún artículo se obtienen dichos intervalos. Intervalos que sí se obtienen en Lee-Carter.
3. Hay bastante evidencia de que las series temporales de los parámetros se comportan de forma irregular, complicando la predicción. Además, en muchos de los artículos los parámetros que se utilizan en las series temporales son no significativos en la ley de mortalidad estimada.
4. Por esta razón en McNown y Rogers (1989) las predicciones son sólo a 15 años vista. Desconocemos como se comportarán a largo término. En Lee-Carter se hacen predicciones desde 1990 a 2065.

En su revisión del método, Lee (2000) resalta el gran logro que supone que un sólo parámetro recoja la variación del tiempo y proporcione el esquema general de la familia de tablas de mortalidad. A pesar de esto también Lee (2000) discute algunos de los problemas y limitaciones del método.

1. La expresión de una tabla requiere un número de coeficientes (a_x, b_x) igual a dos veces el número de grupos de edad considerados.
2. El método se utiliza para extrapolar tendencias históricas en cada nivel y edad de la mortalidad, sin embargo esta extrapolación puede no estar bien reflejada por el procedimiento utilizado. Así pues, la mortalidad no tiene siempre que decrecer a lo largo de la gráfica que produce la representación de los valores de k_t . Si hacemos una retrospectiva podríamos llegar a niveles imposibles de mortalidad, por demasiado altos, a principios del siglo XIX. El comportamiento de las series temporales de k_t durante el periodo de observación desde 1900 a 1996 no es típico del total de la experiencia histórica y no puede posiblemente reflejar la propiedad fundamental del cambio de mortalidad sobre el tiempo. Sin embargo, la investigación médica proporciona algunas razones para creer que la tendencia del siglo XX continuará en el futuro.

3. El método asume que el cambio en la distribución de la mortalidad según la edad, dado por los ratios de decrecimiento a diferentes edades, $b_x(dk_t/dt)$, son constantes a lo largo del tiempo. Pero en la práctica, la velocidad relativa de decrecimiento a diferentes edades puede variar. El problema estaría en la predicción cuando estos cambios comiencen.
4. El método no está preparado para acomodarse a información inesperada sobre futuras tendencias, basa sus predicciones en que las tendencias presentes continuarán en un futuro.
5. El método incorpora incertidumbre a la predicción a través del modelo de k_t , sin embargo la incertidumbre puede provocar errores en la estimación de los coeficientes b_x . Estos errores pueden violar la hipótesis de que los errores estimados son incorrelados a través de la edad.
6. Los intervalos de confianza no reflejan incertidumbre sobre si la especificación del modelo es correcta, ni tampoco acerca de si el futuro se parece al pasado. Mientras unos piensan que mayores avances en la tecnología médica acelerarán el descenso de la mortalidad, otros esperan enfermedades extrañas que proliferarán y causarán un retardo del descenso.
7. El método puede dar lugar a diferencias no plausibles entre los sexos.

Finalizado el análisis de los métodos estructurales veamos el siguiente enfoque para trabajar con la dinamicidad de la mortalidad.

4.3. Métodos no estructurales

Una representación de la evolución de la mortalidad, diferente de la anterior, es la que plantean los métodos no estructurales, que suponen que la medida de mortalidad elegida, q_{xt} , μ_{xt} o m_{xt} , es función de la edad x y del tiempo t , $f(x, t)$, de forma que la función f o ley de mortalidad a lo largo de los años depende explícitamente del tiempo. Pasemos a describir cada una de las propuestas.

4.3.1. El método logarítmico

Para su introducción haremos uso del texto de Benjamin y Soliman (1995) donde tenemos la descripción, obtención de la estimación de sus parámetros para unos datos y evaluación de sus proyecciones. Nos ayudaremos también del trabajo de Felipe y Guillén (1999) cuyo capítulo 6 hace una recopilación de varios métodos con una breve descripción de los mismos mostrando ventajas e inconvenientes de cada uno.

El método logarítmico es un modelo paramétrico que representa la mortalidad q_{xt} en función del tiempo t y de la edad x mediante la expresión,

$$q_{xt} = \alpha_x + (\beta_x)^t,$$

de forma que al tomar logaritmos se obtiene

$$\ln(q_{xt}) = \ln(\alpha_x) + \ln(\beta_x)t.$$

En la práctica, el periodo de años para el que el decremento de $\ln(q_{xt})$ parece ser relativamente constante no es indefinido y debe ser determinado, por ejemplo, gráficamente. Los parámetros que dependen de la edad se han de estimar por algún criterio de optimización, como mínimos cuadrados. Posteriormente se pueden determinar las proyecciones de q_{xt} en el futuro utilizando la ecuación con los parámetros estimados y sustituyendo t por un valor posterior al periodo de estimación. En algunos casos las probabilidades brutas son previamente suavizadas, normalmente mediante aplicación de splines. Este el método utilizado para las tablas alemanas DAV1994R (Shütz, 1997, y Pannenberg y Shütz, 1997), según Felipe y Guillén (1999), donde además pueden encontrarse los comentarios detallados sobre el método y sus inconvenientes.

En el libro de Benjamin y Soliman (1995) hay una aplicación a los datos para grupos de edad de 5 años y periodos de tiempo del calendario de 5 en 5 años, desde 1901 hasta 1960.

Un método de planteamiento diferente al anterior pero de igual naturaleza (Felipe y Guillén, 1999) es el propuesto por Nolfi (1960). Este método se ha utilizado en el cálculo de las tablas de mortalidad suizas GRM/GRF 1995 y en el cálculo de las tablas españolas PERM/F 2000. El método considera, para cada edad x , que la mortalidad decrece de forma exponencial, de forma que para cierto tiempo $t > t_0$,

$$q_{xt} = q_{xt_0} \exp(-\lambda_x(t - t_0)),$$

donde q_{xt_0} es la mortalidad base, $e^{\lambda_x(t-t_0)}$, el factor de proyección, y λ_x tasa media anual de mejora. Para comprobar que es un método logarítmico basta con llevar a cabo algunas sencillas transformaciones,

$$\begin{aligned} \frac{q_{xt}}{q_{xt_0}} &= \exp(-\lambda_x(t - t_0)), \\ \ln\left(\frac{q_{xt}}{q_{xt_0}}\right) &= -\lambda_x(t - t_0), \\ \ln(q_{xt}) &= \ln(q_{xt_0}) - \lambda_x(t - t_0) \\ &= (\ln(q_{xt_0}) + \lambda_x t_0) - \lambda_x t \\ &= \ln(\alpha_x) + \ln(\beta_x)t. \end{aligned}$$

Felipe y Guillén describen el parámetro λ_x , distinto para cada edad, como la intensidad del decrecimiento relativo de la mortalidad, cuya expresión es

$$\lambda_x = \frac{\ln 2}{T_x},$$

donde T_x es el tiempo necesario para que la tasa de mortalidad a la edad x disminuya a la mitad del valor considerado en el tiempo inicial, t_0 .

4.3.2. Método del CMI BUREAU y modelos lineales generalizados

Esta sección está dedicada a modelos paramétricos para la proyección de tantos de mortalidad cuyo esquema básico son los modelos lineales generalizados a partir de las funciones Gompertz-Makeham, ya descritos en un capítulo anterior, y que se amplían ahora mediante la inclusión del tiempo cronológico como variable. Su descripción completa y aplicación podemos encontrarla en el artículo de Sithole, Haberman y Verrall (2000) donde se hace referencia al origen de esta metodología, el trabajo de Renshaw, Haberman y Hatzopoulos (1996).

Describiremos primero la práctica habitual del CMI Bureau, expondremos luego los GLM, siguiendo a Renshaw, Haberman y Hatzopoulos (1996), y a continuación las conclusiones que se extraen de la comparación de ambos.

Método CMI BUREAU

En UK las continuas investigaciones sobre mortalidad del CMI efectúan proyecciones de las futuras mejoras en la mortalidad de los pensionistas y anualidades. El procedimiento se lleva a cabo en dos pasos,

1. En primer lugar, para un periodo de investigación dado, los datos son graduados y producidas las tablas de mortalidad.
2. En segundo lugar, se construyen las tablas de mortalidad proyectadas por aplicación de factores de reducción derivados de la consideración de mejoras pasadas y probables mejoras futuras en los tantos de mortalidad. Los factores de reducción se estudian más detenidamente en la última sección del presente capítulo.

Para obtener las tablas base la práctica habitual del CMI es graduar la fuerza de mortalidad a edad x , μ_x , ajustando la fórmula clásica de Gompertz-Makeham (Forfar, McCutcheon y Wilkie, 1988),

$$\mu_x = GM(r, s) = \sum_{i=1}^{r-1} \alpha_i x^i + \exp\left(\sum_{j=1}^{s-1} \beta_j x^j\right).$$

Posteriormente estas tablas son proyectadas aplicando los factores de mejora de la mortalidad $RF(x, t)$, para un individuo que ha alcanzado la edad x en el tiempo t , donde t está medido en años a partir de un origen apropiado. Del concepto de factor de mejora nos ocuparemos con detalle en un párrafo posterior.

La fuerza de mortalidad proyectada al tiempo t viene dada por

$$\mu_{xt} = \mu_{x0} RF(x, t),$$

o, equivalentemente, la probabilidad de mortalidad proyectada al tiempo t viene dada por

$$q_{xt} = q_{x0} RF(x, t),$$

donde q_{x0} es el tanto de mortalidad de la tabla base. La sección 4 del Informe 10 del CMI Bureau (1990) contiene una descripción completa de este método de tantos de mortalidad

proyectados. Posteriormente el CMI Bureau ha propuesto un nuevo modelo para la proyección de la mortalidad para pensionistas y anualidades (Informe 17, 1999) con tablas de mortalidad basadas en experiencias del periodo 1991-1994. La forma del modelo supone que, a cada edad, el tanto de mortalidad límite es no nulo y que el ratio de mortalidad decrece exponencialmente hasta su valor límite. Se supone además que un porcentaje dado del decrecimiento total futuro en mortalidad ocurrirá en los primeros 20 años, con dicho porcentaje variando con la edad.

El modelo de mejora de la mortalidad adoptado por el CMI es

$$RF(x, t) = \alpha(x) + (1 - \alpha(x))(1 - f(x))^{t/20},$$

donde

$$\alpha(x) = \begin{cases} c, & x < 60, \\ 1 + \frac{(1-c)(x-110)}{50}, & 60 \leq x \leq 110, \\ 1, & x > 110, \end{cases}$$

$$f(x) = \begin{cases} p, & x < 60, \\ \frac{(110-x)p + (x-60)q}{50}, & 60 \leq x \leq 110, \\ q, & x > 110, \end{cases}$$

con $c = 0,13$, $p = 0,55$ y $q = 0,29$.

El nuevo modelo es tal que asume que el factor de mejora de la mortalidad se supone dependiente de ambos, edad y tiempo del calendario, para individuos con edad entre 60 y 110 años solamente. A edades por debajo de 60 años, el tanto de mejora es que depende solo del año, mientras que no hay mejora para vidas de 110 años o superior. Los mismos factores son aplicados para todas las experiencias, hombres y mujeres, y para datos basados sobre individuos o importes.

Modelos lineales generalizados con respecto a la edad y el tiempo

El modelo que Renshaw, Haberman y Hatzopoulos (1996) proponen es el siguiente,

$$\mu_{xt} = \exp \left[\beta_0 + \sum_{j=1}^s \beta_j L_j(x') + \sum_{i=1}^r \alpha_i t'^i + \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij} L_j(x') t'^i \right] \quad (4.7)$$

sujeto a la convención que algunos de los términos γ_{ij} pueden ser cero. En (4.7) x' y t' son transformaciones de las edad y el año de calendario, respectivamente, de forma que los valores están dentro de intervalo $[-1, 1]$. $L_j(x')$ son los polinomios de Legendre generados por

$$L_0(x) = 1, \quad L_1(x) = x, \quad L_{n+1}(x) = xL_n(x) - nL_{n-1}(x),$$

donde n es un entero mayor o igual a que 1.

Reescribiendo la ecuación(4.7) de la forma

$$\mu_{xt} = \exp \left[\beta_0 + \sum_{j=1}^s \beta_j L_j(x') \right] \exp \left[\sum_{i=1}^r \left(\alpha_i + \sum_{j=1}^s \gamma_{ij} L_j(x') \right) t'^i \right],$$

el primer término puede interpretarse como una función Gompertz-Makeham $GM(0, s+1)$ correspondiente a la graduación. El segundo término puede ser interpretado como término de ajuste del efecto del año del calendario, de forma que cuando al menos uno de los γ_{ij} es no nulo depende también de la edad. Un ejemplo de esto sería el incremento de muertes de hombres adultos y jóvenes debido al SIDA.

Para estimar los parámetros desconocidos α_i , β_j y γ_{ij} , se supone que las muertes observadas, d_{xt} , son valores correspondientes a variables aleatorias D_{xt} , Poisson independientes de media y varianza dadas por

$$E(D_{xt}) = \lambda_{xt} = E_{xt}\mu_{xt}, \quad Var(D_{xt}) = \phi\lambda_{xt}.$$

Siguiendo la notación ya introducida en el Capítulo 1, E_{xt} , son los expuestos centrales al riesgo, y ϕ un parámetro de escala que hay que tener en cuenta cuando los datos están basados en pólizas más que en individuos (pues a un mismo individuo le pueden corresponder varias pólizas), de forma que puedan aparecer duplicidades. Si esto ocurriera trabajaríamos con una Poisson con sobredispersión.

Dichas variables Poisson se suponen variables respuesta de un modelo lineal generalizado cuyo link es

$$\log \lambda_{xt} = \log E_{xt} + \log \mu_{xt} = \eta_{xt},$$

por tanto

$$\log \lambda_{xt} = \log E_{xt} + \beta_0 + \sum_{j=1}^s \beta_j L_j(x') + \sum_{i=1}^r \alpha_i t'^i + \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij} L_j(x') t'^i.$$

La estimación de los parámetros se realiza entonces usando la cuasi-log-verosimilitud aproximada, que en este caso tiene la expresión

$$\frac{1}{\phi} \sum_{xt} (-\lambda_{xt} + d_{xt} + \log \lambda_{xt}),$$

y es el núcleo de la función correspondiente a la log-verosimilitud de variables Poisson independientes D_{xt} con media $E(D_{xt}) = \lambda_{xt}$.

La bondad del ajuste se evalúa mediante la *Unscaled Deviance* correspondiente a los valores ajustados $\hat{\mu}_{xt}$. Su expresión es

$$D(d_{xt}; \hat{\lambda}_{xt}) = 2 \sum_{xt} \left[d_{xt} \log \left(\frac{d_{xt}}{\hat{\lambda}_{xt}} \right) - (d_{xt} - \hat{\lambda}_{xt}) \right]$$

donde $\hat{\lambda}_{xt} = E_{xt}\hat{\mu}_{xt}$, es el número de muertes que predice el modelo.

La *Scaled Deviance* se define como la *Unscaled Deviance* pero dividida por el parámetro de escala ϕ ,

$$S(d_{xt}; \hat{\lambda}_{xt}) = \frac{D(d_{xt}; \hat{\lambda}_{xt})}{\phi}$$

con ϕ estimado por el cociente de la deviance del modelo óptimo y el número de grados de libertad.

La forma de proceder para determinar los parámetros es la siguiente:

1. Los valores óptimos de r y s se obtienen calculando la mejora de la *Deviance* para sucesivos incrementos de r y s y comparándola con el valor de una variable χ^2 con un grado de libertad. Los valores elegidos de r y s son aquellos máximos a partir de los cuales los incrementos de la deviance no resultan estadísticamente significativos.
2. El paso siguiente es completar el modelo mediante la elección de los coeficientes γ_{ij} que determinan si el factor de actualización temporal depende de la edad. Para ello, partiendo de los valores predeterminados de r y s , se analiza si el incremento de la *Deviance* al introducir cada parámetro resulta significativo.
3. Paralelamente a los puntos anteriores se determinan los errores estándar de la estimaciones de los parámetros y su significación mediante la prueba usual t-Student.
4. La validación final del modelo se realiza analizando dos tipos de residuos, los Deviance y los de Pearson. Se aplicaran también los test de Forfar, McCutcheon y Wilkie (1988) a los diferentes años.

Sithole, Haberman y Verrall (2000) aplican este modelo, pero el objetivo principal de su trabajo no sólo es encontrar un modelo que proporcione un buen ajuste de los datos, sino que también proporcione un buen esquema para realizar proyecciones. Para ello proceden de la siguiente forma,

1. Determinan un modelo que proporcione el mejor ajuste a los datos.
2. Consideran las proyecciones basadas en dicho modelo para un periodo de 20 años.
3. Finalmente, utilizando la información obtenida por ajuste del modelo y las fuerzas de mortalidad proyectadas, revisan el modelo.

Consideran adecuado el modelo para realizar proyecciones si todas las fuerzas de mortalidad predichas progresan suavemente con respecto a la edad y el tiempo, y predice una reducción en los tantos de mejora en la mortalidad a edades muy avanzadas. En definitiva, se trata de alcanzar un equilibrio entre el buen ajuste y el esquema predictivo.

Por último, utilizando la expresión

$$RF(x, t) = \frac{\hat{\mu}_{xt}}{\mu_{x0}},$$

en la que $\hat{\mu}_{xt}$ son los valores estimados con el modelo, obtienen los distintos factores de mejora, $RF_{GLM}(x, t)$, que comparan con los $RF_{CMI}(x, t)$ obtenidos mediante el método del CMI Bureau, para analizar la consistencia de unos con otros.

4.4. Métodos no paramétricos

La metodología aquí descrita corresponde al artículo de Felipe, Guillen y Nielsen (2001). Son los propios autores los que la clasifican como una técnica de análisis exploratorio de los datos, especialmente diseñada para diagnosticar el comportamiento de la mortalidad según dos dimensiones (edad, tiempo) y previa a la graduación mediante cualquier ley de mortalidad. En el modelo se combinan las ideas básicas de procesos estocásticos donde el tiempo cronológico es la escala de tiempo, mientras que la edad se trata como una covariable del proceso. Para un tratamiento extenso de la importancia que los actuarios están dando a la aplicación de los procesos estocásticos para modelos de supervivencia en análisis estadístico de la mortalidad puede consultarse los trabajos de (MacDonald, 1996a, 1996b y 1996c).

El método utilizado es no paramétrico, concretamente utilizan el smoother kernel ya descrito para el caso univariante en la sección 3.2.2. Es éste un aspecto interesante puesto que los resultados no dependen de hipótesis sobre la distribución de los datos, a diferencia de otros métodos que basan sus conclusiones en distribuciones como la binomial o la poisson.

En el artículo mencionado se analiza el desarrollo de la mortalidad en dos poblaciones (española y danesa) y se señalan las diferencias existentes, a continuación se determina dónde estas diferencias son más prominentes y, finalmente, se identifica la forma en que esas diferencias se han desarrollado a lo largo del tiempo. Todo ello con el mínimo número de hipótesis posibles, para lo cuál hay que desarrollar el estimador bidimensional de la fuerza de mortalidad y después estimadores constantes locales para cada componente, edad y tiempo cronológico.

4.4.1. El modelo

En esta sección vamos a describir el modelo de mortalidad dentro del esquema de procesos estocásticos como una función bidimensional. Sea t el tiempo cronológico y sea $X(t)$ el proceso de edad. Observamos n individuos, $i = 1, \dots, n$. Ahora, sea N_i el número de fallos observados para el i -ésimo individuo en el intervalo de tiempo $[t_1, t_2]$. Supongamos que $N = (N_1, \dots, N_n)$ es un proceso estocástico que obedece a las condiciones habituales, ver Andersen et al. (1994). Modelamos la intensidad aleatoria del proceso $\lambda = (\lambda_1, \dots, \lambda_n)$ de N dependiente de los valores,

$$\lambda_i(t) = \alpha(X_i(t), t)Y_i(t),$$

sin restringir la forma funcional de $\alpha(\cdot)$. Además, en la expresión anterior, $X_i(t)$ es el proceso predecible unidimensional de edad, mientras que Y_i es un proceso predecible que toma el valor 1 cuando el i -ésimo individuo está bajo riesgo y 0 en caso contrario.

4.4.2. Estimación constante local de α

La metodología de estimación constante local consiste, dados el kernel unidimensional K , y $b = (b_1, b_2)$, un vector de bandwidth bidimensional, en definir el estimador sugerido

por Fusaro, Nielsen y Scheike (1993) y Nielsen y Linton (1995) según Felipe, Guillen y Nielsen (2001),

$$\hat{\alpha}(x, t) = \frac{D_{xt}}{E_{xt}},$$

donde

$$D_{xt} = \sum_{i=1}^n \int_{t_1}^{t_2} K_{b_1}(x - X_i(s)) K_{b_2}(t - s) dN_i(s)$$

y

$$E_{xt} = \sum_{i=1}^n \int_{t_1}^{t_2} K_{b_1}(x - X_i(s)) K_{b_2}(t - s) Y_i(s) ds$$

son, respectivamente, las ocurrencias y expuestos suavizados, por lo que el estimador kernel constante local corresponde, en este caso, al bien conocido ratio muertos-expuestos para el intervalo de tiempo $[t_1, t_2]$ y las edades de $[n_1, n_2]$.

4.4.3. El submodelo multiplicativo y el método de estimación de integración marginal

Para poder separar el efecto de las dos componentes, edad y tiempo cronológico, restringiremos nuestra atención a los submodelos multiplicativos bidimensionales,

$$\alpha(x, t) = \alpha_1(x)\alpha_2(t).$$

El modelo puede ser interpretado como un ratio de mortalidad continuo y cuya única hipótesis funcional es la interacción multiplicativa entre las dos componentes. Pasemos a la identificación de $\tilde{\alpha}_1(x)$ y $\tilde{\alpha}_2(t)$ que son las componentes univariantes de interés en la estructura multiplicativa y están definidas por

$$\begin{aligned} \tilde{\alpha}_1(x) &= \int_{t_1}^{t_2} \alpha(x, t) dt \\ &= \int_{t_1}^{t_2} \alpha_1(x)\alpha_2(t) dt \\ &= \alpha_1(x) \int_{t_1}^{t_2} \alpha_2(t) dt \\ &= \alpha_1(x)c_1, \end{aligned} \tag{4.8}$$

$$\begin{aligned} \tilde{\alpha}_2(t) &= \int_{n_1}^{n_2} \alpha(x, t) dx \\ &= \int_{n_1}^{n_2} \alpha_1(x)\alpha_2(t) dx \\ &= \alpha_2(t) \int_{n_1}^{n_2} \alpha_1(x) dx \\ &= \alpha_2(t)c_2. \end{aligned} \tag{4.9}$$

Obsérvese además que, en este caso, $c = \int_{t_1}^{t_2} \int_{n_1}^{n_2} \alpha(x, t) dx dt = c_1 c_2$. Los valores de (4.8) y (4.9) son estimados por

$$\hat{\alpha}_1(x) = \int_{t_1}^{t_2} \hat{\alpha}(x, t) dt,$$
$$\hat{\alpha}_2(x) = \int_{n_1}^{n_2} \hat{\alpha}(x, t) dx.$$

mientras que la constante es estimada mediante

$$\hat{c} = \int_{t_1}^{t_2} \int_{n_1}^{n_2} \hat{\alpha}(x, t) dx dt.$$

En el estudio de los datos del artículo de Felipe, Guillen y Nielsen (2001) el proceso de edad $X_i(t)$ toma valores en $[0, 90]$ y el tiempo cronológico pertenece al intervalo $I_2 = [1975, 1993]$.

Algunas ventajas interesantes de este método son, que puede aplicarse cuando hay pocos datos y que además permite una rápida primera impresión del desarrollo de la mortalidad, todo ello sin hipótesis de partida de distribución de los datos.

Por otra parte se ha tenido que renunciar a ventajas de los métodos paramétricos como son la proyección, la revisión de los parámetros y su interpretación. Además la hipótesis multiplicativa implica que en los años siguientes se modelan los ratios de edad a todas las edades en igual proporción, una hipótesis demasiado fuerte en este contexto. Por ello se estudia incorporar interacciones entre la edad y el tiempo. Otro problema detectado es la mala estimación en los extremos, este hecho es propio de la estimación kernel y algunas soluciones se apuntan en el Capítulo 3.

4.5. Métodos basados en factores de reducción de la mortalidad

En base a lo ocurrido en las experiencias pasadas de mejora de la mortalidad, el CMI considera imprudente no incorporar en las tablas con las que trabaja habitualmente, una medida que nos permita proyectarlas de forma que se tengan en cuenta las mejoras en la mortalidad a lo largo del tiempo. Estas medidas son los *factores de mejora de la mortalidad* $RF(x, t)$. Hasta donde sea posible, los factores de mejora de la mortalidad deben reflejar tendencias recientes, pero el modelo adoptado debe ser relativamente simple de forma que se eviten anomalías obvias (por ejemplo, se debe plantear si estaría justificado un modelo de mejora de la mortalidad que conduzca en tiempo futuro a probabilidades de mortalidad más bajas para mujeres que para hombres). En esta sección describiremos los diferentes modelos propuestos para obtener valores para las mejoras futuras en mortalidad. Los modelos adoptados por el CMI de mejora de la mortalidad se hallan descritos en la sección 4.3 del Informe 10 del CMI Bureau (1990). Posteriormente en el Informe 17 del CMI Bureau (1999) se considera un modelo esencialmente igual, pues para cada edad el ratio de mortalidad se supone que disminuye exponencialmente hasta un valor límite, pero

en el que la velocidad de convergencia al límite de los nuevos factores depende de la edad. El procedimiento que se sigue es comparar ratios de mortalidad, con un mismo rango de edad, a lo largo de un periodo del calendario con el objetivo de analizar y construir los factores de reducción de la mortalidad con los que identificar una tendencia que permita predecir su evolución futura. En primer lugar veremos de forma esencial como caracterizar los factores de mejora de la mortalidad y, después, analizaremos los modelos descritos por Renshaw y Haberman (2000) y Renshaw y Haberman (2001).

4.5.1. Caracterización del factor de reducción para q_x

Los factores de reducción de la mortalidad en la práctica habitual colectiva de UK se aplican a los ratios de mortalidad estándar, validados en el centro de un periodo base, para ajustarlos a futuros años. Sea t el tiempo, con origen en $t = 0$ situado en el centro del periodo base, y q_{xt} la probabilidad de muerte entre la edad x y $x + 1$ en el tiempo $t \geq 0$. El factor de reducción está caracterizado por la ecuación

$$q_{xt} = q_{x0}RF(x, t), \quad t \geq 0, \quad (4.10)$$

sujeito a las restricciones

$$\begin{aligned} RF(x, 0) &= 1 & \forall x \geq 0 \\ 0 < RF(x, t) &\leq 1 & \forall x \geq 0 \quad \forall t \geq 0. \end{aligned} \quad (4.11)$$

Los q_{x0} se determinan por graduación de los tantos de mortalidad del periodo base de forma objetiva. Posteriormente se establece una opinion acerca de los $RF(x, t)$ dado que su naturaleza es eminentemente subjetiva y, finalmente, pueden obtenerse los valores q_{xt} a partir de la ecuación (4.10). En USA, el Group Annuity Valuation Table (GAVT) Task Force (1995) plantea una aproximación similar, definiendo el factor de mejora mediante la expresión

$$RF(x, t) = (1 - AA_x)^t,$$

donde AA_x se denomina, para cada edad específica, *factor de mejora anual*.

4.5.2. Caracterización del factor de reducción para μ_x

Análogamente esta metodología puede extenderse a factores de reducción de la fuerza de mortalidad en lugar de probabilidades de muerte. Si μ_{xt} es la fuerza de mortalidad a edad x , en el tiempo $t \geq 0$, entonces podemos definir el factor de reducción a partir de la ecuación

$$\mu_{xt} = \mu_{x0}RF(x, t) \quad t \geq 0, \quad (4.12)$$

sujeito a las restricciones

$$\begin{aligned} RF(x, 0) &= 1 & \forall x \geq 0 \\ 0 < RF(x, t) &\leq 1 & \forall x \geq 0 \quad \forall t \geq 0. \end{aligned} \quad (4.13)$$

Los μ_{x0} se determinan por graduación Forfar, McCutcheon y Wilkie (1988) y Renshaw (1991) de las fuerzas de mortalidad del periodo base de forma objetiva. Posteriormente se establece una opinion acerca de los $RF(x, t)$ dada su naturaleza eminentemente subjetiva y, finalmente, se obtienen los valores μ_{xt} a partir de la ecuación (4.12).

4.5.3. Modelización de los factores de reducción

Los modelos utilizados para representar el comportamiento de los factores de reducción que nos permitan estimar y predecir para futuros años, han sido recopilados en varios trabajos (Renshaw y Haberman, 2000, 2001 y 2003). Mientras que el primero se centra en desarrollar una metodología que nos permita trabajar dentro de los GLM para describir el método de modelización de los $RF(x, t)$, en el segundo se adapta la metodología de Lee-Carter para poder predecir el futuro comportamiento de los factores de reducción. La comparación y aplicación de estos métodos para datos de Inglaterra y Gales podemos encontrarla en el último de los trabajos.

Modelos lineales generalizados

Nos ocupamos ahora de describir el esquema utilizado por Renshaw y Haberman (2000) para modelizar y predecir factores de reducción de mortalidad basado en GLM. Sean D_{xt} las variables aleatorias que representan el número real de muertes a edades individuales y E_{xt} los correspondientes expuestos iniciales en los tiempos $t = t_1, t_1 + 1, \dots, t_{max}$.

Comenzaremos considerando el modelo propuesto para q_{xt} . Las variables D_{xt} son independientes con distribución $D_{xt} \sim Bi(E_{xt}, q_{xt})$ y parámetro de sobredispersión ϕ . Podemos definir el GLM, cuyas variables respuesta son

$$Y_{xt} = \frac{D_{xt}}{E_{xt}},$$

con

$$E(Y_{xt}) = q_{xt} \quad y \quad Var(Y_{xt}) = \phi \frac{q_{xt}(1 - q_{xt})}{E_{xt}},$$

con un predictor lineal η_{xt} y una función monótona y diferenciable g , tal que

$$\eta_{xt} = g(q_{xt}) \iff q_{xt} = g^{-1}(\eta_{xt}).$$

Considerando la ecuación (4.10) se obtiene

$$RF(x, t) = \frac{g^{-1}(\eta_{xt})}{q_{x0}}.$$

Descomponiendo el predictor lineal

$$\eta_{xt} = \eta_{x0} + \eta'_{xt} \quad \text{con} \quad \eta'_{x0} = 0,$$

se sigue entonces que

$$RF(x, t) = \frac{g^{-1}(\eta_{x0} + \eta'_{xt})}{q_{x0}}.$$

Como la primera condición de (4.11) implica que $g^{-1}(\eta_{x0}) = q_{x0}$, entonces

$$\eta_{x0} = g(q_{x0}).$$

El sumando $g(q_{x0})$ es un término “offset” del predictor lineal η_{xt} y se considera conocido. Hay que proporcionar una estructura paramétrica para η'_{xt} de forma que $\eta'_{x0} = 0$, de la que obtener

$$RF(x, t) = \frac{g^{-1}(\eta_{x0} + \eta'_{xt})}{g^{-1}(\eta_{x0})} \quad (4.14)$$

que cumpla las condiciones (4.11).

Las estructuras paramétricas consideradas pueden ser:

- Una recta que parte del origen y con pendiente dependiendo de la edad,

$$\eta'_{xt} = \beta_x t.$$

- Otra estructura alternativa es el caso general de splines con $k + 1$ nodos

$$\sum_{i=1}^r \sum_{j=1}^k \beta_{ij} (x - x_j)^i + t, \quad \forall (x, t).$$

Como caso particular los splines cúbicos ($r = 3$ y $\beta_{ij} = 0, \forall i = 1, 2$) son particularmente satisfactorios.

Consideran todos los link habituales ya descritos en anteriores capítulos, log-log, complementario log-log, logit y probit. Con todo ello describimos a continuación como obtener los distintos modelos.

Estimación

Para ajustar la recta de la opción primera el predictor es

$$\eta_{xt} = g(q_{x0}) + \beta_x t,$$

donde $g(q_{x0})$ se calcula aplicando el link a los valores graduados del periodo base. Si llamamos $\hat{\beta}_x$ a las estimaciones de los parámetros, los valores ajustados del predictor lineal son

$$\hat{\eta}_{xt} = g(q_{x0}) + \hat{\beta}_x t$$

con error estándar

$$se(\hat{\eta}_{xt}) = se(\hat{\beta}_x) t.$$

Aplicando a la función la función inversa del link obtenemos

$$\hat{q}_{xt} = g^{-1}(\hat{\eta}_{xt}) = g^{-1}(g(q_{x0}) + \hat{\beta}_x t),$$

que podemos sustituir en la expresión de los factores de reducción para calcular sus valores ajustados,

$$\hat{RF}(x, t) = \frac{\hat{q}_{xt}}{q_{x0}} = \frac{g^{-1}(g(q_{x0}) + \hat{\beta}_x t)}{q_{x0}}, \quad (4.15)$$

o sus logaritmos,

$$\log(\hat{RF}(x, t)) = \log(g^{-1}(g(q_{x0}) + \hat{\beta}_x t)) - \log(q_{x0}).$$

Intervalos de confianza

Por analogía con la sección 4.4.6 pp 122-24 de McCullagh y Nelder (1989), bajo la hipótesis de aproximación Normal,

$$\frac{\hat{\eta}_{xt} - g(q_{xt})}{\text{se}(\hat{\eta})} \sim N(0, 1),$$

podemos obtener el intervalo de confianza para $g(q_{xt})$,

$$\hat{\eta}_{xt} \pm z_{\alpha/2} \text{se}(\hat{\eta}_{xt}) = g(q_{x0}) + \{\hat{\beta}_x \pm z_{\alpha/2} \text{se}(\hat{\beta}_x)\}t,$$

donde $z_{\alpha/2}$ es el percentil $\alpha/2$ de la distribución normal tipificada. Podemos obtener para el factor de reducción de la mortalidad el intervalo

$$g^{-1}(g(q_{x0}) + \{\hat{\beta}_x \pm z_{\alpha/2} \text{se}(\hat{\beta}_x)\}t) / q_{x0}, \quad (4.16)$$

y para su logaritmo

$$\log(g^{-1}(g(q_{x0}) + \{\hat{\beta}_x \pm z_{\alpha/2} \text{se}(\hat{\beta}_x)\}t)) - \log(q_{x0}),$$

Es posible extrapolar a valores no observados de t sin más que sustituirlos en las expresiones (4.15) y (4.16). Los peligros asociados a estas extrapolaciones están discutidas en la referencia citada de McCullagh y Nelder (1989). De forma similar podemos encontrar las estimaciones si utilizamos la fórmula de splines.

Representación de los resultados

Es conveniente representar nuestros resultados en escala logarítmica. Específicamente el objetivo de nuestro análisis son los factores de reducción $RF(x, t)$ que pueden escribirse

$$\frac{q_{xt}}{q_{x0}} = RF(x, t)$$

$$\log(q_{xt}) - \log(q_{x0}) = \log(RF(x, t)) \quad (4.17)$$

Supongamos además que conocemos d_{xt} , número observado de muertes para edades x a tiempos $t = t_1, t_1 + 1, \dots, t_{max}$, y sus correspondientes expuestos iniciales, E_{xt} . Podemos realizar estudios retrospectivos para años posteriores al periodo base. Para tales datos podemos obtener los ratios brutos

$$\dot{q}_{xt} = \frac{d_{xt}}{E_{xt}}$$

que al sustituir en la ecuación (4.17), nos proporcionan los valores

$$z_{xt} = \log(\dot{q}_{xt}) - \log(q_{x0})$$

que podemos representar gráficamente respecto de t , separadamente para cada edad. Además, la ecuación (4.14) implica que

$$\log(RF(x, t)) = \log(g^{-1}(\eta_{x0} + \eta'_{xt})) - \log(g^{-1}(\eta_{x0})),$$

uno de cuyos posibles ajustes es, por ejemplo, (4.15). Podemos comparar gráficamente los valores de z_{xt} observados y ajustados. Es posible, utilizar también estudios prospectivos mediante métodos gráficos que representan el perfil de $\log(RF(x, t))$ para futuros valores de t , determinados por la ecuación (4.15) junto con sus intervalos de confianza determinados por la ecuación (4.16).

Factores de reducción para la fuerza de mortalidad

Es posible adaptar el modelo a los factores de reducción de μ_{xt} , fuerza de mortalidad a edad x y tiempo $t \geq 0$.

A partir de la distribución asociada a las $D_{xt} \sim Po(E_{xt}\mu_{xt})$, variables independientes con parámetro de sobredispersión ϕ , donde E_{xt} son los expuestos centrales al riesgo, podemos definir el GLM, cuyas variables respuesta son

$$Y_{xt} = \frac{D_{xt}}{E_{xt}}$$

con

$$E(Y_{xt}) = \mu_{xt} \quad \text{y} \quad Var(Y_{xt}) = \phi \frac{\mu_{xt}}{E_{xt}},$$

con un predictor lineal η_{xt} y una función monótona y diferenciable g , tal que

$$\eta_{xt} = g(\mu_{xt}) \iff \mu_{xt} = g^{-1}(\eta_{xt}).$$

Considerando la ecuación (4.12) se obtiene

$$RF(x, t) = \frac{g^{-1}(\eta_{xt})}{\mu_{x0}}.$$

Descomponiendo el predictor lineal,

$$\eta_{xt} = \eta_{x0} + \eta'_{xt} \quad \text{con} \quad \eta'_{x0} = 0,$$

se sigue entonces que

$$RF(x, t) = \frac{g^{-1}(\eta_{x0} + \eta'_{xt})}{\mu_{x0}},$$

y teniendo en cuenta la primera condición de (4.13),

$$\mu_{x0} = g^{-1}(\eta_{x0}) \iff \eta_{x0} = g(\mu_{x0}).$$

Podríamos continuar como antes reemplazando los q_{xt} por μ_{xt} y ajustando un modelo para $RF(x, t)$ referido a la fuerza de mortalidad.

Modelo de Lee-Carter modificado

En el párrafo anterior hemos descrito un método para modelizar factores de reducción de la mortalidad dentro del esquema de los GLM, cuyos resultados son, en ciertas circunstancias, similares a los obtenidos con el método de Lee-Carter. En su artículo, Renshaw y Haberman (2001), sugieren formas para las cuales la metodología de Lee-Carter de ajustes y predicción de tendencias de mortalidad puede ser adaptada a la construcción de factores de reducción de mortalidad. En primer lugar formulan dicha metodología en función de la fuerza de mortalidad, pues aunque el método de Lee-Carter se basa en los tantos centrales de mortalidad, los modelos probabilísticos en los que se basan ambos son de idéntica forma. Sin embargo, como la práctica habitual del CMI utiliza los factores de reducción basados en las probabilidades de muerte, nos ocuparemos también de ellos. El modelo de Lee-Carter modificado para predecir factores de reducción de la mortalidad se introduce de la siguiente forma. Partimos de la caracterización de los factores de reducción de las fuerzas de mortalidad dada por la ecuación,

$$\log(\mu_{xt}) = \log(\mu_{x0}) + \log(RF(x, t))$$

sujeto a las restricciones (4.13), y la adaptamos al esquema básico de Lee-Carter, obteniendo

$$\log(\mu_{xt}) = \alpha_x + \beta_x k_t,$$

sujeto a las restricciones $k_0 = 0$, donde

$$\begin{aligned} \alpha_{x0} = \log(\mu_{x0}) &\implies \mu_{x0} = \exp(\alpha_{x0}) \\ \beta_x = \log(RF(x, t)) &\implies RF(x, t) = \exp(\beta_x k_t). \end{aligned} \quad (4.18)$$

Los parámetros estarán unívocamente determinados si además se cumple,

$$k_0 = 0, \quad \sum_x \beta_x = 1, \quad \alpha_{x0} = \log(\mu_{x0}) .$$

Considerando los datos para tiempos $t = t_1, t_1 + 1, \dots, t_m, \dots, t_n$, denotamos por h el rango de años de calendario de los datos, $h = t_n - t_1 + 1$, y g el rango de un subintervalo, $g = t_n - t_m + 1$, que especifica el periodo de datos que han sido graduados y en cuyo centro se va a fijar el origen t_0 , $t_0 = t_n - \left[\frac{g}{2}\right]$, donde $\left[\frac{g}{2}\right]$ denota la parte entera de $\frac{g}{2}$. Obviamente se cumple que $t_m \leq t_0 \leq t_n$, y con igualdad si $g = 1$.

El modelo puede ajustarse siguiendo los siguientes pasos,

1. Estimar α_{x0} de cualquiera de las siguientes formas:
 - i) a partir de la tabla de mortalidad estándar basada en un grupo de g años, se calcula

$$\hat{\alpha}_x = \log(\mu_{x0}),$$

ii) a partir de la tabla de mortalidad estándar basada en un grupo de g años,

$$\hat{\alpha}_{x0} = \log \left(\frac{\sum_{t=t_m}^{t_n} d_{xt}}{\sum_{t=t_m}^{t_n} E_{xt}} \right).$$

iii) A partir de la media geométrica de los datos de mortalidad brutos $r_{xt} = \frac{d_{xt}}{E_{xt}}$ para los años de t_m hasta t_n , ambos inclusive,

$$\hat{\alpha}_{x0} = \log \left(\prod_{t=t_m}^{t_n} (r_{xt})^{1/g} \right).$$

2. Suavizar, si fuera necesario, el $\hat{\alpha}_{x0}$ obtenido en los apartados ii) o iii) del primer paso.
3. Estimar (β_x, k_t) aplicando SVD, o su aproximación, a la matriz formada por los estadísticos $z_{xt} = \log(r_{xt}) - \hat{\alpha}_{x0}$, como se describe en el método de Lee-Carter
4. Ajustar \hat{k}_t de forma que el total de muertes observadas $\sum_x d_{xt}$ sea igual al total de muertes ajustadas $\sum_x \hat{d}_{xt}$, $\forall t$, con $\hat{d}_{xt} = E_{xt} \exp(\hat{\alpha}_{x0} + \hat{\beta}_x \hat{k}_t)$.
5. Transformar k_t en $\hat{k}_t - \hat{k}_{t_0}$.
6. Suavizar $\hat{\beta}_x$ si fuera necesario.
7. Contrastar la bondad del ajuste calculando los residuos

$$\frac{d_{xt} - \hat{d}_{xt}}{\sqrt{\hat{d}_{xt}}}.$$

Varios comentarios hacen los autores respecto al método.

- En el caso $g = 1$, los pasos 1(ii) y 1(iii) son idénticos, y la predicción coincide con los ratios observados más recientes, como se señala en la pag. 88 de Lee (2000).
- El paso 1(iii) se sugiere, como una alternativa viable a 1(ii), dentro del esquema del método de Lee-Carter.
- El paso 1(ii), emparejado con el paso 2, es potencialmente equivalente a el paso 1(i).
- Está previsto que los datos estén clasificados por años individuales de edad (así como por año de calendario), por tanto la introducción de la opción de smoothing del paso 6 es esencial para que las predicciones resultantes están bien ordenadas con respecto a años individuales de edad y no haya irregularidades.

- Como se comenta más tarde, el paso 4 es opcional, y se lleva a cabo sólo cuando el valor obtenido a partir del 3 no es satisfactorio.
- Estos métodos también pueden aplicarse a q_{x0} .

Predecir siguiendo el modelo de series temporales de los k_t , según el paso 5, con el proceso ARIMA(0,1,0) juega un papel central. Para este particular proceso

$$k_t = k_{t-1} + \lambda + \epsilon_t,$$

donde ϵ_t denota el error. Por tanto k_t se predice linealmente por incrementos de λ . Denotando la predicción como $\{\dot{k}_{t_n+s} : s > 0\}$, el factor de reducción se predice mediante

$$\log(RF(x, t_n + s)) = \hat{\beta}_x \dot{k}_{t_n+s}.$$

Para el proceso ARIMA(0,1,0)

$$\dot{k}_{t_n+s} = \hat{k}_{t_n} + s\hat{\lambda}$$

y de aquí,

$$RF(x, t_n + s) = \exp\left(\hat{\beta}_x(\hat{k}_{t_n} + \hat{\lambda}s)\right)$$

Para estimar la predicción se utiliza metodología estándar y la predicción asociada de errores por ejemplo, (Hamilton, 1994).

Posteriormente los autores establecen el paralelismo existente entre la metodología descrita en Renshaw y Haberman (2000) mediante GLM, expuesta en la sección anterior, y la metodología de Lee-Carter puesto que aquella puede ser escrita en términos de esta como sigue

$$RF(x, t) = \frac{g^{-1}(g(\exp \alpha_{x0}) + \beta_x k_t)}{\exp \alpha_{x0}}, \quad k_{t_0} = 0,$$

donde g es la función link especificada con modelos binomial o Poisson sobredispersos descritos en la sección anterior. En particular para los factores de reducción de μ_{xt} seleccionan la estructura

$$RF(x, t) = \exp(\beta_x k_t),$$

que es idéntica a la estructura de Lee-Carter de la que partimos en (4.18).

Para concluir, los autores apuntan varias reflexiones respecto a la necesidad de estos métodos y sus ventajas comparativas:

1. Predecir los ratios de mortalidad (particularmente para edades avanzadas) tiene importantes implicaciones para la financiación de programas públicos de pensiones de jubilación y soporte a los ancianos, y para la financiación de sistemas de seguridad privados cuyo propósito es proporcionar pensiones, anualidades y, a largo plazo, seguros médicos. Es pues importante disponer de métodos y modelos dinámicos que nos permitan la predicción.
2. El método de Lee-Carter de predicción de la mortalidad es flexible. Puede ser modificado, como se ha descrito, para proporcionar una alternativa al método GLM.

3. Ambos métodos, el de Lee-Carter y la aproximación GLM, proporcionan un modelo formal estadístico para los cambios de la mortalidad y por tanto pueden ser determinados intervalos de predicción. El cálculo de los errores de predicción bajo el método de Lee-Carter aparece más complejo que bajo su equivalente GLM.
4. Mientras que en el método original de Lee-Carter los grupos de edad eran de 5 años, en los trabajos de Renshaw y Haberman son de un sólo año, por lo que se hace necesario el smoothing de los valores $\{\alpha_x\}$ y $\{\beta_x\}$ para que las predicciones estén bien ordenadas respecto a la edad.
5. El ajuste de k_t , involucrando el total de muertes observadas y esperadas en cada año del calendario (paso 4 del método), es necesaria para establecer la tendencia de las series temporales $\{k_t\}$ y consistencia entre los diferentes métodos de determinar $\{k_t\}$. Esto no es necesario en la aproximación GLM (Benjamin y Pollard, 1992).

Capítulo 5

Ajustes paramétricos de tablas de mortalidad de la Comunidad Valenciana

5.1. Introducción

El objetivo del presente capítulo es mostrar la aplicación a datos de mortalidad reales de los diferentes métodos de graduación paramétricos. Objetivo que formalizamos proporcionando soporte estadístico a los métodos y al análisis de su bondad de ajuste. Realizaremos también la comparación de las diferentes graduaciones mediante los contrastes no paramétricos y gráficos descritos en la última sección del Capítulo 2.

Los datos que hemos utilizado son las cifras de población y muertes de la Comunidad Valenciana, éstas han sido publicadas por el Instituto Nacional de Estadística (INE) y tanto la población total como las defunciones están clasificadas por edad y sexo. Ambas cifras se refieren a la Comunidad Valenciana como lugar de residencia, lo que proporciona coherencia a la correspondencia entre dichas cantidades.

Respecto a las cifras de población hay que decir que se realiza un censo cada 10 años de forma que son verdaderos recuentos de población los años 1981, 1991 y 2001. El resto de cifras son estimaciones intercensales, obtenidas de sendas publicaciones del INE: “Evolución de la población de España entre los Censos de 1981 y 1991” (INE, 1997), y “Estimaciones intercensales entre los Censos de 1991 y 2001 (cifras provisionales)”. Las edades contempladas tienen un rango que va de 0 a 100 años o más, para hombres y mujeres.

Para cubrir el objetivo del presente capítulo, lo hemos estructurado en 6 secciones. La primera la dedicamos a la obtención de las estimaciones brutas de las probabilidades correspondientes a los tres últimos años. Las tres siguientes secciones se refieren a la aplicación de los diferentes métodos para la construcción de una tabla de momento, comentado sus ventajas e inconvenientes así como su adecuación a la experiencia de mortalidad analizada. Finalmente se comparan los diferentes ajustes mediante los test no paramétricos habituales descritos en la sección 2.7.

5.2. Estimación de las q_x

El primer paso lo constituye el cálculo de las estimaciones brutas de la medida de mortalidad elegida para posteriormente proceder a su graduación. Dicha graduación nos conduce a la expresión de la tabla de mortalidad correspondiente, en la que fundamentalmente nos interesa el valor de las probabilidades q_x que aplicaremos a los cálculos actuariales del último capítulo.

Dado el interés que tienen las q_x procedemos a explicar detalladamente los diferentes métodos para la obtención de sus estimaciones brutas a partir la expresión (2.3), cociente entre las defunciones y la población inicialmente expuesta al riesgo a una determinada edad x . Ambos, numerador y denominador se calculan a partir de la cifras de población y defunciones, clasificadas según edad y sexo, de la Comunidad Valenciana para varios años a fin de favorecer la estabilidad de los cocientes.

Existen varias propuestas,

- La Unión Española de Entidades Aseguradoras y Reaseguradoras (UNESPA) tradicionalmente utiliza las tablas de mortalidad elaboradas por el INE como datos base para ajustar la ley de Makeham, procedimiento que han seguido utilizando Prieto y Fernández (1994). Los individuos expuestos al riesgo a la edad x , E_x , el INE lo determina considerando la suma de los habitantes que a finales de año tiene esa edad y la mitad de los fallecidos con edad x durante el año considerado, ya que se supone que las muertes se distribuyen uniformemente a través del mismo. Por otro lado, los fallecidos en un año con edad x proceden de individuos que contaban a principio de año con edad $x - 1$ o x y, admitiendo la hipótesis de uniformidad, la mitad correspondería a una edad y la otra mitad a la siguiente. Según esto, la estimación máximoverosimil de q_x es

$$\hat{q}_x = \frac{1/2(D_{xt} + D_{x(t+1)})}{P_{xt} + 1/2D_{xt}}, \quad (5.1)$$

donde D_{xt} son los fallecidos en el año t a la edad x , $D_{x(t+1)}$ son los fallecidos en el año $t + 1$ a la edad x , y P_{xt} población que a 31 de diciembre del año t tienen edad x . La fórmula puede ser aplicada para todas las edades salvo para la edad cero y para la edad uno por la concentración de defunciones en los primeros meses de vida. Para la primera edad el INE aplica,

$$\hat{q}_0 = \frac{\alpha_0 D_{xt} + \beta_0 D_{x(t+1)}}{P_x^t + \alpha_0 D_{xt}},$$

donde α_0 corresponde a la proporción de fallecidos en el año t con menos de un año de edad, entre los nacidos ese mismo año, y β_0 la proporción de fallecidos en el año $t + 1$ con menos de un año de edad entre los nacidos en el año t . De forma análoga se calcula

$$\hat{q}_1 = \frac{\alpha_1 D_{xt} + \beta_1 D_{x(t+1)}}{P_x^t + \alpha_1 D_{xt}},$$

donde α_1 corresponde a la proporción de fallecidos en el año t con un año cumplido entre los nacidos en el año $t - 1$, y β_1 la proporción de fallecidos en el año $t + 1$ con un

año cumplido entre los nacidos en el año $t - 1$. Prieto y Fernández plantean también que no es posible utilizar la fórmula para edades superiores a 95 por dos razones: la primera, que los fallecidos y los expuestos al riesgo son reducidos, y no permiten hacer inferencias válidas, y la segunda, el número de errores en las declaraciones de edad que provocan distorsiones en las series de probabilidades.

- Otra posibilidad para la estimación de los cocientes es la que utiliza Rue (1992) a partir de cuatro años consecutivos. Siguiendo el esquema de Lexis, plantea la fórmula siguiente

$$\hat{q}_x = \frac{D_{x(t-2)} + D'_{x(t-1)} + D_{x(t-1)} + D'_{xt} + D_{xt} + D'_{x,t-1}}{E_{x,t-1} + E_{x,t} + E_{x,t+1}},$$

donde, t es un año censal y

$$E_{x(t-1)} = P_{xt} + D_{x(t-1)},$$

$$E_{x(t+1)} = P_{(x-1)t} - D'_{(x-1)t},$$

$$E_{x(t+1)} = P_{(x+1)t} + D_{(x+1)(t-1)} + D'_{x(t-1)} + D_{x(t-2)},$$

excepto para el primer grupo de edad, donde $E_{0(t+1)}$ son los nacimientos entre t y $t + 1$. En las anteriores expresiones, P_{xt} , es la población de edad x censada en el año t , D_{xt} es el número de muertes de edad x durante el año t a $t + 1$, tales que si hubiesen llegado vivos a $t + 1$ todavía tendrían x años y D'_{xt} es el número de muertes de edad x durante el año t a $t + 1$, tales que si hubiesen llegado vivos a $t + 1$ tendrían $x + 1$. De forma que los E_{xt} , los expuestos al riesgo entre $t - 1$ y t , se pueden considerar los supervivientes de una generación de nacidos en el año $t - x$ que ha llegado a cumplir los x años de edad. Cada sumando $D'_{xt} + D_{xt}$ corresponde a los muertos de una misma generación que han muerto a una misma edad pero en dos años de calendario diferentes. Esta metodología la aplica para obtener tablas de mortalidad de Cataluña y España.

- Por último Navarro (1991) publica unas tablas de mortalidad de la población española correspondientes a los años 1981-1982. Los tantos brutos de mortalidad sin ajustar, como él llama a las probabilidades brutas, los obtiene como sigue

$$\hat{q}_x = \frac{D_{xt} + D_{x(t+1)}}{P_{xt} + P_{x(t+1)} + 1/2(D_{xt} + D_{x(t+1)})},$$

donde P_{xt} es la población con edades entre x y $x + 1$ años a 1 de julio del año t , y D_{xt} es el número de defunciones de individuos con edades entre x y $x + 1$ años durante el año t . En un trabajo posterior, Navarro et al. (1995), con datos de la Comunidad Valenciana, obtiene los tantos de mortalidad a partir del cociente siguiente

$$\hat{q}_x = \frac{D_{x(t-1)} + D_{xt}}{1/2P_{x(t-1)} + P_{xt} + 1/2P_{x(t+1)} + 1/2(D_{x(t-1)} + D_{xt})}, \quad (5.2)$$

la nomenclatura seguida es la misma, pero utiliza población a 1 de enero y referida a 3 años. Interpreta el denominador, que son los expuestos iniciales al riesgo, como

el periodo de tiempo potencialmente vivido por el colectivo objeto del estudio con edades comprendidas entre los x y $x + 1$ años durante el periodo de observación, es decir, los años de $t - 1$ a t .

Nosotros hemos utilizado la propuesta de Navarro et al. (1995) expresada mediante (5.2). Dado que no poseíamos las defunciones clasificadas según el año de nacimiento, sino sólo la edad y el sexo, necesitábamos hacer uso de la hipótesis de uniformidad y la sencillez de esta técnica nos ha parecido adecuada. Como señalaba Rue (1992) al comparar su ajuste con el de Navarro (1991), los resultados son similares si la estimaciones son buenas, cosa que ocurre en nuestro caso por tener un número elevado de expuestos al riesgo. La representación gráfica de los logaritmos de las estimaciones brutas, Figura 5.1, conduce a tomar un rango de edades de 0 a 96 años, que nos parece compatible con la utilización del máximo posible y de un comportamiento relativamente estable.

A continuación aplicaremos los distintos métodos expuestos en el Capítulo 2 a la graduación de las probabilidades brutas estimadas a partir de los datos de la Comunidad Valenciana correspondientes a los años 1999, 2000 y 2001 y utilizando el rango de edades mencionado, 0 a 96 años.

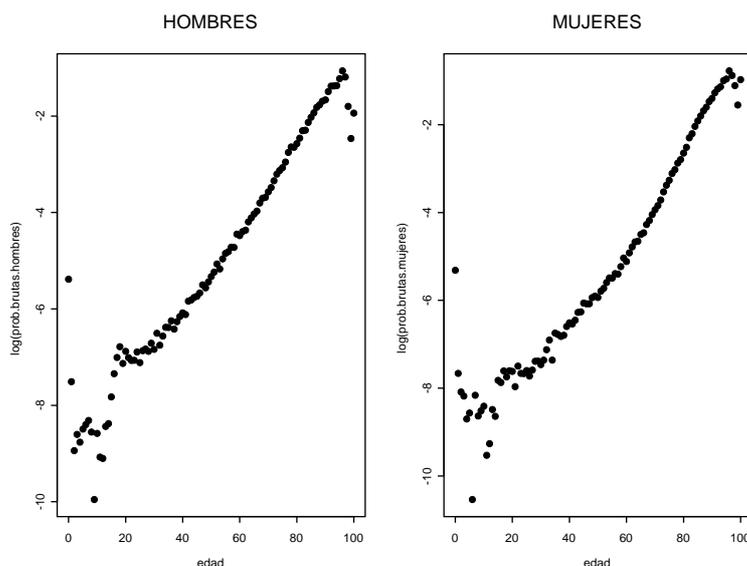


Figura 5.1: Estimaciones brutas de las probabilidades q_x

5.3. Modelos lineales generalizados para q_x

Hemos utilizado modelos lineales generalizados realizando ajustes para polinomios hasta grado 11, lo que significa que hemos ajustado las funciones $LGM(0, 2)$, $LGM(0, 3)$, $LGM(0, 4)$, $LGM(0, 5)$, $LGM(0, 6)$, $LGM(0, 7)$, $LGM(0, 8)$, $LGM(0, 9)$, $LGM(0, 10)$,

$LGM(0, 11)$, $LGM(0, 12)$. La Tabla 5.1 recoge los estadísticos asociados a cada uno de estos modelos para hombres y mujeres y para el rango de edades de 0 a 96 años. El modelo Null consiste en ajustar una constante a los $\text{logit}(q_x)$ y su log-verosimilitud es la correspondiente a las probabilidades brutas o modelo saturado. Como era de esperar se observa una mejora paulatina de la bondad del ajuste a medida que aumenta el número de parámetros. Pero necesitamos contrastar si la disminución de la Deviance es significativa y si el estadístico C_p indica una mejora significativa del ajuste en relación al incremento de complejidad del modelo. El resultado de ambos contrastes se recoge en la Tabla 5.2, en la que la Deviance representa ahora la disminución de la Deviance de un modelo respecto del inmediato anterior, puesto que se trata de comparar modelos.

Se deduce de la tabla que un modelo adecuado para ambos sexos, teniendo en cuenta el compromiso entre la bondad del ajuste y la complejidad del modelo, es el correspondiente al polinomio de grado 10, $LGM(0, 11)$. Los coeficientes de los modelos para cada sexo se muestran en la Tabla 5.3. Todos ellos resultan ser significativos puesto que $t_{86}(0,025) = 2,2813$. Las Figuras 5.2 y 5.3 recogen la comparación gráfica de los modelos para cada sexo. Se presentan ambas en escala logit puesto que el ajuste se ha realizado para esta transformación de las probabilidades brutas de muerte. Para las edades extremas inferiores puede observarse una mayor distancia entre el modelo y los datos observados.

El diagnóstico del modelo lo hemos evaluado a partir de los gráficos de residuos, Figuras 5.4 y 5.5. Cada conjunto de gráficos muestra el comportamiento de los valores observados y de los residuos frente a los valores ajustados. En los correspondientes gráficos Q-Q puede observarse que para ambos sexos, aunque es más acentuado en los hombres, hay estimaciones con desviaciones relativas elevadas, lo que puede suponer un alejamiento de la normal.

5.4. Modelos para μ_x

En esta sección analizamos los ajustes sobre la fuerza de mortalidad, μ_x . Elegiremos un modelo adecuado, como lo hicimos en el anterior apartado, mediante criterios de bondad de ajuste y de no excesiva complejidad del modelo. Primero comenzamos con los modelos lineales generalizados Poisson y Gamma, descritos en la sección 2.5.2. y a continuación los modelos no lineales generalizados descritos en la sección 2.5.3.

5.4.1. Modelos lineales generalizados

Comenzamos por analizar el ajuste de las funciones $GM(0, s)$ con modelos lineales generalizados de la familia Poisson, realizando ajustes para polinomios de hasta grado 11, lo que significa que hemos ajustado las funciones $GM(0, 2)$, $GM(0, 3)$, $GM(0, 4)$, $GM(0, 5)$, $GM(0, 6)$, $GM(0, 7)$, $GM(0, 8)$, $GM(0, 9)$, $GM(0, 10)$, $GM(0, 11)$, $GM(0, 12)$. La Tabla 5.4 recoge los resultados de estos ajustes para hombres y mujeres y la Tabla 5.5 la comparación entre los distintos modelos. Se deduce de ambas tablas que al igual que en el apartado anterior, un modelo adecuado para ambos sexos, teniendo en cuenta el compromiso entre la bondad del ajuste y la complejidad del modelo, es el correspondiente al polinomio de grado 10, $GM(0, 11)$. Los ajustes siguiendo un modelo Gamma, en el que se considera

HOMBRES				
	Null	LGM(0,2)	LGM(0,3)	LGM(0,4)
Deviance	102248	2332.71	890.299	794.336
g.l.	96	95	94	93
log-verosim.	-169352.1	-170518.5	-16797.3	-16749.3
χ^2		14214.66	2042.378	1226.102
	LGM(0,5)	LGM(0,6)	LGM(0,7)	LGM(0,8)
Deviance	769.871	635.951	517.095	375.222
g.l.	92	91	90	89
log-verosim.	-169737.1	-169670.1	-169610.7	-169539.7
χ^2	1000.902	683.4484	527.719	371.5069
	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)
Deviance	225.161	173.946	166.139	165.587
g.l.	88	87	86	85
log-verosim.	-169464.7	-169439.1	-169435.2	-169434.9
χ^2	221.3204	175.5385	169.2215	168.3063

MUJERES				
	Null	LGM(0,2)	LGM(0,3)	LGM(0,4)
Deviance	172974.9	5080.834	886.030	806.223
g.l.	96	95	94	93
log-verosim.	-190038.2	-192578.6	-190481.2	-190441.3
χ^2		121027.5	1839.7	1195.728
	LGM(0,5)	LGM(0,6)	LGM(0,7)	LGM(0,8)
Deviance	792.986	408.348	323.936	279.361
g.l.	92	91	90	89
log-verosim.	-190434.7	-190242.3	-190200.1	-190177.8
χ^2	1047.876	411.9397	317.108	268.026
	LGM(0,9)	LGM(0,10)	LGM(0,11)	LGM(0,12)
Deviance	186.841	155.369	114.160	114.097
g.l.	88	87	86	85
log-verosim.	-190131.6	-190115.8	-190095.2	-190095.2
χ^2	179.1533	151.6369	111.7697	111.825

Tabla 5.1: Medidas de bondad de ajuste para $LGM(0, s)$ con diferentes grados

HOMBRES			
	Deviance	p-value	C_p
LGM(0,2)			2340.630
LGM(0,3)	1442.411	0	902.179
LGM(0,4)	95.962	0	810.177
LGM(0,5)	24.466	0	789.671
LGM(0,6)	133.919	0	659.712
LGM(0,7)	118.856	0	544.816
LGM(0,8)	141.873	0	406.904
LGM(0,9)	150.061	0	260.803
LGM(0,10)	51.215	0	213.548
LGM(0,11)	7.807	0.0052042	209.701
LGM(0,12)	0.552	0.4574606	213.109

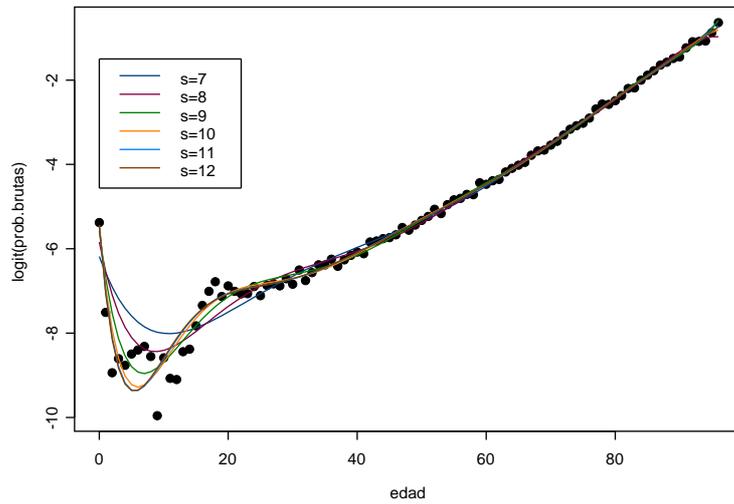
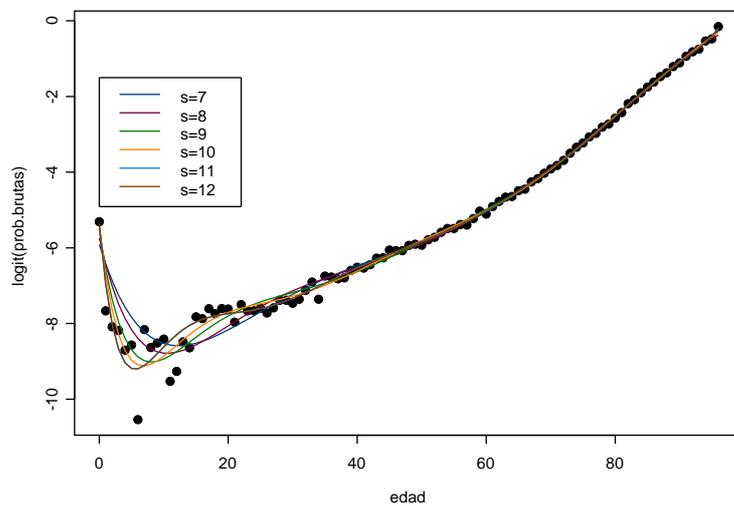
MUJERES			
	Deviance	p-value	C_p
LGM(0,2)			5086.096
LGM(0,3)	4194.804	0	893.923
LGM(0,4)	79.8060	0	816.748
LGM(0,5)	13.238	0.0002744	806.141
LGM(0,6)	384.638	0	424.135
LGM(0,7)	84.412	0	342.354
LGM(0,8)	44.575	0	300.410
LGM(0,9)	92.520	0	210.522
LGM(0,10)	31.473	0	181.680
LGM(0,11)	41.208	0	143.103
LGM(0,12)	0.063	0.8019238	145.671

Tabla 5.2: Comparación de los modelos $LGM(0, s)$ con diferentes grados

HOMBRES			
	coef	std error	t-value
const.	-5.4383e00	7.4566e-02	-72.9324
edad	-1.8744e00	1.1450e-01	-16.3694
edad ²	3.1671e-01	2.6933e-02	11.7589
edad ³	-2.4301e-02	2.7160e-03	-8.9473
edad ⁴	1.0625e-03	1.4991e-04	7.0871
edad ⁵	-2.8896e-05	5.0007e-06	-5.7784
edad ⁶	5.0772e-07	1.0525e-07	4.8239
edad ⁷	-5.7823e-09	1.4075e-09	-4.1081
edad ⁸	4.1300e-11	1.1603e-11	3.5594
edad ⁹	-1.6839e-13	5.3782e-14	-3.1310
edad ¹⁰	2.9938e-16	1.0725e-16	2.7913

MUJERES			
	coef	std error	t-value
const.	-5.3669e00	7.4199e-02	-72.3313
edad	-1.9129e00	1.1578e-01	-16.5227
edad ²	3.4664e-01	2.7820e-02	12.4598
edad ³	-2.9620e-02	2.8311e-03	-10.4624
edad ⁴	1.4456e-03	1.5621e-04	9.2544
edad ⁵	-4.3602e-05	5.1783e-06	-8.4201
edad ⁶	8.4205e-07	1.0792e-07	7.8023
edad ⁷	-1.0445e-08	1.4260e-09	-7.3251
edad ⁸	8.0587e-11	1.1601e-11	6.9468
edad ⁹	-3.5214e-13	5.3023e-14	-6.6413
edad ¹⁰	6.6616e-16	1.0423e-16	6.3910

Tabla 5.3: Coeficientes del modelo $LGM(0, 11)$

Figura 5.2: Comparación Modelos $LGM(0, s)$ para los hombresFigura 5.3: Comparación Modelos $LGM(0, s)$ para las mujeres

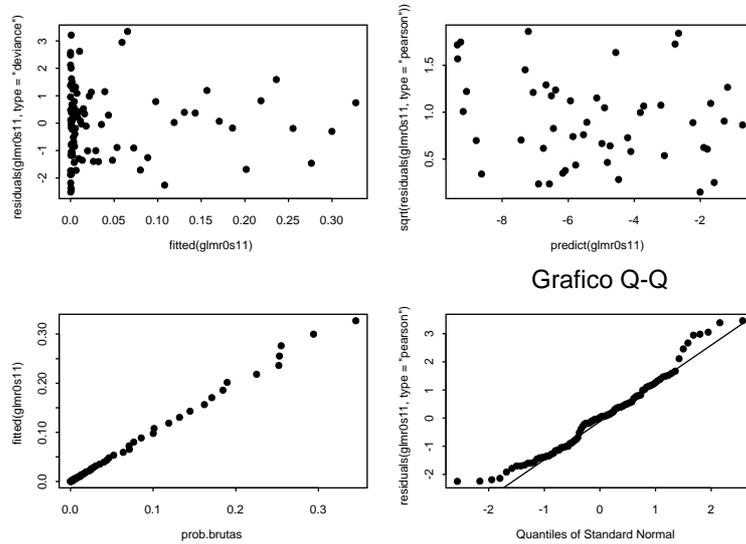


Figura 5.4: $LGM(0, 11)$ para los hombres

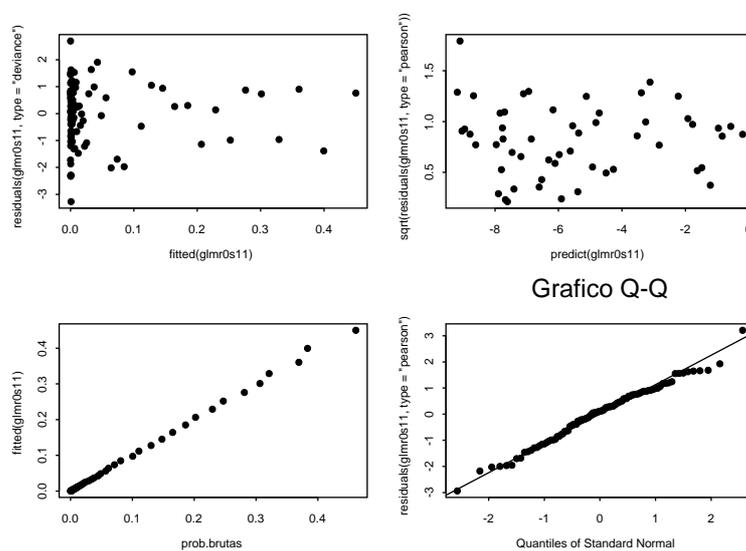


Figura 5.5: $LGM(0, 11)$ para las mujeres

HOMBRES				
	Null	GM(0,2)	GM(0,3)	GM(0,4)
Deviance	102212	2144.978	960.442	793.612
g.l.	96	95	94	93
log-verosim.	218846.9	217774.4	218366.7	218450.1
χ^2		13552.89	2348.766	1202.753
	GM(0,5)	GM(0,6)	GM(0,7)	GM(0,8)
Deviance	782.453	636.609	519.506	379.364
g.l.	92	91	90	89
log-verosim.	218455.7	218528.6	218587.1	218657.2
χ^2	1058.717	685.1312	530.3593	380.6582
	GM(0,9)	GM(0,10)	GM(0,11)	GM(0,12)
Deviance	227.253	174.139	165.855	165.349
g.l.	88	87	86	85
log-verosim.	218733.3	218759.8	218764	218764.2
χ^2	224.1449	176.5361	170.0661	169.2193

MUJERES				
	Null	GM(0,2)	GM(0,3)	GM(0,4)
Deviance	172821.4	4481.879	1133.217	863.537
g.l.	96	95	94	93
log-verosim.	312737.5	310496.6	312170.9	312305.8
χ^2		100664.9	2573.698	1151.411
	GM(0,5)	GM(0,6)	GM(0,7)	GM(0,8)
Deviance	862.634	409.179	324.468	283.970
g.l.	92	91	90	89
log-verosim.	312306.2	312532.9	312575.3	312595.6
χ^2	1198.487	414.03	317.7293	276.3795
	GM(0,9)	GM(0,10)	GM(0,11)	GM(0,12)
Deviance	185.990	155.383	113.458	113.458
g.l.	88	87	86	85
log-verosim.	312644.5	312659.8	312680.8	312680.8
χ^2	178.9773	152.1323	111.5396	111.5346

Tabla 5.4: Medidas de bondad de ajuste para $GM(0, s)$ con diferentes grados. Distribución Poisson

	HOMBRES		
	Deviance	p-value	C_p
GM(0,2)			2152.942
GM(0,3)	1184.537	0	972.387
GM(0,4)	166.829	0	809.539
GM(0,5)	11.159	0.0008362	802.361
GM(0,6)	145.844	0	660.499
GM(0,7)	117.103	0	547.378
GM(0,8)	140.143	0	411.217
GM(0,9)	152.111	0	263.087
GM(0,10)	53.114	0	213.955
GM(0,11)	8.284	0.0039990	209.653
GM(0,12)	0.506	0.477682	213.129

	MUJERES		
	Deviance	p-value	C_p
GM(0,2)			4487.127
GM(0,3)	3348.662	0	1141.090
GM(0,4)	269.680	0	874.034
GM(0,5)	0.903	0.3420526	875.756
GM(0,6)	453.456	0	424.925
GM(0,7)	84.710	0	342.839
GM(0,8)	40.498	0	304.965
GM(0,9)	97.981	0	209.609
GM(0,10)	30.607	0	181.627
GM(0,11)	41.925	0	142.326
GM(0,12)	0.000	0.9920078	144.950

Tabla 5.5: Comparación de los modelos $GM(0, s)$ con diferentes grados. Distribución Poisson

como variable los expuestos al riesgo (Renshaw, Haberman y Hatzopoulos, 1997) y se predice la fuerza de vitalidad, inversa de la fuerza de mortalidad, constituyen la segunda parte de este apartado. Los resultados de los modelos Gamma están medidos en términos de las predicciones de los expuestos al riesgo, mientras que los modelos Poisson lo están en función de las defunciones que es lo que realmente nos interesa. Por este motivo en la Tabla 5.6 presentamos la función de verosimilitud Poisson y la Deviance evaluada en las inversas de las predicciones del modelo Gamma, y la χ^2 calculada para las defunciones, a fin de poder comparar con los resultados anteriores. Estos modelos con los anteriores Poisson, proporcionan peores resultados que los obtenidos con la familia Poisson en todas las medidas de bondad de ajuste y para cada una de las funciones $GM(0, s)$ consideradas.

5.4.2. Modelos no lineales generalizados

Consideramos ahora los modelos correspondientes a las funciones $GM(r, 2)$ con $r \neq 0$ que poseen tanto parte polinómica como parte exponencial, de forma que no es posible transformarlas en una función lineal aplicando logaritmos como en los modelos anteriores. El procedimiento para linealizar la parte exponencial con un máximo de dos parámetros ($s = 2$) y añadirle la parte polinómica según el método iterativo se encuentra descrito en la sección 2.5.3, los resultados que se obtienen después de su ajuste están en la Tabla 5.7. Observamos también en este caso según la Tabla 5.7, donde hemos considerado funciones del tipo $GM(r, 2)$ hasta un máximo de 12 parámetros, que todos los modelos dan peores resultados que en el modelo Poisson.

5.4.3. Resultados

El mejor modelo de cuantos hemos ajustado en esta sección es el $GM(0, 11)$ obtenido mediante modelos lineales generalizados utilizando la familia Poisson. Sus coeficientes para ambos sexos, se muestran en la Tabla 5.8, siendo todos ellos significativos ($t_{86}(0,025) = 2,2813$). Las gráficas de las Figuras 5.6 a 5.9 recogen los resultados para ambos sexos de todos los modelos obtenidos mediante Poisson. Las dos primeras se refieren al ajuste de la fuerza de mortalidad, μ_x , y las dos últimas corresponden al ajuste de q_x obtenidas a partir de aquella mediante la transformación $q_x = 1 - \exp(-\mu_{x+\frac{1}{2}})$. En todas la gráficas se ha utilizado escala logarítmica. También se observa para las edades extremas inferiores una mayor distancia entre el modelo y las observaciones. El diagnóstico del modelo lo hemos realizado también en este caso a partir de los gráficos de residuos, Figuras 5.10 y 5.11. Cada conjunto de gráficos muestra el comportamiento de los valores observados y de los residuos frente a los valores ajustados. También para este modelo podemos observar en las gráficas Q-Q, sobre todo para el sexo masculino, que hay estimaciones en las que se producen desviaciones relativas elevadas, con el consecuente alejamiento de la normal.

HOMBRES				
	Null	GM(0,2)	GM(0,3)	GM(0,4)
Deviance	102212	2311.178	1426.804	1423.071
g.l.	96	95	94	93
log-verosim.	218846.9	217691.3	218133.5	218135.4
χ^2		10391.08	7198.958	7846.742
	GM(0,5)	GM(0,6)	GM(0,7)	GM(0,8)
Deviance	1547.641	1392.247	1156.787	731.034
g.l.	92	91	90	89
log-verosim.	218073.1	218150.8	218268.5	218481.4
χ^2	6854.49	3499.793	1885.866	933.8152
	GM(0,9)	GM(0,10)	GM(0,11)	GM(0,12)
Deviance	323.0882	271.3671	279.2248	278.1222
g.l.	88	87	86	85
log-verosim.	218685.4	218711.2	218707.3	218707.8
χ^2	381.4288	318.1182	331.7748	331.5815

MUJERES				
	Null	GM(0,2)	GM(0,3)	GM(0,4)
Deviance	172821.4	11628.81	1694.593	1691.821
g.l.	96	95	94	93
log-verosim.	312737.5	306923.1	311890.2	311891.6
χ^2		29977.59	7965.665	9101.772
	GM(0,5)	GM(0,6)	GM(0,7)	GM(0,8)
Deviance	2014.139	1080.164	1018.142	897.1084
g.l.	92	91	90	89
log-verosim.	311730.5	312197.5	312228.5	312289
χ^2	8058.538	2837.938	1600.561	1048.8
	GM(0,9)	GM(0,10)	GM(0,11)	GM(0,12)
Deviance	408.2796	203.1785	173.3137	159.9273
g.l.	88	87	86	85
log-verosim.	312533.4	312635.9	312650.9	312657.6
χ^2	438.6796	220.3409	188.9953	180.2086

Tabla 5.6: Medidas de bondad de ajuste para $GM(0, s)$ con diferentes grados. Distribución Gamma

HOMBRES					
	Null	GM(1,2)	GM(3,2)	GM(4,2)	GM(5,2)
Deviance	102212	862.7652	736.7106	605.2137	342.1944
g.l.	96	95	94	93	92
log-verosim.	218846.9	218415.5	21478.5	218544.3	218675.8
χ^2		1378.717	1066.584	803.3898	347.6592
	GM(6,2)	GM(7,2)	GM(8,2)	GM(9,2)	GM(10,2)
Deviance	311.1831	339.0331	330.6377	318.0727	274.3146
g.l.	91	90	89	88	87
log-verosim.	218691.3	218677.4	218681.6	218687.9	218709.7
χ^2	280.7295	338.2825	324.1219	309.911	268.9084

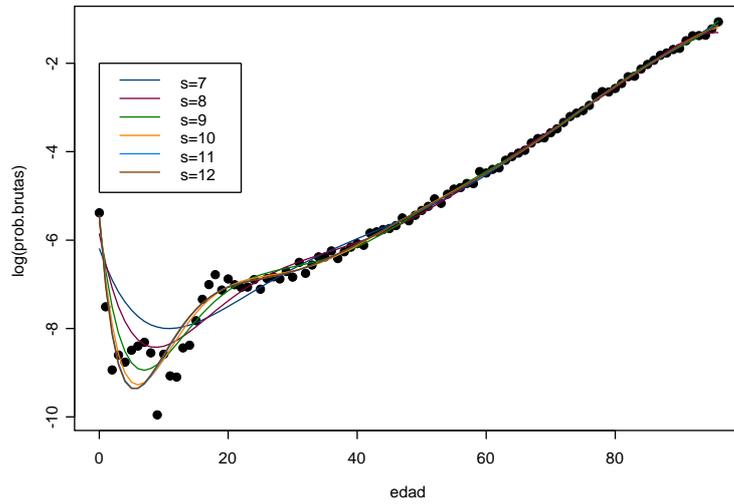
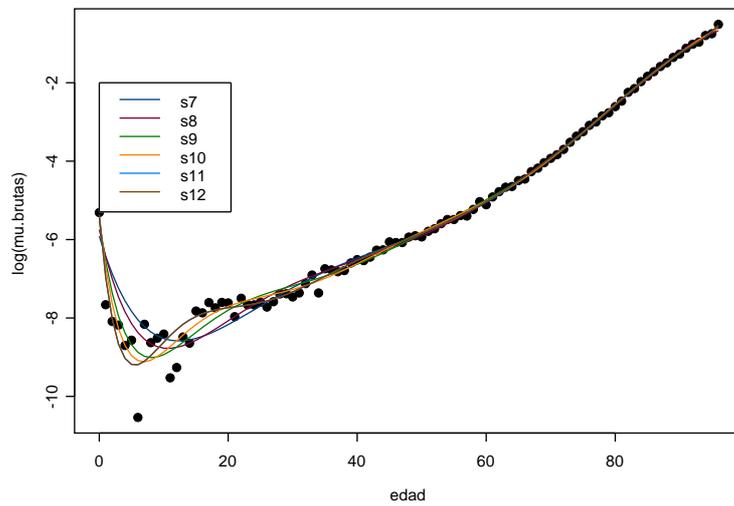
MUJERES					
	Null	GM(1,2)	GM(3,2)	GM(4,2)	GM(5,2)
Deviance	172821.4	1129.709	736.711	611.269	603.067
g.l.	96	95	94	93	92
log-verosim.	312737.5	312172.7	310226.6	312431.9	312436
χ^2		1782.388	1066.584	787.8738	771.4147
	GM(6,2)	GM(7,2)	GM(8,2)	GM(9,2)	GM(10,2)
Deviance	352.417	337.966	270.229	262.898	215.508
g.l.	91	90	89	88	87
log-verosim.	312561.3	312568.6	312602.4	312606.1	312629.8
χ^2	390.4903	362.174	274.3989	261.4965	206.5404

Tabla 5.7: Medidas de bondad de ajuste para $GM(r, 2)$ con diferentes grados

HOMBRES			
	coef	std error	t-value
const.	-5.4388e00	7.43e-02	-73.105
edad	-1.8739e00	1.1472e-01	-16.3344
edad ²	3.1652e-01	2.6846e-02	11.7901
edad ³	-2.4276e-02	2.6933e-03	-9.0133
edad ⁴	1.0607e-03	1.4795e-04	7.1691
edad ⁵	-2.8825e-05	4.9134e-06	-5.8666
edad ⁶	5.0599e-07	1.0298e-07	4.9134
edad ⁷	-5.7562e-09	1.3717e-09	-4.1964
edad ⁸	4.1062e-11	1.1264e-11	3.6452
edad ⁹	-1.6718	5.2021e-14	-3.2137
edad ¹⁰	2.9677e-16	1.0337e-16	2.8709

MUJERES			
	coef	std error	t-value
const.	-5.4388e00	7.43e-02	-73.105
edad	-1.8739e00	1.1472e-01	-16.3344
edad ²	3.1652e-01	2.6846e-02	11.7901
edad ³	-2.4276e-02	2.6933e-03	-9.0133
edad ⁴	1.0607e-03	1.4795e-04	7.1691
edad ⁵	-2.8825e-05	4.9134e-06	-5.8666
edad ⁶	5.0599e-07	1.0298e-07	4.9134
edad ⁷	-5.7562e-09	1.3717e-09	-4.1964
edad ⁸	4.1062e-11	1.1264e-11	3.6452
edad ⁹	-1.6718	5.2021e-14	-3.2137
edad ¹⁰	2.9677e-16	1.0337e-16	2.8709

Tabla 5.8: Coeficientes del modelo $GM(0, 11)$. Distribución Poisson

Figura 5.6: Comparación Modelos $GM(0, s)$ para los hombresFigura 5.7: Comparación Modelos $GM(0, s)$ para las mujeres

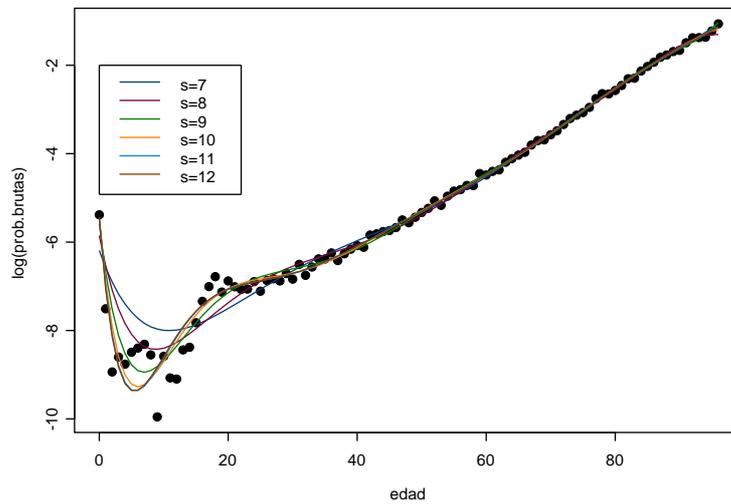


Figura 5.8: Comparación de los q_x correspondientes a los Modelos $GM(0, s)$ para hombres

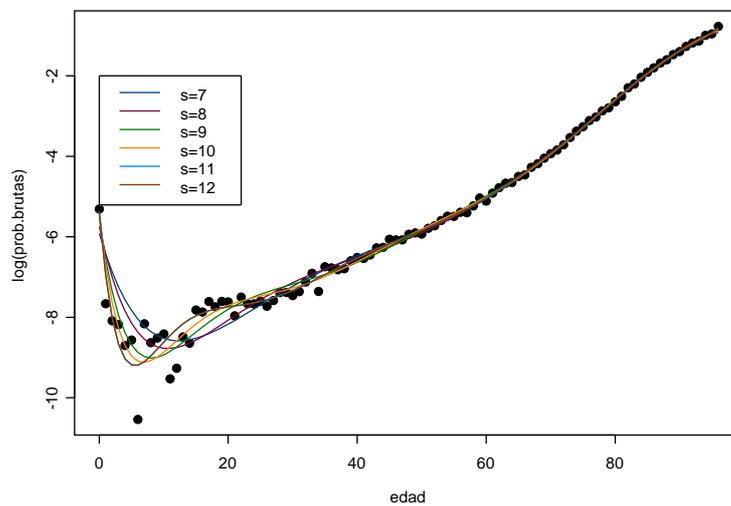


Figura 5.9: Comparación de los q_x correspondientes a los Modelos $GM(0, s)$ para mujeres

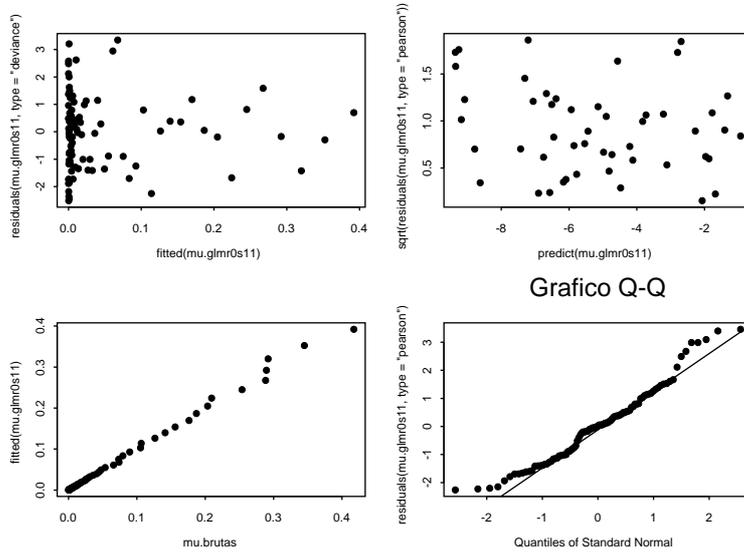


Figura 5.10: $GM(0, 11)$ para los hombres

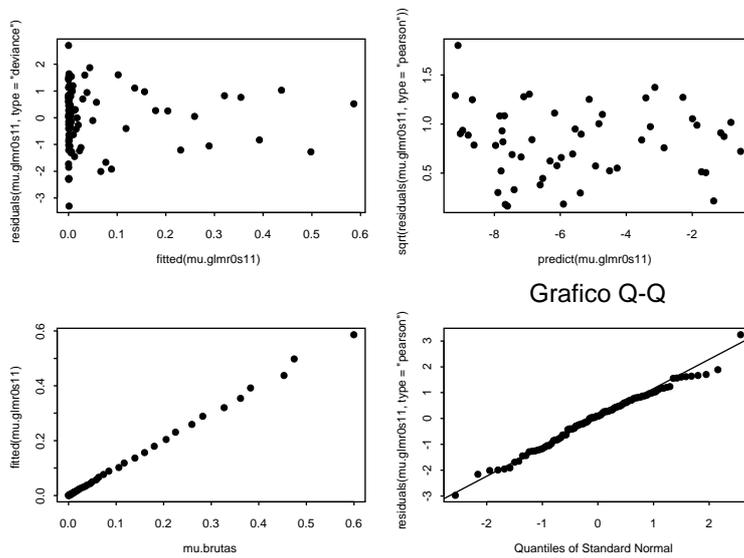


Figura 5.11: $GM(0, 11)$ para las mujeres

5.5. Modelo de Heligman y Pollard

Otra alternativa para la graduación paramétrica de la mortalidad es la que presentan las leyes de Heligman y Pollard, su descripción la encontramos en la sección 2.6. Entre los distintos trabajos que aplican estas leyes podemos encontrar la tesis Rue (1992), trabajo riguroso donde se ajustan las diferentes leyes a datos de España y Cataluña, concluyendo que la segunda es la que mejores resultados proporciona. Para estimar los coeficientes, dada la heterocedasticidad de los datos, se recurre al criterio de mínimos cuadrados ponderados $\sum_{x=0}^{96} \omega_x (\dot{q}_x - f_\alpha(x))^2$, donde $f_\alpha(x)$ denota la función correspondiente a las segunda leyes de Heligman y Pollard. Para ponderar se toman valores que resultan inversamente proporcionales a la varianza $\left(\text{var}(q_x) = \frac{q_x(1-q_x)}{E_x} \right)$ que al ser estimadas por $\text{var}(\dot{q}_x) = \frac{\dot{q}_x(1-\dot{q}_x)}{E_x}$ y puesto que $(1-q_x)$ es aproximadamente 1, dan lugar a los mínimos cuadrados siguientes, entre los que hemos incluido los mínimos cuadrados sin ponderar.

1.

$$\sum_{x=0}^{96} \frac{E_x}{\dot{q}_x} (\dot{q}_x - f_\alpha(x))^2$$

2.

$$\sum_{x=0}^{96} (\dot{q}_x - f_\alpha(x))^2$$

3.

$$\sum_{x=0}^{96} \frac{E_x}{\dot{q}_x^2} (\dot{q}_x - f_\alpha(x))^2$$

4.

$$\sum_{x=0}^{96} \frac{1}{\dot{q}_x^2} (\dot{q}_x - f_\alpha(x))^2$$

La comparación gráfica de los ajustes obtenidos con los diferentes criterios de ponderación (Figuras 5.12 y 5.13) y los resultados de la Tabla 5.9 en la que se muestran el número de desviaciones relativas mayores que 2 y 3 y el valor de la χ^2 para la bondad del ajuste, nos llevan a elegir el primer criterio como el que mejor resultados presenta para nuestros datos.

Elegimos los coeficientes del primer criterio, pero dado que la matriz de correlaciones correspondiente a estos ajustes resultaba ser singular no podíamos analizar la significatividad de los coeficientes. Por ello, hemos recurrido a mínimos cuadrados no lineales generalizados (generalized non-linear least squares, gnls). Los coeficientes correspondientes al modelo Heligman-Pollard en el caso de los hombres no ha presentado mayor dificultad, hemos partido del punto inicial proporcionado por el ajuste anterior y como estructura de varianzas hemos utilizado una potencia de la variable de la ponderación. Para las mujeres

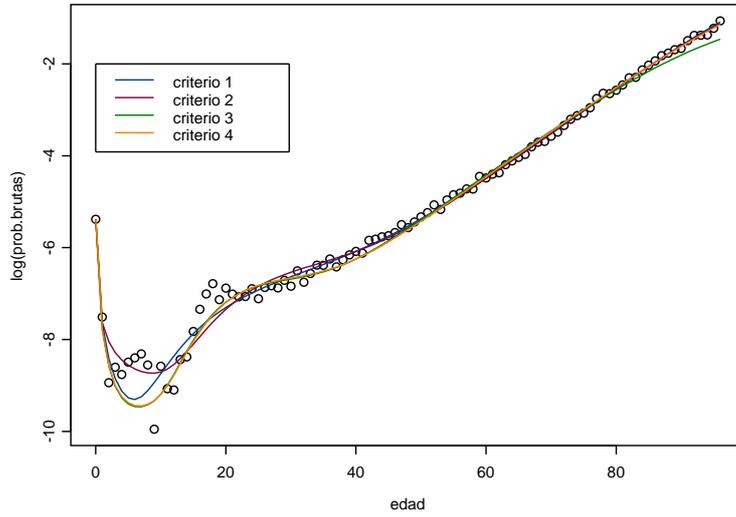


Figura 5.12: Comparación Modelos Heligman y Pollard para los hombres

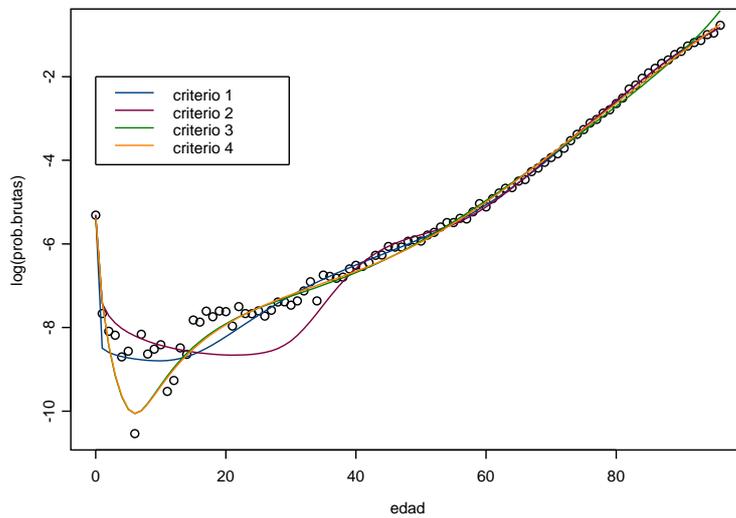


Figura 5.13: Comparación Modelos Heligman y Pollard para las mujeres

criterio	HOMBRES			MUJERES		
	$2 \leq z_x \leq 3$	$z_x > 3$	χ^2	$2 \leq z_x \leq 3$	$z_x > 3$	χ^2
1	8	6	229.9192	9	8	218.7686
3	14	25	703.145	15	24	1267.317
4	17	8	336.4597	11	21	497.1236

Tabla 5.9: Comparación de los resultados según los diferentes criterios

sí se han presentado problemas porque la población femenina española, como ya señalan Felipe y Guillén (1999), no presenta la sobremortalidad que la ley presupone. La sobremortalidad que presentan las mujeres españolas es mucho menor que la de los hombres y se extiende a lo largo de muchos más años, debido a lo cual hemos decidido centrar la joroba de los accidentes para que no tubiera efecto en 96 años, es decir fijar el parámetro $F = 96$, y estimar el resto de parámetros, además de utilizar como estructura de varianza la proporcionada por los propios valores ajustados ($\text{fitted}(\cdot)$). Esta técnica de fijar valores de unos parámetros y ajustar el resto viene sugerida por el trabajo de Congdon (1993). Los valores resultantes para los parámetros de ambos sexos se recogen en la Tabla 5.10.

5.6. Comparación de los diferentes métodos

Para contrastar si los ajustes obtenidos son una razonable representación del comportamiento de la mortalidad que estamos estudiando, hemos aplicado los tests descritos en Forfar, McCutcheon y Wilkie (1988), que también aplican en su trabajo Navarro (1991) y Navarro et al. (1995) y que hemos descrito en la sección 2.7. Los test no paramétricos para los modelos de μ_x han sido aplicados sobre los correspondientes valores, q_x , calculados a partir del mejor modelo para las μ_x a través de la relación

$$q_x = 1 - \exp(-\mu_{x+\frac{1}{2}}).$$

En algunos casos hemos necesitado agrupar las edades hasta tener un número de muertes esperadas superior a 5 con la consecuente disminución del número de grados de libertad. La agrupación ha sido necesaria en el ajuste de las funciones $LGM(0, 11)$ y $GM(0, 11)$ en los valores ajustados correspondientes a las edades de 4 a 9 años para los hombres y a las edades de 4 a 7 años para las mujeres, mientras que en los ajustes de la ley de Heligman y Pollard solo ha sido necesaria en los hombres para las mismas edades.

Hemos obtenido también los valores del error porcentual absoluto medio (EPAM) y R^2 que utilizan Felipe y Guillén (1999) en su trabajo. El valor de R^2 se ha obtenido como 1 menos la proporción de varianza que queda sin explicar, pues si lo calculamos directamente como porcentaje de varianza explicada en algunos casos superaba la unidad, esto puede ocurrir cuando los modelos no son lineales.

De los tests, cuyos resultados se presentan en la Tabla 5.11, podemos deducir que los modelos Heligman y Pollard ajustan mal, proporcionando bastantes desviaciones superiores a 2. En el caso de los modelos lineales generalizados el número de desviaciones

HOMBRES				
	coef	std error	t-value	p-value
A	0.00054	0.00035	1.5670	0.1207
B	0.12921	0.21037	0.6142	0.5407
C	0.16301	0.09059	1.7993	0.0754
D	0.00138	0.00065	2.1289	0.0361
E	0.74764	0.44753	1.6706	0.0984
F	63.03293	37.14621	1.6969	0.0933
G	0.00002	0	3.4307	0.0009
H	1.11313	0.00425	262.0285	< ,0001
K	0.91755	0.26233	3.4977	0.0007

MUJERES				
	coef	std error	t-value	p-value
A	0.000335	0.0000746	4.4886	<0.0001
B	0	0.0000030	0.1073	0.9148
C	0.027444	0.0153661	1.7860	0.0775
D	0.002757	0.0003135	8.7939	<0.0001
E	1.140099	0.1308293	8.7144	<0.0001
F	96	—	—	—
G	0.000001	0.0000001	4.5996	<.0001
H	1.159430	0.0031811	364.4753	<.0001
K	1.108379	0.0822712	13.4723	<.0001

Tabla 5.10: Coeficientes de los modelos Heligman y Pollard

		LGM(0,11) para q_x		GM(0,11) para μ_x		H-P	
		Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres
Desv. relativas	> 2	8	4	8	3	6	10
	> 3	3	0	3	0	6	2
Test de signos	pos.(neg.)	46 (48)	53 (42)	46 (48)	53 (42)	50 (44)	47 (50)
	p-value	0.4589	0.8909	0.4589	0.8909	0.7647	0.4196
Test de rachas	rachas	44	50	44	50	41	35
	p-value	0.4319	0.5372	0.4319	0.5372	0.3839	0.2731
Test de K-S	K-S	0.0426	0.0316	0.0433	0.0316	0.0532	0.0825
	p-value	1	1	1	1	0.9994	0.8987
Test χ^2	χ^2	164.18	101.07	164.32	102.44	224.31	165.56
	g.l.	83	84	83	84	85	88
	p-value	2.80e-07	0.0989	2.69e-07	0.0836	1.66e-04	1.11e-06
R²		0.9972	0.9991	0.9972	0.9991	0.9967	0.9981
EPAM		16.35	16.44	16.34	16.44	15.13	21.17

Tabla 5.11: Resultados de los contrastes no paramétricos sobre los modelos paramétricos

superiores a 2 es bastante menor, aunque la χ^2 continua resultando significativa para los hombres. Para las mujeres el valor de χ^2 es aceptable y nos conduce a aceptar las funciones $LGM(0, 11)$ y $GM(0, 11)$. El resto de contrastes son positivos en ambos casos aunque mejores para $LGM(0, 11)$, por lo que pensamos que podemos aceptar los modelos $LGM(0, 11)$ como buenos. Los tests correspondientes a las autocorrelaciones de las desviaciones relativas pueden observarse en las Figuras 5.14 a 5.19. Muestran pequeñas correlaciones aisladas en todos los casos. Aunque para el modelo H-P de las mujeres muestran además un patrón reconocible, lo que corrobora el mal comportamiento de este modelo al que antes aludíamos.

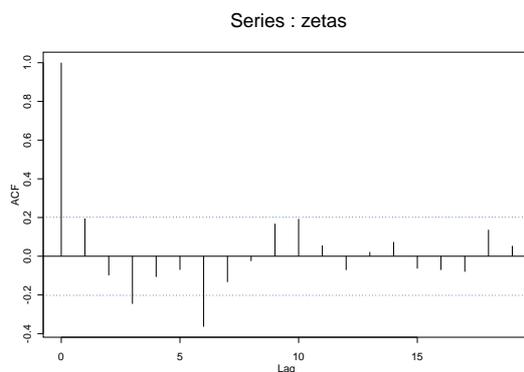


Figura 5.14: Autocorrelaciones del modelo $LGM(0, 11)$ para los hombres

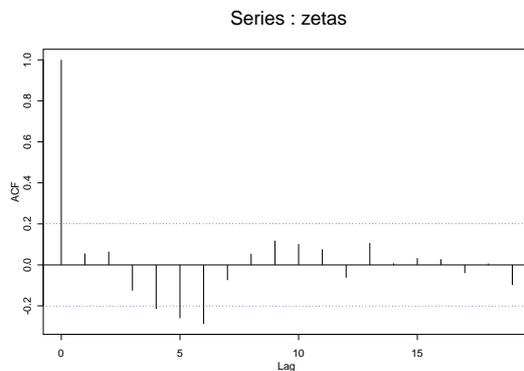


Figura 5.15: Autocorrelaciones del modelo $LGM(0, 11)$ para las mujeres

En resumen, de nuestros resultados deducimos que el modelo $LGM(0, 11)$ proporciona ajustes razonablemente buenos para ambos sexos sin necesidad de recurrir a una división en tramos de edad. Navarro et al. (1995) consiguen mejores ajustes, pero fraccionando el rango de edades, lo que conlleva una disminución de los grados de libertad de las

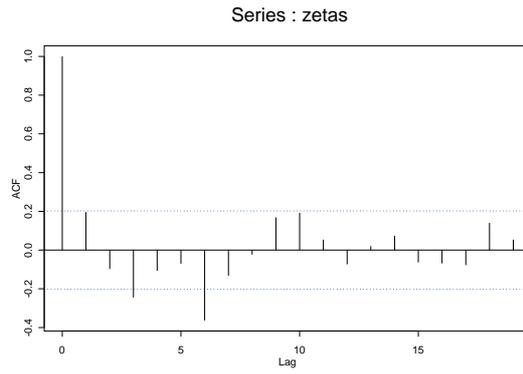


Figura 5.16: Autocorrelaciones de las q_x correspondientes al modelo $GM(0, 11)$ para los hombres

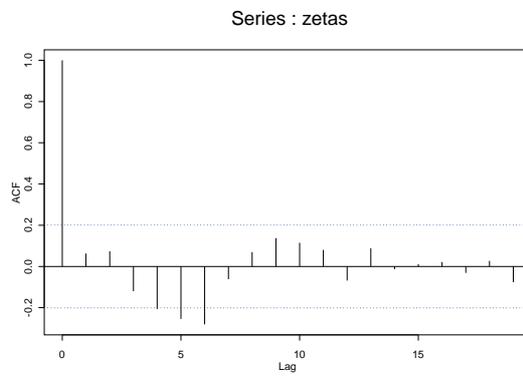


Figura 5.17: Autocorrelaciones de las q_x correspondientes al modelo $GM(0, 11)$ para las mujeres

correspondientes χ^2 . Además, debido a las restricciones impuestas a las funciones en los puntos de empalme, se producen irregularidades (picos) en dichos puntos. Respecto a los Modelos Heligman y Pollard debemos desestimarlos debido al mal ajuste que presentan.

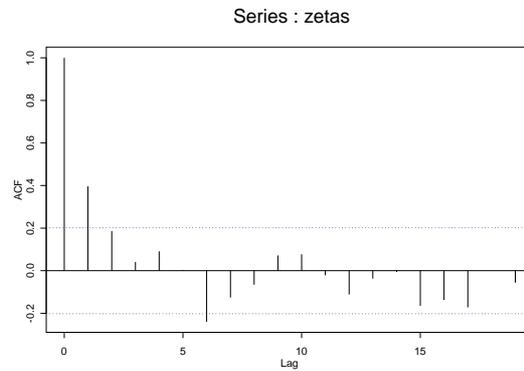


Figura 5.18: Autocorrelaciones del modelo H-P para los hombres

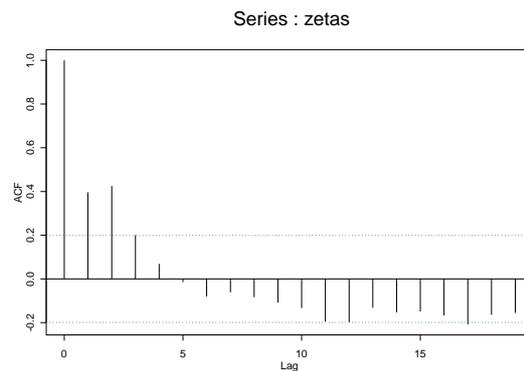


Figura 5.19: Autocorrelaciones del modelo H-P para las mujeres

Capítulo 6

Ajustes no paramétricos de tablas de mortalidad de la Comunidad Valenciana

6.1. Introducción

Los métodos no paramétricos para la graduación de los datos de mortalidad constituyen una buena alternativa a los métodos paramétricos cuando no se tiene información sobre la función adecuada. Además, son un necesario complemento para éstos, cuando los datos presentan formas muy irregulares, susceptibles de ser suavizadas previamente al ajuste de la función.

En la sección 2 analizamos los métodos kernel, en la tercera sección aplicamos la graduación mediante splines, la cuarta se ocupa de la graduación local polinomial y la quinta de los modelos aditivos generalizados (GAM) todos ellos aplicados a los mismos datos que en el anterior capítulo. Posteriormente comparamos los resultados de todos estos métodos.

6.2. Graduación mediante kernel

Consideramos los estimadores kernel descritos en la sección 3.2, que aplicaremos primero sobre las probabilidades brutas y después sobre sus transformaciones logaritmo, logit y complementario log-log. Una primera comparación de los resultados la llevamos a cabo gráficamente, analizando después para cada una de estas transformaciones el comportamiento de sus residuos, r_i , de la ecuación (3.6) correspondientes a diferentes bandwidth. Posteriormente los comparamos mediante los tests no paramétricos descritos al final de Capítulo 2 con el fin de estudiar la incidencia de la elección del bandwidth sobre el comportamiento de los valores ajustados. Por último, planteamos también la posibilidad de calcular el valor de bandwidth mediante validación cruzada.

Comenzamos nuestro trabajo en esta sección graduando las probabilidades de muerte, q_x , según los estimadores descritos en la sección 3.2.2, el estimador Nayadara-Watson y el

Copas-Haberman, con diferentes elecciones del bandwidth. Las Figuras 6.1 y 6.2 muestran los ajustes para hombres y mujeres respectivamente.

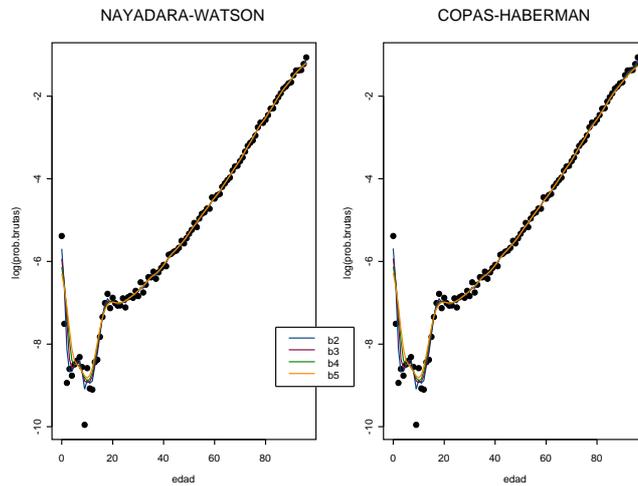


Figura 6.1: Comparación estimación kernel para los hombres

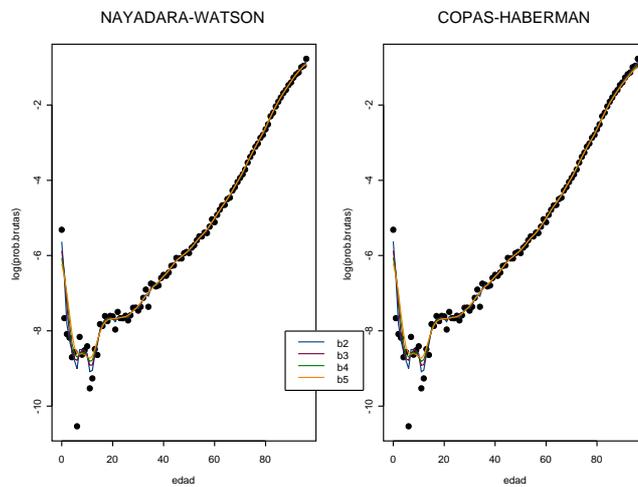


Figura 6.2: Comparación estimación kernel para las mujeres

Los resultados para ambos sexos se recogen en la Tabla 6.1 donde se presentan las principales medidas de bondad de ajuste. Para la validación de las hipótesis de los residuos mostramos las gráficas de los residuos de los ajustes frente a los valores estimados, Figuras 6.3 a 6.6, gráficas de las que se deduce existencia de heterocedasticidad, lo que nos lleva a

HOMBRES				
Nayadara-Watson	b=2	b=3	b=4	b=5
Deviance	77.90238	160.0798	228.3591	287.7463
log-verosim.	-169391.1	-169432.2	-169466.3	-169496
χ^2	73.05089	156.7288	235.9519	313.3537
Copas-Haberman	b=2	b=3	b=4	b=5
Deviance	79.18682	164.5619	240.7729	315.5821
log-verosim.	-169391.7	-169434.4	-169472.5	-169509.9
χ^2	74.05068	160.5707	247.4864	340.3919

MUJERES				
Nayadara-Watson	b=2	b=3	b=4	b=5
Deviance	84.23872	158.8281	222.7136	290.6831
log-verosim.	-190080.3	-190117.6	-190149.5	-190183.5
χ^2	76.52356	154.7207	231.8736	318.9326
Copas-Haberman	b=2	b=3	b=4	b=5
Deviance	87.24015	170.5856	254.5687	359.5939
log-verosim.	-1900081.8	-190123.5	-190165.4	-190218
χ^2	79.18068	165.746	262.8249	387.6561

Tabla 6.1: Medidas de bondad de ajuste para suavización kernel sobre las probabilidades brutas

proponer las transformaciones ya citadas sobre las probabilidades brutas.

De acuerdo a los resultados de la Tabla 6.1 ajusta mejor el estimador Nayadara-Watson que Copas-Haberman para ambos sexos. Estimador que hemos analizado con diferentes valores del bandwidth, a los que hemos aplicado los test no paramétricos para que nos ayuden en la difícil tarea de elección del valor de b . En la Tabla 6.2 puede apreciarse que el número de desviaciones relativas mayores de 2 y 3 aumentan conforme aumentamos el suavizado y también lo hace la χ^2 . Por otra parte, el test de signos no se supera para los bandwidth altos. El valor $b = 3$ parece una buena elección para ambos sexos, siendo también recomendable $b = 4$ para el caso de las mujeres. Hay que resaltar también que este suavizado deja una ondulación en la edades muy próximas a cero difícil de explicar. Hemos analizado también el comportamiento de las desviaciones y las desviaciones relativas para cada uno de los bandwidth. Lo mostramos en las Figuras 6.7 y 6.8 para el bandwidth elegido $b = 3$. Posteriormente realizaremos la elección del bandwidth mediante validación cruzada con el fin de compararla con la anterior y discutirla.

Puesto que la hipótesis de homocedasticidad no se cumple, aplicaremos ahora la suavización mediante el método de Nayadara-Watson a las probabilidades brutas transformadas. Las correspondientes medidas de bondad de ajuste se muestran en la Tabla 6.3 y son en general peores que los obtenidos con los valores brutos de los q_x sin transformar, resultando un poco mejores para la transformación logit. Pero estas transformaciones se han realizado con el fin de eliminar la heterocedasticidad, veamos si esto es así representando los residuos frente a las estimaciones de los valores transformados en las Figuras 6.9 y 6.10. Como los

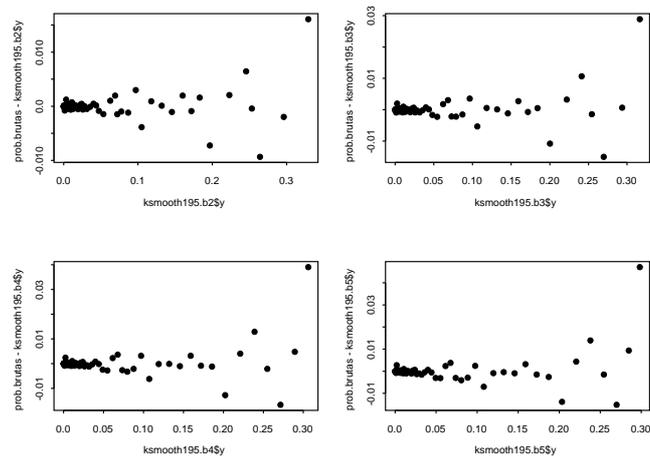


Figura 6.3: Residuos frente a estimación Nadaraya-Watson para los hombres

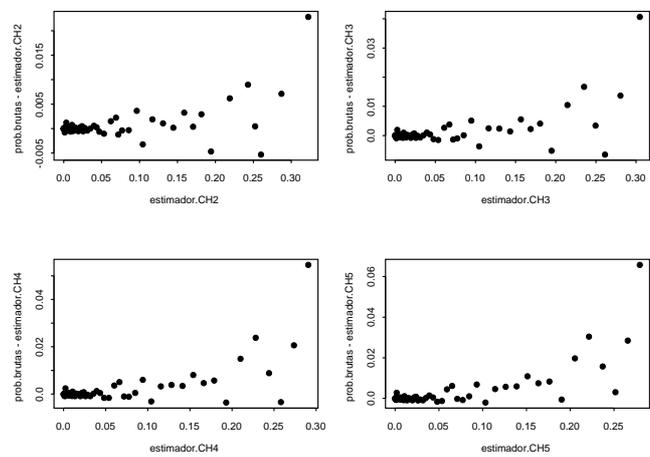


Figura 6.4: Residuos frente a estimación Copas-Haberman para los hombres

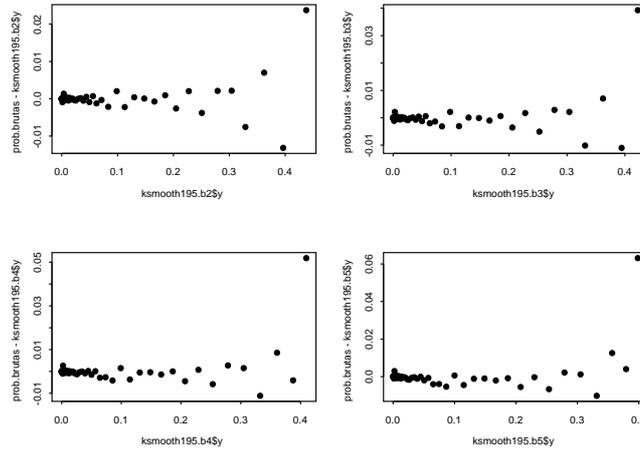


Figura 6.5: Residuos frente a estimación Nadaraya-Watson para las mujeres

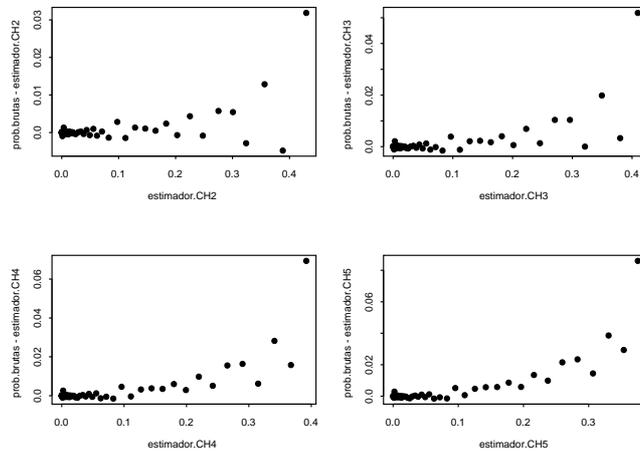


Figura 6.6: Residuos frente a estimación Copas-Haberman para las mujeres

		HOMBRES			
		b=2	b=3	b=4	b=5
Desv. relativas	≥ 2	0	0	2	6
	> 3	2	3	3	3
Test de signos	pos. (neg.)	49 (48)	45 (52)	40 (57)	39 (58)
	p-value	0.58	0.27	0.052	0.033
Test de rachas	rachas	66	56	48	46
	p-value	0.75	0.61	0.5	0.47
Test de K-S	K-S	0.031	0.041	0.041	0.051
	p-value	1	1	1	0.999
χ^2		73.05	156.73	235.95	313.35
R^2		0.99	0.98	0.96	0.95
EPAM		8.57	13.34	16.86	19.66

		MUJERES			
		b=2	b=3	b=4	b=5
Desv. relativas	mayores de 2	0	3	3	7
	mayores de 3	2	2	4	4
Test de signos	pos. (neg.)	48 (49)	44 (53)	42 (55)	36 (61)
	p-value	0.50	0.208	0.111	0.007
Test de rachas	rachas	70	66	60	40
	p-value	0.80	0.76	0.68	0.38
Test de K-S	K-S	0.031	0.041	0.041	0.041
	p-value	1	1	1	1
χ^2		76.52	154.72	231.87	318.93
R^2		0.99	0.97	0.96	0.95
EPAM		11.80	16.16	18.69	21.06

Tabla 6.2: Contrastes no paramétricos para el estimador Nayadara-Watson sobre las probabilidades brutas

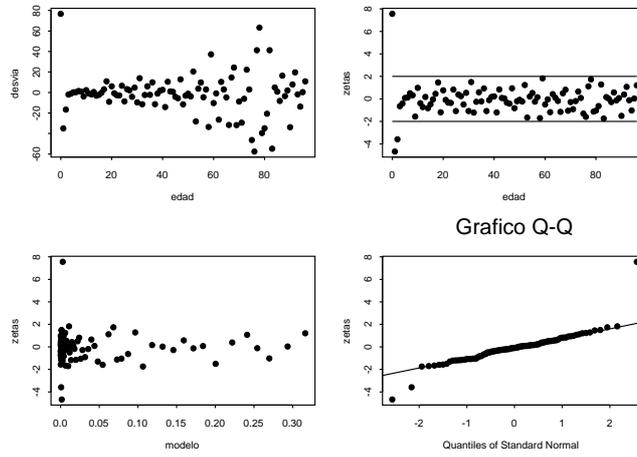


Figura 6.7: Nayadara-Watson con $b = 3$ para los hombres

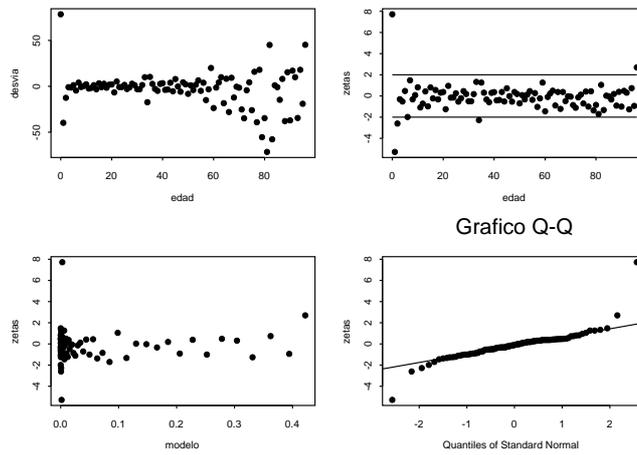


Figura 6.8: Nayadara-Watson con $b = 3$ para las mujeres

HOMBRES				
log(q_x)	b=2	b=3	b=4	b=5
Deviance	100.4441	240.6098	362.3080	457.6372
log-verosim.	-169402.3	-169472.4	-169533.3	-169580.9
χ^2	118.4661	343.0014	610.4596	879.7216
logit(q_x)	b=2	b=3	b=4	b=5
Deviance	100.2475	240.2169	361.7316	456.8064
log-verosim.	-169402.2	-169472.2	-169472.2	-169533
χ^2	118.1821	342.256	609.2218	877.9678
clog(q_x)	b=2	b=3	b=4	b=5
Deviance	100.3425	240.4077	361.9959	457.1701
log-verosim.	-169402.3	-169472.3	-169472.3	-169533.1
χ^2	118.3204	342.6216	609.8023	878.7643

MUJERES				
log(q_x)	b=2	b=3	b=4	b=5
Deviance	113.4832	238.3679	334.6651	417.6035
log-verosim.	-190094.9	-190157.3	-190205.5	-190247
χ^2	136.5452	340.5108	548.2968	762.4333
logit(q_x)	b=2	b=3	b=4	b=5
Deviance	113.1109	237.378	332.5785	413.4449
log-verosim.	-190094.7	-190156.9	-190156.9	-190204.5
χ^2	136.0494	339.1464	545.5995	757.4457
clog(q_x)	b=2	b=3	b=4	b=5
Deviance	113.2812	237.8081	333.4597	415.1706
log-verosim.	-190094.8	-190157.1	-190157.1	-190204.9
χ^2	136.2812	339.7663	546.79	759.569

Tabla 6.3: Medidas de bondad de ajuste para kernel sobre las transformadas de las probabilidades brutas

gráficos son similares para las diferentes transformaciones, sólo mostramos la logarítmica y la logit. A la vista de los resultados comprobamos que ha mejorado la situación y aun-

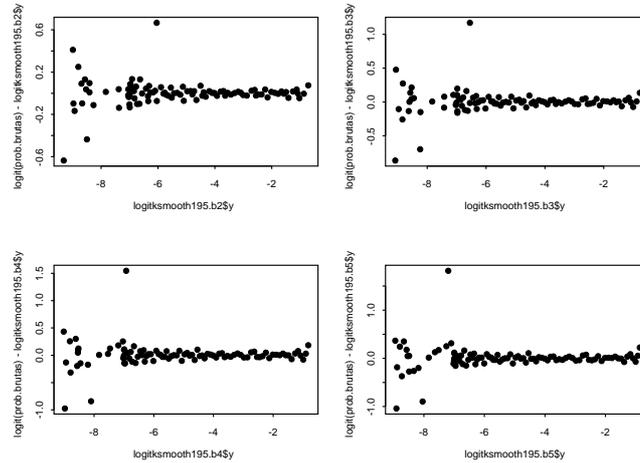


Figura 6.9: Residuos frente a estimación del $\text{logit}(q_x)$ para los hombres

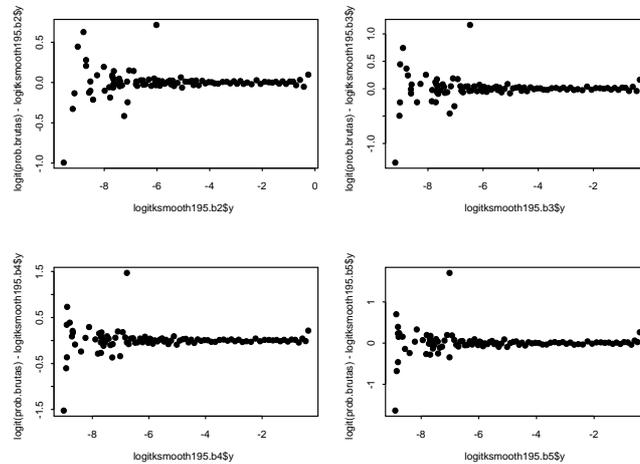


Figura 6.10: Residuos frente a estimación del $\text{logit}(q_x)$ para las mujeres

que parece que se invierte la relación de la varianza con las estimaciones, se observa que los residuos se estabilizan a medida que aumenta la estimación, obteniéndose desviaciones bastante elevadas en las estimaciones para edades tempranas. Tomaremos pues en la comparación final con los diferentes ajustes no paramétricos, el kernel sobre la transformación

	HOMBRES	MUJERES
Bandwidth	5.218196	4.554322
Deviance	113.2749	96.087
log-verosim.	-169408.8	-190086.2
χ^2	117.979	96.01836
Df	20.44698	23.35616

Tabla 6.4: Medidas de bondad de ajuste para kernel sobre las transformadas logit de las probabilidades brutas para el bandwidth obtenido mediante validación cruzada

logit. Hemos realizado el suavizado con los diferentes valores del bandwidth, resultando recomendable el suavizado con $b = 5$ que supera todos los test no paramétricos para ambos sexos. En la Tabla 6.3 hay que tener en cuenta que el elevado valor de χ^2 para los hombres y las mujeres se debe a la desviaciones relativas en cero, 27.82 y 25.82, respectivamente, ambas incluidas en el valor de la tabla. De no considerar estas desviaciones, resultarían valores de $\chi^2 = 103,8895$ y $\chi^2 = 90,81021$, respectivamente, ambos muy razonables para el ajuste. Además esta elección nos proporciona valores de edades cercanas a cero sin oscilaciones.

Finalmente hemos calculado el valor del bandwidth por validación cruzada para las transformaciones logit, sin tener en cuenta la edad cero pues si la consideráramos nos proporcionaba un valor muy bajo de b que apenas producía suavizado. En la Tabla (6.4) se recogen los resultados. De acuerdo con ellos la validación cruzada proporciona un valor cercano al que habíamos seleccionado con los test no paramétricos, lo que reafirma aquella elección.

Finalmente seleccionamos el modelo kernel con bandwidth elegido según validación cruzada para comparar con el resto de ajustes no paramétricos en la última sección.

6.3. Graduación mediante splines

En la sección 3.3 introdujimos el criterio de Whittaker-Henderson que conducía a los splines cúbicos como mejor solución de equilibrio entre la bondad de ajuste y la suavidad. Para el ajuste utilizamos la función *smooth.spline* del S-plus, en la que podemos ir variando el número de grados de libertad, que se interpretan como el número aproximado de parámetros que corresponderían al modelo. Hemos utilizado validación cruzada para la elección de la cantidad de suavizado aplicada a diferentes transformaciones de las probabilidades brutas. Se observa que en la edad cero el sesgo aumenta mucho a medida que aumentamos el suavizado, de forma que el valor óptimo se alcanza para un valor que proporciona un ajuste con muchas fluctuaciones. Por esta razón que hemos aplicado también el cálculo del smoothing dejando fuera la observación en cero, la comparación podemos observarla en la Figuras 6.11 y 6.12, en las que observamos que el suavizado resultante sin tener en cuenta el cero proporciona una gráfica razonablemente suave. Los residuos resultantes de los ajustes frente a las estimaciones se presentan a continuación en las Figuras

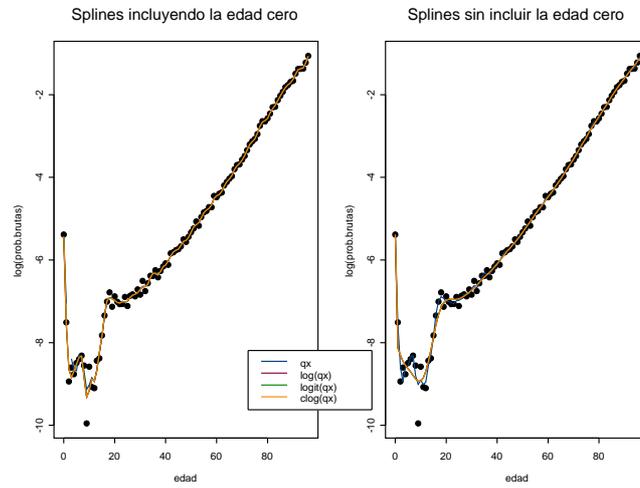


Figura 6.11: Smoothing mediante splines para los hombres

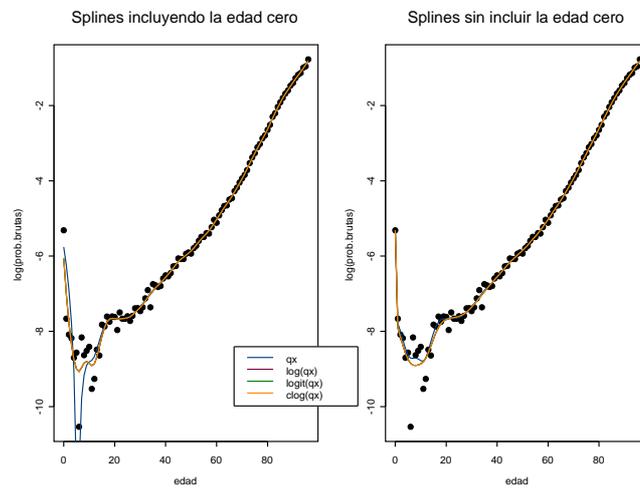


Figura 6.12: Smoothing mediante splines para las mujeres

6.13 y 6.14 con el fin de observar la corrección de la heterocedasticidad. El ajuste, según

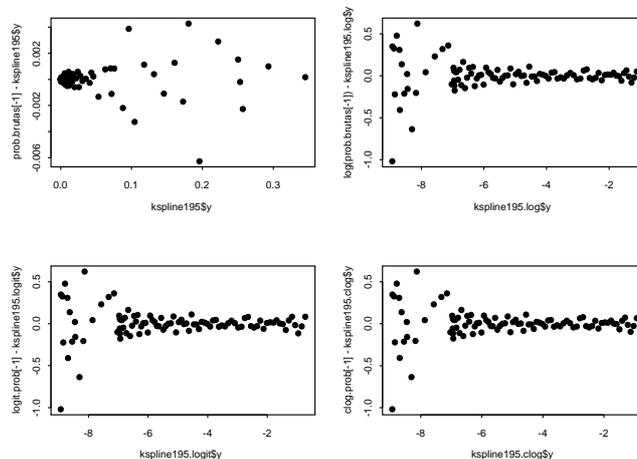


Figura 6.13: Residuos frente a estimación de las transformaciones de q_x para los hombres

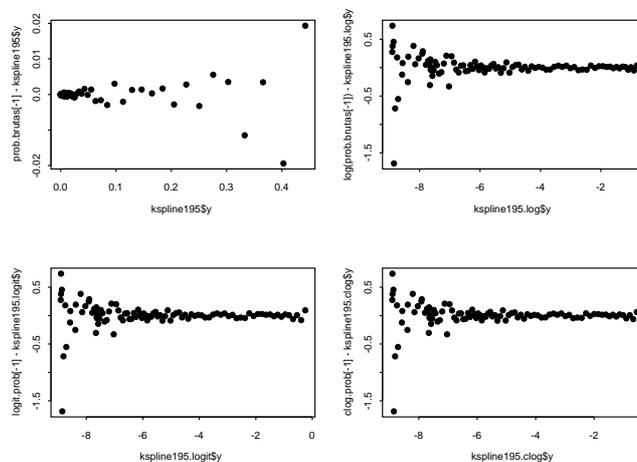


Figura 6.14: Residuos frente a estimación de las transformaciones de q_x para las mujeres

puede observarse en las gráficas, es muy parecido para las diferentes transformaciones, entre las cuales elegiremos de acuerdo con las medidas de bondad del ajuste que tenemos en la Tabla 6.5. Aunque las diferencias son mínimas, la transformación logarítmica es la que mejores resultados de bondad de ajuste presenta. Hemos analizado las desviaciones y las desviaciones relativas, Figuras 6.15 y 6.16. También hemos realizado los test no

HOMBRES				
con el cero	q_x	log(q_x)	logit(q_x)	clog(q_x)
Deviance	NA	41.1925	41.125	41.16039
log-verosim.	NA	-169372.7	-169372.7	-169372.7
χ^2	NA	41.65626	41.58871	41.6241
Df	57.67209	52.53577	52.59802	52.56444
sin el cero	q_x	log(q_x)	logit(q_x)	clog(q_x)
Deviance	40.07903	116.9773	117.1566	117.0651
log-verosim.	-169372.2	-169410.6	-169410.7	-169410.7
χ^2	39.53347	119.8757	120.0578	119.9645
Df	56.7864	17.78958	17.76874	17.77948

MUJERES				
con el cero	q_x	log(q_x)	logit(q_x)	clog(q_x)
Deviance	235.7353	160.4945	160.7995	160.6622
log-verosim.	-190156	-190118	-190118.6	-190118.5
χ^2	391.5805	186	186.2121	186.1318
Df	22.80457	24.98834	24.98567	24.98643
sin el cero	q_x	log(q_x)	logit(q_x)	clog(q_x)
Deviance	69.75433	88.70336	89.68427	89.15601
log-verosim.	-190074.5	-190082.6	-190083.1	-190082.8
χ^2	69.75433	88.70336	89.68427	89.15601
Df	22.53969	15.92764	15.91088	15.916

Tabla 6.5: Medidas de bondad de ajuste para el ajuste mediante splines sobre las transformadas de las probabilidades brutas

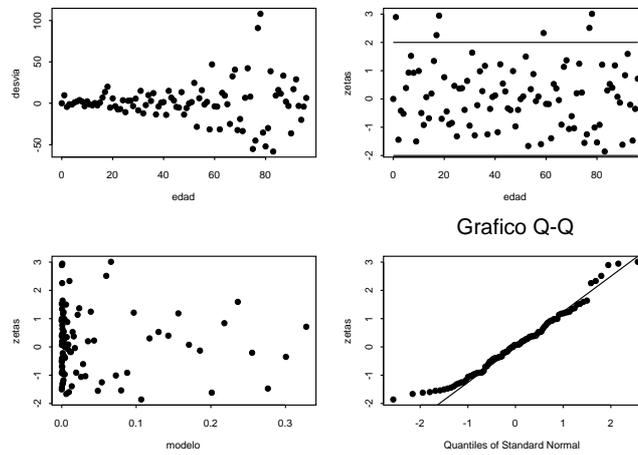


Figura 6.15: Transformación log para los hombres

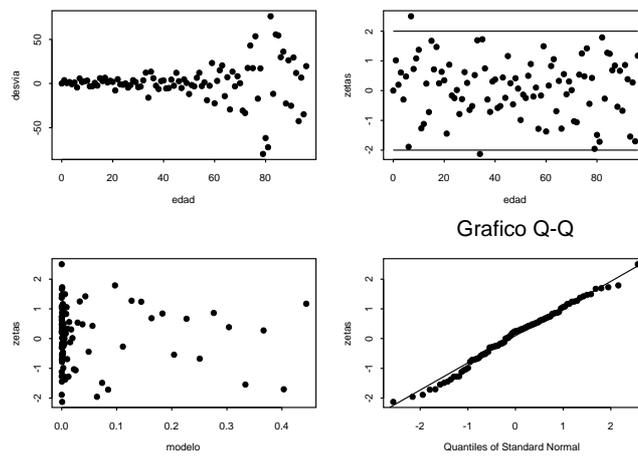


Figura 6.16: Transformación log para las mujeres

paramétricos cuyos resultados se presentaran más adelante con el fin de comparar todos los ajustes no paramétricos. Este método tiene una ventaja frente al kernel porque nos permite obtener predicciones para valores elevados de la edad, prolongando la función spline en el último tramo. Esto se consigue en el S-plus a partir del comando *predict*.

6.4. Graduación polinomial local

Una alternativa prometedora según Verrall (1996) es la regresión local ponderada. Hemos descrito este método en la sección 3.4 y como allí decíamos se encuentra implementado en S-plus con el comando *loess*. Las posibilidades de este suavizado son varias al permitir actuar sobre su parámetro *span*, que representa la proporción del número de elementos de la vecindad frente al número total de puntos, y sobre el grado del polinomio local utilizado, que puede ser lineal o cuadrático. En primer lugar vamos a probar las distintas posibilidades y después utilizaremos validación cruzada para elegir los parámetros.

Comenzamos nuestro trabajo en esta sección graduando las probabilidades de muerte, q_x , según diferentes elecciones del *span*, cuya comparación gráfica muestran las Figuras 6.17 y 6.18.

Los resultados para ambos sexos se recogen en la Tabla 6.6 donde presentamos sus principales medidas de bondad de ajuste. Se observa que es difícil la elección de los parámetros puesto que a medida que aumenta el suavizado mejora el ajuste, pero sin embargo se observa gráficamente que los valores pequeños del *span* proporcionan funciones con oscilaciones en edades bajas.

Además, como en el ajuste anterior, las gráficas de los residuos de los ajustes frente a los valores estimados muestran existencia de heterocedasticidad, por lo que se hacen necesarias las transformaciones sobre las probabilidades brutas. En las Tablas 6.7 y 6.8 mostramos la comparación de los estimadores *loess* realizados sobre las probabilidades brutas transformadas de hombres y mujeres, respectivamente. Según las Tablas 6.7 y 6.8 los resultados en cuanto a las medidas de bondad de ajuste empeoran respecto a los obtenidos con los valores de q_x sin transformar, resultando un poco mejor la transformación *logit*. No obstante, estas transformaciones se han realizado con el fin de eliminar la heterocedasticidad, hecho que puede observarse representando los residuos frente a las estimaciones de los valores transformados según el *logit* (Figuras 6.19 y 6.20). Los resultados muestran que la situación ha mejorado, se aprecia que los residuos se estabilizan a medida que aumenta la estimación, aunque se observan desviaciones bastante elevadas en las estimaciones para edades tempranas.

También se observa gráficamente en las Figuras 6.21 y 6.22 una notable mejoría respecto a la suavidad del ajuste resultante que ahora no muestra los picos que se observaban al ajustar las probabilidades brutas. Para la comparación final con los diferentes ajustes no paramétricos la regresión polinomial local la realizaremos sobre la transformación *logit*. Elegimos el valor de sus parámetros, grado y *span*, por validación cruzada sin tener en cuenta el cero por las razones anteriormente expuestas. En las Figuras 6.23 y 6.24 puede observarse la comparación entre los resultados de la regresión polinomial local lineal y cuadrática. La solución cuadrática en los hombres forma una curva en las edades cercanas

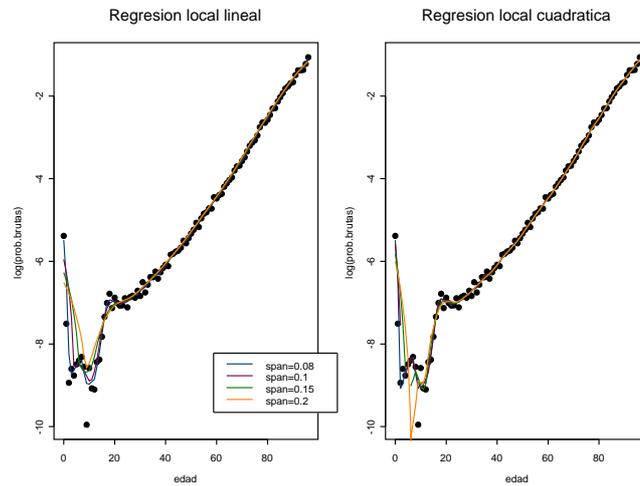


Figura 6.17: Comparación de la regresión local polinomial para los hombres

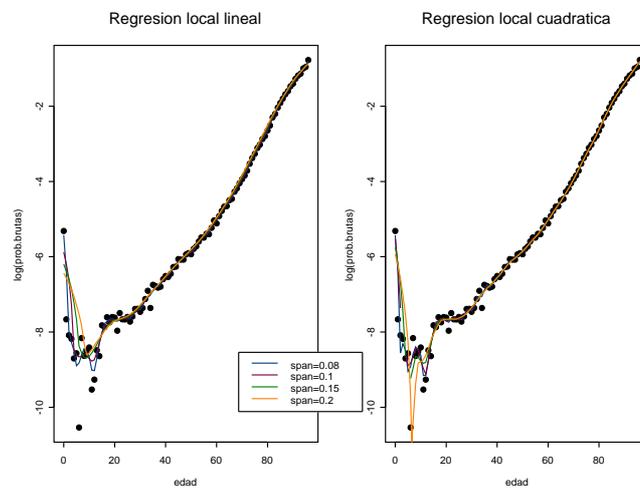


Figura 6.18: Comparación de la regresión local polinomial para las mujeres

HOMBRES				
Lineal	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	93.14998	283.3013	432.7743	575.9492
log-verosim.	-169398.7	-169493.8	-169568.5	-169640.1
χ^2	83.02004	248.4826	416.9678	604.1553
Df	36.8	18.2	11	8.7
Cuadrático	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	77.62113	NA	NA	425.5066
log-verosim.	-169390.9	NA	NA	-169564.9
χ^2	69.51507	NA	NA	3299.606
Df	47.9	35.2	20.2	15.5

MUJERES				
Lineal	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	100.8607	254.9217	480.9268	785.1758
log-verosim.	-190088.6	-190165.6	-190278.6	-190430.8
χ^2	85.64981	225.2538	469.2766	802.7191
Df	36.8	18.2	11	8.7
Cuadrático	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	95.40759	NA	NA	343.3059
log-verosim.	-190085.9	NA	NA	-190209.8
χ^2	82.13673	NA	NA	969.7938
Df	47.9	35.2	20.2	15.5

Tabla 6.6: Medidas de bondad de ajuste para loess sobre las probabilidades brutas

HOMBRES				
Lineal				
log(q_x)	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	56.62978	301.8783	473.7375	595.4351
log-verosim.	-169380.4	-169503.1	-169589	-169649.8
χ^2	56.65942	418.8795	840.0031	1273.735
Df	36.8	18.2	11	8.7
logit(q_x)	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	56.64018	301.7796	472.7916	593.6001
log-verosim.	-169380.4	-169503	-169588.5	-169648.9
χ^2	56.67367	418.3344	838.1498	1270.643
Df	36.8	18.2	11	8.7
clog(q_x)	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	56.62826	301.8193	473.1494	594.1681
log-verosim.	-169425.5	-169503	-169588.7	-169649.2
χ^2	56.66477	418.5967	838.9591	1271.836
Df	36.8	18.2	11	8.7
Cuadrático				
log(q_x)	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	46.1665	80.20533	243.12	319.2327
log-verosim.	-169375.2	-169392.2	-169473.7	-169511.7
χ^2	46.57409	81.73283	301.0464	441.6318
Df	47.9	35.2	20.2	15.5
logit(q_x)	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	46.17714	80.20717	243.0492	318.9744
log-verosim.	-169375.2	-169392.2	-169473.6	-169511.6
χ^2	46.58416	81.72405	300.7145	440.993
Df	47.9	35.2	20.2	15.5
clog(q_x)	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	46.17089	80.20361	243.0852	319.1001
log-verosim.	-169419.5	-169392.2	-169473.7	-169511.7
χ^2	46.57819	81.72579	300.8809	441.3089
Df	47.9	35.2	20.2	15.5

Tabla 6.7: Medidas de bondad de ajuste para loess sobre las transformadas de las probabilidades brutas correspondientes a los hombres

MUJERES				
Lineal				
log(q_x)	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	69.22913	214.0445	408.0304	560.6064
log-verosim.	-190072.8	-190145.2	-190242.2	-190318.5
χ^2	70.92847	266.328	701.9709	1200.237
Df	36.8	18.2	11	8.7
logit(q_x)	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	69.41955	214.5479	404.4701	546.0535
log-verosim.	-190072.9	-190145.4	-190240.4	-190311.2
χ^2	71.0986	266.6114	697.5935	1184.238
Df	36.8	18.2	11	8.7
clog(q_x)	span=0.05	span=0.1	span=0.15	span=0.2
Deviance	69.32244	214.2047	405.2559	550.0049
log-verosim.	-190244.7	-190145.3	-190240.8	-190313.2
χ^2	71.01146	266.3776	698.7801	1188.879
Df	36.8	18.2	11	8.7
Cuadrático				
log(q_x)	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	66.41704	106.2099	152.4717	249.9436
log-verosim.	-190071.4	-190091.3	-190114.4	-190163.1
χ^2	69.04315	112.8767	169.6698	328.7343
Df	47.9	35.2	20.2	15.5
logit(q_x)	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	66.6862	106.2823	152.7888	250.0574
log-verosim.	-190071.5	-190091.3	-190114.6	-190163.2
χ^2	69.29907	112.8974	169.9057	328.6001
Df	47.9	35.2	20.2	15.5
clog(q_x)	span=0.08	span=0.1	span=0.15	span=0.2
Deviance	66.55736	106.2476	152.6213	249.9395
log-verosim.	-190242.3	-190091.3	-190114.5	-190163.1
χ^2	69.1768	112.8885	169.7786	328.606
Df	47.9	35.2	20.2	15.5

Tabla 6.8: Medidas de bondad de ajuste para loess sobre las transformadas de las probabilidades brutas correspondientes a las mujeres

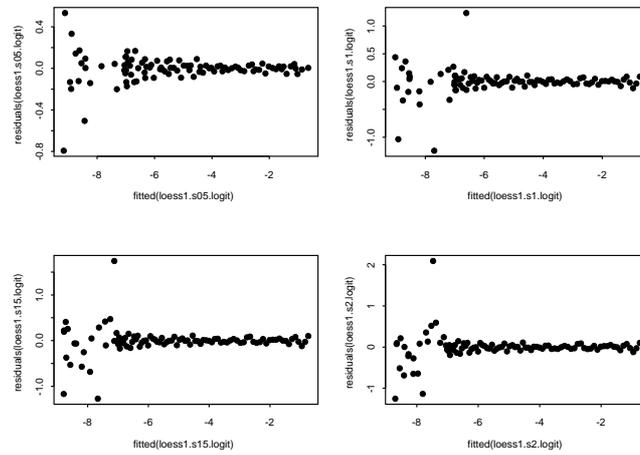


Figura 6.19: Residuos frente a estimación del $\text{logit}(q_x)$ para los hombres

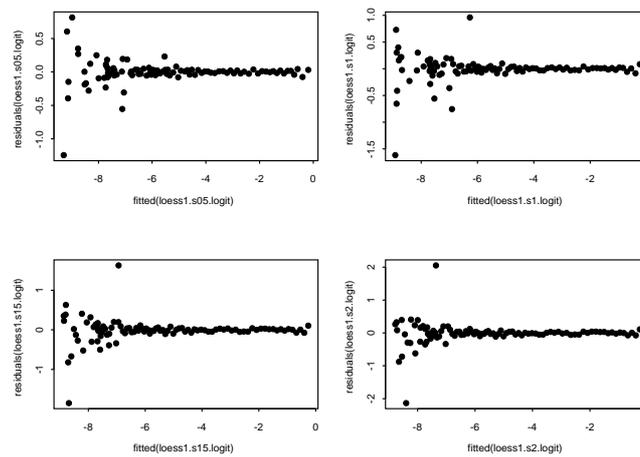


Figura 6.20: Residuos frente a estimación del $\text{logit}(q_x)$ para las mujeres

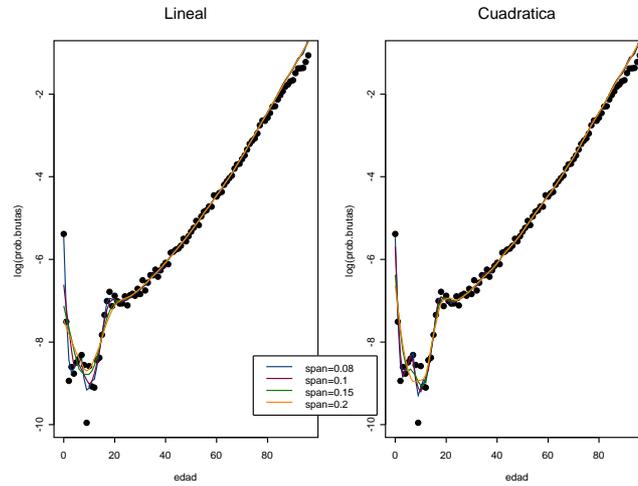


Figura 6.21: Comparación regresión polinomial local sobre la transformación logit para los hombres

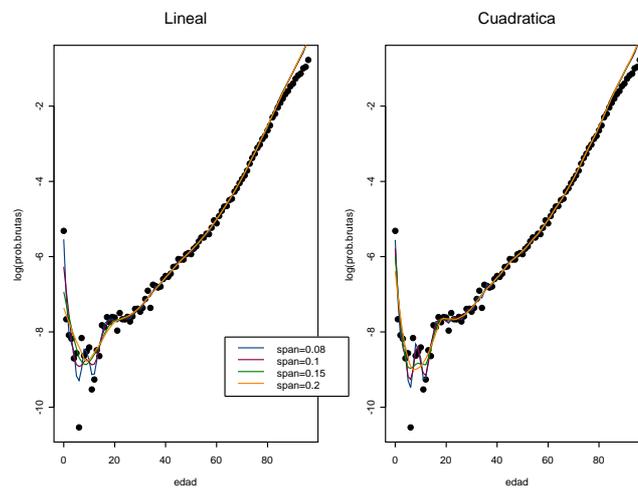


Figura 6.22: Comparación regresión polinomial local sobre la transformación logit para las mujeres

a cero difícil de explicar, por ello decidimos elegir la lineal. Sin embargo en las mujeres aparece como mucho mejor la cuadrática por su suavidad y menor número de parámetros. En la Tabla 6.9 se recogen los resultados para ambas regresiones locales.

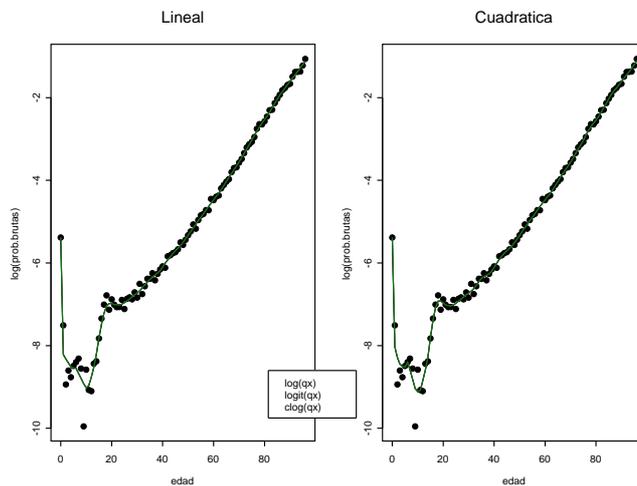


Figura 6.23: Comparación loess lineal y cuadrática con span elegido por validación cruzada para los hombres

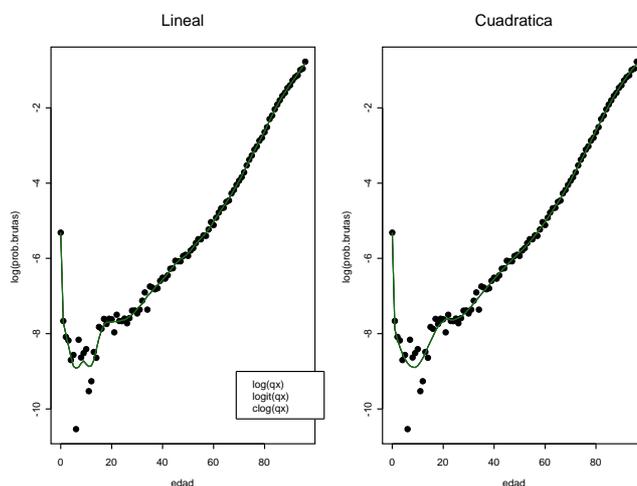


Figura 6.24: Comparación loess lineal y cuadrática con span elegido por validación cruzada para las mujeres

Para comparar con el resto de ajustes hemos realizado también los correspondientes tests no paramétricos.

	HOMBRES		MUJERES	
	Lineal	Cuadrático	Lineal	Cuadrático
span	0.10	0.12	0.10	0.26
Deviance	108.4945	87.83237	73.38883	100.3294
log-verosim.	-169406.4	-169396	-190074.9	-190088.3
χ^2	110.7888	88.36024	72.86834	99.40046
Df	18	27.2	18	11.9

Tabla 6.9: Bandwidth y medidas de bondad de ajuste para loess sobre las transformadas logit de las probabilidades brutas

6.5. Modelos aditivos generalizados

Vamos a reformular dos de los métodos de graduación no paramétricos, splines y regresión local polinomial como modelos aditivos generalizados (GAM). Los GAM mejoran la formulación que ya tenían en la teoría actuarial estos dos métodos no paramétricos utilizando la misma estructura que los GLM y ajustando a la transformación de las probabilidades brutas (link) un predictor cuyas variables son los splines o loess de las variables explicativas, en nuestro caso la edad. Estos modelos aportan una mejora respecto a las otras técnicas no paramétricas porque permiten utilizar la distribución exacta binomial.

6.5.1. GAM con splines

En nuestro trabajo hemos comenzado utilizando modelos aditivos generalizados realizando ajustes para funciones splines con grados de libertad desde 10 hasta 24. Las Tablas 6.10 y 6.11 recogen los resultados de estos ajustes para hombres y mujeres y para un rango de edades de 1 a 96 años, pues como en los otros contrastes no paramétricos el cero influía excesivamente. En las Tablas 6.10 y 6.11 el modelo NULL consiste en ajustar una constante a los $\text{logit}(q_x)$ y la log-verosimilitud que aparece en su misma columna es la correspondiente a las probabilidades brutas o modelo saturado. Como era de esperar, se observa una mejora paulatina de la bondad del ajuste a medida que aumenta el número de parámetros, es decir con menos suavizado. Pero necesitamos contrastar si la disminución de la Deviance es significativa y si el estadístico C_p indica una mejora del ajuste en relación al incremento de complejidad del modelo como en el capítulo anterior hacíamos con los modelos lineales generalizados. El resultado de ambos contrastes se recoge en la Tabla 6.12, en la que la Deviance representa ahora la disminución de la Deviance de un modelo respecto del inmediato anterior, puesto que se trata de comparar modelos. No se deduce de la Tabla un modelo adecuado que sirva simultáneamente para ambos sexos teniendo en cuenta el incremento significativo de la χ^2 y el valor de C_p . Por lo que hemos realizado contrastes no paramétricos para grados de 16 a 21, en el caso de las mujeres tomaríamos el modelo con $df = 19$ pues no mejoraban los resultados después de este valor según se muestra en la Tabla 6.14. En el caso de los hombres teníamos una situación similar a partir de $df = 21$ lo que puede observarse en la Tabla 6.13. No ha sido necesario agrupar

HOMBRES	Null	df=10	df=11	df=12
Deviance	102104.3	190.8475	184.2989	177.2764
g.l.	95	85.05043	84.04519	83.03993
log-verosim.	-169352.1	-169447.5	-169444.3	-169440.8
χ^2		198.7128	191.8373	184.0873
	df=13	df=14	df=15	df=16
Deviance	169.9128	162.1247	154.4828	146.8157
g.l.	82.04274	81.02351	80.03405	79.02981
log-verosim.	-169437.1	-169433.2	-169429.4	-169425.5
χ^2	175.7408	166.8283	158.0983	149.4128
	df=17	df=18	df=19	df=20
Deviance	139.2888	132.2492	125.7667	119.5375
g.l.	78.01231	77.01376	76.00902	75.00502
log-verosim.	-169421.8	-169418.2	-169415	-169411.9
χ^2	140.9874	133.2119	125.6355	118.9816
	df=21	df=22	df=23	df=24
Deviance	113.8336	108.5536	103.6553	99.1255
g.l.	74.01156	73.01474	72.01108	71.00440
log-verosim.	-169409	-169406.4	-169403.9	-169401.7
χ^2	112.9597	107.4447	102.378	97.73242

Tabla 6.10: Medidas de bondad de ajuste para los modelos GAM con splines para hombres

MUJERES	Null	df=10	df=11	df=12
Deviance	172750.6	116.3650	111.9010	107.5807
g.l.	95	85.05043	84.04519	83.03993
log-verosim.	-190038.2	-190096.3	-190094.1	-190092
χ^2		113.3006	108.3498	103.6425
	df=13	df=14	df=15	df=16
Deviance	103.3710	99.3890	95.5273	91.8164
g.l.	82.04274	81.02351	80.03405	79.02981
log-verosim.	-190089.8	-190087.9	-190085.9	-190084.1
χ^2	99.11696	94.88754	90.83509	86.99044
	df=17	df=18	df=19	df=20
Deviance	88.3967	85.2008	82.2687	79.6000
g.l.	78.01231	77.01376	76.00902	75.00502
log-verosim.	-190082.4	-190080.8	-190079.3	-190078
χ^2	83.4943	80.27152	77.35625	74.74101
	df=21	df=22	df=23	df=24
Deviance	77.2114	74.9974	73.0595	71.3439
g.l.	74.01156	73.01474	72.01108	71.00440
log-verosim.	-190076.8	-190075.7	-190074.7	-190073.8
χ^2	72.43442	70.32838	68.5125	66.92777

Tabla 6.11: Medidas de bondad de ajuste para los modelos GAM con splines para mujeres

HOMBRES	Deviance	p-value	C _p
df=10			220.9901
df=11	6.548663	0.01058940	217.2087
df=12	7.022484	0.00812207	212.9536
df=13	7.363618	0.00662331	208.3351
df=14	7.788055	0.00543903	203.3528
df=15	7.641898	0.00559812	198.4348
df=16	7.667124	0.00566549	193.5322
df=17	7.526931	0.00626556	188.8062
df=18	7.039553	0.00795313	184.5156
df=19	6.482518	0.01098105	180.7989
df=20	6.229185	0.01265020	177.3336
df=21	5.703929	0.01674692	174.3645
df=22	5.280008	0.02146202	171.8286
df=23	4.898225	0.02703913	169.6934
df=24	4.529877	0.03365094	167.9347

MUJERES	Deviance	p-value	C _p
df=10			136.9567
df=11	4.464024	0.0351068	134.3942
df=12	4.320270	0.0377528	131.9611
df=13	4.209749	0.0406561	129.6499
df=14	3.981954	0.0453137	127.5334
df=15	3.861722	0.0496853	125.5632
df=16	3.710849	0.0554345	123.7691
df=17	3.419737	0.0642346	122.2298
df=18	3.195889	0.0741188	120.9234
df=19	2.932118	0.0868248	119.8753
df=20	2.668716	0.1021578	119.0883
df=21	2.388593	0.1203971	118.5621
df=22	2.213953	0.1404014	118.2715
df=23	1.937877	0.1639100	118.2179
df=24	1.715655	0.1880832	118.3683

Tabla 6.12: Comparación de los modelos GAM con diferentes grados de libertad

HOMBRES									
df	χ^2	$z > 2$	$z > 3$	pos	p(pos)	rachas	p(rachas)	K-S	p(K-S)
16	149.01	6	2	46	0.3424	45	0.4407	0.0619	0.993
17	140.68	5	2	46	0.3424	45	0.4407	0.0619	0.993
18	132.98	5	2	45	0.2713	47	0.4770	0.0619	0.993
19	125.57	5	2	43	0.1550	45	0.4472	0.0619	0.993
20	118.98	5	2	44	0.2084	45	0.4444	0.0619	0.993
21	112.99	3	2	45	0.2713	49	0.5119	0.0619	0.993
22	107.50	3	1	45	0.2713	49	0.5119	0.0619	0.993

Tabla 6.13: Contrastes no paramétricos del modelo GAM con splines para los hombres

MUJERES									
df	χ^2	$z > 2$	$z > 3$	pos	p(pos)	rachas	p(rachas)	K-S	p(K-S)
16	85.56	5	0	46	0.3424	55	0.6128	0.0825	0.899
17	82.25	4	0	47	0.4196	57	0.6444	0.0825	0.899
18	79.19	4	0	48	0.5000	57	0.6439	0.0825	0.899
19	76.42	2	0	48	0.5000	57	0.6439	0.0619	0.993
20	73.92	2	0	49	0.5804	57	0.6444	0.0619	0.993
21	71.71	2	0	49	0.5804	57	0.6444	0.0619	0.993
22	69.69	2	0	49	0.5804	57	0.6444	0.0619	0.993

Tabla 6.14: Contrastes no paramétricos del modelo GAM con splines para las mujeres

edades en ninguno de los dos casos luego el total de observaciones es 96 y solo mostramos el número de positivas.

Las Figuras 6.25 y 6.26 recogen la comparación gráfica de los modelos para cada sexo. Para poder compararlas con el resto de gráficas, se presentan ambas en escala log aun cuando el ajuste se ha realizado para la transformación logit. Para las edades extremas inferiores puede observarse una mayor distancia entre el modelo y los datos observados.

6.5.2. GAM con loess

A continuación consideramos modelos aditivos generalizados con regresión local polinomial con parámetro span inicial de 0.05 para la lineal y 0.08 para la cuadrática. La Tabla 6.15 recoge los resultados de estos ajustes para hombres y mujeres y para un rango de edades de 1 a 96 años, pues como en los anteriores, el cero influía excesivamente. En las Tablas 6.15 no aparece el modelo NULL y la log-verosimilitud del modelo saturado pues es el mismo que en las Tablas 6.10 y 6.11. Como en todos los casos hay un empeoramiento de la bondad del ajuste a medida que aumenta el parámetro de suavizado y consecuentemente aumentan el número grados de libertad residuales. Pero necesitamos contrastar si la disminución de la Deviance es significativa y si el estadístico C_p indica

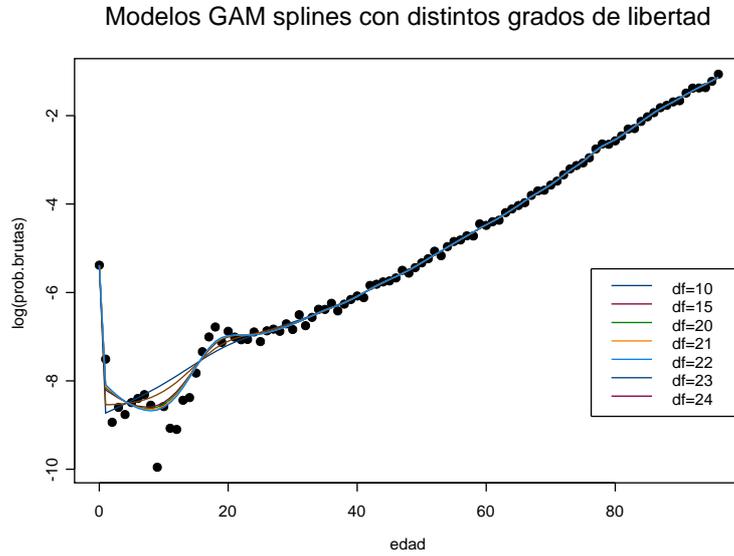


Figura 6.25: Comparación de los GAM con splines para los hombres

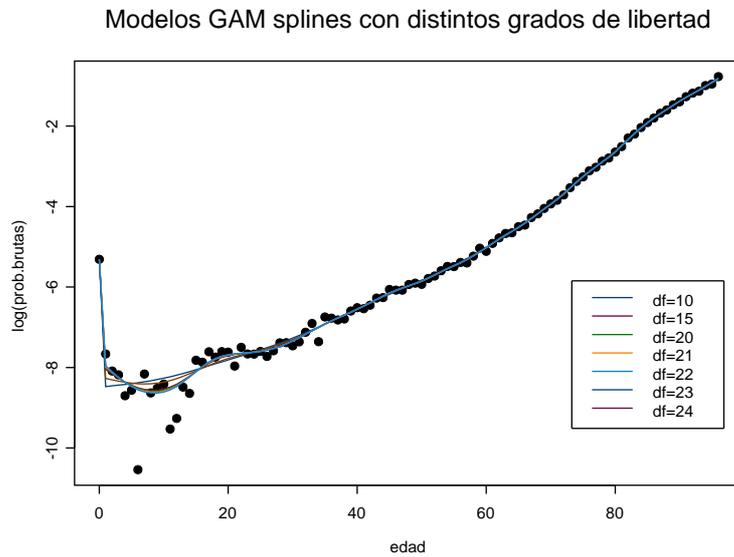


Figura 6.26: Comparación de los GAM con splines para las mujeres

HOMBRES									
Lineal	span=0.05	span=0.08	span=0.09	span=0.1	span=0.15	span=0.2			
Deviance	54.76074	84.0133	102.0081	103.3721	126.6746	145.9003			
log-verosim.	-169379.5	-169394.1	-169403.1	-169403.8	-169415.5	-169425.1			
χ^2	53.81335	82.24093	100.4185	101.924	126.2677	148.6677			
g.l.	53.85111	67.64111	74.21044	74.43913	82.95368	85.66408			
Cuadrático	span=0.08	span=0.1	span=0.13	span=0.14	span=0.15	span=0.2			
Deviance	45.49009	64.6869	96.7787	97.93632	104.6561	116.3091			
log-verosim.	-169410.3	-169404.4	-169400.5	-169401.1	-169384.5	-169374.9			
χ^2	44.91986	63.52954	95.29789	96.4964	103.2891	115.4633			
g.l.	44.77002	58.06511	70.73423	70.98848	74.24463	79.46306			
MUJERES									
Lineal	span=0.05	span=0.08	span=0.09	span=0.1	span=0.15	span=0.2			
Deviance	54.76074	84.0133	102.0081	103.3721	126.6746	145.9003			
log-verosim.	-169379.5	-169394.1	-169403.1	-169403.8	-169415.5	-169425.1			
χ^2	53.81335	82.24093	100.4185	101.924	126.2677	148.6677			
g.l.	53.85111	67.64111	74.21044	74.43913	82.95368	85.66408			
Cuadrático	span=0.08	span=0.1	span=0.13	span=0.14	span=0.15	span=0.2			
Deviance	45.49009	64.6869	96.7787	97.93632	104.6561	116.3091			
log-verosim.	-169410.3	-169404.4	-169400.5	-169401.1	-169384.5	-169374.9			
χ^2	44.91986	63.52954	95.29789	96.4964	103.2891	115.4633			
g.l.	44.77002	58.06511	70.73423	70.98848	74.24463	79.46306			

Tabla 6.15: Medidas de bondad de ajuste para GAM con loess para cada sexo

una mejora del ajuste en relación al incremento de complejidad del modelo, como en el capítulo anterior hacíamos con los modelos lineales generalizados. El resultado de ambos contrastes se recoge en la Tabla 6.16.

En el caso lineal no se deduce de la Tabla 6.16 un modelo adecuado para los hombres pues mejora a medida que el suavizado disminuye, no obstante observando el gráfico vemos que el menor span produce ondulaciones en las edades inferiores por lo que no es adecuado, siendo el span 0.1 el menor de los que no produce estas oscilaciones. En el caso de las mujeres elegiríamos los modelos de span 0.1 o 0.09 de acuerdo con el valor del estadístico C_p . Para los modelos cuadráticos tomamos para ambos sexos el span 0.15 pues aunque en los hombres parece adecuado 0.1, gráficamente parece mejor el anterior. Las Figuras 6.27 y 6.28 recogen la comparación gráfica de los modelos para cada sexo. Se presentan ambas en escala log por la razón antes expuesta. Para las edades extremas inferiores puede observarse una mayor distancia entre el modelo y los datos observados. Realizamos los contrastes no paramétricos sobre los modelos elegidos, obteniendo resultados muy similares para el lineal con span=0.1 y cuadrático con span=0.15. No obstante, presentan autocorrelaciones ligeramente significativas. A este respecto los GAM con splines aunque empeoran un poco la bondad de ajuste puesto que tienen más grados de libertad residuales, superan el test de autocorrelaciones para la mujeres y las autocorrelaciones en los hombres son casi inapreciables. A consecuencia de estos motivos se utilizarán los modelos GAM con splines en la comparación con el resto de ajustes no paramétricos en la siguiente sección.

6.6. Comparación de los diferentes métodos

Al igual que en el caso de los modelos paramétricos hemos aplicado los tests descritos en Forfar, McCutcheon y Wilkie (1988). En los métodos no paramétricos, a diferencia de los paramétricos, no hemos necesitado en ningún caso agrupar las edades hasta tener un número de muertes esperadas superior a 5. Hemos obtenido los valores del error porcentual absoluto medio (EPAM) y R^2 que utilizan Felipe y Guillén en su trabajo, y que nos pueden servir para comparar con sus ajustes.

De los tests cuyos resultados se presentan en la Tabla 6.17 podemos deducir que los modelos kernel y splines proporcionan un número razonable y similar de desviaciones superiores a 2 y a 3. Para las mujeres el estadístico χ^2 para el modelo de splines resulta no significativo lo que indica mejor ajuste que para los hombres. En el caso de los modelos kernel el número de desviaciones superiores a 3 es 1 frente a 0 en los splines para las mujeres y resulta un valor del estadístico χ^2 menor. En el caso del contraste para los signos los splines para el caso de las mujeres están en el límite de proporcionar demasiadas predicciones positivas, el kernel da una probabilidad ligeramente menor proporcionando pues mejores resultados en este sentido. La R^2 sin embargo indica mejores resultados para los splines y el EPAM es mejor para el kernel. Los tests correspondientes a las autocorrelaciones de las desviaciones relativas pueden observarse en las Figuras 6.29, 6.30, 6.31 y 6.32 que muestran la no existencia de autocorrelaciones para ambos modelos en los dos sexos.

La Tabla 6.17 muestra los resultados para la regresión local polinomial, la lineal en el

HOMBRES				
Lineal	Deviance	g.l.	p-value	C_p
0.2				166.5577
0.15	19.22576	2.71040	0.000173560	152.7489
0.1	23.30245	8.51454	0.00414878	146.4636
0.09	1.36401	0.22870	0.04772252	145.5567
0.08	17.96677	6.56933	0.00925169	140.7193
0.05	29.28059	13.790000	0.00869744	138.9994
Cuadrático	Deviance	g.l.	p-value	C_p
0.2				149.4937
0.15	11.65302	5.21843	0.0455145	148.3125
0.14	6.71976	3.25615	0.0971300	148.1268
0.13	1.15762	0.25424	0.0645897	147.4793
0.1	32.09180	12.66913	0.0019479	140.8106
0.05	19.19680	13.29509	0.1283133	148.2930

MUJERES				
Lineal	Deviance	g.l.	p-value	C_p
0.2			130.2318	
0.15	14.85992	2.66092	0.0013413	120.3797
0.1	22.89332	8.47831	0.0047259	113.4424
0.09	0.99774	0.22988	0.0678963	112.8773
0.08	6.82684	6.59411	0.4022416	118.4604
0.05	12.58668	13.80488	0.5441139	131.8542
Cuadrático	Deviance	g.l.	p-value	C_p
0.2			119.2488	
0.15	13.88188	5.18788	0.0186139	115.9776
0.14	4.38888	3.25499	0.2549880	118.2461
0.13	0.61014	0.25117	0.1138081	118.1497
0.1	9.39783	12.69337	0.7208414	134.7134
0.05	9.49391	13.29026	0.7542906	152.4018

Tabla 6.16: Comparación de los modelos GAM con diferentes span

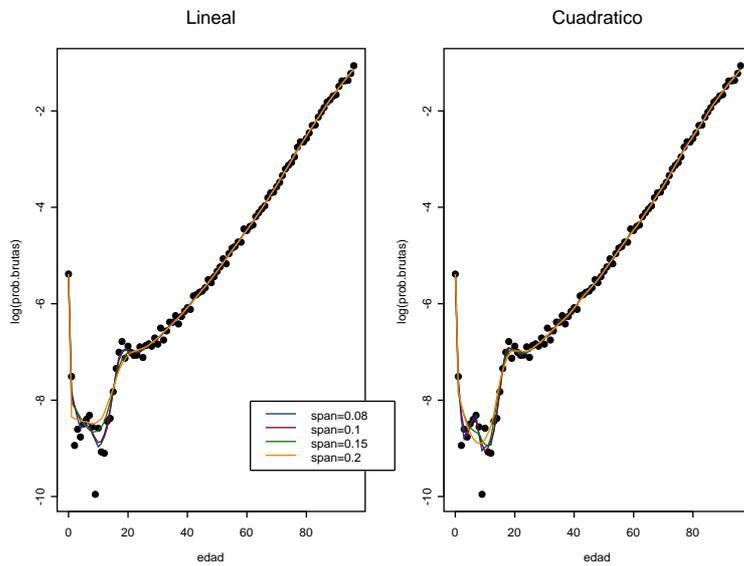


Figura 6.27: Comparación de los GAM con loess para los hombres

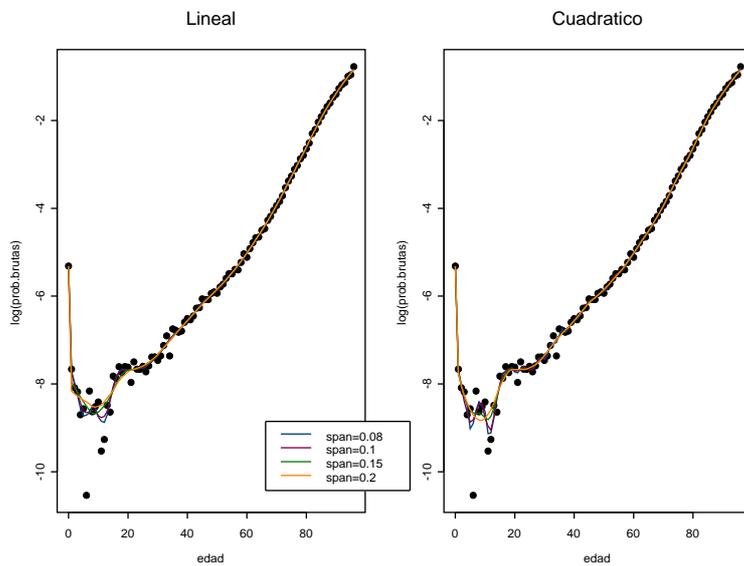


Figura 6.28: Comparación de los GAM con loess para las mujeres

	Kernel		Splines		LOESS		GAM			
	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres		
Desv. relativas	> 2 1	> 3 1	52 (44) 0.7916	54 (42) 0.8886	50 (46) 0.6576	56 (40) 0.9401	49 (47) 0.5804	53 (43) 0.8450	45 (51) 0.2713	49 (47) 0.5804
Test de signos	pos. (neg.) p-value		51 0.5495	57 0.6607	51 0.5495	57 0.6741	49 0.5009	57 0.6555	49 0.5119	57 0.6444
Test de rachas	rachas p-value		0.0412	0.0515	0.0412	0.0515	0.0412	0.0515	0.0619	0.0619
Test de K-S	K-S p-value		1	0.9996	1	0.9996	1	0.9996	0.993	0.993
Test χ^2	χ^2 g.l. p-value		117.98 75.53 0.0011	96.03 72.64 0.03085	119.87 78.21 0.0017	88.70 80.07 0.2385	110.79 78 0.0087	99.40 84.1 0.1219	112.99 74.01 0.0024	71.71 74.05 0.5554
R²			0.9346	0.9247	0.9946	1	0.9960	0.9996	0.9852	0.9910
EPAM			10.93	13.44	10.93	14.12	10.17	14.82	13.94	17.33

Tabla 6.17: Resultados de los contrastes no paramétricos sobre el kernel y splines

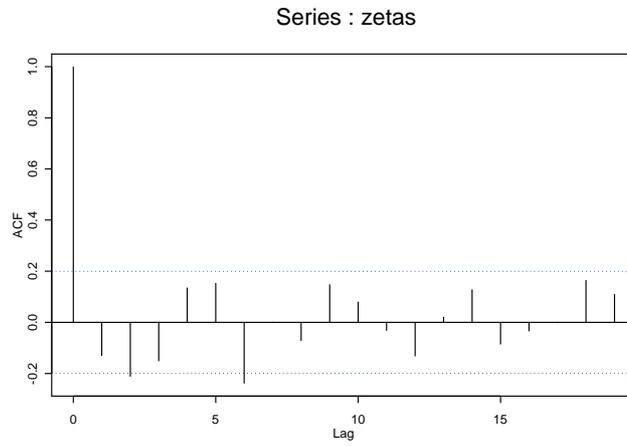


Figura 6.29: Autocorrelaciones del modelo kernel para los hombres

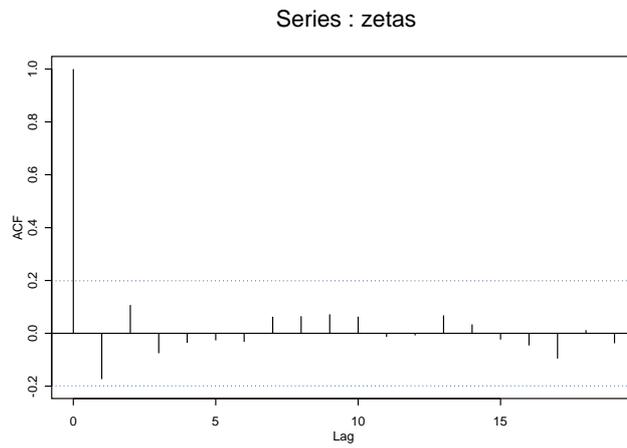


Figura 6.30: Autocorrelaciones del modelo kernel para las mujeres

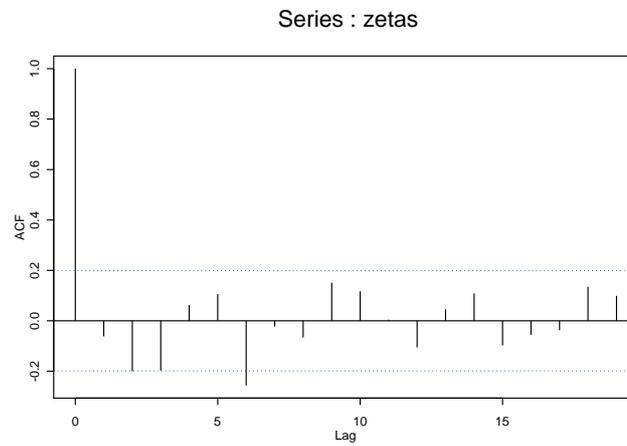


Figura 6.31: Autocorrelaciones del modelo splines para los hombres

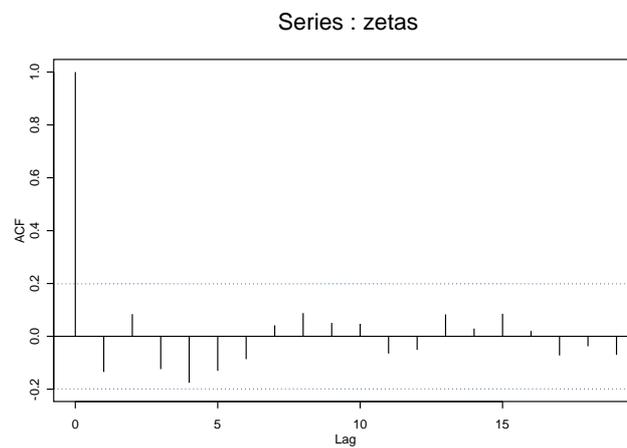


Figura 6.32: Autocorrelaciones del modelo splines para las mujeres

caso de los hombres y cuadrática para las mujeres, y los resultados para los modelos aditivos generalizados sobre splines que según argumentamos resultaban los más adecuados. De la Tabla 6.17 se deduce que los métodos no paramétricos loess y GAM conducen a valores ajustados que superan todos los test. De la comparación de ambos hay que resaltar que el GAM tiene mejor bondad de ajuste aunque lo consigue con un menor número de grados de libertad. Respecto a los contrastes para las autocorrelaciones, las Figuras 6.33 y 6.34 muestran la posible existencia de algunas significativas para la regresión local polinomial, mientras que la 6.35 y la 6.36 muestran autocorrelaciones menores en significatividad y número para los modelos aditivos generalizados. Como reflexión final a la comparación

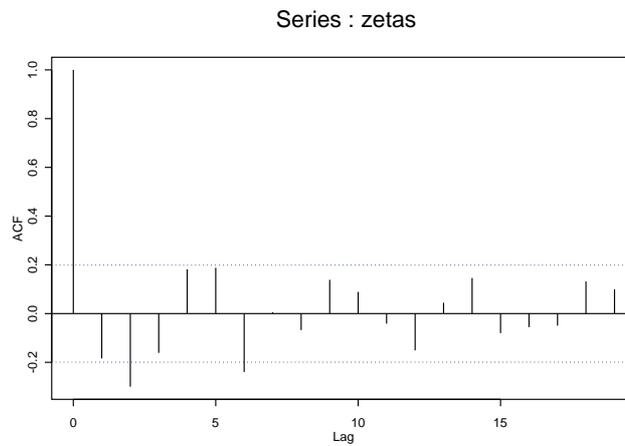


Figura 6.33: Autocorrelaciones del modelo loess lineal para los hombres

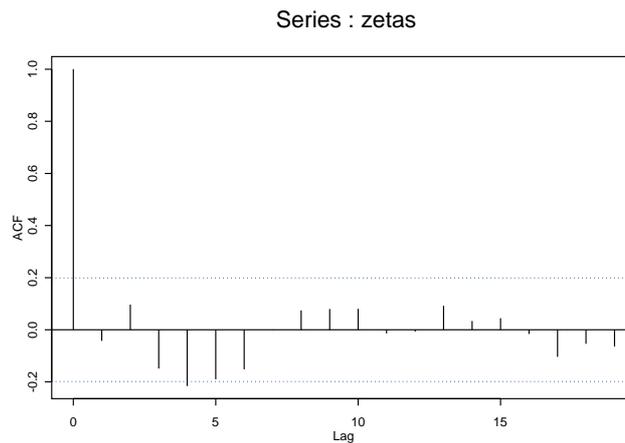


Figura 6.34: Autocorrelaciones del modelo loess cuadrático para las mujeres

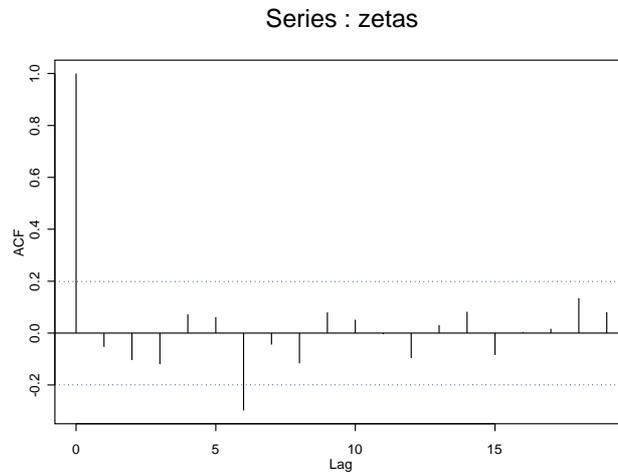


Figura 6.35: Autocorrelaciones del modelo GAM para los hombres

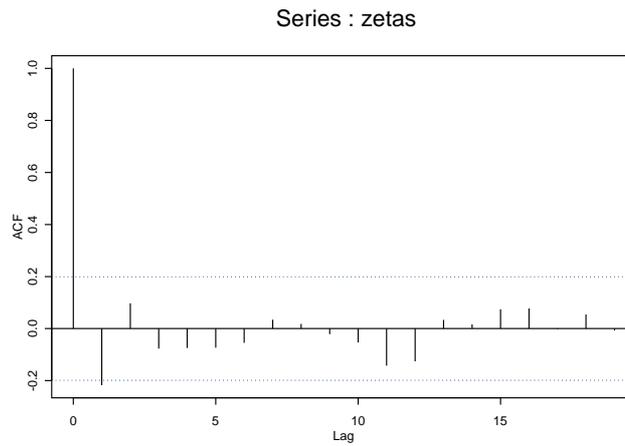


Figura 6.36: Autocorrelaciones del modelo GAM para las mujeres

de todos los métodos no paramétricos, hay que resaltar que todos los modelos presentan muy buenos resultados. La elección de uno de ellos resulta difícil. No obstante conviene recordar que los modelos aditivos generalizados resultan muy satisfactorios para ambos sexos utilizando la verdadera distribución de los datos, la binomial.

Capítulo 7

Tablas de mortalidad dinámicas para la Comunidad Valenciana

7.1. Introducción

En los anteriores capítulos mostramos como aplicar la metodología de obtención de tablas de mortalidad a partir de modelos paramétricos y no paramétricos. No obstante, el uso de dichos modelos basados en la hipótesis de estacionariedad y por tanto la perpetuación del comportamiento de la mortalidad en un periodo largo de tiempo, suele ser frágil, por este motivo presentamos los modelos dinámicos en el Capítulo 4 cuya aplicación a los datos de mortalidad de la Comunidad Valenciana constituye el eje central del presente capítulo. Aunque originalmente los métodos dinámicos se han aplicado a diferentes medidas de mortalidad, en este capítulo hemos decidido aplicarlos a las probabilidades brutas q_x , a fin de facilitar las comparaciones con otros métodos sin necesidad de posteriores transformaciones.

En la sección 2 comenzamos analizando las diferencias existentes entre los distintos años, para ver si resultan significativas y después estudiar su perfil. En la sección 3 aplicamos los métodos estructurales y en la cuarta los no estructurales.

7.2. Evolución de la mortalidad a lo largo del tiempo

Antes de comenzar la obtención de los diferentes ajustes dinámicos, consideramos muy importante poner de relieve las diferencias existentes entre las experiencias de mortalidad correspondientes a diferentes periodos. Los datos que hemos utilizado son las cifras de población y muertes de la Comunidad Valenciana de periodo comprendido entre 1980 y 2000, ambos años incluidos. En el cálculo de las estimaciones brutas de cada uno de los años recurrimos a la expresión del INE (5.1).

En primer lugar hemos utilizado los test clásicos para la comparación de mortalidades propuestos por Forfar, McCutcheon y Wilkie (1988), aplicados a las correspondientes estimaciones brutas de las probabilidades de muerte. Aunque originalmente dichos test se describen para aplicarse sobre las fuerzas de mortalidad, ya han sido utilizados en el

trabajo de Felipe y Guillén (1999) para datos de la población española adaptándolos a las estimaciones brutas de sus probabilidades q_x . Concluyendo, para la población española entre 1975 y 1993, que la evolución que sufre la probabilidad de fallecimiento a lo largo de los años del calendario es de tal intensidad que de un año a otro es significativa, salvo para los años 1988, 1989 y 1990.

Comentaremos brevemente los contrastes aplicados para la comparación, y la notación seguida. Supongamos que tenemos dos experiencias de mortalidad I y II, que en nuestro caso corresponden a dos periodos de tiempo, medidas con sus correspondientes probabilidades de muerte,

$$q_x^I = \frac{D_x^I}{E_x^I}$$

$$q_x^{II} = \frac{D_x^{II}}{E_x^{II}},$$

a partir de las cuales podemos construir la experiencia de mortalidad agrupada según la probabilidad de muerte,

$$q_x = \frac{D_x^I + D_x^{II}}{E_x^I + E_x^{II}}.$$

Si los dos experiencias son similares se parecerán a la resultante de su agrupación y podrán por tanto representarse según esta. La hipótesis nula correspondiente es que las medidas de mortalidad son iguales para las dos experiencias. Los contrastes propuestos por Forfar, McCutcheon y Wilkie (1988) suponen, inicialmente, que los fallecidos en cada edad siguen una distribución normal y requieren que el número esperado de muertes sea mayor que 5 ($D_x^I + D_x^{II} > 5$), si no fuera así se agruparían las edades necesarias para conseguirlo. Estos test ya fueron brevemente descritos en la sección 2.7 pero ahora los aplicamos a las diferencias.

Test de signos

Consiste en aplicar el test de signos a las diferencias $q_x^I - q_x^{II}$, puesto que bajo la hipótesis nula de igualdad entre ambas, el número de diferencias positivas se distribuye $Bi(N, 0.5)$, con N número de edades consideradas. Una alternativa es el test de rango de signos de Wilcoxon, que da más peso a los pares que tienen una diferencia de magnitud más elevada.

Test de rachas

Consiste en aplicar el test de rachas a las diferencias anteriores, puesto que si ambas mortalidades son semejantes los signos positivos y negativos deben aparecer de manera aleatoria.

Test χ^2

Con las desviaciones relativas de cada mortalidad, calculadas a partir de la estima-

ción conjunta de la probabilidad de muerte,

$$z_x^I = \frac{D_x^I - E_x^I q_x}{\sqrt{E_x^I q_x (1 - q_x)}},$$

$$z_x^{II} = \frac{D_x^{II} - E_x^{II} q_x}{\sqrt{E_x^{II} q_x (1 - q_x)}},$$

calculamos,

$$\chi^2 = \sum_x ((z_x^I)^2 + (z_x^{II})^2),$$

que se distribuye como una χ^2 con N grados de libertad. Valores pequeños de χ^2 indicarían que ambas mortalidades tienen un comportamiento similar. Para valores altos de χ^2 no necesariamente llegamos a la conclusión contraria, deberíamos inspeccionar detenidamente las z_x^I y z_x^{II} , para determinar si las diferencias no se deben a unas pocas edades.

En nuestro caso, sin embargo, hemos obtenido que las diferencias año a año no eran significativas, de forma que este análisis para cada par de años no era lo suficientemente concluyente para incorporar la influencia del tiempo del calendario en nuestro estudio. Esto puede ser debido a que la estimación de las probabilidades brutas la realizamos aplicando la hipótesis de uniformidad lo que supone calcular la media entre dos años consecutivos y por tanto suavizar de alguna forma las diferencias. Este problema podría resolverse si pudiéramos utilizar la doble clasificación de las muertes según edad y años de nacimiento como Felipe y Guillén (1999) hacen en su trabajo. Para tratar de detectar las diferencias, si realmente existen, hemos recurrido a un nuevo análisis que consiste en comparar tres periodos de tiempo a distancia tal que la suavización mencionada no tenga efecto. Los resultados de la comparación gráfica se muestran en la Figura 7.1, mientras que en las Tablas 7.1 y 7.2 se recogen los resultados de los tests para hombres y mujeres, respectivamente. No ha sido necesario llevar a cabo ninguna agrupación de edades porque el número esperado de defunciones era superior a 5 en todas ellas.

De las Tablas 7.1 y 7.2 se deduce el distinto comportamiento de los pares de años estudiados. Las diferencias se han calculado restando las probabilidades de muerte del año mayor a las del menor. En la mayoría de los casos son positivas, lo que pone de manifiesto la disminución de la mortalidad a medida que avanza el calendario. Todos los tests conducen al rechazo de la hipótesis de igualdad, excepto el test de rachas que la acepta para las dos primeras diferencias en el caso de las mujeres y para las dos últimas diferencias en el caso de los hombres. A señalar que en el caso del test de la χ^2 las diferencias relativas para cada año son elevadas para la mayoría de las edades.

Para analizar si las diferencias se comportan igual a lo largo de todas las edades, Felipe y Guillén (1999) utilizan un ratio que denominan E75 y que compara las diferentes mortalidades con la del año 1975. En el trabajo citado Felipe y Guillén recogen claramente las diferentes alternativas para la comparación de las experiencias de mortalidad a través de los valores brutos de las probabilidades y de sus logaritmos graduados o sin graduar. El método que hemos utilizado se basa en el análisis gráfico mediante el suavizado, con

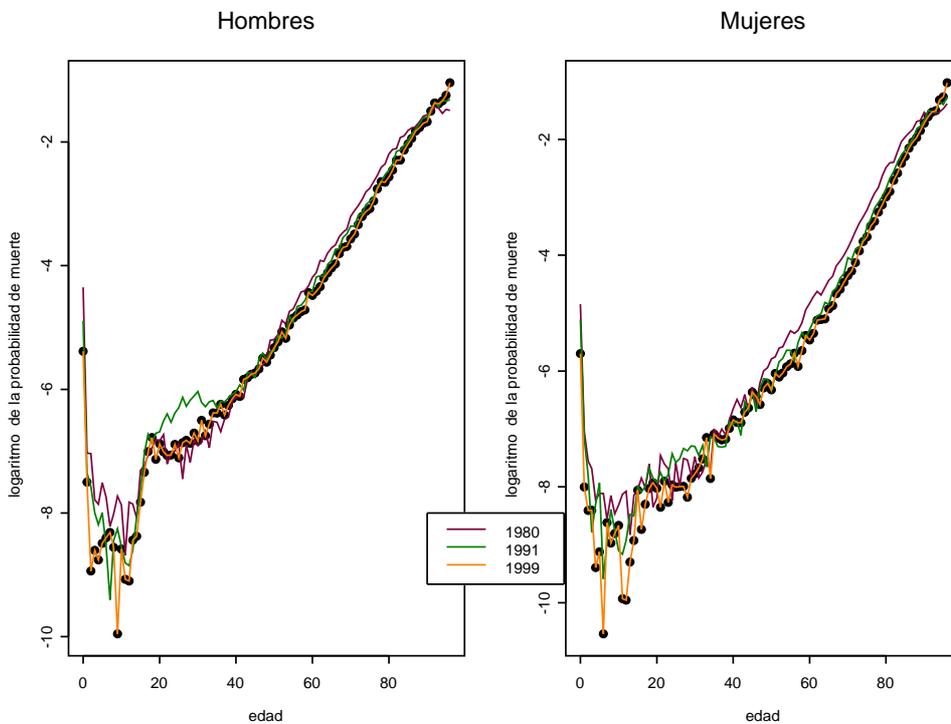


Figura 7.1: Comparación de las probabilidades de muerte brutas

	HOMBRES		
	1980-1991	1991-1999	1980-1999
Test de signos significación	60 0.9928	85 1	69 0.9999
Test Wilcoxon significación	2.9649 3.0273e-03	6.7539 1.4394e-11	4.6669 3.0576e-06
Test de rachas significación	7 0.03122	19 0.2449	17 0.0695
Test χ^2 significación	453.35 0	268.99 0	789.29 0

Tabla 7.1: Contrastes clásicos para la comparación de experiencias de mortalidad para los hombres

	MUJERES		
	1980-1991	1991-1999	1980-1999
Test de signos significación	75 1	76 1	84 1
Test de Wilcoxon significación	5.6096 2.0274e-08	5.8723 4.2974e-09	6.5164 7.2014e-11
Test de rachas significación	19 0.0855	27 0.2638	13 0.0186
Test χ^2 significación	603.54 0	259.60 1.1102e-16	1385.58 0

Tabla 7.2: Contrastes clásicos para la comparación de experiencias de mortalidad para las mujeres

splines de las diferencias de las estimaciones brutas de las transformaciones de las probabilidades de muerte. Haberman y Renshaw (1999) proponen graduar las diferencias entre los logaritmos de las fuerzas de mortalidad, es decir,

$$Z_x = \log\left(\mu_{x+\frac{1}{2}}^I\right) - \log\left(\mu_{x+\frac{1}{2}}^{II}\right)$$

mediante splines con ponderaciones,

$$\omega_x = \frac{D_x^I D_x^{II}}{D_x^I + D_x^{II}}.$$

Como nosotros utilizamos las q_x , plantemos graduar la transformación $\log(-\log(1 - q_x))$ y elegimos los grados de libertad mediante validación cruzada. La elección de dicha transformación viene justificada por la relación entre ambas medidas de mortalidad,

$$q_x \simeq 1 - \exp(-\mu_{x+\frac{1}{2}}),$$

de forma que despejando

$$\mu_{x+\frac{1}{2}} \simeq -\log(1 - q_x),$$

y por tanto,

$$\log\left(\mu_{x+\frac{1}{2}}\right) \simeq \log(-\log(1 - q_x)).$$

Las gráficas de las diferencias y su correspondiente graduación mediante splines se muestran en la Figura 7.2 para los hombres y en la Figura 7.3 para las mujeres. Alternativamente pueden graduarse las probabilidades y después calcular las diferencias. Esta segunda forma tiene la ventaja de no presentar problemas cuando el número de muertes observadas es 0. En nuestro caso hemos graduado las probabilidades brutas mediante modelos lineales generalizados, con distribución Binomial y link complementario log-log, utilizando un polinomio de grado 10. Después hemos calculado las diferencias entre las predicciones de las transformaciones de las probabilidades de muerte, obteniendo las gráficas de las Figuras 7.4 y 7.5 muy similares a las anteriores.

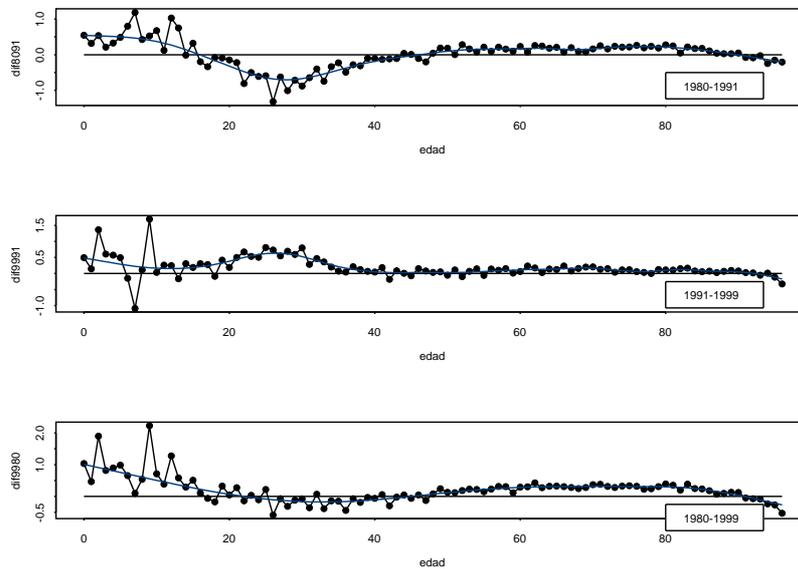


Figura 7.2: Comparación de estimaciones brutas de q_x para los hombres

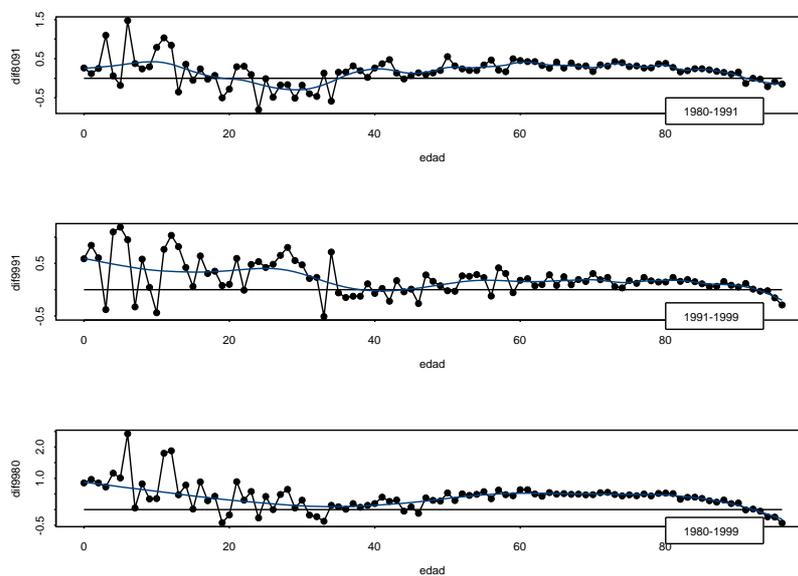


Figura 7.3: Comparación de estimaciones brutas de q_x para las mujeres

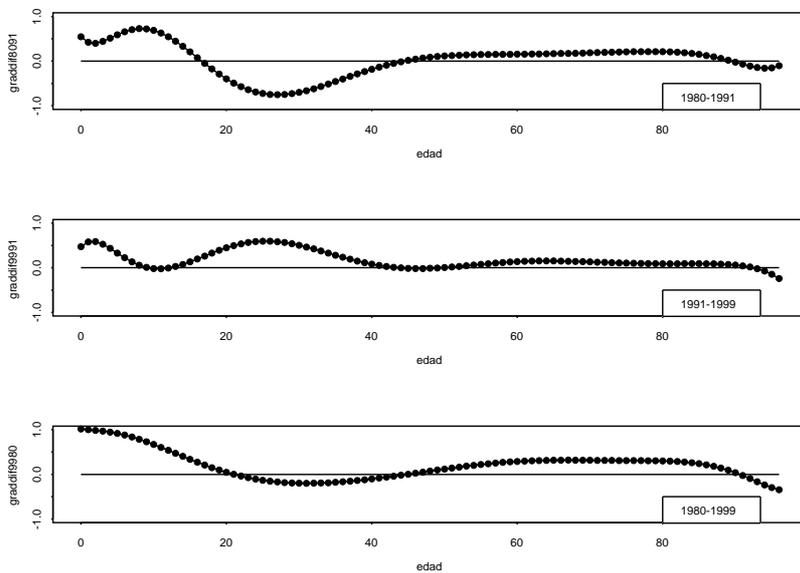


Figura 7.4: Comparación de estimaciones graduadas de q_x para los hombres

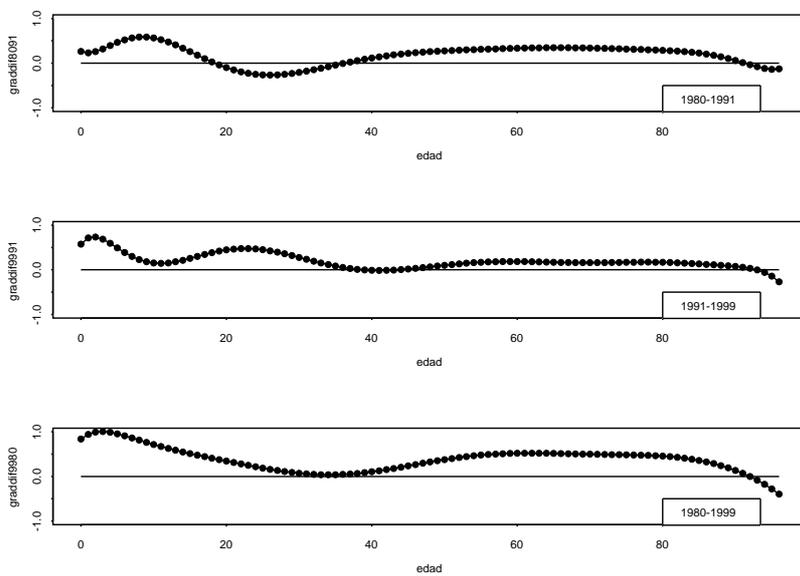


Figura 7.5: Comparación de estimaciones graduadas de q_x para las mujeres

En las gráficas de los hombres, Figuras 7.2 y 7.4, puede verse que la mortalidad de 1991 respecto a 1980 ha disminuido salvo para las edades entre 20 y 40 años. Mientras que la de 1999 respecto a 1991 ha disminuido para el conjunto de edades pero sobre todo para el intervalo de 20 a 40. En términos globales las probabilidades de muerte de 1991 han disminuido respecto a las de 1980, aunque hay un pequeño aumento para el mismo intervalo de edades.

En el caso de las mujeres, Figuras 7.3 y 7.5, la mortalidad ha disminuido en 1991 respecto a 1980, salvo para el rango de edades de entre 20 y 35 años donde parece aumentar aunque en menor grado que para el sexo masculino. En el siguiente periodo, 1991-1999, parecen existir disminuciones apreciables para edades inferiores a 35, manteniéndose parecida para el resto de edades. En el periodo global, 1980-1999, la mortalidad ha disminuido para el total de edades manteniéndose la igualdad sólo entre 30 y 40.

Como conclusión, hemos obtenido medidas para verificar las diferencias entre las experiencias de mortalidad correspondientes a periodos de calendario distintos para cada uno de los sexos. Además hemos comprobado la existencia de diferencias entre dichas experiencias tanto para los hombres como para las mujeres. La comparación gráfica de las diferencias suavizadas muestra un comportamiento distinto a lo largo del rango de edades consideradas, lo que justifica introducir dinamicidad en nuestro estudio.

7.3. Métodos estructurales

Comenzaremos aplicando los métodos estructurales que se basan en la hipótesis de que la medida de mortalidad elegida, q_{xt} , esta representada por una misma función de la edad a lo largo de todos los años cuyos parámetros dependen del tiempo. Estos métodos requieren una primera fase de estimación de los parámetros para cada año y una segunda fase donde los parámetros son modelados ajustándoles una serie temporal que nos permitirá predecir sus valores en futuros años. Pasemos a obtener y evaluar cada una de las propuestas.

7.3.1. Modelo de Heligman y Pollard

Estimación del Modelo

En esta sección estimamos la ley de Heligman y Pollard para los diferentes años de 1980 a 1999, y estudiamos la evolución de los parámetros para poder llevar a cabo predicciones. La estimación de los coeficientes correspondientes a cada año, según el criterio de mínimos cuadrados ponderados, no ha presentado mayor dificultad en el caso de los hombres. Para las mujeres sí se han vuelto a presentar problemas y hemos decidido centrar la joroba de los accidentes en 96 años, es decir, fijar el parámetro $F = 96$, además, la estimación del parámetro K era repetidamente no significativa y no permitía realizar el ajuste, por lo que finalmente hemos ajustado a los datos de las mujeres la primera ley y no la segunda. Los valores resultantes para los parámetros de cada sexo se recogen en la Tabla 7.3 y 7.4 respectivamente. Los valores de los parámetros A y D aparecen en las tablas multiplicados por 10^4 , los de B y C por 10^3 y el de G por 10^5 en los hombres y por 10^6 en las mujeres.

HOMBRES									
Año	A	B	C	D	E	F	G	H	K
1980	8.822	1.977	77.028	6.657	13.624	19.713	3.025	1.109	1.713
1981	8.314	5.334	91.504	7.175	11.074	20.304	3.003	1.109	1.601
1982	8.141.	4.515	81.856	6.742	8.150	20.961	2.817	1.110	1.449
1983	8.219	3.211	78.365	6.814	8.253	20.372	2.803	1.109	1.446
1984	7.816	6.937	84.765	7.284	8.738	20.4217	3.513	1.105	0.894
1985	8.939	18.625	101.681	8.195	8.288	21.161	3.410	1.105	0.911
1986	8.880	33.903	113.035	10.556	7.551	21.425	3.409	1.105	0.932
1987	7.080	15.404	98.755	12.471	8.032	21.943	3.850	1.103	0.587
1988	7.617	27.996	108.520	13.361	7.264	22.824	3.879	1.103	0.575
1989	9.441	58.107	125.426	14.142	5.195	23.849	3.467	1.104	0.727
1990	7.647	33.053	106.831	14.908	4.844	25.412	3.249	1.105	0.770
1991	7.541	54.179	131.646	15.138	5.076	26.156	3.319	1.104	0.802
1992	8.032	51.364	129.734	13.150	4.721	27.370	3.156	1.105	0.952
1993	5.931	17.578	99.356	12.261	3.555	30.398	2.604	1.108	1.271
1994	4.181	12.163	89.632	12.859	3.245	31.422	2.528	1.108	1.204
1995	4.041	3.237	74.451	13.432	3.113	33.180	2.342	1.109	1.014
1996	6.211	44.790	112.666	13.740	2.049	37.709	1.773	1.113	1.159
1997	4.643	13.190	90.973	11.863	1.547	40.205	1.483	1.115	1.139
1998	4.149	19.673	100.192	11.662	1.153	45.768	1.523	1.114	0.858
1999	5.451	130.390	163.234	13.164	0.790	59.569	1.648	1.113	0.888

Tabla 7.3: Coeficientes de la ley de Heligman y Pollard para los hombres

Año	MUJERES						
	A	B	C	D	E	G	H
1980	8.499	16.707	92.239	5.660	0.460	4.530	1.130
1981	7.096	8.948	96.083	3.089	0.994	5.369	1.130
1982	6.225	4.436	87.564	2.955	1.0328	4.529	1.130
1983	8.647	34.238	124.120	2.850	0.659	4.692	1.130
1984	9.097	53.553	128.266	3.638	0.899	3.672	1.131
1985	8.640	79.269	134.814	4.013	0.831	3.181	1.133
1986	10.186	14.905	182.492	4.503	0.676	3.203	1.133
1987	8.829	12.718	148.237	5.223	0.884	2.8154	1.134
1988	4.876	6.506	80.126	5.215	1.153	2.389	1.137
1989	5.975	17.905	90.162	5.113	1.237	2.252	1.137
1990	7.434	63.416	117.895	5.306	0.998	2.011	1.138
1991	8.487	111.981	148.364	6.140	1.100	1.717	1.140
1992	8.539	74.338	130.237	6.524	1.645	1.681	1.140
1993	5.981	20.431	92.537	7.069	2.278	1.695	1.140
1994	5.981	20.431	92.537	7.069	2.278	1.695	1.140
1995	3.291	0.896	45.982	6.989	3.2156	1.434	1.141
1996	5.770	57.756	102.212	6.831	2.157	1.191	1.144
1997	5.086	84.698	119.564	6.029	1.723	1.019	1.146
1998	2.687	5.191	73.650	5.969	2.422	9.758	1.147
1999	2.834	5.921	70.223	6.442	2.852	9.750	1.146

Tabla 7.4: Coeficientes de la ley de Heligman y Pollard para las mujeres

	HOMBRES								
	χ^2	$z > 2$	$z > 3$	pos	neg	p(pos)	rachas	p(rachas)	EPAM
1980	81.14	3	0	50	47	0.6576	42	0.3778	10.71
1981	72.75	2	0	49	48	0.5804	34	0.2592	9.26
1982	82.45	2	0	51	46	0.7287	42	0.3787	9.53
1983	84.67	2	0	49	48	0.5804	43	0.3933	10.05
1984	57.61	1	0	50	47	0.6576	35	0.2731	7.96
1985	63.04	1	0	46	51	0.3424	51	0.5272	8.24
1986	65.58	3	0	47	50	0.4196	41	0.3621	8.05
1987	59.70	1	0	50	47	0.6576	27	0.1742	8.50
1988	82.92	1	0	51	46	0.7287	39	0.3319	8.59
1989	85.26	5	0	53	44	0.8450	38	0.3192	9.40
1990	65.70	2	0	51	46	0.7287	43	0.3948	8.71
1991	77.34	2	0	47	46	0.5821	45	0.4568	10.61
1992	62.67	3	0	54	40	0.9393	37	0.3270	9.30
1993	98.16	3	2	53	40	0.9269	29	0.2130	11.56
1994	114.47	5	1	49	43	0.7671	45	0.4681	14.75
1995	75.86	2	1	45	47	0.4585	41	0.3962	10.47
1996	102.27	3	2	48	45	0.6607	43	0.4231	12.26
1997	104.96	7	2	48	44	0.6988	49	0.5367	11.18
1998	100.17	4	1	48	44	0.6988	32	0.2552	11.90
1999	109.86	5	2	46	46	0.5415	39	0.3625	16.00

Tabla 7.5: Contrastes no paramétricos del modelo de Heligman y Pollard para los hombres

Para el diagnóstico del modelo hemos realizado un gráfico de tallo y hojas de las desviaciones relativas del modelo, en el que hemos apreciado su comportamiento aproximadamente normal. La bondad de ajuste del modelo para cada uno de los años lo hemos analizado a través de los test habituales de Forfar, McCutcheon y Wilkie (1988), los resultados están recogidos en las Tablas 7.5 y 7.6 para hombres y mujeres, respectivamente. Hay pocas observaciones superiores a 2 para el ajuste de los hombres y un número mayor en el caso de las mujeres. En general, el ajuste para los hombres presenta mejores resultados. Los test de autocorrelaciones que hemos llevado a cabo resultan significativos para el ajuste de las mujeres en prácticamente todos los años, mientras que para los hombres no es así, lo que viene a corroborar que el modelo es mejor en el caso masculino.

El perfil de la evolución de las probabilidades de muerte a lo largo de los años así como su ajuste al modelo lo mostramos gráficamente en las Figuras 7.6 a 7.8 para los hombres y en las Figuras 7.9 a 7.11 para las mujeres.

Predicción

Una vez estimados los valores de los parámetros de la segunda ley de Heligman y Pollard para los hombres, y la primera, para las mujeres, en cada uno de los años. El

MUJERES									
	χ^2	$z > 2$	$z > 3$	pos	neg	p(pos)	rachas	p(rachas)	EPAM
1980	194.82	7	6	57	40	0.9665	34	0.2670	15.46
1981	178.98	10	6	56	41	0.9481	23	0.1337	13.60
1982	147.11	7	4	53	44	0.8450	27	0.1745	13.05
1983	166.80	8	4	52	45	0.7916	35	0.2742	13.30
1984	147.37	9	2	55	40	0.9499	27	0.1814	14.57
1985	127.03	10	0	53	41	0.9102	26	0.1725	13.45
1986	123.68	8	1	49	44	0.7330	32	0.2508	13.31
1987	125.39	7	0	53	40	0.9269	33	0.2693	13.55
1988	152.65	8	1	55	39	0.9605	25	0.1616	15.37
1989	148.20	9	1	52	41	0.8934	25	0.1641	14.86
1990	139.69	6	3	52	41	0.8934	27	0.1872	12.57
1991	148.33	13	0	53	39	0.9413	32	0.2603	14.18
1992	145.13	9	2	49	43	0.7671	41	0.3984	13.85
1993	137.13	7	2	51	40	0.8959	35	0.3099	12.01
1994	128.64	4	2	43	48	0.3376	22	0.1379	12.54
1995	115.97	5	2	54	36	0.9778	22	0.1392	13.74
1996	112.70	4	2	51	39	0.9149	32	0.2692	15.70
1997	122.06	4	1	51	38	0.9313	23	0.1531	18.99
1998	156.35	12	1	50	38	0.9173	23	0.1558	22.88
1999	160.78	10	3	50	38	0.9173	25	0.1798	22.68

Tabla 7.6: Contrastes no paramétricos del modelo de Heligman y Pollard para las mujeres

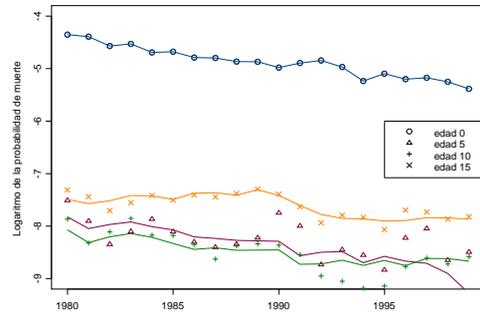


Figura 7.6: Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Heligman y Pollard para los hombres

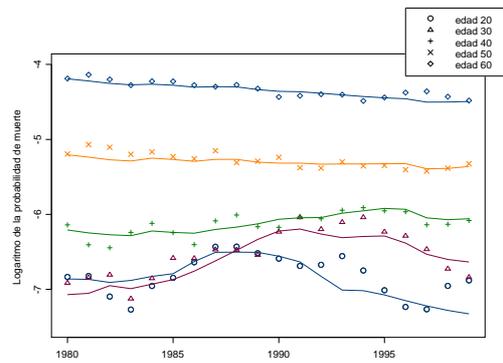


Figura 7.7: Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Heligman y Pollard para los hombres

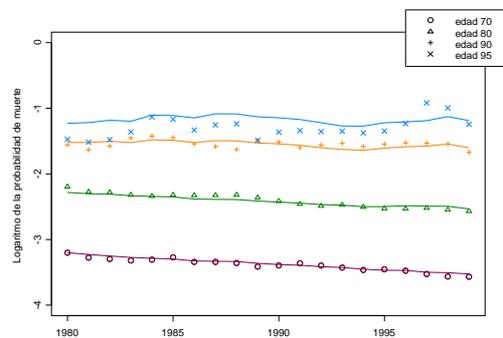


Figura 7.8: Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Heligman y Pollard para los hombres

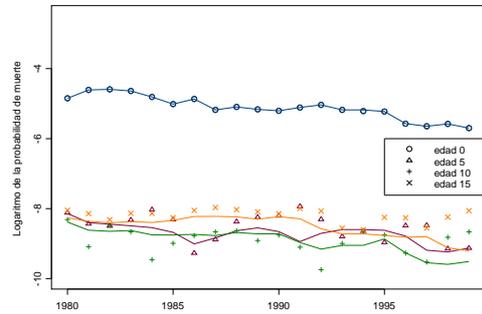


Figura 7.9: Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Heligman y Pollard para las mujeres

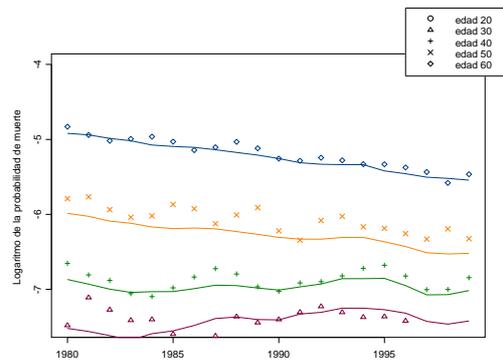


Figura 7.10: Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Heligman y Pollard para las mujeres

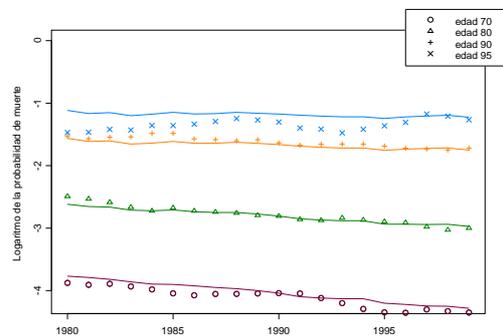


Figura 7.11: Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Heligman y Pollard para las mujeres

Parámetro	Modelo Hombres	Modelo Mujeres
A	ARIMA(0,1,1)	ARIMA(0,1,1)
B	ARIMA(0,2,2)	ARIMA(0,1,2)
C	ARIMA(0,1,0)	ARIMA(0,2,1)
D	ARIMA((0,1,1)	ARIMA(1,1,0)
E	ARIMA(0,1,1)	ARIMA(0,1,1)
F	ARIMA(0,2,1)	—
G	ARIMA(2,2,0)	ARIMA(0,1,0)
H	ARIMA(2,2,0)	ARIMA(0,1,0)
K	ARIMA(0,2,1)	—

Tabla 7.7: Modelos obtenidos para los parámetros del modelo de Heligman y Pollard

siguiente paso es determinar el modelo que sigue cada uno de los parámetros para poder obtener predicciones.

Aplicamos un análisis univariante de series temporales a cada uno de los parámetros de la ley según la metodología habitual. Los modelos obtenidos se presentan en la Tabla 7.7 y nos proporcionan predicciones e intervalos de confianza para los valores de los parámetros en años posteriores, que no reproducimos aquí, pero es importante reseñar que dichos intervalos de confianza muestran amplitud creciente a medida que nos alejamos en las predicciones por lo que no parece recomendable efectuar predicciones para años muy distantes.

7.3.2. Método de Lee-Carter

Estimación del Modelo

La propuesta de Lee y Carter es un nuevo modelo proporcionado por la ecuación (4.1) descrita en la sección 4.2.3, que ahora nosotros adaptamos a la estimación de los q_{xt} ,

$$\ln(q_{xt}) = a_x + b_x k_t + \epsilon'_{xt}. \quad (7.1)$$

Según este modelo hay que encontrar un único parámetro que dependa del tiempo del calendario, k_t , y dos parámetros a_x y b_x que dependen de la edad.

Para estimar los parámetros del modelo recurrimos al método de descomposición de valor singular (SVD) aplicado a la matriz de los logaritmos de las probabilidades de muerte

menos los valores estimados de las a_x , $\frac{\sum_t \ln(q_{xt})}{T}$, resolviendo el sistema

$$\ln(q_{xt}) - \frac{\sum_t \ln(q_{xt})}{T} = b_x k_t. \quad (7.2)$$

Los valores estimados de los parámetros a_x y b_x los mostramos en la Tabla 7.8 y su comportamiento gráfico en las Figuras de 7.12 a 7.14.

edad	HOMBRES		MUJERES	
	a_x	b_x	a_x	b_x
0	-4.877916	0.067319945	-5.113717	0.0295074751
1	-7.316237	0.067703543	-7.425365	0.0217740050
2	-7.695727	0.070174811	-7.873099	0.0189233826
3	-8.017089	0.036325678	-8.193740	0.0208081643
4	-8.113204	0.077498988	-8.523536	0.0192975403
5	-8.237719	0.050918724	-8.522656	0.0157277729
6	-8.286935	0.067523772	-8.990516	0.0648382892
7	-8.422949	0.050670598	-8.725170	-0.0043617434
8	-8.388135	0.068384031	-8.767054	0.0261753652
9	-8.410996	0.075654373	-8.810377	0.0006582203
10	-8.505554	0.086276959	-8.913930	0.0077428269
11	-8.531218	0.025736949	-9.319181	0.1080305709
12	-8.482226	0.063310969	-8.679441	0.0381543521
13	-8.263542	0.082110461	-8.630212	0.0054176150
14	-7.961448	0.018441748	-8.662408	0.0281033813
15	-7.610589	0.040341100	-8.195489	0.0052678286
16	-7.166518	0.005141068	-8.092523	0.0079590841
17	-6.905882	0.013891890	-8.025976	0.0001795500
18	-6.935136	0.025343773	-7.980057	0.0056098471
19	-6.846572	0.026483107	-7.929224	0.0007858730
20	-6.821792	-0.001624714	-7.995515	0.0050388168
21	-6.805630	0.025124671	-7.996760	0.0140711738
22	-6.799032	0.008369161	-7.897715	0.0013403127
23	-6.799963	0.011246161	-7.844392	0.0070050437
24	-6.720291	-0.001061435	-7.766820	-0.0013898925
25	-6.690216	-0.015257020	-7.820182	0.0092628995
26	-6.744981	-0.057227888	-7.679594	-0.0012575741
27	-6.617500	-0.037297763	-7.782821	-0.0074724923
28	-6.589744	-0.063965483	-7.622748	-0.0008084706
29	-6.560573	-0.062435569	-7.645935	-0.0131330952
30	-6.514739	-0.059289019	-7.472025	0.0116972303
31	-6.484176	-0.060845826	-7.458411	-0.0011000982
32	-6.472586	-0.062730086	-7.481730	-0.0132005141
33	-6.490741	-0.064808401	-7.420157	-0.0073291450
34	-6.416663	-0.058252967	-7.416797	-0.0001969037
35	-6.406274	-0.054889997	-7.308928	0.0031169523
36	-6.378795	-0.046225094	-7.122605	0.0005342592
37	-6.309521	-0.034470314	-7.187105	0.0005454137
38	-6.244897	-0.036551099	-7.158209	0.0017463850
39	-6.162874	-0.025726589	-6.977758	-0.0019444046
40	-6.131987	-0.027941799	-6.875661	0.0031603535

Tabla 7.8: Estimaciones de los parámetros del modelo de Lee-Carter

edad	HOMBRES		MUJERES	
	a_x	b_x	a_x	b_x
41	-6.041891	-0.021217332	-6.816800	0.0134368655
42	-5.983089	-0.003894709	-6.758100	0.0047950100
43	-5.880453	-0.013533172	-6.637603	0.0006129553
44	-5.794776	-0.004328977	-6.621683	0.0040793296
45	-5.748922	-0.012496318	-6.470545	0.0034526440
46	-5.645067	-0.007772606	-6.465754	0.006918980
47	-5.532937	0.006302121	-6.351409	0.006862816
48	-5.480689	0.009668463	-6.264362	0.010757132
49	-5.325193	0.012019076	-6.129266	0.008899943
50	-5.275238	0.022386512	-6.075124	0.015142630
51	-5.189940	0.007414079	-5.978993	0.015540260
52	-5.139671	0.015675106	-5.871424	0.015933524
53	-5.005730	0.012680779	-5.801530	0.013748492
54	-4.915276	0.014112411	-5.733941	0.011750620
55	-4.842140	0.014459986	-5.654921	0.016224475
56	-4.723685	0.008745404	-5.595931	0.014980891
57	-4.628812	0.024990722	-5.481055	0.014110966
58	-4.539263	0.019379743	-5.357000	0.012394800
59	-4.423143	0.013605249	-5.268710	0.011401553
60	-4.332122	0.024183112	-5.187177	0.019489306
61	-4.240918	0.020757049	-5.062206	0.013739087
62	-4.168840	0.029366245	-4.988076	0.014929788
63	-4.064417	0.027980160	-4.881255	0.016963304
64	-3.956063	0.022214617	-4.798840	0.013847493
65	-3.860443	0.019142262	-4.700925	0.014455162
66	-3.775745	0.017127228	-4.584750	0.015184751
67	-3.676891	0.023904516	-4.453636	0.015991595
68	-3.588265	0.016518302	-4.343544	0.012724352
69	-3.509034	0.022648121	-4.242541	0.0140617902
70	-3.387954	0.021923452	-4.111827	0.0135562261
71	-3.305879	0.023021375	-3.997452	0.0129402872
72	-3.190125	0.022778656	-3.843867	0.0148981513
73	-3.094410	0.019106953	-3.727952	0.0136925111
74	-2.982000	0.020828341	-3.613166	0.0134979065
75	-2.904351	0.020641614	-3.456580	0.0132181946
76	-2.816413	0.024683913	-3.320094	0.0132425093
77	-2.702451	0.020898457	-3.195419	0.0128004068
78	-2.609540	0.022979320	-3.043310	0.0121993341
79	-2.498277	0.024405867	-2.920846	0.0155434876
80	-2.402067	0.026399736	-2.784207	0.0137289301
81	-2.292508	0.023568122	-2.651932	0.0126040366

Tabla 7.8: Estimaciones de los parámetros del modelo de Lee-Carter

edad	HOMBRES		MUJERES	
	a_x	b_x	a_x	b_x
82	-2.187305	0.020501870	-2.515768	0.0101143732
83	-2.126693	0.024098994	-2.400776	0.0095988901
84	-2.011189	0.019108764	-2.270876	0.0096625223
85	-1.942092	0.017622295	-2.157824	0.0090260670
86	-1.846707	0.013587471	-2.033795	0.0078905550
87	-1.764192	0.009620014	-1.941734	0.0068926120
88	-1.685931	0.013600119	-1.834037	0.0075666382
89	-1.605512	0.009314188	-1.725804	0.0054491653
90	-1.547438	0.004348319	-1.618989	0.0070745466
91	-1.509171	-0.000336958	-1.557883	0.0006026095
92	-1.414125	-0.0030052996	-1.492463	0.002334726
93	-1.388621	-0.0007342304	-1.417618	-0.003288043
94	-1.321780	-0.0175693589	-1.373921	-0.006058445
95	-1.298290	-0.0135417474	-1.348778	-0.005513596
96	-1.392868	-0.0032798089	-1.294306	-0.005990769

Tabla 7.8: Estimaciones de los parámetros del modelo de Lee-Carter

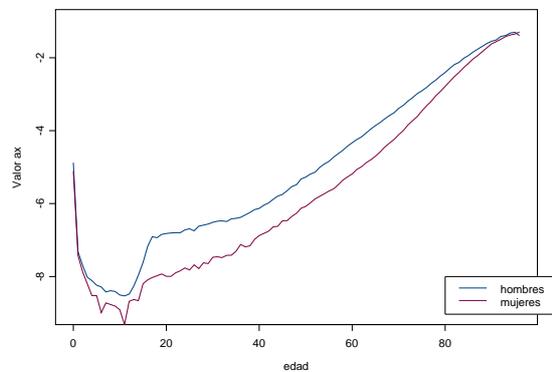


Figura 7.12: Comparación de las estimaciones del parámetro a_x

En la Figura 7.12, correspondiente al parámetro a_x , mostramos el perfil general de la mortalidad, mientras que las Figuras 7.13 y 7.14, correspondientes al parámetro b_x , muestran la velocidad de decrecimiento de los ratios en respuesta a cambios en k_t . En general, la mortalidad de las mujeres es inferior a la de los hombres como muestran los valores de a_x . El parámetro b_x muestra valores negativos para algunas edades intermedias de los hombres y para ambos sexos en las edades finales. Estos valores negativos indican que la mortalidad no decrece al variar el tiempo.

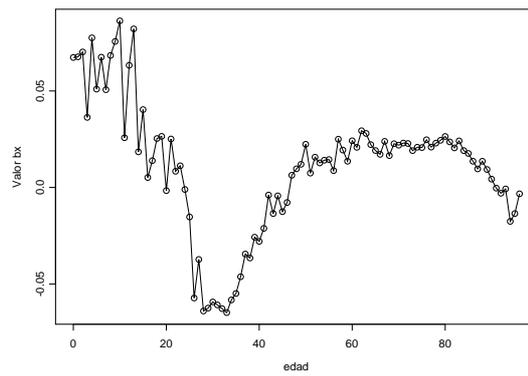


Figura 7.13: Estimación del parámetro b_x del modelo de Lee-carter para los hombres

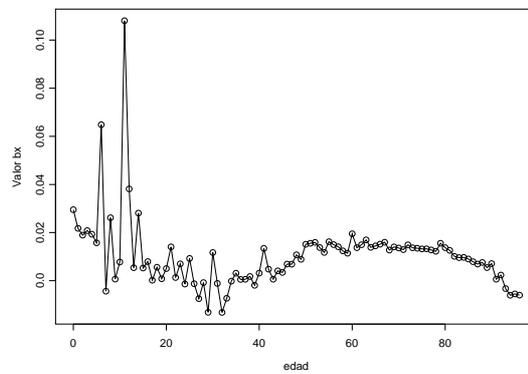
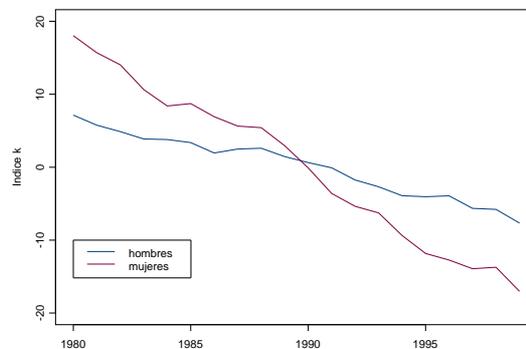


Figura 7.14: Estimación del parámetro b_x del modelo de Lee-carter para las mujeres

año	HOMBRES	MUJERES	año	HOMBRES	MUJERES
1980	7.1330	18.0329	1990	0.6368	-0.10106
1981	5.7658	15.6983	1991	-0.0897	-3.59475
1982	4.8774	14.0319	1992	-1.7636	-5.37425
1983	3.8773	10.6471	1993	-2.6937	-6.26828
1984	3.7885	8.3928	1994	-3.9066	-9.36136
1985	3.3714	8.7193	1995	-4.0542	-11.8346
1986	1.9477	6.9166	1996	-3.9271	-12.7326
1987	2.4925	5.6423	1997	-5.6419	-13.9263
1988	2.6038	5.4154	1998	-5.7794	-13.7145
1989	1.4447	2.9437	1999	-7.6607	-17.0152

Tabla 7.9: Estimaciones del parámetro k_t del modelo de Lee-carter

Hemos de recurrir a reestimar k_t utilizando la ecuación (4.5) puesto que ha sido estimado minimizando los errores en los logaritmos y no en los mismos q_x y por tanto no conducen al número real de muertes para el total de edades. Las soluciones finales las mostramos en la Tabla 7.9 y su comparación gráfica para hombres y mujeres en la Figura 7.15, en la que se observa claramente que la disminución de la mortalidad es más pronunciada para las mujeres.

Figura 7.15: Comparación de las estimaciones del parámetro k_t

Para el diagnóstico del modelo hemos realizado un gráfico de tallo y hojas para las desviaciones relativas del modelo observando un comportamiento aproximadamente normal. La bondad de ajuste del modelo para cada uno de los años lo hemos analizado a través de los test habituales de Forfar, McCutcheon y Wilkie (1988), los resultados están recogidos en las Tablas 7.10 y 7.11 para hombres y mujeres, respectivamente. El número de desviaciones relativas es menor en el caso femenino y en general los resultados de los

	HOMBRES								
	χ^2	$z > 2$	$z > 3$	pos	neg	p(pos)	rachas	p(rachas)	EPAM
1980	64.20	1	0	49	48	0.5804	43	0.3933	10.22
1981	54.28	0	0	45	52	0.2713	40	0.3485	9.90
1982	44.19	1	0	48	49	0.5000	44	0.4094	8.44
1983	64.32	1	0	42	55	0.1114	39	0.3386	10.91
1984	40.58	0	0	45	52	0.2713	36	0.2885	7.99
1985	44.67	0	0	45	52	0.2713	41	0.3642	8.50
1986	54.29	1	0	49	48	0.5804	47	0.4586	8.50
1987	81.15	5	0	54	43	0.8886	33	0.2485	9.81
1988	135.64	10	0	55	42	0.9226	33	0.2496	12.13
1989	181.88	12	5	62	35	0.9979	35	0.2991	12.16
1990	197.43	7	6	55	42	0.9226	33	0.2497	11.80
1991	175.50	6	5	49	48	0.5804	33	0.2459	13.03
1992	72.25	1	0	50	44	0.7648	23	0.1410	9.63
1993	74.26	3	0	43	49	0.3012	39	0.3646	9.36
1994	67.98	3	0	47	45	0.6227	38	0.3464	12.13
1995	57.48	0	0	48	44	0.6988	39	0.3635	10.19
1996	77.89	3	1	46	46	0.5415	40	0.3791	12.54
1997	128.843	7	2	35	57	0.0140	25	0.1687	13.72
1998	191.38	6	7	36	56	0.0235	32	0.2670	15.88
1999	267.79	7	11	44	47	0.4170	32	0.2599	22.64

Tabla 7.10: Contrastes no paramétricos del modelo de Lee-Carter para los hombres

test son mejores. Respecto a las autocorrelaciones resultan significativas para el ajuste de los hombres en prácticamente todos los años, mientras que para las mujeres no es así, lo que viene a corroborar que el modelo es mejor en el caso femenino.

El perfil de la evolución de las probabilidades de muerte a lo largo de los años así como su ajuste al modelo lo mostramos gráficamente en las Figuras 7.16 a 7.18 para los hombres y en las Figuras 7.19 a 7.21 para las mujeres.

Predicción

El último paso del método de Lee-Carter consiste en encontrar un modelo para los valores de los índices de mortalidad, $\{k_t\}$, utilizando metodología Box-Jenkins que permita obtener predicciones. Comenzamos observando el perfil de su evolución en la Figura 7.15 que muestra una clara tendencia. Para conseguir la serie estacionaria en el caso de los hombres tomamos una diferencia, observando en el correspondiente correlograma que es un ruido blanco. Un buen modelo para k_t es

$$k_t = -0,778623 + k_{t-1} + u_t,$$

MUJERES									
	χ^2	$z > 2$	$z > 3$	pos	neg	p(pos)	rachas	p(rachas)	EPAM
1980	71.63	1	1	52	43	0.8476	45	0.4470	12.44
1981	40.87	1	0	45	50	0.3409	39	0.3442	11.49
1982	53.55	1	0	45	50	0.3409	46	0.4596	13.78
1983	54.00	2	0	42	53	0.1524	41	0.3820	11.52
1984	50.66	0	0	45	50	0.3409	43	0.4091	11.96
1985	53.71	2	0	47	48	0.5000	39	0.3430	12.66
1986	57.16	1	0	54	41	0.9247	42	0.4020	13.14
1987	70.67	3	0	53	42	0.8910	43	0.4158	12.56
1988	68.78	2	0	50	44	0.7647	43	0.4174	12.04
1989	63.65	2	1	46	47	0.5000	42	0.4056	11.60
1990	65.49	2	0	51	42	0.8501	44	0.4460	284.49
1991	70.18	2	0	51	41	0.8743	46	0.4918	12.86
1992	70.95	1	0	50	41	0.8528	43	0.4450	13.74
1993	75.33	4	0	57	34	0.9943	37	0.3688	10.99
1994	79.84	3	1	52	39	0.9291	37	0.3456	12.61
1995	70.18	4	0	55	36	0.9823	31	0.2554	11.66
1996	43.56	1	0	45	46	0.5000	36	0.3204	10.35
1997	37.53	0	0	39	52	0.1041	44	0.4710	13.05
1998	49.86	0	1	34	57	0.0103	43	0.4880	16.12
1999	71.22	1	1	44	46	0.4581	43	0.4467	15.81

Tabla 7.11: Contrastes no paramétricos del modelo de Lee-Carter para las mujeres

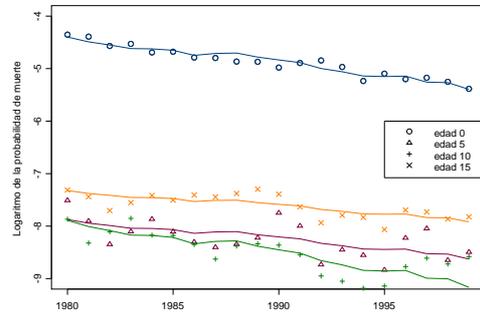


Figura 7.16: Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Lee-Carter para los hombres

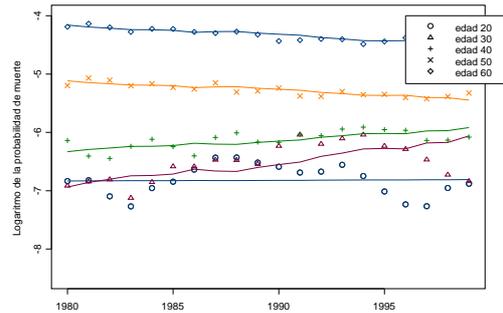


Figura 7.17: Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Lee-Carter para los hombres

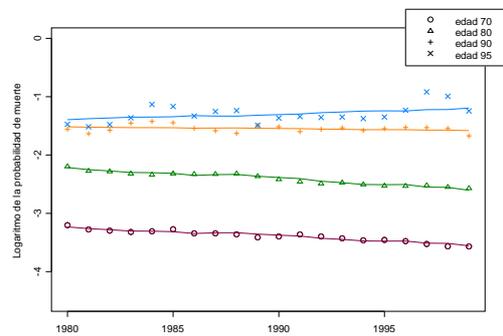


Figura 7.18: Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Lee-Carter para los hombres

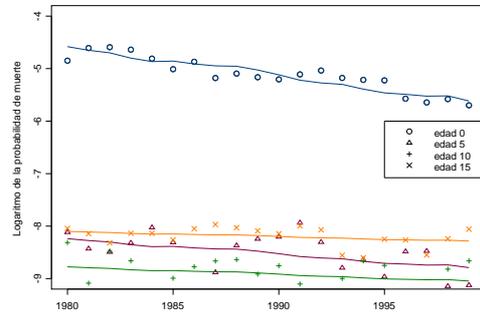


Figura 7.19: Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo de Lee-Carter para las mujeres

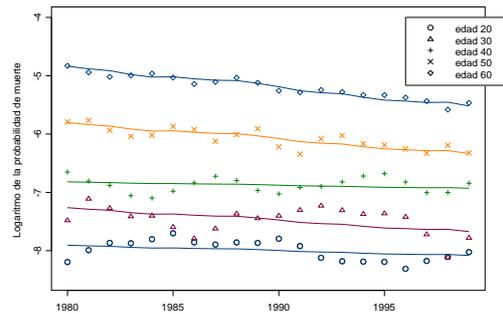


Figura 7.20: Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo de Lee-Carter para las mujeres

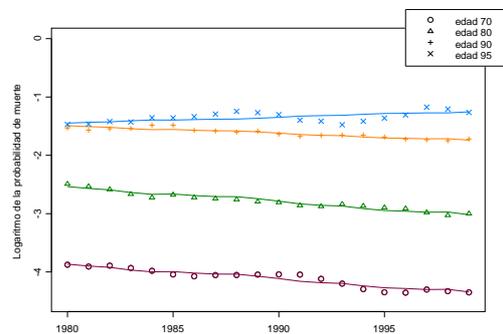


Figura 7.21: Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo de Lee-Carter para las mujeres

HOMBRES						
	Serie Temporal			Recta de Regresión		
	Predicción	Intervalo al 95 %		Predicción	Intervalo al 95 %	
2000	-8.43939	-9.9131	-6.96567	-7.38707	-9.02927	-5.74488
2001	-9.21801	-11.3022	-7.13386	-8.10214	-9.76652	-6.43775
2002	-9.99663	-12.5492	-7.44408	-8.8172	-10.5055	-7.12893
2003	-10.7753	-13.7227	-7.82783	-9.53226	-11.246	-7.8185
2004	-11.5539	-14.8492	-8.25855	-10.2473	-11.9881	-8.50653
2005	-12.3325	-15.9424	-8.72265	-10.9624	-12.7317	-9.19309
2006	-13.1111	-17.0102	-9.21204	-11.6775	-13.4767	-9.87824
2007	-13.8897	-18.058	-9.72145	-12.3925	-14.223	-10.5621
2008	-14.6684	-19.0895	-10.2472	-13.1076	-14.9706	-11.2446
2009	-15.447	-20.1073	-10.7867	-13.8226	-15.7193	-11.926

Tabla 7.12: Predicciones del índice de mortalidad k_t del modelo de Lee-Carter para los hombres

donde la constante del modelo, con error estándar 0.160926, es significativa y u_t es un ruido blanco. La predicción de k_t se muestra en la Figura 7.22 y en la Tabla 7.12 mostramos las predicciones de 2000 a 2009.

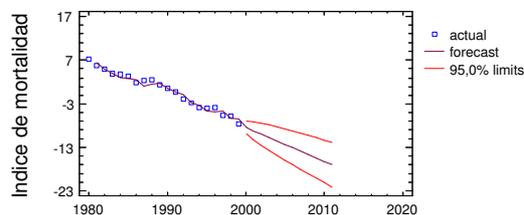


Figura 7.22: Evolución de k_t para los hombres

Como alternativa sencilla a la metodología habitual, dado que la evolución del índice de mortalidad era bastante lineal, hemos pensado en ajustarle una recta de regresión y comparar ambos resultados. El modelo en este caso resulta ser,

$$k_t = 6,38051(se = 0,479807) - 0,529584(se = 0,0400535)t$$

donde ambos coeficientes significativos. El error estándar figura entre paréntesis.

Una primera comparación entre ambos modelos puede hacerse a través del valor de R^2 que es 0.9967 para la serie temporal y 0.972647 para la recta de regresión. También para las predicciones los intervalos de confianza obtenidos mediante la serie temporal son

MUJERES						
	Serie Temporal			Recta de Regresión		
	Predicción	Intervalo al 95 %		Predicción	Intervalo al 95 %	
2000	-18.8599	-21.3631	-16.3567	-19.2315	-22.1227	-16.3402
2001	-20.7045	-24.2446	-17.1645	-21.075	-24.0054	-18.1447
2002	-22.5492	-26.8848	-18.2135	-22.9186	-25.891	-19.9462
2003	-24.3938	-29.4002	-19.3874	-24.7621	-27.7794	-21.7448
2004	-26.2384	-31.8357	-20.6412	-26.6057	-29.6706	-23.5408
2005	-28.0831	-34.2146	-21.9516	-28.4492	-31.5643	-25.3342
2006	-29.9277	-36.5505	-23.3049	-30.2928	-33.4605	-27.1251
2007	-31.7724	-38.8524	-24.6923	-32.1363	-35.3591	-28.9136
2008	-33.617	-41.1265	-26.1075	-33.9799	-37.2599	-30.6999
2009	-35.4616	-43.3774	-27.5459	-35.8235	-39.1628	-32.4841

Tabla 7.13: Predicciones del índice de mortalidad k_t del modelo de Lee-Carter para las mujeres

de menor amplitud que los obtenidos a partir de la recta de regresión, al menos para los primeros años.

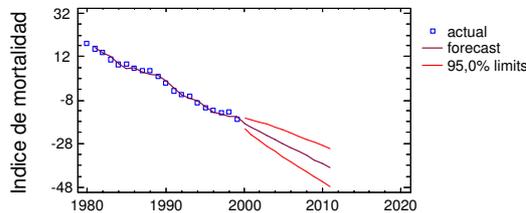


Figura 7.23: Evolución de k_t para las mujeres

En el caso de las mujeres también hemos tomado una diferencia, observando que se obtiene un buen modelo análogo al anterior mediante la expresión

$$k_t = -1,84464 + k_{t-1} + u_t,$$

donde la constante del modelo, con error estándar 0.273341, es significativa y u_t un ruido blanco. La predicción de k_t gráficamente se muestra en la Figura 7.23 y en la Tabla 7.13 mostramos las predicciones de 2000 a 2009.

También ahora hemos ajustado una recta de regresión cuya expresión es,

$$k_t = 19,4832(se = 0,579782) - 1,84356(se = 0,0483993) t = 0, \dots 20$$

resultando ambos coeficientes significativos, y cuyo error estándar figura entre paréntesis.

Una primera comparación puede hacerse a través del valor de R^2 que es prácticamente 1 para la serie temporal mientras que es 0.9871 para la recta de regresión. En cuanto a las predicciones también obtenemos intervalos de confianza de menor amplitud al menos para los primeros años. Además las rectas de regresión muestran una tendencia decreciente más elevada y poco prudente en ambos casos, por ello las series temporales nos parecen un modelo más adecuado.

7.4. Métodos no estructurales

Los métodos no estructurales a diferencia de los anteriores suponen que la medida de mortalidad elegida, q_{xt} , es función de la edad x y del tiempo t , simultáneamente. Pasemos a obtener y evaluar los modelos lineales generalizados propuestos según este esquema.

7.4.1. Modelos lineales generalizados con respecto a la edad y el tiempo

La descripción de este modelo así como la obtención de los coeficientes del ajuste se encuentran descritos en la sección 4.3.2. Como hemos decidido obtener todos los ajustes sobre las probabilidades de muerte, q_x , utilizaremos la familia binomial con link logit, de forma que ajustamos,

$$\log\left(\frac{q_x}{1-q_x}\right) = \beta_0 + \sum_{j=1}^s \beta_j L_j(x') + \sum_{i=1}^r \alpha_i t'^i + \sum_{i=1}^r \sum_{j=1}^s \gamma_{ij} L_j(x') t'^i.$$

Para obtener un modelo adecuado a los criterios de bondad de ajuste y de no excesiva complejidad del modelo, seleccionamos inicialmente los valores óptimos de r y s calculando la mejora de la *Deviance* para sucesivos incrementos de r y s y comparándola con el valor de una variable χ^2 con un grado de libertad. Eligiendo aquellos valores a partir de los cuales el incremento ya no es significativo. Mostramos en la Tabla 7.14 los valores de deviance para el ajuste de las funciones $GM(r, s)$, para $r = 0, 1, 2, 3$, $s = 1, 2, 3, \dots, 20$ y donde los coeficientes γ_{ij} se suponen cero. A continuación, partiendo de los valores predeterminados de $r = 1$ y $s = 17$, procedemos a completar el modelo mediante la elección de los coeficientes γ_{ij} . En la Tabla 7.15 se analiza si el incremento de la *Deviance* al introducir cada uno de estos parámetro resulta ser significativo. Finalmente el modelo elegido para ambos sexos es el mismo

$$\text{logit}(q_{xt}) = \beta_0 + \sum_{j=1}^{17} \beta_j L_j(x') + \alpha_1 t' + \sum_{j=1}^3 \gamma_{1j} L_j(x') t' \quad (7.3)$$

con 1918 grados de libertad.

Se concluye al completar el modelo que la elección de los coeficientes γ_{ij} indican que el factor de actualización temporal depende de la edad. Factor que puede apreciarse en la formula (7.3) y que corresponde a la segunda parte del exponente, la primera sería el resultado de graduar $\frac{q_x}{1-q_x}$ mediante una función $GM(0, 17)$. Sus coeficientes para ambos

HOMBRES				
	r=0	r=1	r=2	r=3
s=0	934242.9			
s=1	43411.384	41886.324	41884.818	41883.938
s=2	22099.817	20333.486	20330.920	20330.642
s=3	17626.226	15997.455	15994.656	15994.065
s=4	17080.456	15464.402	15461.171	15460.734
s=5	14452.050	12841.227	12837.261	12836.591
s=6	11644.771	9951.687	9945.409	9945.097
s=7	7426.265	5666.843	5664.226	5664.126
s=8	5896.119	4203.746	4201.567	4201.246
s=9	5619.034	3914.284	3911.199	3910.924
s=10	5577.255	3866.231	3863.552	3863.327
s=11	5507.789	3807.843	3805.213	3804.893
s=12	5475.779	3774.286	3771.402	3771.104
s=13	5350.175	3642.865	3640.314	3640.070
s=14	5229.336	3526.014	3523.537	3523.196
s=15	5045.511	3343.088	3340.273	3339.948
s=16	4947.208	3243.416	3240.651	3240.384
s=17	4889.701	3185.348	3182.611	3182.320
s=18	4888.448	3183.998	3181.258	3180.964
s=19	4875.242	3170.165	3167.410	3167.118

MUJERES				
	r=0	r=1	r=2	r=3
s=0	1044087			
s=1	53538.598	50015.330	50007.630	50005.242
s=2	19978.974	15689.621	15684.989	15680.318
s=3	15378.198	11353.818	11348.592	11344.884
s=4	14825.962	10807.093	10802.841	10798.767
s=5	10034.615	5967.920	5963.250	5959.137
s=6	9196.543	5116.303	5112.352	5108.199
s=7	8072.618	3918.108	3913.498	3908.619
s=8	7003.322	2917.358	2911.481	2907.187
s=9	6672.966	2580.459	2575.745	2571.446
s=10	6619.890	2516.035	2510.970	2506.485
s=11	6519.143	2428.779	2423.512	2419.279
s=12	6468.553	2378.581	2373.578	2369.373
s=13	6433.036	2339.920	2334.880	2330.587
s=14	6365.525	2273.299	2268.118	2263.888
s=15	6327.659	2236.830	2231.748	2227.596
s=16	6288.451	2197.120	2192.089	2187.812
s=17	6272.154	2180.314	2175.254	2170.987
s=18	6266.363	2174.517	2169.448	2165.178
s=19	6266.337	2174.438	2169.371	2165.100

Tabla 7.14: Deviance para los modelos lineales generalizados dinámicos

HOMBRES				
	Deviance	D.f.	Deviance (dif)	p-value
	3185.348	1921		
γ_{11}	3161.857	1920	23.4910	0.0000013
γ_{12}	3136.431	1919	25.4253	0.0000005
γ_{13}	2327.507	1918	808.9241	0.0000000
γ_{14}	2327.131	1917	0.3763	0.5396142

MUJERES				
	Deviance	D.f.	Deviance (dif)	p-value
	2180.314 2	1921		
γ_{11}	2104.487	1920	75.8271	0.0000000
γ_{12}	2006.316	1919	98.1711	0.0000000
γ_{13}	1434.066	1918	572.2504	0.0000000
γ_{14}	1433.980	1917	0.0855	0.7699675

Tabla 7.15: Diferencia de Deviance cuando introducimos interacciones en los modelos generalizados dinámicos

sexos se muestran en las Tablas 7.16 y 7.17, siendo todos ellos significativos ($t_{1918}(0,025) = 1,96$), salvo el parámetro β_9 para los hombres. Para el diagnóstico del modelo hemos realizado un gráfico de tallo y hojas para las desviaciones relativas como en los anteriores modelos, en él hemos podido apreciar un comportamiento aproximadamente normal. La bondad de ajuste del modelo para cada uno de los años lo hemos analizado a través de los test habituales de Forfar, McCutcheon y Wilkie (1988), los resultados están recogidos en las Tablas 7.18 y 7.19 para hombres y mujeres, respectivamente. También en este modelo resultan mejores los resultados en el caso femenino. Los test de autocorrelaciones resultan significativos para el ajuste de los hombres prácticamente en todos los años, mientras que para las mujeres no es así, lo que viene a corroborar que el modelo es mejor en el caso femenino.

El perfil de la evolución de las probabilidades de muerte a lo largo de los años así como su ajuste al modelo lo mostramos gráficamente en las Figuras 7.24 a 7.26 para los hombres y en las Figuras 7.27 a 7.29 para las mujeres.

Las edades entre 20 y 40 para los hombres presentan oscilaciones a lo largo de los años difíciles de recoger por el modelo. Para los años 1991 en adelante la sobremortalidad de la joroba de los accidentes no se ajusta según este modelo, motivo por el cual el ajuste de los hombres resulta peor que el de las mujeres donde aquella no aparece.

Respecto a la predicción para futuros años como es un método no estructural a diferencia de los métodos del apartado anterior, cuyo procedimiento se dividía en dos partes estimación y predicción, en éste obtenemos una función del tiempo, en la que basta con sustituir los valores correspondientes a t para calcular el valor de los futuros q_{xt} .

HOMBRES			
	coef	std error	t-value
β_0	-4.94521369	0.004619532	-1070.501007
β_1	3.71990286	0.009390396	396.139066
β_2	0.61510683	0.009794262	62.802772
β_3	-0.18388860	0.009696302	-18.964818
β_4	0.09273791	0.010851371	8.546193
β_5	-0.48680832	0.012588050	-38.672257
β_6	0.55705379	0.013352790	41.718157
β_7	-0.51392563	0.013442666	-38.230931
β_8	0.18137914	0.013863505	13.083210
β_9	0.01881685	0.014995850	1.254804
β_{10}	-0.06791302	0.015418257	-4.404714
β_{11}	-0.03481558	0.015286954	-2.277470
β_{12}	0.11803521	0.014730069	8.013215
β_{13}	-0.25407786	0.014500125	-17.522460
β_{14}	0.27528373	0.015105096	18.224559
β_{15}	-0.28840660	0.015831600	-18.217148
β_{16}	0.17875211	0.014335972	12.468782
β_{17}	-0.09115035	0.012108648	-7.527707
α_1	-0.08384075	0.005368708	-15.616559
γ_{11}	0.06235338	0.009306104	6.700267
γ_{12}	-0.11834738	0.008784937	-13.471625
γ_{13}	0.25300014	0.008947747	28.275289

Tabla 7.16: Coeficientes del modelo lineal generalizado para los hombres

MUJERES			
	coef	std error	t-value
β_0	-5.54571929	0.005878505	-943.389437
β_1	3.89646720	0.011294122	344.999575
β_2	1.15083526	0.011444426	100.558585
β_3	-0.31377174	0.011721128	-26.769757
β_4	0.23109134	0.013155943	17.565548
β_5	-0.51340117	0.015162574	-33.859764
β_6	0.33767181	0.016278685	20.743188
β_7	-0.40078907	0.016803606	-23.851373
β_8	0.27577582	0.017674168	15.603327
β_9	-0.10465703	0.018983633	-5.513014
β_{10}	0.04681200	0.019604939	2.387766
β_{11}	-0.08281728	0.019754146	-4.192400
β_{12}	0.11982939	0.019319198	6.202607
β_{13}	-0.16925571	0.018960900	-8.926565
β_{14}	0.20776250	0.019099950	10.877646
β_{15}	-0.19166078	0.019907711	-9.627464
β_{16}	0.13985630	0.019008885	7.357418
β_{17}	-0.05655189	0.013781578	-4.103441
α_1	-0.15694347	0.007226550	-21.717620
γ_{11}	0.03837568	0.011304907	3.394603
γ_{12}	-0.03299597	0.010414035	-3.168413
γ_{13}	0.24141211	0.010126704	23.839160

Tabla 7.17: Coeficientes del modelo lineal generalizado para las mujeres

HOMBRES									
	χ^2	$z > 2$	$z > 3$	pos	neg	p(pos)	rachas	p(rachas)	EPAM
1980	105.74	7	0	46	51	0.3424	35	0.2735	14.33
1981	96.50	4	0	37	60	0.0125	34	0.2748	14.20
1882	93.13	2	0	37	60	0.0125	24	0.1430	13.28
1983	113.67	7	0	36	61	0.0072	37	0.3284	14.72
1984	89.07	3	0	45	52	0.2713	36	0.2884	12.68
1985	81.83	2	0	45	52	0.2713	39	0.3330	12.57
1986	81.80	3	1	39	58	0.0335	37	0.3157	10.41
1987	85.16	4	1	57	40	0.9665	37	0.3125	10.87
1988	117.88	7	1	61	36	0.9960	35	0.2946	12.05
1989	148.39	7	5	60	37	0.9928	41	0.3934	11.98
1990	146.99	9	3	57	40	0.9665	37	0.3125	11.76
1991	150.63	5	4	54	42	0.9079	40	0.3599	14.82
1992	93.62	6	0	49	46	0.6591	31	0.2284	12.83
1993	116.30	8	1	39	56	0.0501	29	0.2069	13.58
1994	136.22	6	3	43	51	0.2352	36	0.3054	16.63
1995	122.80	8	1	42	52	0.1767	29	0.2081	14.33
1996	142.21	7	2	48	46	0.6214	35	0.2881	16.36
1997	128.45	9	1	32	60	0.0023	45	0.5482	14.57
1998	144.22	8	2	36	56	0.0235	28	0.2073	17.13
1999	126.99	8	1	26	66	0.00002	28	0.2443	19.51

Tabla 7.18: Contrastes no paramétricos del modelo lineal generalizado para los hombres

MUJERES									
	χ^2	$z > 2$	$z > 3$	pos	neg	p(pos)	rachas	p(rachas)	EPAM
1980	90.80	0	1	55	42	0.9226	44	0.4207	14.07
1981	59.73	1	0	48	49	0.5000	35	0.2729	11.44
1982	57.57	2	0	42	55	0.1114	42	0.3872	13.40
1983	69.34	3	0	36	61	0.0072	43	0.4376	13.65
1984	69.68	1	0	33	64	0.0011	39	0.3869	13.08
1985	62.45	3	0	44	52	0.2376	43	0.4044	12.79
1986	54.36	1	0	48	47	0.5812	48	0.4916	12.54
1987	67.47	2	0	52	43	0.8476	41	0.3796	12.44
1988	78.68	3	0	64	30	0.9999	41	0.4806	13.41
1989	67.89	2	0	58	35	0.9938	42	0.4478	12.32
1990	68.25	3	0	53	39	0.9413	40	0.3922	INF
1991	68.40	0	0	47	45	0.6227	44	0.4479	13.88
1992	68.61	2	0	43	49	0.3012	43	0.4330	14.60
1993	75.67	5	0	48	44	0.6988	35	0.2995	11.74
1994	74.42	2	1	43	48	0.3377	33	0.2750	12.43
1995	71.28	2	0	41	50	0.2009	39	0.3745	13.08
1996	55.15	0	0	44	47	0.4170	39	0.3700	14.20
1997	60.30	3	0	35	56	0.0177	31	0.2581	19.55
1998	82.34	2	1	43	47	0.3760	33	0.2800	25.93
1999	71.54	1	1	44	46	0.4581	37	0.3431	23.60

Tabla 7.19: Contrastes no paramétricos del modelo lineal generalizado para las mujeres

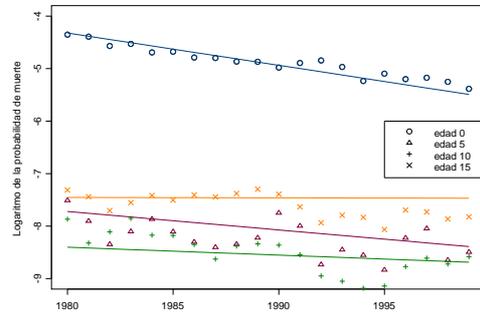


Figura 7.24: Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo lineal generalizado para los hombres

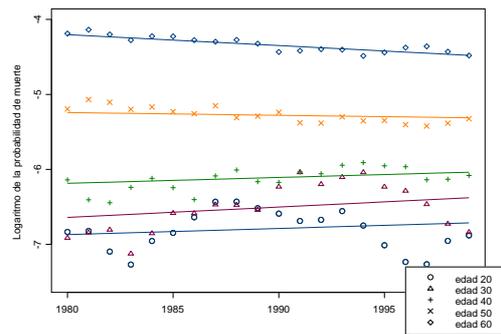


Figura 7.25: Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo lineal generalizado para los hombres

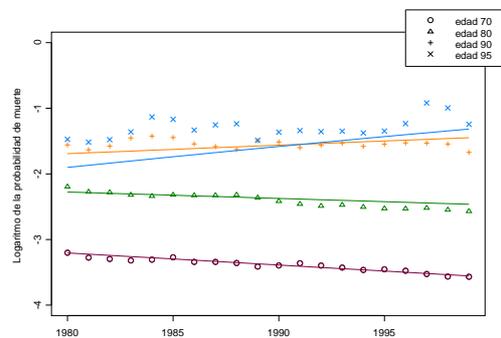


Figura 7.26: Evolución de $\log(q_x)$ para edades de 60 a 95 del modelo lineal generalizado para los hombres

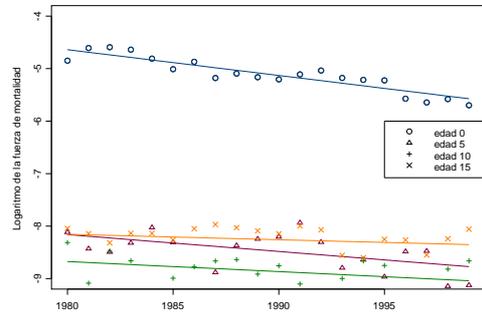


Figura 7.27: Evolución de $\log(q_x)$ para edades de 0 a 15 del modelo lineal generalizado para las mujeres

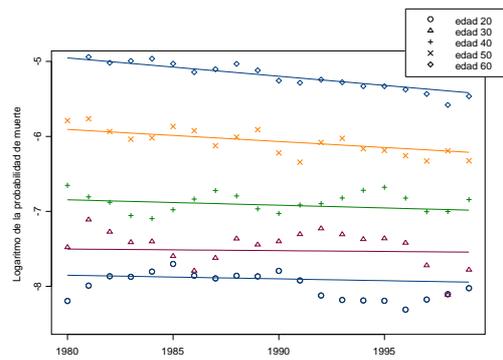


Figura 7.28: Evolución de $\log(q_x)$ para edades de 20 a 60 del modelo lineal generalizado para las mujeres

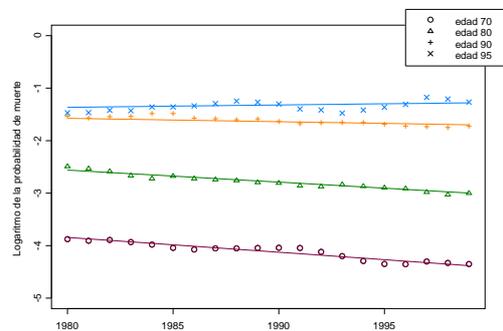


Figura 7.29: Evolución de $\log(q_x)$ para edades de 70 a 95 del modelo lineal generalizado para las mujeres

7.5. Comparación de los modelos

Para la comparación de los modelos utilizaremos los resultados de los contrastes no paramétricos de cada uno de ellos en cada uno de los años. En el caso de los hombres dichos resultados se hallan recogidos en las Tablas 7.5, 7.10 y 7.18 para los modelos Heligman y Pollard, Lee-Carter y lineal generalizado, respectivamente. El valor del estadístico χ^2 muestra, de 1980 a 1986, para el modelo Lee-Carter un mejor ajuste que para los otros dos modelos, después de 1987 a 1992 es mejor Heligman y Pollard, de nuevo de 1993 a 1996 vuelve a ser mejor Lee-Carter pero después para 1997 y 1999 alcanza valores muy superiores a Heligman y Pollard e incluso al modelo lineal generalizado. El modelo lineal generalizado tiene la ventaja de proporcionar los valores de χ^2 equilibrados pero son los más altos salvo para los años 1998 y 1999. El número de desviaciones relativas también se comporta de forma similar para los tres modelos. En cuanto al número de diferencias positivas presenta problemas el modelo lineal generalizado no superando el test para algunos años, en los que se muestra clara tendencia a valores positivos lo que supone que el modelo subestima las defunciones para esos años en una gran parte de edades. Dado lo cual concluimos que el modelo de Heligman y Pollard es preferible en cuanto que proporcionan valores de χ^2 más equilibrados para los diferentes años, manteniendo mayor regularidad en la bondad del ajuste.

En el caso de las mujeres los resultados de los tres modelos se hallan recogidos en las Tablas 7.6, 7.11 y 7.19. Para modelo Heligman y Pollard los resultados de χ^2 y el número de desviaciones relativas es más elevado que para los otros dos. Según estos estadísticos el modelo Lee-Carter resulta mejor que el lineal generalizado para todos los años. Respecto al signo de las diferencias, en el caso del modelo lineal generalizado se muestra tendencia a valores negativos lo que significa que el modelo sobreestima las defunciones en algunos años para un número considerable de edades. El test de rachas no plantea ningún inconveniente para ningún modelo en cada uno de los diferentes años.

Aunque el modelo Heligman y Pollard obtiene buenos resultados para los hombres no los consigue para la mujeres, de lo que se deduce que el modelo Lee-Carter muestra buenos resultados para ambos sexos y es por tanto más recomendable para la construcción de tablas de mortalidad dinámicas.

Sin embargo en el capítulo siguiente utilizaremos el modelo lineal generalizado dinámico con el que hemos obtenido las predicciones para el año 2000 con el fin de obtener valores actuariales que compararemos con la tabla obtenida a partir del modelo lineal generalizado seleccionada en el Capítulo 5. La razón es comparar modelos obtenidos con una metodología similar.

7.6. Métodos basados en factores de reducción de la mortalidad

En esta sección aplicamos el cálculo y modelización de los factores de reducción de la mortalidad que explicamos en la sección 5 del Capítulo 3. Tomamos 1991 como origen ($t = 0$) por estar situado en el centro del periodo base (1990-1992) que vamos a graduar

según la metodología de modelos lineales generalizados, ajustando una función $LGM(0, s)$ como hicimos en el Capítulo 4. Para ambos sexos tomamos la función $LGM(0, 10)$ según las medidas que muestra la Tabla 7.20 donde a partir del polinomio de grado 9 no se produce incremento significativo de la Deviance y además aumenta C_p .

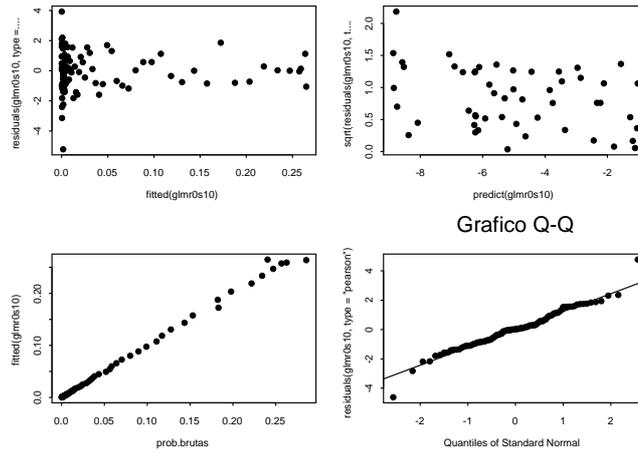


Figura 7.30: Bondad del ajuste $LGM(0, 10)$ para los hombres

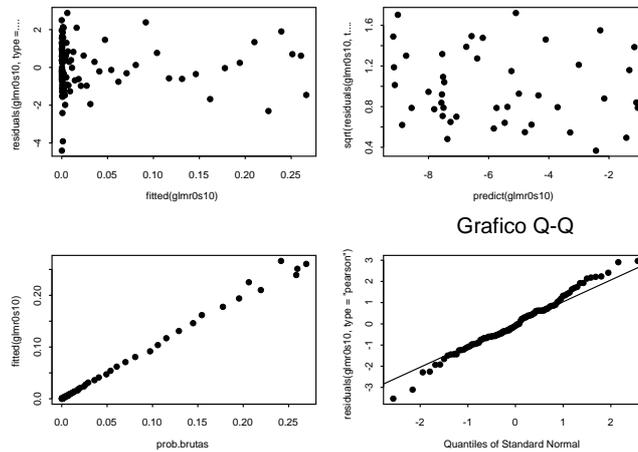


Figura 7.31: Bondad del ajuste $LGM(0, 10)$ para las mujeres

El diagnóstico del modelo según sus gráficos, Figura 7.30 y 7.31, es muy aceptable para ambos sexos. El aspecto de ambos ajustes, así como su comparación gráfica se muestra en la Figura 7.32.

HOMBRES				
	Deviance	Deviance (dif)	p-value	C_p
LGM(0,2)	3992.332			3999.772
LGM(0,3)	1577.065	2415.267	0.000000	1588.226
LGM(0,4)	1404.059	173.006	0.000000	1418.940
LGM(0,5)	1403.952	0.107	0.7437651	1422.553
LGM(0,6)	1238.946	165.006	0.000000	1261.268
LGM(0,7)	863.6133	375.333	0.000000	889.655
LGM(0,8)	288.698	574.915	0.000000	318.460
LGM(0,9)	174.8433	113.855	0.000000	208.326
LGM(0,10)	158.1266	16.717	0.0000434	195.329
LGM(0,11)	158.126	0.001	0.9800434	199.049
LGM(0,12)	157.002	1.124	0.2890478	201.645

MUJERES				
	Deviance	Deviance (dif)	p-value	C_p
LGM(0,2)	4562.565			4569.191
LGM(0,3)	1328.106	3234.459	0.0000	1338.046
LGM(0,4)	908.522	419.584	0.0000	921.775
LGM(0,5)	899.7313	8.791	0.0030276	916.297
LGM(0,6)	502.6986	397.033	0.0000	522.578
LGM(0,7)	442.2805	60.418	0.0000	465.473
LGM(0,8)	292.1144	150.166	0.0000	318.620
LGM(0,9)	174.1442	117.970	0.0000	203.963
LGM(0,10)	156.5909	17.553	0.0000279	189.723
LGM(0,11)	156.5891	0.002	0.9660576	193.034
LGM(0,12)	149.7283	6.861	0.0088108	189.487

Tabla 7.20: Comparación de los modelos $LGM(0, s)$ con diferentes grados para el periodo base 1990-1992

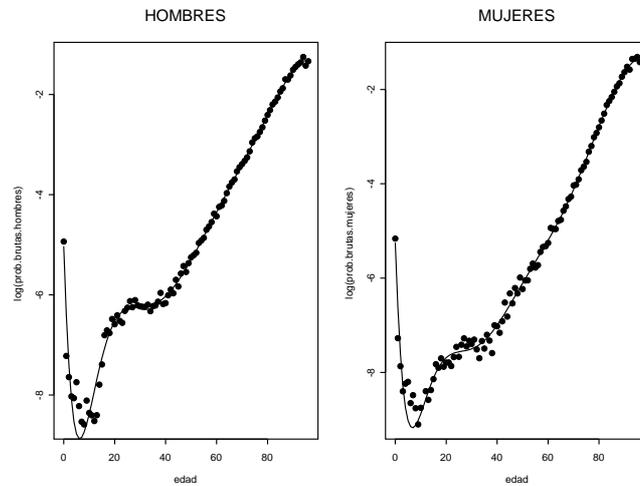


Figura 7.32: Modelos $LGM(0, 10)$ para ambos sexos

Consideramos ahora, la probabilidad de muerte q_{xt} para los años de 1980 a 1999 y sea el tiempo $t = \text{año} - 1991$. En general, los factores de reducción está caracterizado por la ecuación

$$q_{xt} = q_{x0}RF(x, t),$$

para los que vamos a encontrar un modelo en la siguiente sección.

7.6.1. Modelos lineales generalizados

Para modelizar los factores de reducción de la mortalidad hemos de considerar el modelo binomial con link “logit” propuesto para q_{xt} que son las variables respuesta, a las que vamos a ajustar una recta según el predictor,

$$\eta_{xt} = g(q_{x0}) + \beta_x t,$$

donde $g(q_{x0}) = \log\left(\frac{q_{x0}}{1 - q_{x0}}\right)$ pues se calcula aplicando el link a los valores graduados del periodo base. Mostramos los valores de q_{x0} para cada edad, así como los coeficientes β_x en la Tabla 7.21 para cada sexo.

edad	HOMBRES		MUJERES	
	q_{x0}	β_x	q_{x0}	β_x
1	0.0065322035	-0.05636	0.0052510030	-0.05616
2	0.0016625016	0.02013	0.0013528153	0.01685
3	0.0006142660	-0.02057	0.0005023138	-0.00302
4	0.0003082688	-0.02962	0.0002509619	-0.03666

Tabla 7.21: Estimaciones de los q_{x0} y factores de reducción

edad	HOMBRES		MUJERES	
	q_{x0}	β_x	q_{x0}	β_x
5	0.0001980050	-0.08238	0.0001588394	-0.05042
6	0.0001546581	-0.08320	0.0001209864	-0.06952
7	0.0001406393	-0.09232	0.0001062151	-0.08757
8	0.0001435236	-0.06632	0.0001036766	-0.04524
9	0.0001593995	-0.07503	0.0001092298	-0.07277
10	0.0001878301	-0.06235	0.0001212507	-0.01492
11	0.0002300072	-0.03691	0.0001390921	-0.01857
12	0.0002878417	0.00803	0.0001623962	-0.03503
13	0.0003632932	-0.00066	0.0001907265	-0.04027
14	0.0004577326	-0.00439	0.0002233621	0.00411
15	0.0005713331	0.01285	0.0002592133	-0.00925
16	0.0007025905	-0.00242	0.0002968532	-0.00624
17	0.0008481057	0.00016	0.0003346463	-0.00878
18	0.0010027274	-0.00996	0.0003709425	0.00328
19	0.0011600625	-0.00312	0.0004042834	-0.00556
20	0.0013132578	-0.00993	0.0004335744	0.00290
21	0.0014558763	0.00460	0.0004581867	-0.00147
22	0.0015826738	-0.00357	0.0004779778	-0.01024
23	0.0016901255	0.00336	0.0004932420	0.00606
24	0.001776634	0.00403	0.0005046129	-0.00354
25	0.001842436	0.00742	0.0005129493	0.00294
26	0.001889287	0.01070	0.0005192248	-0.00772
27	0.001920034	0.02277	0.0005244392	0.00633
28	0.001938162	0.01788	0.0005295563	0.01126
29	0.001947408	0.02713	0.0005354690	0.00389
30	0.001951453	0.02705	0.0005429864	0.01909
31	0.001953731	0.02441	0.0005528367	-0.01579
32	0.001957325	0.02923	0.0005656811	0.00152
33	0.001964938	0.02891	0.0005821317	0.02794
34	0.001978918	0.02985	0.0006027713	0.02096
35	0.002001310	0.03198	0.0006281720	0.00624
36	0.002033933	0.02896	0.0006589111	-0.00136
37	0.002078452	0.02463	0.0006955832	0.00094
38	0.002136449	0.01967	0.0007388090	0.00386
39	0.002209490	0.01962	0.0007892398	0.00954
40	0.002299184	0.01653	0.0008475585	0.00192
41	0.002407229	0.01602	0.0009144776	-0.00376
42	0.002535451	0.01076	0.0009907344	-0.01869
43	0.002685839	0.00382	0.0010770843	-0.00926
44	0.002860570	0.00867	0.0011742942	-0.00715
45	0.003062027	0.00075	0.0012831370	-0.00415

Tabla 7.21: Estimaciones de los q_{x0} y factores de reducción

edad	HOMBRES		MUJERES	
	q_{x0}	β_x	q_{x0}	β_x
46	0.003292820	0.00806	0.0014043883	-0.01307
47	0.003555794	0.00244	0.001538829	-0.01291
48	0.003854040	-0.00465	0.001687253	-0.01728
49	0.004190907	-0.00638	0.001850486	-0.01805
50	0.004570008	-0.01039	0.002029415	-0.02078
51	0.004995236	0.01535	0.002225025	-0.02485
52	0.005470772	-0.00611	0.002438455	-0.02750
53	0.006001115	-0.00835	0.002671060	-0.02593
54	0.006591102	-0.00995	0.002924496	-0.02588
55	0.007245946	-0.01079	0.003200805	-0.02224
56	0.007971288	-0.00800	0.003502523	-0.02414
57	0.008773250	-0.00756	0.003832800	-0.02085
58	0.009658512	-0.01661	0.004195529	-0.02343
59	0.010634396	-0.01275	0.004595498	-0.02260
60	0.011708967	-0.00986	0.005038556	-0.02056
61	0.012891148	-0.01588	0.005531807	-0.03019
62	0.014190854	-0.01487	0.006083826	-0.02983
63	0.015619135	-0.01968	0.006704915	-0.02744
64	0.017188338	-0.01839	0.007407391	-0.02994
65	0.018912286	-0.01611	0.008205923	-0.02444
66	0.020806477	-0.01494	0.009117917	-0.02414
67	0.022888293	-0.01282	0.010163952	-0.02624
68	0.025177243	-0.01645	0.011368267	-0.02865
69	0.027695213	-0.01231	0.012759285	-0.02280
70	0.03046676	-0.01603	0.01437018	-0.02481
71	0.03351941	-0.01635	0.01623943	-0.02560
72	0.03688402	-0.01643	0.01841136	-0.02285
73	0.04059508	-0.01639	0.02093659	-0.02583
74	0.04469115	-0.01355	0.02387234	-0.02555
75	0.04921517	-0.01538	0.02728246	-0.02373
76	0.05421479	-0.01424	0.03123711	-0.02290
77	0.05974262	-0.01575	0.03581202	-0.02506
78	0.06585620	-0.01351	0.04108703	-0.02339
79	0.07261779	-0.01523	0.04714396	-0.02352
80	0.08009359	-0.01801	0.05406363	-0.02830
81	0.08835218	-0.01991	0.06192194	-0.02444
82	0.09746197	-0.01714	0.07078508	-0.02320
83	0.10748704	-0.01564	0.08070394	-0.020345
84	0.11848093	-0.01807	0.09170786	-0.02030
85	0.13047772	-0.01484	0.10379797	-0.01977
86	0.14347985	-0.01338	0.11694041	-0.01843

Tabla 7.21: Estimaciones de los q_{x0} y factores de reducción

edad	HOMBRES		MUJERES	
	q_{x0}	β_x	q_{x0}	β_x
87	0.15744198	-0.00994	0.13105963	-0.01563
88	0.17225068	-0.00841	0.14603187	-0.01441
89	0.18770016	-0.01084	0.16167895	-0.01470
90	0.20346489	-0.00787	0.17776209	-0.01250
91	0.21907086	-0.00404	0.19397549	-0.01613
92	0.23386848	-0.00045	0.20993919	-0.00380
93	0.2470107	0.00152	0.2251910	-0.00236
94	0.2574414	0.00188	0.2391771	0.00463
95	0.2639002	0.01488	0.2512420	0.01293
96	0.2649574	0.02123	0.2606188	0.01131
97	0.2590998	0.01101	0.2664234	0.01521

Tabla 7.21: Estimaciones de los q_{x0} y factores de reducción

A continuación representamos gráficamente en las Figuras 7.33, 7.34, 7.35 y 7.36 el logaritmo del factor de mejora para algunas de las edades. En éstos podemos apreciar los dos principales problemas de este método,

- Primero, como sucede en la edad 15 si los valores de la graduación del periodo base difieren de los valores brutos porque la función no se ajusta bien para esas edades, los factores de mejora que necesariamente han de pasar por los graduados, no representan bien la tendencia, alejándose claramente de los valores brutos.
- Segundo si la tendencia varía de un año a otro de forma que no es lineal, por ejemplo primero crece y luego decrece, como en las edades 20 a 40, el factor de mejora tampoco será adecuado pues se ajustará a la tendencia predominante o mostrará estabilidad.

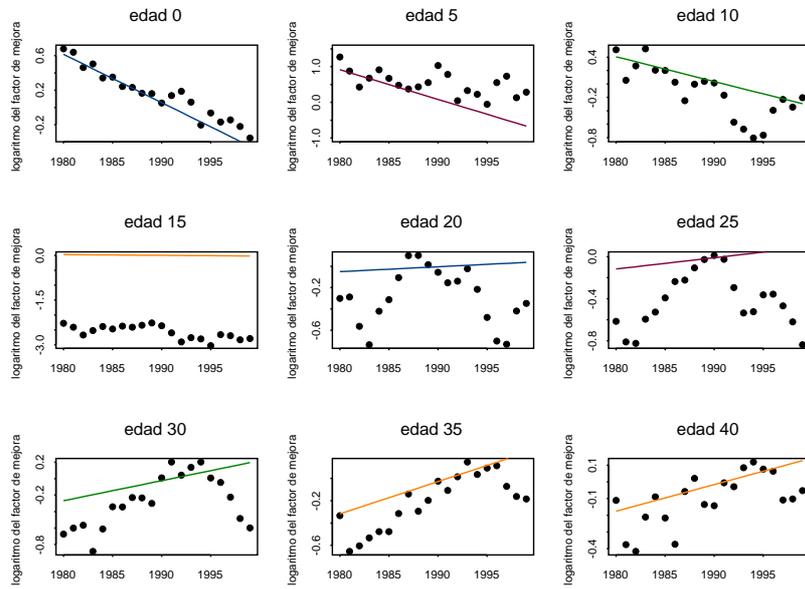


Figura 7.33: Factores de mejora de la mortalidad de edades de 0 a 40 para los hombres

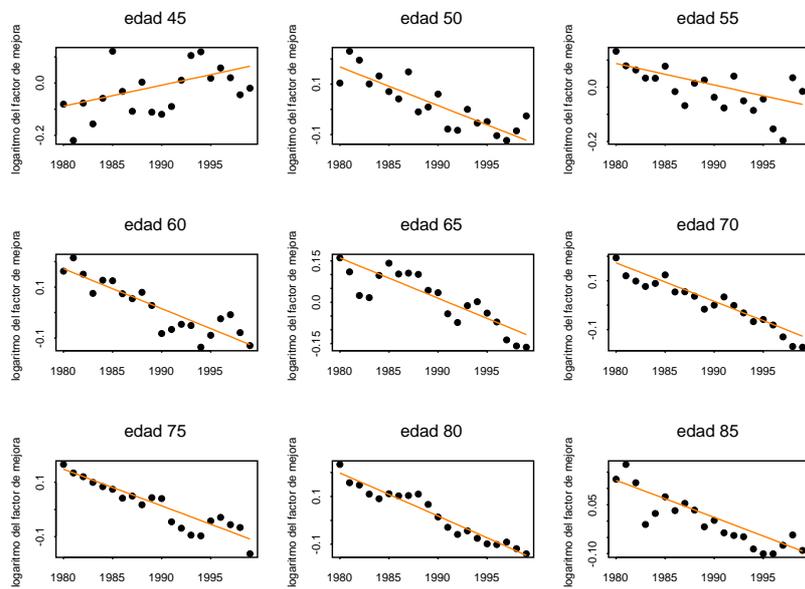


Figura 7.34: Factores de mejora de la mortalidad de edades de 45 a 85 para los hombres

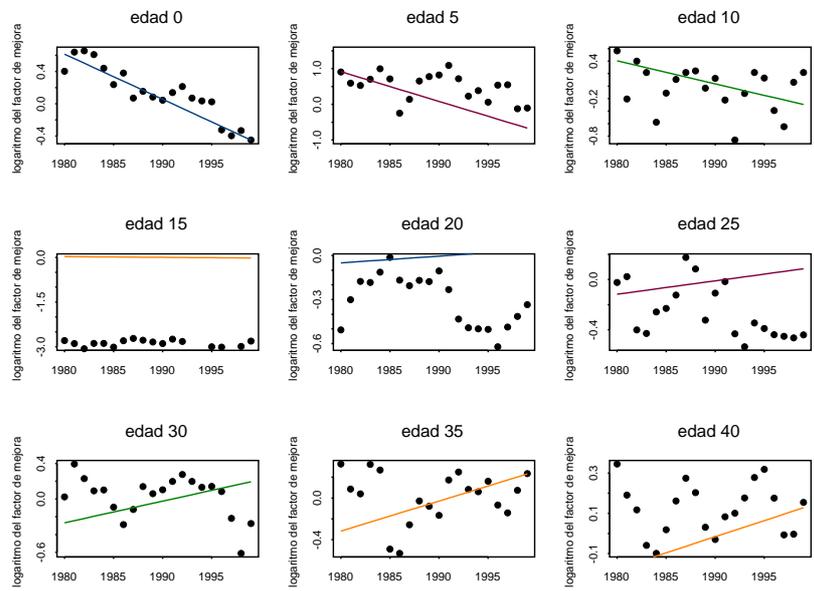


Figura 7.35: Factores de mejora de la mortalidad de edades de 0 a 40 para las mujeres

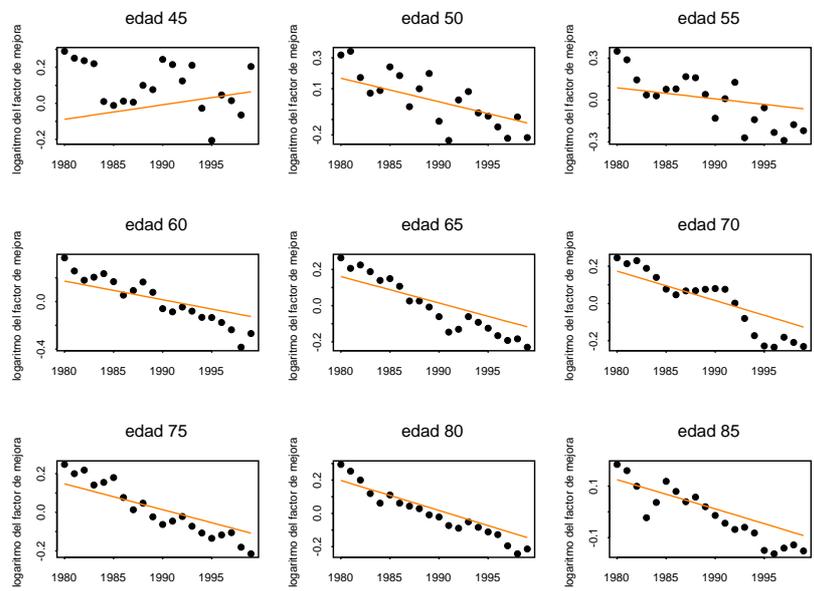


Figura 7.36: Factores de mejora de la mortalidad de edades de 45 a 85 para las mujeres

Capítulo 8

Uso de las tablas de mortalidad en aplicaciones actuariales

8.1. Introducción

El objetivo fundamental de este capítulo es estudiar algunas de las aplicaciones de las tablas de mortalidad construidas en los capítulos precedentes en el campo actuarial, tanto en lo referente al caso individual como al colectivo. En primer lugar se analiza su incidencia sobre las rentas y los seguros individuales y, en segundo lugar su aplicabilidad al caso colectivo. Para ello se ha considerado como sería el comportamiento del cálculo de un hipotético plan de pensiones para el personal de la Universitat de València. Se ha elegido este colectivo por dos razones. La primera de ellas es la posibilidad de disponer de datos de un colectivo real, ya que la simulación de un colectivo ficticio podría resultar difícil puesto que depende de la propia estructura del mismo. En segundo lugar, y aunque se trata de una simplificación de las hipótesis habituales, es importante deducir algunas características en el comportamiento de un colectivo heterogéneo, como es del personal de la Universitat de València.

En la sección 2 se resumen algunos de los principales conceptos sobre rentas y seguros individuales, para después estudiar la variabilidad de su valor en función factores que inciden sobre ellos. En particular el del tipo de interés y los cambios en las probabilidades de supervivencia.

En la sección 3 se realiza una breve presentación de los conceptos fundamentales que intervienen en el cálculo de un plan de pensiones, centrándose su estudio en caso de la capitalización colectiva para realizar una aplicación simulada sobre su afectación al personal de la Universitat de València, así como un análisis de las desviaciones que provocarían los cambios en el uso de unas tablas de mortalidad estáticas por otras dinámicas.

Finalmente, en la sección 4 se exponen las conclusiones más relevantes del estudio.

8.2. Rentas y seguros individuales

En esta sección se resumirá brevemente una parte de la extensa bibliografía que recoge los conceptos básicos para analizar las operaciones demográfico-financieras, que son necesarios a lo largo de todo este capítulo. Para estos conceptos básicos se usará el texto de Villalón (1997).

Las rentas actuariales, a diferencia de las rentas financieras, incorporan el elemento de supervivencia como clave en su operativa. Las rentas se pueden considerar como una sucesión de pagos o cobros periódicos de importe constante o variable. Si los cobros son anuales se les denomina anualidades. Además los pagos pueden hacerse a principio de cada periodo, y se llaman prepagables, o al final de cada periodo, y en este caso reciben el nombre de pospagables.

Las rentas vitalicias, nombre que reciben las rentas actuariales, se aplican sobre el caso de vida, mientras que los seguros son capitales pagaderos por el acaecimiento de un suceso, básicamente el fallecimiento, y por tanto calculan las primas a pagar para una persona en caso de muerte. Una clasificación muy general para ambos conceptos es,

Vitalicia Temporal, cuando se paga o cobra durante un determinado número de periodos, pero mientras que se esta vivo.

Vitalicia de Vida Entera, cuando se paga o cobra hasta el fallecimiento, es decir, hasta la edad ω .

Cada una de las cuales se clasifica a su vez en,

Inmediata, cuando se empieza a pagar o cobrar al iniciarse el pago o cobro al principio de la contratación.

Diferida, cuando ha de pasar un determinado tiempo hasta que se empieza a pagar o cobrar. Existen dos etapas, una primera m , llamada plazo del diferimiento, y las subsiguientes n , denominadas etapa de pagos.

Toda operación de rentas supone la consideración del tiempo, y por tanto de un tipo de interés i para la operación. En caso de que el interés sea pagado por el deudor, dicho intereses está incluido en las cuotas de amortización (constantes o variables) A ; mientras que si el intereses es a favor de quien deposita cuotas constantes A de ahorro, son éstas las que lo perciben.

8.2.1. Determinación del valor de las rentas vitalicias

Se iniciará este repaso con el caso más sencillo, que consiste en obtener lo que se denomina *capital diferido*. Se trata de una capital K , cuyo pago es un suceso aleatorio, condicionado a que una persona de edad x este viva al cabo de n años para recibirlo. Por tanto, el precio (según el principio de equivalencia de la operación) para esta eventualidad vendrá dado por la esperanza matemática, es decir,

$$U = K(1 + i)^{-n} {}_n p_x.$$

Donde ${}_n p_x$ es la probabilidad de supervivencia introducida en el Capítulo 1 según la expresión (1.1), aplicada para $t = n$. El factor $(1+i)^{-n}$ actualiza el capital depositado en un momento futuro (n) hasta un momento presente (0) y no se consideran los intereses acumulados de los n periodos recorridos. Se le suele denotar $v = (1+i)^{-1}$, de forma que nos queda,

$$U = K v^n {}_n p_x.$$

Supuesto que $K = 1$ y además **constante**, entonces aparece el denominado *factor de actualización demográfico financiero*

$${}_n E_x = (1+i)^{-n} {}_n p_x = (1+i)^{-n} \frac{l_{x+n}}{l_x}.$$

Es muy importante señalar que esta operación, y en general todas las operaciones actuariales, solamente tienen sentido desde el punto de vista técnico si se realizan para una colectividad grande de personas, que en aplicación de la teoría de los grandes números producen la compensación de las desviaciones (positivas y negativas) sobre estos sucesos.

Este factor de actualización permite deducir su inverso, el *factor de capitalización demográfico financiero*, es decir, cuando queremos calcular un capital futuro después de n periodos

$${}_n I_x = (1+i)^n \frac{l_x}{l_{x+n}}.$$

Para facilitar los cálculos, como ya avanzábamos en el Capítulo 1, se utilizan los *símbolos de conmutación*, que incluyen los dos factores de forma simultánea, el demográfico y el financiero. Por ello en las tablas aparecen valores calculados según un *interés técnico* (cuanto mayor sea el interés menor será el valor).

Los símbolos de conmutación aparecieron a finales del siglo XIX (obra de Tiens, Barret Bayly), vease Nieto y Vegas (1993), a fin de facilitar la labor del cálculo de los valores actuariales.

El primer factor a considerar será el símbolo, D_x cuya expresión es

$$D_x = l_x v^x = l_x (1+i)^{-x},$$

de cuya expresión se deduce que representará el número de supervivientes l_x descontados al tipo de interés por un tiempo equivalente a su edad (v^x). Con este símbolo podremos expresar el factor de actualización como

$$\begin{aligned} {}_n E_x &= (1+i)^{-n} \frac{l_{x+n}}{l_x} \frac{v^x}{v^x} \\ &= \frac{l_{x+n} v^{x+n}}{l_x v_x} \\ &= \frac{D_{x+n}}{D_x}. \end{aligned}$$

En el epígrafe anterior ya se realizó una somera clasificación de las rentas y por ello solamente nos referiremos aquí a las expresiones para las rentas constantes y unitarias

en términos de los símbolos de conmutación. El cálculo del valor actual de una renta es similar al cálculo del valor actual de un renta financiera, es decir, trasladar hasta el momento actual el valor de cada una de las cuotas anuales utilizando para ello el factor de actualización demográfico financiero en lugar del factor financiero (rentas financieras).

Sea una renta vida entera, inmediata y prepagable \ddot{a}_x , su valor actual se obtiene,

$$\begin{aligned} \ddot{a}_x &= {}_0E_x + {}_1E_x + \dots + {}_{\omega-x}E_x \\ &= \frac{D_x + D_{x+1} + \dots + D_{\omega}}{D_x}. \end{aligned}$$

Para simplificar estos cálculos se introduce el símbolo,

$$N_x = D_x + D_{x+1} + \dots + D_{\omega}, \tag{8.1}$$

de forma que la expresión ¹ queda como

$$\ddot{a}_x = \frac{N_x}{D_x} \tag{8.2}$$

Análogamente se pueden obtener las expresiones, en función de los símbolos de conmutación, para el resto de rentas con pagos constantes,

De vida entera	inmediata	prepagable	$\ddot{a}_x = \frac{N_x}{D_x}$
		pospagable	$a_x = \frac{N_{x+1}}{D_x}$
	diferida	prepagable	${}_m/\ddot{a}_x = \frac{N_{x+m}}{D_x}$
		pospagable	${}_m/a_x = \frac{N_{x+m+1}}{D_x}$
Temporal	inmediata	prepagable	$\ddot{a}_{x:\overline{n} } = \frac{N_x - N_{x+n}}{D_x}$
		pospagable	$a_{x:\overline{n} } = \frac{N_{x+1} - N_{x+n+1}}{D_x}$
	diferida	prepagable	${}_m/n\ddot{a}_x = \frac{N_{x+m} - N_{x+m+n}}{D_x}$
		pospagable	${}_m/n a_x = \frac{N_{x+m+1} - N_{x+m+n+1}}{D_x}$

Si se considera el caso con capitales variables, también es sencillo obtener las expresiones de las rentas. Concretamente dos casos comunes, rentas cuyos términos varían periodo a periodo en progresión aritmética y aquellas que lo hacen en progresión geométrica.

¹A veces se utiliza la notación \ddot{a} no obstante ésta suele estar reservada para rentas financieras y \ddot{a} para actuariales

Si los capitales a pagar (cobrar) a partir de la edad x varían en progresión aritmética definida por,

$$\alpha + tr, \text{ donde } t = 0, 1, \dots, \omega - x,$$

introducimos un nuevo símbolo, S_x , para la simplificación de las expresiones de las rentas,

$$S_x = N_x + N_{x+1} + \dots + N_\omega.$$

Las rentas correspondientes a los pagos **variables en progresión aritmética** son,

De vida entera	inmediata	prepagable	$(V\ddot{a})_x = \alpha \frac{N_x}{D_x} + r \frac{S_{x+1}}{D_x}$
		pospagable	$(Va)_x = \alpha \frac{N_{x+1}}{D_x} + r \frac{S_{x+2}}{D_x}$
	diferida	prepagable	${}_m/(V\ddot{a})_x = \alpha \frac{N_{x+m}}{D_x} + r \frac{S_{m+1}}{D_x}$
		pospagable	${}_m/(Va)_x = \alpha \frac{N_{x+m+1}}{D_x} + r \frac{S_{m+2}}{D_x}$
Temporal	inmediata	prepagable	$(V\ddot{a})_{\overline{x:n} } = \alpha \frac{N_x - N_{x+n}}{D_x}$ $+ r \frac{S_{x+1} - S_{x+n}}{D_x} - \frac{(n-1)N_{x+n}}{D_x}$
		pospagable	$(Va)_{\overline{x:n} } = \alpha \frac{N_{x+1} - N_{x+n+1}}{D_x}$ $+ r \frac{S_{x+2} - S_{x+n+1}}{D_x} - \frac{(n-1)N_{x+n+1}}{D_x}$
	diferida	prepagable	${}_m/n(V\ddot{a})_x = \alpha \frac{N_{x+m} - N_{x+m+n}}{D_x}$ $+ r \frac{S_{x+m+1} - S_{x+m+n}}{D_x} - \frac{(n-1)N_{x+m+n}}{D_x}$
		pospagable	${}_m/n(Va)_x = \alpha \frac{N_{x+m+1} - N_{x+m+n+1}}{D_x}$ $+ r \frac{S_{x+m+2} - S_{x+m+n+1}}{D_x} - \frac{(n-1)N_{x+m+n+1}}{D_x}$

Si consideramos $\alpha = r = 1$ en la renta de vida entera inmediata, tenemos que

$$(I\ddot{a})_x = \frac{N_x}{D_x} + \frac{S_{x+1}}{D_x} = \frac{S_x}{D_x}$$

$$(Ia)_x = \frac{N_{x+1}}{D_x} + \frac{S_{x+2}}{D_x} = \frac{S_{x+1}}{D_x}.$$

En el caso en que los capitales a pagar (cobrar) a partir de la edad x varían en progresión geométrica definida por,

$$\alpha r^t, \quad \text{donde } t = 0, 1, \dots, \omega - x,$$

considerando $\alpha = 1$ y $r = 1 + \beta$, se pueden sustituir los símbolos N_x y D_x por N_x^* y D_x^* , donde el * indica un nuevo tipo de interés,

$$j = \frac{1+i}{1+\beta} - 1,$$

y con estos símbolos se obtienen los valores actuales de las rentas variables con pagos en progresión geométrica equivalentes a los de las rentas constantes.

8.2.2. Seguros o capitales pagaderos por fallecimiento

Al igual que para las rentas se iniciará este repaso con un caso sencillo de cálculo de prima única, pagadera al comienzo del año a la edad x , que garantiza el pago a un beneficiario de un capital si el deceso de una persona se produce dentro del año, es decir, antes de alcanzar la edad $x + 1$. Cuando el capital unitario sea pagadero sólo al final del año, cualquiera que sea el momento del fallecimiento del asegurado, se calcula mediante

$${}_1A_x = q_x(1+i)^{-1} = \frac{d_x}{l_x}v.$$

Introducimos un nuevo símbolo para simplificar su expresión

$$C_x = d_x v^{x+1},$$

de forma que la expresión se simplifica,

$${}_1A_x = \frac{d_x}{l_x} v \frac{v^x}{v^x} = \frac{C_x}{D_x}.$$

Si se generaliza este caso para más años, el cálculo de la prima única o valor actual que debe pagar una persona de edad actual x para el caso en que su muerte se produzca en cualquier momento a partir de la edad x , a la edad x o a la $x + 1$ o a la $x + 2 \dots$, se lleva a cabo teniendo en cuenta que los sucesos son disjuntos y por tanto dicha prima es la suma de las esperanzas o primas anuales de cada una de las alternativas

$$A_x = \frac{d_x}{l_x}v + \frac{d_{x+1}}{l_x}v^2 + \dots + \frac{d_\omega}{l_x}v^{\omega+1-x}.$$

Para simplificar esta expresión se recurre a unos nuevos símbolos de conmutación. El primero de ellos es,

$$M_x = C_x + C_{x+1} + \dots + C_\omega$$

De forma análoga se pueden obtener las expresiones de los seguros en función de los símbolos de conmutación del resto de seguros valorados a final del periodo año en el que se produce la muerte,

De vida entera	inmediato	$A_x = \frac{M_x}{D_x}$
	diferido	${}_m/A_x = \frac{M_{x+m}}{D_x}$
Temporal	inmediato	${}_nA_x = \frac{M_x - M_{x+n}}{D_x}$
	diferido	${}_{m/n}A_x = \frac{M_{x+m} - M_{x+m+n}}{D_x}$

Si se consideran ahora seguros para el caso de fallecimiento a capitales variables en progresión aritmética o geométrica, al igual que en el caso anterior se hace necesario introducir un nuevo símbolo

$$R_x = M_x + M_{x+1} + \dots + M_\omega,$$

y con ello se tienen las expresiones para los seguros, como por ejemplo de vida entera

De vida entera	en progresión aritmética	$(VA)_x = \alpha \frac{M_x}{D_x} + r \frac{R_{x+1}}{D_x}$
	en progresión geométrica	$(V_rA)_x = \frac{1}{1+r} \frac{M_x^*}{D_x^*}$

Si $\alpha = r = 1$ en el seguro de vida entera variable en progresión aritmética, se tiene el denominado *seguro increasing*

$$(IA)_x = \frac{M_x}{D_x} + \frac{R_{x+1}}{D_x} = \frac{R_x}{D_x}.$$

De las expresiones anteriores solamente se ha analizado el caso de la prima única que debería pagar un asegurado de edad x al principio de la operación. Para la generalización de este pago mediante primas periódicas P a pagar mientras sobreviva el asegurado, será suficiente con combinar los conceptos de rentas y de seguros, igualando las expresiones de la prestación y la contraprestación según el llamado principio de equivalencia.

Así, por ejemplo, la prima periódica vitalicia a pagar por un asegurado de edad x se obtendría al igualar la prima a satisfacer, P , por la renta vitalicia prepagable y la contraprestación A_x ,

$$P\ddot{a}_x = A_x.$$

que sería equivalente a,

$$P = \frac{A_x}{\ddot{a}_x} = \frac{M_x}{N_x}$$

8.2.3. Rentas fraccionadas

Hasta este momento solamente se han considerado las rentas con periodo anual, pero en el mercado del seguro se utiliza el término fraccionado. Si las bases de cálculo (anual,

trimestral, mensual, etc.) coinciden con los periodos de la operación, no habrá problema, pero cuando no coinciden se presenta el problema de su adaptación puesto que las tablas actuariales vienen en bases anuales. La solución más técnica, sin necesidad de volver a calcular la tabla, es la de reducir la operación o el valor fraccionario a operaciones o valores equivalentes. No obstante, la aparición de los ordenadores ha facilitado enormemente los cálculos cayendo en desuso las expresiones de aproximación que se utilizaban.

Por ejemplo, si se consideran h periodos en el año para aproximar la renta fraccionaria pospagable de vida entera a partir de la anual, tenemos la expresión²

$$\ddot{a}_x^{(h)} \simeq \ddot{a}_x - \frac{h-1}{2h}.$$

8.2.4. Análisis de sensibilidad y estabilidad

Las expresiones que se han obtenido para las rentas y seguros individuales están fuertemente influenciadas por el tipo de interés y las probabilidades de supervivencia. Y aunque han sido ampliamente estudiadas estas expresiones, no lo ha sido tanto su variación según el tipo de interés y su variación según las probabilidades de muerte e incluso el tiempo del calendario. En este sentido se pueden citar los trabajos de Betzuen (1989) y más recientemente el trabajo de Olivieri (2001).

En esta sección se estudiara precisamente la influencia del tipo de interés y de las probabilidades de muerte sobre las rentas y seguros. Desde el punto de vista matemático la forma de analizar la influencia que tienen las variaciones infinitesimales de una variable sobre una función de la misma, es calcular la expresión de su derivada. En cada uno de los epígrafes siguientes, vamos a calcular las derivadas justificando posteriormente el uso de métodos aproximados de diferencias finitas.

Respecto a la variación del interés técnico i

Para obtener las derivadas de las rentas respecto al tipo de interés i , comenzaremos por calcular la del símbolo D_x . La expresión de D_x en función del tipo de interés es,

$$D_x = l_x v^x = l_x (1+i)^{-x}$$

de forma que su derivada es,

$$\begin{aligned} \frac{\partial D_x}{\partial i} &= -x l_x (1+i)^{-x-1} \\ &= \frac{-x l_x}{(1+i)^{x+1}} \end{aligned}$$

$$\begin{aligned} \Delta D_x &= -x l_x (1+i)^{-x-1} \Delta i \\ &= -x D_x (1+i)^{-1} \Delta i \end{aligned}$$

²La fórmula es la de Woolhouse cuya expresión es $\ddot{a}_x^{(h)} \simeq \ddot{a}_x - \frac{h-1}{2h} - \frac{h^2-1}{12h^2}(\mu_x + \delta)$ aunque por aproximación se utiliza la simplificada que ponemos en el texto

es decir,

$$D_x^{i'} = D_x \left(1 + \frac{x(i - i')}{1 + i} \right)$$

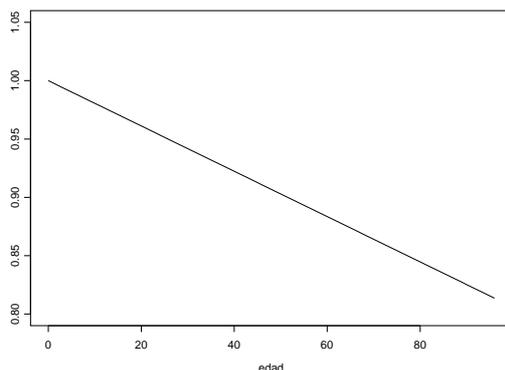


Figura 8.1: Variación de D_x por cambio de interés del 3 al 3.2%

En la Figura 8.1 aparece la variación de D_x con relación a un cambio en el tipo de interés. En particular al variar desde el 3% al 3.2%. Como se desprende de la expresión anterior la variación de D_x solamente depende de la edad x y de la relación entre $\frac{i - i'}{1 + i}$ y no de p_x , por tanto es indiferente que estas variaciones afecten a las tablas de hombre como de mujeres.

Como se desprende de forma inmediata del gráfico de la Figura 8.1 y de la inspección de la expresión de la variación de D_x , hay decrecimiento lineal de su valor debido a esa dependencia de la diferencia de los tipos de interés, pero amplificada por la edad, por tanto a mayor edad mayor variación del valor de D_x .

De la expresión anterior se puede obtener la derivada del símbolo N_x dado por la expresión (8.1) pues no es más que la suma de las correspondientes derivadas,

$$\begin{aligned} \frac{\partial N_x}{\partial i} &= \frac{\partial D_x}{\partial i} + \frac{\partial D_{x+1}}{\partial i} + \dots + \frac{\partial D_\omega}{\partial i} \\ &= -\frac{x}{(1+i)} D_x - \frac{x+1}{(1+i)} D_{x+1} - \frac{x+2}{(1+i)} D_{x+2} - \dots - \frac{\omega}{(1+i)} D_\omega \\ &= -\frac{1}{(1+i)} (xD_x + xD_{x+1} + D_{x+1} + xD_{x+2} + D_{x+2} + D_{x+2} + \dots) \\ &= -\frac{1}{(1+i)} [x(D_x + D_{x+1} + xD_{x+2} + \dots + D_\omega) + (D_{x+1} + \dots + D_\omega) \\ &\quad (D_{x+2} + \dots + D_\omega) + \dots] \\ &= -\frac{1}{(1+i)} (xN_x + N_{x+1} + N_{x+2} + \dots + N_\omega), \end{aligned} \tag{8.3}$$

si a esta expresión se le resta y suma N_x , teniendo en cuenta la expresión de S_x considerada anteriormente, se tiene

$$\begin{aligned}\frac{\partial N_x}{\partial i} &= -\frac{1}{(1+i)}(xN_x - N_x + N_x + N_{x+1} + N_{x+2} + \dots + N_\omega) \\ &= -\frac{1}{(1+i)}(N_x(x-1) + S_x).\end{aligned}\tag{8.4}$$

Si se consideran dos tipos de interés i e i' , y sus correspondientes símbolos de comutación N_x y N'_x el incremento vendría dado por $N'_x - N_x \simeq \frac{\partial N_x}{\partial i}(i' - i)$, expresión de la cual que al despejar N'_x y sustituir (8.3) se obtiene,

$$\begin{aligned}N'_x &\simeq N_x + \frac{i-i'}{(1+i)}(xN_x + N_{x+1} + N_{x+2} + \dots + N_\omega) \\ &= N_x \left(1 + x \frac{(i-i')}{(1+i)}\right) + \frac{(i-i')}{(1+i)}(N_{x+1} + N_{x+2} + \dots + N_\omega) \\ &= N_x \left(1 + x \frac{(i-i')}{(1+i)}\right) + \frac{(i-i')}{(1+i)}S_{x+1}\end{aligned}$$

y sustituyendo (8.4) ,

$$\begin{aligned}N'_x &\simeq N_x + \frac{(i-i')}{(1+i)}(N_x(x-1) + S_x) \\ &= N_x \left(1 + \frac{(i-i')}{(1+i)}(x-1)\right) + \frac{(i-i')}{(1+i)}S_x\end{aligned}$$

Las derivadas de las rentas se calculan a partir de la derivada de un cociente. Tomando por ejemplo la expresión de la renta de vida entera inmediata prepagable (8.2) se tiene que

$$\begin{aligned}\frac{\partial \ddot{a}_x}{\partial i} &= \frac{D_x \frac{\partial N_x}{\partial i} - N_x \frac{\partial D_x}{\partial i}}{D_x^2} \\ &= \frac{D_x \left(-\frac{1}{(1+i)}(N_x(x-1) + S_x)\right) - N_x \left(\frac{-x}{1+i}\right) D_x}{D_x^2} \\ &= \frac{-\frac{D_x}{1+i}(N_x(x-1) + S_x - xN_x)}{D_x^2} \\ &= -\frac{1}{1+i} \left(\frac{N_x x - N_x + S_x - xN_x}{D_x}\right) \\ &= -\frac{1}{1+i} \frac{S_{x+1}}{D_x}.\end{aligned}$$

A partir de ésta se obtiene, considerando los incrementos como en el caso anterior,

$$\ddot{a}'_x = \ddot{a}_x + \frac{(i-i')}{(1+i)} \frac{S_{x+1}}{D_x}$$

La complicada expresión obtenida para esta derivada, justifica el hecho de calcular la renta para varios valores del interés i y analizar el incremento sufrido. Es decir, sustituimos su expresión por una aproximación discreta.

Respecto a la variación de las probabilidades de supervivencia p_x

Para obtener las derivadas de las rentas respecto a las probabilidades de supervivencia p_x se actuará de forma análoga. La expresión de D_x en función de dichas probabilidades es,

$$\begin{aligned} D_x &= l_x(1+i)^{-x} \\ &= l_{x-1}p_{x-1}(1+i)^{-x} \end{aligned} \quad (8.5)$$

$$\begin{aligned} &\vdots \\ &= l_0p_0p_1 \dots p_{x-2}p_{x-1}(1+i)^{-x}. \end{aligned} \quad (8.6)$$

A partir de (8.6) se deduce que el símbolo D_x depende de todas las probabilidades, $p_{x-1}, p_{x-2}, \dots, p_1, p_0$, de forma que la variación de cualquiera de ellas le afectaría. Deberíamos pues calcular su derivada respecto de todas y cada una de ellas, teniendo en cuenta que las probabilidades dependen unas de otras,

$$\begin{aligned} \frac{\partial D_x}{\partial p_{x-1}} &= l_{x-1}(1+i)^{-x} \\ \frac{\partial D_x}{\partial p_{x-2}} &= l_{x-1}(1+i)^{-x} \frac{\partial p_{x-1}}{\partial p_{x-2}} \\ &\vdots \end{aligned}$$

La conclusión es que obtener las derivadas de esta forma se convierte en una tarea que además de no ser fácil, conduce a una expresión difícil de interpretar y manejar. Con mayor motivo en este caso, calcularemos las rentas y seguros para valores diferentes de los p_x y analizaremos sus diferencias. Para obtener la derivada de D_x se consideran las probabilidades independientes y se deriva 8.5 respecto de p_{x-1} que hemos denotado p , obteniendo

$$\frac{\partial D_x}{\partial p} = l_{x-1}(1+i)^{-x},$$

y aproximando por incrementos,

$$\Delta D_x \simeq l_{x-1}(1+i)^{-x} \Delta p,$$

de donde

$$\begin{aligned} D'_x &\simeq D_x + l_{x-1}(1+i)^{-(x-1)}(1+i)^{-1} \Delta p \\ &= D_x + D_{x-1}(1+i)^{-1} \Delta p. \end{aligned}$$

A partir de aquí obtenemos

$$\begin{aligned}\frac{\partial N_x}{\partial p} &= \frac{\partial D_x}{\partial p} + \frac{\partial D_{x+1}}{\partial p} + \dots + \frac{\partial D_\omega}{\partial p} \\ \frac{\partial N_x}{\partial p} &= l_{x-1}(1+i)^{-x} + l_x(1+i)^{-(x+1)} + l_{x+1}(1+i)^{-(x+2)} + \dots \\ &= (1+i)^{-1}(D_x - 1 + D_x + D_x + 1 + \dots) = (1+i)^{-1}N_{x-1}.\end{aligned}$$

Así, transformando la derivada en incrementos queda,

$$\Delta N_x \simeq (1+i)^{-1}N_{x-1}\Delta p.$$

Con ello, para determinar el nuevo valor del símbolo N_x frente a una variación de la probabilidad de supervivencia, se tiene que

$$N'_x \simeq N_x + (1+i)^{-1}N_{x-1}\Delta p.$$

En las aplicaciones siguientes las variaciones sobre las probabilidades han sido aditivas y multiplicativas, considerando, respectivamente, un cambio constante en cada una de las probabilidades de +0.01 % o corrigiendo todas las probabilidades por un factor 1.0001 que supone un incremento relativo de 0.01 % .

Finalmente, la derivada de la renta inmediata \ddot{a}_x respecto a una variación en la probabilidad, tendrá la siguiente expresión

$$\begin{aligned}\frac{\partial \ddot{a}_x}{\partial p} &= \frac{D_x \frac{\partial N_x}{\partial p} - N_x \frac{\partial D_x}{\partial p}}{D_x^2} \\ &= \frac{D_x(1+i)^{-1}N_{x-1} - N_x D_{x-1}}{D_x^2} \\ &= (1+i)^{-1} \frac{(D_x N_{x-1} - N_x D_{x-1})}{D_x^2}.\end{aligned}\tag{8.7}$$

Y transformando la derivada en incrementos queda,

$$\Delta \ddot{a}_x \simeq (1+i)^{-1} \frac{(D_x N_{x-1} - N_x D_{x-1})}{D_x^2} \Delta p.$$

Con ello, para determinar el nuevo valor del símbolo \ddot{a}_x frente a una variación de la probabilidad de supervivencia, se tiene que

$$\ddot{a}'_x \simeq \ddot{a}_x + \simeq (1+i)^{-1} \frac{(D_x N_{x-1} - N_x D_{x-1})}{D_x^2} \Delta p.$$

8.2.5. Aplicación

En este apartado se presentarán una serie de gráficas que tratan de explicar el comportamiento de las rentas y de las primas únicas de los seguros ante las variaciones de los tipos de interés y el uso de los cambios en las probabilidades de supervivencia, y una comparación de los valores mediante el uso de las tablas estáticas y dinámicas calculadas mediante modelos lineales generalizados en los capítulos anteriores. Las figuras que se presentan tienen en común que todas están realizadas sobre la base de un tipo de interés del 3% y una variación de 0.2%, es decir para valores del 2.8 y 3.2%.

Las tablas utilizadas han sido tanto las estáticas como las dinámicas construidas en los capítulos anteriores. En cada una de las figuras se especifica el tipo de tabla. Además estas figuras están referidas tanto a hombres como a mujeres, para ver la incidencia de estas variaciones sobre cada tipo de tabla. Las Figuras 8.2 y 8.3 analizan la variación del valor de

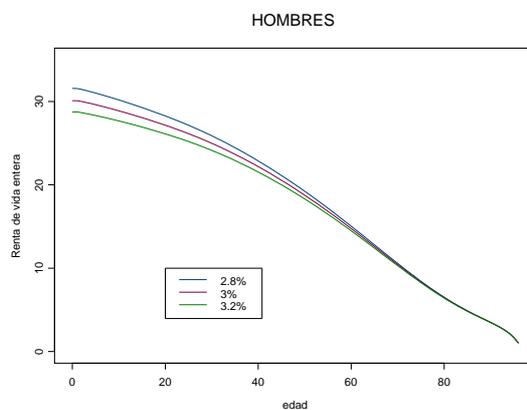


Figura 8.2: Variación del valor de la renta actual ante cambios en el tipo de interés con tabla estática para los hombres

una renta vitalicia ante los cambios en los tipos de interés con tabla *estática*. Como puede observarse estos cambios afecta en la misma medida a hombres y mujeres. Las Figuras 8.4 y 8.5 analizan la variación del valor de una renta vitalicia ante los cambios en los tipos de interés con tabla *dinámica*. Como puede observarse el valor inicial (para la edad 0) es mucho mayor para las mujeres que para los hombres, con lo que la tendencia decreciente y cóncava es más pronunciada en el caso de los hombres que para las mujeres. Las Figuras 8.6 y 8.7, recogen la variación de la prima única de un seguro vida entera para diferentes edades ante variaciones de los tipos de interés cuando se han utilizado tablas *estáticas*. La primera idea relevante es que el valor de la prima es ligeramente mayor en el caso de los hombres que en el de las mujeres, como era de esperar, pero a partir de ahí la incidencia de los cambios es ligeramente mayor en el caso de las mujeres hasta alcanzar un valor máximo próximo a la unidad. Las Figuras 8.8 y 8.9 muestran las variaciones en el valor de una prima única del seguro vida entera cuando se usan tablas *dinámicas*. Los resultados que se pueden comentar en este caso son similares a los realizados anteriormente, destacando

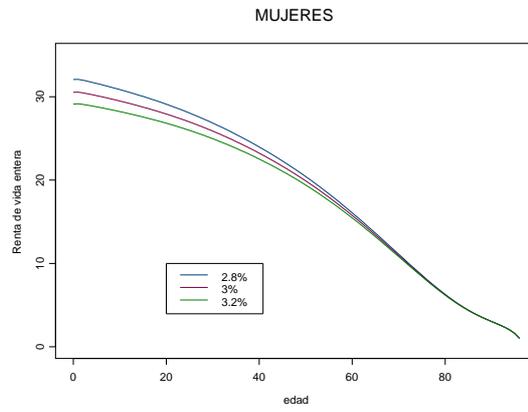


Figura 8.3: Variación del valor de la renta actual ante cambios en el tipo de interés con tabla estática para las mujeres

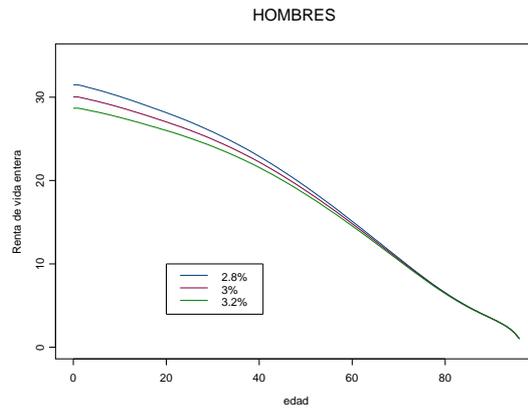


Figura 8.4: Variación del valor de la renta actual ante cambios en el tipo de interés con tabla dinámica para los hombres

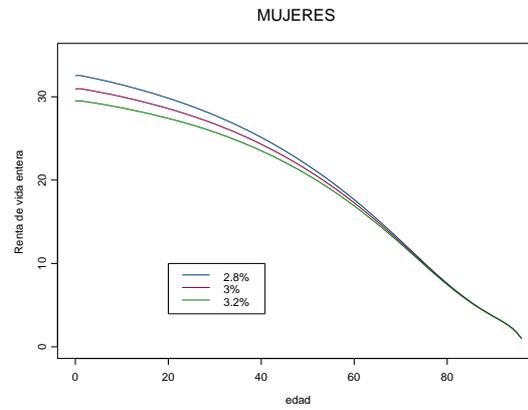


Figura 8.5: Variación del valor de la renta actual ante cambios en el tipo de interés con tabla dinámica para las mujeres

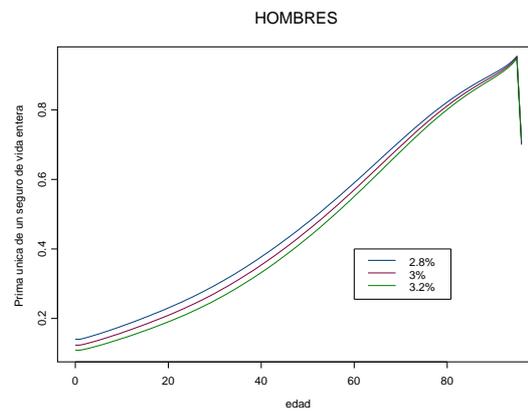


Figura 8.6: Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla estática para los hombres

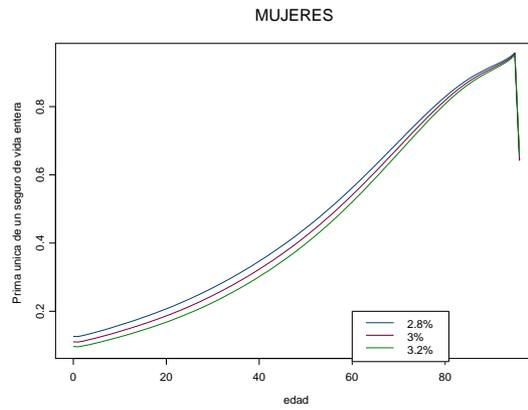


Figura 8.7: Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla estática para las mujeres

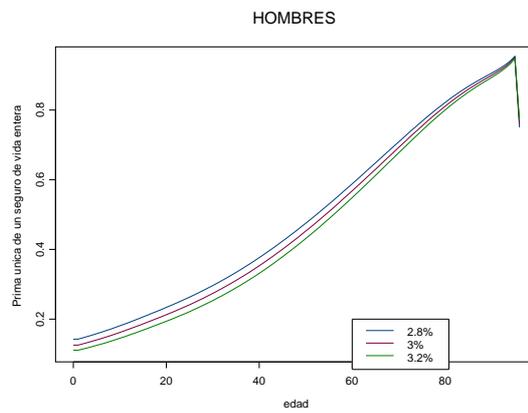


Figura 8.8: Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla dinámica para los hombres

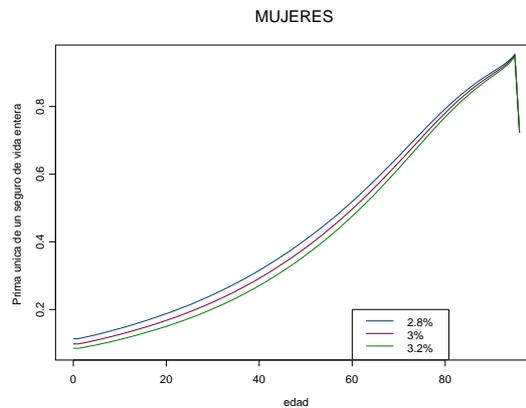


Figura 8.9: Variación del valor de la prima de seguro ante cambios en el tipo de interés con tabla dinámica para las mujeres

nuevamente la suavidad y constancia del máximo en el caso de los hombres frente a las mujeres, donde se aprecia una mayor concavidad y un crecimiento más acentuado en las edades superiores.

En los gráficos siguientes se analiza la incidencia de variaciones constantes sobre las probabilidades de supervivencia, tanto de hombres como de mujeres. Las Figuras 8.10 y

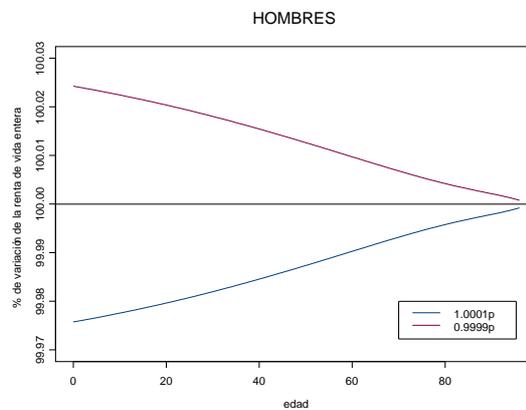


Figura 8.10: Variación de la renta cuando varían la probabilidades de forma constante para los hombres

8.11 muestran la convergencia de las variaciones del valor actual de una renta inmediata vida entera cuando las probabilidades de supervivencia están afectadas por el producto de un factor constante (0.0001), tanto positiva como negativamente, y de ahí el proceso de convergencia. Esta convergencia es prácticamente idéntica en el caso de hombres y mujeres.

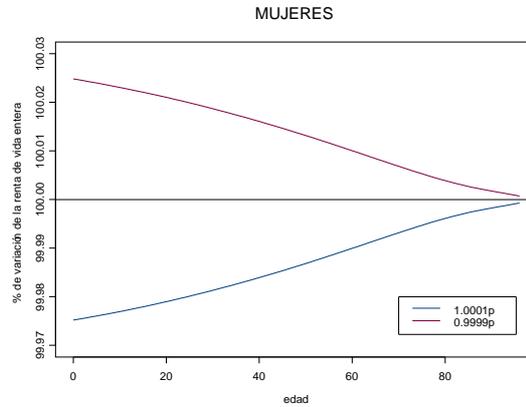


Figura 8.11: Variación de la renta cuando varían la probabilidades de forma constante para las mujeres

Las Figuras 8.12 y 8.13 analizan el mismo proceso de afectación de las probabilidades de

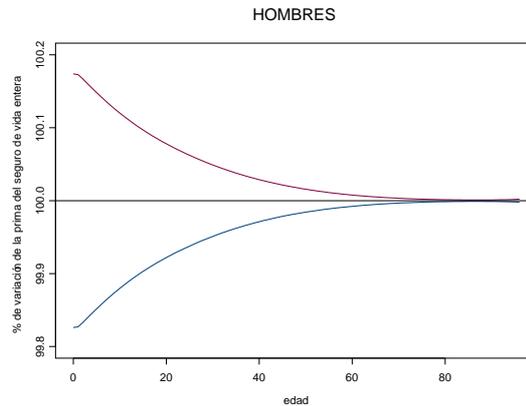


Figura 8.12: Variación de la prima de seguro cuando varían la probabilidades de forma constante para los hombres

supervivencia, pero aplicado al caso de la prima única de un seguro vida entera. Lo más reseñable en este caso es que los valores iniciales de las tablas de mujeres son superiores a los de las tablas de hombres, pero a media que avanza la edad tienden a converger al mismo valor. Para las rentas y los seguros cuando las variaciones son aditivas, es decir, cuando a las probabilidades de la tabla se le añade una cantidad constante (0.0001), tanto en positivo como en negativo, se han obtenido las Figuras 8.14 y 8.15 (para las rentas) y las 8.16 y 8.17 (para los seguros). Si se comparan estas figuras con las anteriores, se puede comprobar una gran similitud entre ellas.

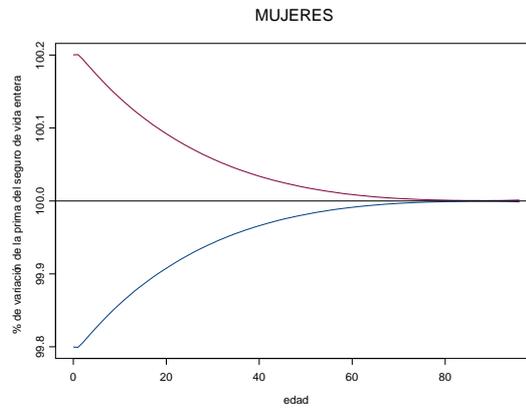


Figura 8.13: Variación de la prima de seguro cuando varían la probabilidades de forma constante para las mujeres

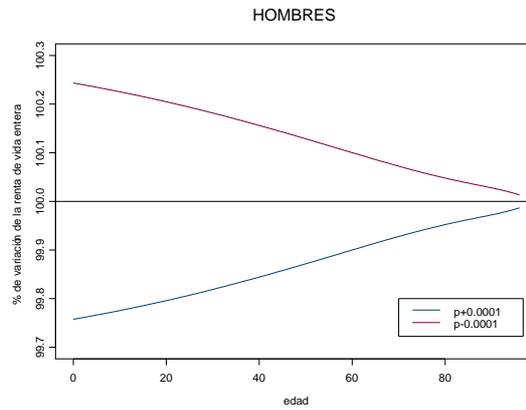


Figura 8.14: Variación de la renta cuando varían la probabilidades sumándoles una constante para los hombres

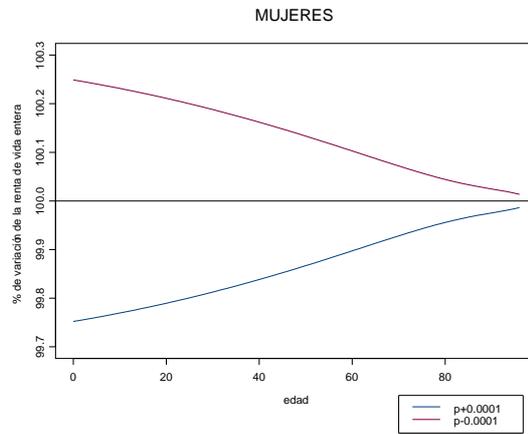


Figura 8.15: Variación de la renta cuando varían la probabilidades sumándoles una constante para las mujeres

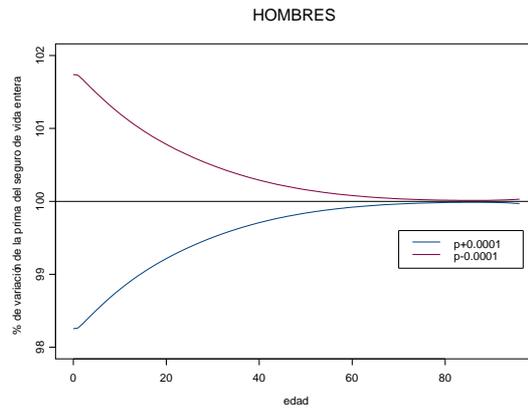


Figura 8.16: Variación de la prima del seguro cuando varían la probabilidades sumándoles una constante para los hombres

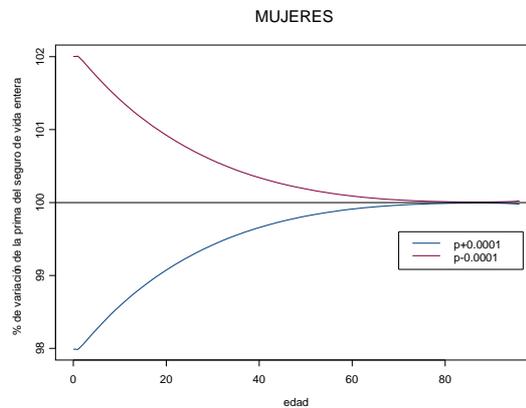


Figura 8.17: Variación de la prima del seguro cuando varían la probabilidades sumándoles una constante para las mujeres

Una vez analizadas las variaciones aditivas y multiplicativas sobre las tablas estáticas, nos queda por analizar las variaciones que se producen en los valores de las rentas y los seguros cuando se sustituye un tipo de tablas por otras. La Figura 8.18 y 8.19 recoge las

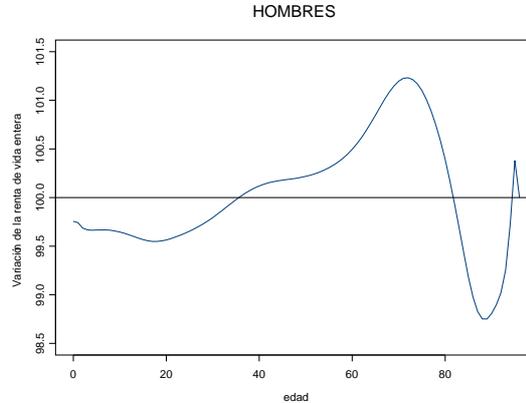


Figura 8.18: Variación de la renta calculada con tabla dinámica respecto a estática para los hombres

diferencias en el cociente del valor actual de una renta inmediata y vida entera cuando esta se calcula con una tabla dinámica frente al mismo valor cuando se utiliza estática. La Figura 8.19 presenta una ligera tendencia creciente del cociente hasta los últimos años donde comienza a decrecer. Más dificultades presenta la Figura 8.18, donde el cociente es menor que la unidad para edades inferiores a los 40 años, alcanzando un mínimo alrededor de los 20 años. A partir de esa edad comienza a crecer el cociente de forma que pasa a

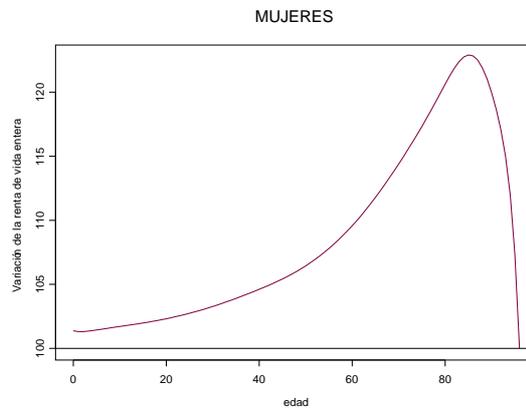


Figura 8.19: Variación de la renta calculada con tabla dinámica respecto a estática para las mujeres

ser mayor que uno hasta alcanzar un máximo en la edad de 75 años. De los 75 a los 90 es decreciente hasta alcanzar otro mínimo local en los 90 años, y después se convierte en creciente. Las Figuras 8.20 y 8.21 están referidas al cociente entre las primas únicas de un

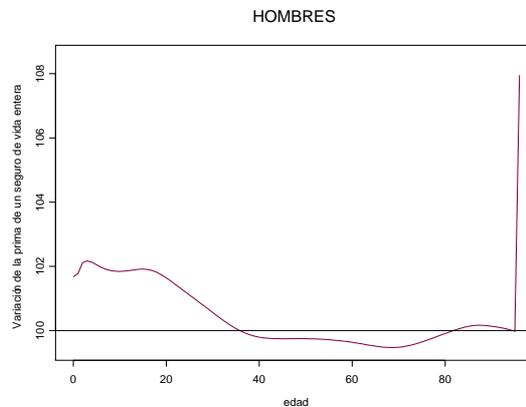


Figura 8.20: Variación de la prima de seguro calculada con tabla dinámica respecto a estática para los hombres

seguro vida entera con tanto con tablas dinámicas como estáticas. La figura 8.20 referida a los hombres, muestra que que la prima dinámica es superior a la estática para todas las edades entre 0 y 36 años, mientras que a partir de esa edad ocurre lo contrario hasta los 57 años, a partir de esta edad se vuelve a invertir la tendencia. Sin embargo en el caso de las mujeres (figura 8.21) el valor de la prima dinámica es notablemente inferior a la estática hasta una edad avanzada (83 años) a partir de la cual la dinámica se hace mayor

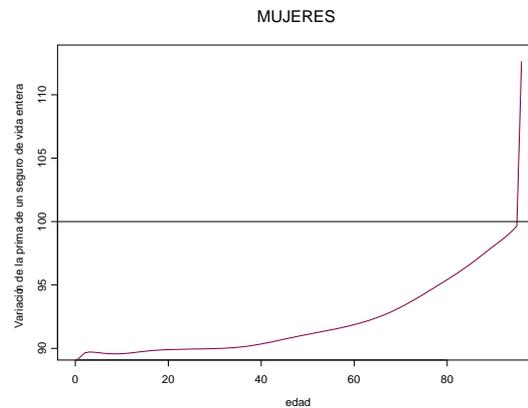


Figura 8.21: Variación de la prima de seguro calculada con tabla dinámica respecto a estática para las mujeres

que la prima estática.

8.3. Planes de pensiones

En esta sección se repasarán brevemente los conceptos básicos de un plan de pensiones, tanto en lo referido al tipo de aportación como a los métodos de cálculo.

Una vez definidos estos conceptos, se pasará a la construcción de una aplicación práctica, la constitución de un hipotético plan de pensiones complementario para el personal de la Universitat de València, distinguiendo los diferentes colectivos que la integran, y se analizará la evolución de las aportaciones colectivas en el caso de tablas estáticas y dinámicas.

8.3.1. Definición y conceptos básicos

Un plan de pensiones no es más que una operación actuarial sobre la vida en el que una serie de individuos pertenecientes a un colectivo aportan una prima durante una parte de su vida activa, aportación, con el fin de cobrar una cantidad, prestación, a partir de su jubilación. Respecto a los colectivos estudiados pueden considerarse abiertos y cerrados Lecina (1989). Un colectivo es cerrado cuando se consideran únicamente los miembros existentes en la fecha de valoración, teniendo en cuenta las posibles salidas por diferentes causas pero asumiendo que no se producirá ninguna entrada. Un colectivo se considera abierto si contempla como comunidad de riesgo no solo el conjunto de miembros existentes en la fecha de valoración del plan sino también los de las generaciones futuras.

El texto de Betzuen (1989) aporta una idea clara de los elementos actuariales que han de fijarse a la hora de diseñar un plan de pensiones. Debemos explicitar las siguientes cuestiones,

- La función de supervivencia o la función que represente el tanto de mortalidad. Utilizaremos las diferentes tablas de mortalidad que hemos construido para la Comunidad Valenciana en los capítulos anteriores.
- El tipo de interés técnico, que primero fijaremos en un 3%.
- El porcentaje de crecimiento anual acumulativo del salario, puesto que influye en las cuantías de las prestaciones.
- Tenemos que definir si se trata de un plan de aportaciones definidas o de prestaciones definidas.
- Si hay aportación o no por parte de la empresa, es decir, si se trata de un Plan de Empresa o no. En nuestro caso no será de empresa, pero sí que se darán algunas ideas en caso de que se desee incorporar una aportación por parte de Universitat de València.
- La edad de jubilación, que en la mayoría de los casos es a los 65 años, en nuestro colectivo es a los 70 para el personal docente e investigador (PDI) y a los 65 para el personal de administración y servicios (PAS).
- Indicar si la pensión se basa en el salario (salario final) o no. Señalando el coeficiente a aplicar sobre el salario final.
- Indicar si la prestación es revalorizable o no.
- Si se incluyen algunas garantías como las prestaciones por invalidez, el reconocimiento de algún derecho adquirido si el partícipe cambia de empresa, . . . En nuestro caso solo consideramos el evento fallecimiento, por lo que las tablas fijadas al principio van a ser fundamentales.
- El sistema financiero, que normalmente es de capitalización individual pero que también puede ser colectiva. En nuestro caso será así.
- Es frecuente considerar un grupo cerrado aunque sería más realista suponer un grupo abierto.

La forma habitual de clasificar los planes de pensiones es según sean de aportación definida o de prestación definida. A continuación presentaremos las ideas básicas de los métodos que se utilizan para el cálculo de las aportaciones.

Para calcular el *coste normal*, que es el coste asignado a cada año desde el comienzo de funcionamiento del plan en base a la edad de entrada, en algunas ocasiones se reconocen obligaciones con los trabajadores que cuando el plan entra en vigor llevan años de servicio en la empresa. El plan reconoce obligaciones contraídas pasadas que tienen un carácter suplementario, en este caso la cuantía asignada a cada año se llama *coste suplementario*.

8.3.2. Plan de pensiones de prestación definida

En este caso lo que se define es la prestación que percibirá el participante del plan a partir de la edad de jubilación y la aportación se obtiene como consecuencia de la igualdad actuarial que garantiza los pagos futuros. La prestación depende de los años de servicio y se calcula como un porcentaje que se acumula por cada año de servicio.

Designaremos por b_x la prestación de un trabajador de edad x , cantidad que puede considerarse constante o variable a lo largo de los años y que puede no depender de la edad. Para su valoración hemos de calcular el coste normal del plan para un empleado de edad x . El coste normal total es la suma de todos los costes individuales de todos los empleados.

El método de las prestaciones devengadas forma parte de la familia de métodos de prestaciones definidas y tiene fácil aplicabilidad a los planes en cuyo diseño se tenga en cuenta el número de años de servicio prestados.

Método de las prestaciones devengadas

Para calcular las aportaciones según este método se procede de la siguiente forma. En primer lugar necesitamos conocer el coste total que representa garantizar el pago de las prestaciones futuras de jubilación. Una vez fijado el valor de la prestación b_x que se devenga cada año, se puede calcular la prestación acumulada B_x sin más que sumar,

$$B_x = \sum_{h=x_e}^{x-1} b_h,$$

donde x_e es la edad de entrada. Pasamos a calcular el coste en base a la edad de entrada. El coste genérico representa el precio que debemos pagar para garantizar la prestación b_x a devengar en el año x ,

$$\text{Coste}_x = b_x {}_{x_j-x_e}E_{x_e} \ddot{a}_{x_j} \quad (8.8)$$

El caso más habitual considerar la prestación b_x fija, pero también puede considerarse como un porcentaje del salario final. Es pues importante, cuando se considera este método, estimar el posible nivel del salario final. En ambos casos, sustituyendo el valor en la ecuación (8.8) tendríamos el coste que puede valorarse en cada edad x .

El coste normal particularizado supone calcular a partir de la prestación anual acumulada total B_{x_j} la acumulativa anual b_x . De forma que si es constante se obtiene

$$b_x = \frac{B_{x_j}}{x_j - x_e},$$

y si depende del salario $b_x = ks_x$

$$k = \frac{B_{x_j}}{S_{x_j}},$$

donde S_{x_j} es el salario acumulado hasta la edad x_j .

Para el cálculo del coste normal genérico hay que tener en cuenta que el participante debe acumular la pensión B_{x_j} durante el periodo de servicio futuro. Dado que el periodo

de servicio futuro es distinto, las cuantías de las prestaciones acumulativas anuales y las acumuladas hasta x_a , edad alcanzada, en este método difieren del anterior. Si se toma como referencia las prestaciones acumulativas anuales del caso anterior b_x , debemos considerar un incremento adicional, Δb_x , que permita recuperar la cuantía no aportada hasta la edad alcanzada. De forma que la nueva cuantía considerada b'_x , se obtiene

$$\begin{aligned} b'_x &= b_x + \Delta b_x \\ &= b_x + b_x \frac{B_{x_a}}{B_{x_j} - B_{x_a}} \\ &= b_x \left(1 + \frac{B_{x_a}}{B_{x_j} - B_{x_a}} \right) \\ &= b_x \left(\frac{B_{x_j}}{B_{x_j} - B_{x_a}} \right), \end{aligned} \tag{8.9}$$

$$\tag{8.10}$$

de forma que la acumulada correspondiente, B'_x , es

$$\begin{aligned} B'_x &= \sum_{h=x_a}^{x-1} b'_h \\ &= \sum_{h=x_a}^{x-1} b_x \left(\frac{B_{x_j}}{B_{x_j} - B_{x_a}} \right) \\ &= \left(\frac{B_{x_j}}{B_{x_j} - B_{x_a}} \right) \sum_{h=x_a}^{x-1} b_x \\ &= \left(\frac{B_{x_j}}{B_{x_j} - B_{x_a}} \right) (B_x - B_{x_a}) \end{aligned} \tag{8.11}$$

8.3.3. Plan de pensiones de aportación definida

En este caso, a diferencia del anterior, se conoce la aportación anual y lo que queda pendiente de averiguar es la cuantía de la prestación anual correspondiente a dicha aportación anual. Ahora se parte del importe total de prestación a la jubilación mientras que antes se definía la prestación que se devengaba cada año. Las diferencias entre los dos métodos según Betzuen (1989) son

1. En las prestaciones acumuladas el coste normal se calculaba a partir de la prestación anual b_x definida para un trabajador de edad x . En el caso de las prestaciones proyectadas el coste se obtiene a partir de la prestación total considerada a la jubilación, por lo tanto el coste normal será la cuantía que permita constituir B_{x_j} .
2. Al tener establecida la prestación total futura que se quiere garantizar, podemos fijar un coste anual normal para el plan. Este coste normal, para una persona de

edad x , puede ser constante en términos monetarios o dependiente del salario. En este caso se puede definir el coste normal como un % constante del salario a la edad alcanzada. La prestación que se acumula anualmente hace de variable dependiente y se puede calcular a partir del coste normal anteriormente definido.

3. En las prestaciones devengadas, por cada año de servicio activo se acredita una fracción de la prestación. En el caso de prestaciones proyectadas se parte de una prestación final prevista que obtendrá el partícipe en activo en la empresa. La prestación proyectada hace realmente de meta.

Método de las prestaciones proyectadas

Se trata de un método de aportaciones definidas que como el anterior utiliza un método de capitalización individual.

Sea B_{x_j} la prestación a la jubilación, el valor total de todas las prestaciones a la edad de jubilación x_j es

$$(Va)_{x_j} = B_{x_j} \ddot{a}_{x_j},$$

que representa el coste inicial. De igual forma se tiene $(Va)_{x_e}$ que representa el coste total del plan para garantizar las prestaciones futuras valorado a la edad de entrada en el plan, x_e , y que se obtiene según

$$(Va)_{x_e} = B_{x_j} {}_{x_j-x_e}E_{x_e} \ddot{a}_{x_j}.$$

Supuesto que la aportación anual sea constante, P_x , la valoración total a la edad del plan es

$$P_x \ddot{a}_{x_e:\overline{x_j-x_e}|}.$$

Para que se garantice la cobertura del plan, las aportaciones y las prestaciones valoradas en el mismo momento deben coincidir según el principio de equivalencia

$$(Va)_{x_e} = P_x \ddot{a}_{x_e:\overline{x_j-x_e}|}$$

de donde,

$$\begin{aligned} P_x &= \frac{(Va)_{x_e}}{\ddot{a}_{x_e:\overline{x_j-x_e}|}} \\ &= \frac{B_{x_j} {}_{x_j-x_e}E_{x_e} \ddot{a}_{x_j}}{\ddot{a}_{x_e:\overline{x_j-x_e}|}} \\ &= \frac{B_{x_j} \frac{N_{x_j}}{D_{x_e}}}{\frac{N_{x_e} - N_{x_j}}{D_{x_e}}} \\ &= \frac{B_{x_j} N_{x_j}}{N_{x_e} - N_{x_j}}. \end{aligned} \tag{8.12}$$

Además se realizaran las operaciones considerando unitaria la prestación, $B_{x_j} = 1$, pues en cualquier otro caso de prestación constante, bastaría con multiplicar por la cantidad el resultado. Utilizando la expresión del símbolo de conmutación $N_x = \sum_{h=x}^{\omega} D_h$, la prima es

$$P_{x_e} = \frac{\sum_{h=x_j}^{\omega} D_h}{\sum_{h=x_e}^{\omega} D_h}. \quad (8.13)$$

Si la aportación en lugar de ser contante es un porcentaje del salario,

$$P_x = K s_h.$$

El valor actual de las aportaciones futuras es

$$\sum_{h=x_e}^{x_j-1} P_h h - x_e E_{x_e} = K s_{x_e} (S\ddot{a})_{x_e:\overline{x_j-x_e}|}.$$

Como ha de cumplirse la igualdad entre las aportaciones y las prestaciones futuras podemos escribir

$$K s_{x_e} (S\ddot{a})_{x_e:\overline{x_j-x_e}|} = (Va)_{x_e},$$

de donde obtenemos el porcentaje

$$K = \frac{(Va)_{x_e}}{s_{x_e} (S\ddot{a})_{x_e:\overline{x_j-x_e}|}},$$

de forma que la aportación es

$$P_x = K s_x.$$

8.3.4. Evaluación del coste de las pensiones mediante el sistema de capitalización colectiva

El cálculo del valor actual de las prestaciones no difiere demasiado del sistema de capitalización individual, pero se realiza para todo el colectivo y no a nivel individual.

Supuesto que la prestación anual por jubilación es B_{x_j} , el valor de la renta, a la edad de jubilación es

$$B_{x_j} \ddot{a}_{x_j}^{(12)}$$

y a la edad actual

$$B_{x_j} {}_{x_j-x}E_x \ddot{a}_x^{(12)}$$

Pero este resultado corresponde a un partícipe y es el mismo para cada uno de los que forman parte del subgrupo de igual edad y categoría. Para otro subgrupo el valor B_{x_j} es diferente. Puesto que B_{x_j} es diferente y la edad actual x también (la diferencia se da

entre los subgrupos), el valor actual total de las prestaciones para todos los partícipes con edades comprendidas entre e_1 y e_2 ($e_2 < x_j$) lo podemos indicar como,

$$V_a^T = n_{e_1} B_{x_j}^{(e_1)} {}_{x_j-e_1} E_{e_1} \ddot{a}_{x_j}^{(12)} + n_{e_1+1} B_{x_j}^{(e_1+1)} {}_{x_j-(e_1+1)} E_{e_1+1} \ddot{a}_{x_j}^{(12)} + \dots \\ \dots + n_{e_2} B_{x_j}^{(e_2)} {}_{x_j-(x_j-1)} E_{e_2} \ddot{a}_{x_j}^{(12)} \quad (8.14)$$

Para el cálculo del valor actual de las aportaciones cuando se utiliza el sistema de capitalización colectiva, la aportación es común para todos los partícipes. Lo que se trata, en realidad, es de calcular el valor actual de todas las rentas temporales a que hubiera lugar en función de la edad actual de los partícipes. Posteriormente este resultado se multiplicará por la cuota de aportación colectiva y el nuevo resultado corresponde al valor actual de las aportaciones de todo el colectivo.

Cuantitativamente el valor actual de todas las aportaciones futuras lo podremos expresar como,

$$C(n_{e_1} \ddot{a}_{x_j} + n_{e_1+1} \ddot{a}_{x_j} + \dots + n_{e_2} \ddot{a}_{x_j}). \quad (8.15)$$

Igualando los resultados (8.14) y (8.15) se puede despejar el valor de C y con ello se obtiene la cuota de aportación, la misma para cada partícipe y para ese año. El coste total para ese año se obtiene multiplicando la cuota anterior C por el número de partícipes del colectivo.

8.3.5. Aplicación a la Universitat de València

Para la aplicación práctica realizada se ha procedido en primer lugar a analizar el colectivo de personal, tanto en volumen como en edad y por sexos.

Se han analizado las retribuciones de este personal. En particular las retribuciones básicas, entendiendo por ello en el caso del PDI, el sueldo base y los complementos de destinos y específico. Se han obviado los complementos de productividad docente e investigadora, así como los otros complemento por cargos, autonómico, etc.

Para el caso del PAS, se ha considerado como retribución de cada grupo la del subgrupo de nivel y complemento con mayor número de efectivos. Así, por ejemplo en el grupo A se ha considerado las retribuciones de un A24B. En la Tabla 8.2 aparecen recogidas las retribuciones de estos cuerpos.

Análisis descriptivo del personal

Para realizar esta aplicación “simulada” de un plan, el primer paso es analizar la estructura del colectivo por edad y sexo, que son las variables determinantes para utilizar las probabilidades de muerte de las tablas que hemos construido. Estudiemos el comportamiento de la variable edad para cada sexo en cada categoría puesto que va a influir muy claramente en el coste del plan. Para ello en la Tabla 8.1 tenemos algunos valores resumen de la variable. Además en las la Figuras 8.22 y 8.23 mostramos el envejecimiento correspondiente a cada una de las categorías. Para mostrar gráficamente el comportamiento del colectivo, hemos considerado conveniente agrupar la edades en intervalos de 5 años, lo

HOMBRES						
PAS	Mínimo	1^{er} Cuartil	Mediana	Media	3^{er} Cuartil	Máximo
A	24	31.25	35	36.82	40.75	67
B	24	35	40	41.01	46.5	63
C	23	34	39.5	40.47	45	63
D	23	34	38	39.31	43	63
E	32	38.5	45.5	45.68	53.25	60
Total	23	34	38	39.77	45	67

MUJERES						
PAS	Mínimo	1^{er} Cuartil	Mediana	Media	3^{er} Cuartil	Máximo
A	24	32	37	38.24	42.75	59
B	24	35	40	40.71	45	60
C	24	36	40	41.28	47	65
D	23	35	39	40.41	45	64
E	28	41.5	52	49.49	59	64
Total	23	35	40	41.23	47	65

HOMBRES						
PDI	Mínimo	1^{er} Cuartil	Mediana	Media	3^{er} Cuartil	Máximo
CU	36	50	54	54.45	59	70
CEU	31	44	52	51.42	58.75	68
TU	30	39	45	45.07	50.5	70
TEU	29	39	46.5	45.87	53	69
Total	29	40	47	47.25	53	70

MUJERES						
PDI	Mínimo	1^{er} Cuartil	Mediana	Media	3^{er} Cuartil	Máximo
CU	42	49	53	53.22	56.5	69
CEU	32	45.75	51.5	52.29	60.75	68
TU	30	38	43	43.84	49	69
TEU	30	37	44.5	45.35	53	69
Total	30	38	44.5	45.27	51	69

Tabla 8.1: Medidas descriptivas de la edad según categorías

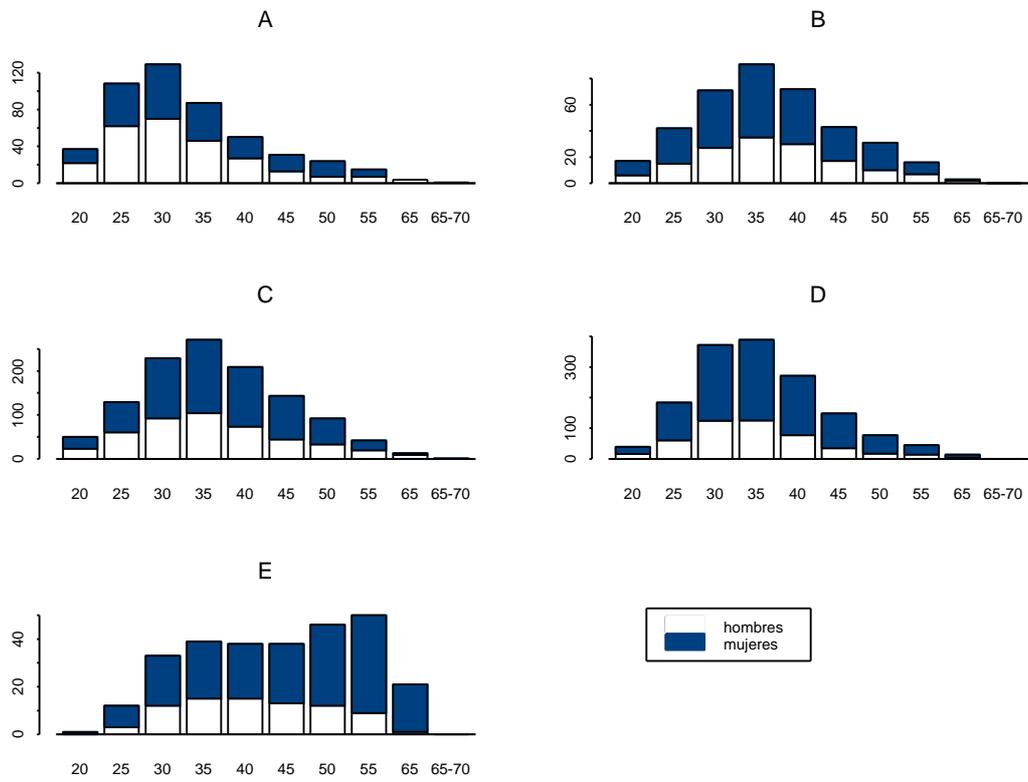


Figura 8.22: Distribución por edad del PAS

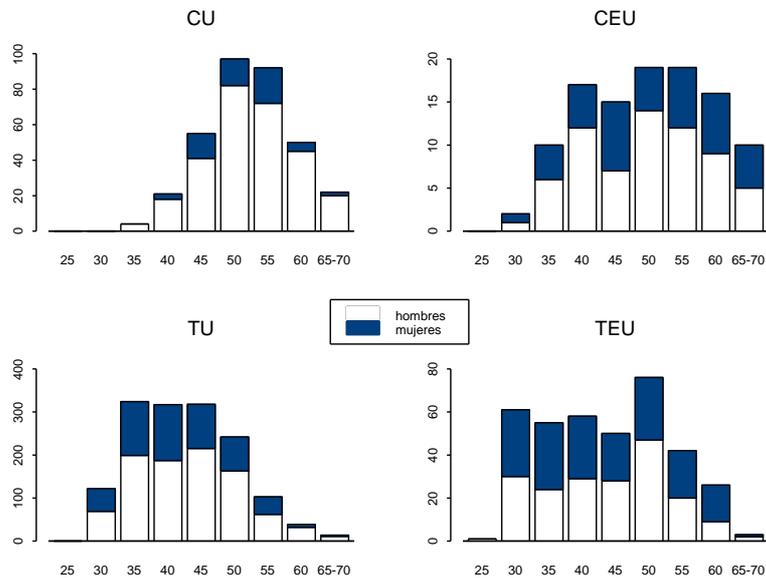


Figura 8.23: Distribución por edad del PDI

que 5 años reduce a un número razonable de intervalos el conjunto de edades y además proporciona el último intervalo de 65 a 70 que contiene las edades de jubilación.

Una conclusión inmediata de la estructura de los dos tipos de colectivos es la presencia mayoritaria de mujeres en el grupo PAS, mientras que en el PDI todavía prevalecen los hombres, aunque en los cuerpos de CEU y TEU y en las edades inferiores a 40, el número de mujeres es muy similar al de hombres.

Además, y puesto que muchos planes se apoyan en el salario, hemos de conocer el sueldo medio anual básico (sin los complementos específicos) de cada uno de ellos. Como ya hemos dicho tenemos dos grupos, que son el personal de administración y servicios (PAS) y el personal docente e investigador (PDI). Cada uno de estos grupos se subdivide en categorías que tienen un salario medio anual que mostramos en la Tabla 8.2.

Simulación de un plan

Para el cálculo del importe de las aportaciones colectivas se ha usado el método de las prestaciones proyectadas. Como se ha comentado se trata de un método de aportaciones definidas, que consiste en “proyectar” una prestación a la jubilación (70 años en el caso del PDI y 65 en el PAS) con el fin de determinar un coste total de la pensión de jubilación evaluada a esa edad, que coincida con el método de las prestaciones acumuladas.

Dado que se usará el sistema de capitalización colectiva, la aportación es común para todos los partícipes de cada grupo, es decir, lo que se trata en realidad es de calcular el

PAS		PDI	
Categoría	Sueldo total	Categoría	Sueldo total
A	33237.01	CU	35065.05
B	25091.16	CEU	28402.82
C	20009.29	TU	28402.82
D	14042.48	TEU	25355.43
E	12795.12		

Tabla 8.2: Tabla de salarios por categoría profesional

valor común de todas las rentas temporales a que hubiera lugar en función de la edad de los partícipes. Por ello, lo importante no es determinar el importe exacto de cada una de las aportaciones, sino su afectación por el uso de unas tablas de mortalidad u otras, razón por la cual se han realizado múltiples hipótesis simplificadoras que en ningún caso afectan al estudio que se pretende realizar.

- Se ha supuesto que el salario de cada uno de los grupos es constante durante todo el análisis. Esta hipótesis significa utilizar el valor de una renta vitalicia temporal constante en lugar de una renta vitalicia creciente en progresión geométrica (establecimiento de las tasas de crecimiento de los salarios en función de la inflación). Tampoco se considera una revalorización de las pensiones.
- Se considera que la prestación a recibir será un 20 % del salario, dado que lo que se trata es de establecer un complemento a las pensiones aportadas por la Seguridad Social o las Clases Pasivas del Estado, y de las posibles cuentas individuales de cada uno de los posibles partícipes. No se consideran otras prestaciones diferentes de la jubilación, como la invalidez, orfandad, etc.
- El tipo de interés técnico es el 3%. Las tablas que se usaran serán las estáticas y dinámicas construidas en las secciones precedentes
- No se produce aportación por parte de la empresa. Es decir, no se trata de un plan de empresa, aunque se darán algunas sugerencias por si se considera oportuno establecer estas aportaciones.
- Las aportaciones se realizará anualmente (no habría problema en suponer las mensualidades) al igualar el valor de la renta temporal del importe de las aportaciones al valor actual de las prestaciones para todos los partícipes por cada uno de los grupos y categorías.
- Se ha considerado oportuno dividir el colectivo de PDI en cuatro grupos, uno por cada una de las categorías académicas (Catedráticos de Universidad, Catedráticos de Escuela Universitaria, Titulares de Universidad y Titulares de Escuela Universitaria). No habría problema de incorporar nuevas figuras: Contratados Doctores, Ayudantes, etc. De la misma forma se han considerado cinco grupos de PAS, por cada una de

las escalas o grupos de Administración General, aunque en ellas se han integrado los contratados laborales, fijos, etc.

Como se ha señalado, y aunque parezcan simplificadoras, estas hipótesis recogen las ideas fundamentales de un plan de pensiones y para el objetivo de este trabajo consideramos que es suficiente. Otra cuestión sería la puesta en marcha de forma práctica y efectiva del plan, pero para ello se requerirían unos acuerdos entre trabajadores, sindicatos, equipo rectoral, etc, que no es posible considerar en este caso. Aunque la puesta en práctica de ello no resultaría excesivamente complejo.

8.3.6. Análisis de estabilidad y sensibilidad de los fondos de pensiones

En esta sección vamos a estudiar la influencia del tipo de tablas de supervivencia sobre el importe de las aportaciones. El análisis podría extenderse también al caso de las variaciones de los tipos de interés.

Tomando como base el método de las prestaciones proyectadas, podríamos analizar la influencia de las variaciones en las probabilidades de supervivencia para lo obtenemos la expresión de la prima en función de las probabilidades y del interés.

Sustituyendo en (8.13) la igualdad $D_x = l_x(1+i)^{-x}$ y $l_{x+1} = l_x p_x$, obtenemos la expresión de la prima en función de las probabilidades de supervivencia p_x y del interés i

$$\begin{aligned}
 P_{x_e} &= \frac{\sum_{h=x_j}^{\omega} l_h(1+i)^{-h}}{x_j-1} & (8.16) \\
 &= \frac{\sum_{h=x_e}^{\omega-1} l_h(1+i)^{-h}}{x_j-2} \\
 &= \frac{\sum_{h=x_j-1}^{\omega-1} l_h p_h (1+i)^{-(h+1)}}{x_j-2} \\
 &= \vdots \\
 &= \frac{\sum_{h=x_j-1}^{\omega-1} l_0 \prod_{k=0}^h p_k (1+i)^{-(h+1)}}{x_j-2} & (8.17) \\
 &= \frac{\sum_{h=x_e-1}^{\omega-1} l_0 \prod_{k=0}^h p_k (1+i)^{-(h+1)}}{x_j-2}
 \end{aligned}$$

Respecto a la variación del interés técnico i

Derivando la igualdad (8.16) respecto de i obtenemos,

$$\frac{\partial P_x}{\partial i} = \frac{\sum_{h=x}^{\omega} \frac{-hl_h}{(1+i)^{h+1}} \sum_{h=x_e}^{x_j-1} l_h(1+i)^{-h} - \sum_{h=x_j}^{\omega} l_h(1+i)^{-h} \sum_{h=x_e}^{x_j-1} \frac{-hl_h}{(1+i)^{h+1}}}{\left(\sum_{h=x_e}^{x_j-1} l_h(1+i)^{-h} \right)^2}$$

Respecto a la variación de las probabilidades de supervivencia p_x

A partir de (8.17) se deduce que las primas van a depender de todas las probabilidades p_x , para $x = 0, 1, \dots, \omega$ de forma que la variación de cualquiera de ellas le afectaría. Además hay que tener en cuenta que cada probabilidad depende de todas las anteriores. Como lo que nos interesa es la comparación de los dos tipos de tablas estáticas y dinámicas, no derivaremos esta expresión evitando así su complejidad.

8.3.7. Aplicación

Las aportaciones de cada una de los grupos y categorías se han evaluado tanto para tablas estáticas como dinámicas. Así, por ejemplo, en el caso de los Catedráticos de Universidad la aportación colectiva de hombres (con tablas estáticas) supondría para cada uno de ellos un 8.48 % del salario. En el caso de las mujeres la aportación supondría el 7.54 % del salario y para todo el colectivo la aportación media sería del 8.32 %. Evidentemente las aportaciones de cada uno de los partícipes se sitúa en el entorno de las aportaciones medias individuales de todos los partícipes, así mismo la suma total de las aportaciones es ligeramente inferior a la que hubiéramos obtenido si se hubiera usado el método de capitalización individual.

De idéntica forma se calcula para el resto de grupos y colectivos. Pero lo que nos interesa resaltar es el comportamiento de estas aportaciones ante los cambios de las tablas de mortalidad. Para ello se presentan una serie de gráficos que comparan las aportaciones de cada uno de los diferentes grupos y colectivos tanto en el caso estático como dinámico.

Lo relevante del análisis es que la estructura de los colectivos (hombres y mujeres) determina en mayor medida los cambios y resultan mucho más ilustrativos a los efectos de este trabajo.

Las gráficas que siguen recogen el importe de las aportaciones de las diferentes categorías tanto para hombre como para mujeres según el tipo de tabla usada para el cálculo. Se ha tomado como base de referencia la aportación de un TEU mujer.

La Figura 8.24 recoge las aportaciones de los Catedráticos de Universidad. En él puede observarse que los cambios debidos al tipo de tabla son poco apreciables en el caso de los hombres (apenas un 2.09 % de variación en las aportaciones), por el contrario, para las mujeres la variación es mucho mayor (un 11.57 %) debido fundamentalmente a la incidencia mayor que en ellas tiene el efecto “calendario”. Otra cuestión que merece detallarse es

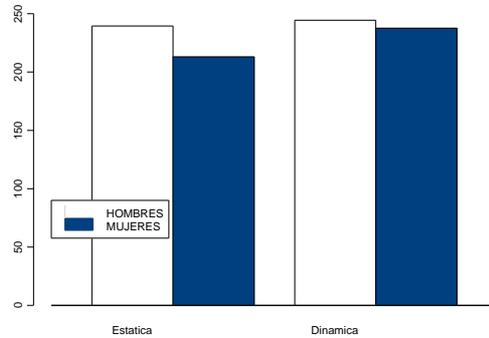


Figura 8.24: Aportaciones para Catedráticos de Universidad

que la aportación de los hombres frente a las mujeres con tablas estáticas es un 11.09% superior, mientras que con tablas dinámicas esta diferencia queda reducida al 2.83%. La explicación se debe al mayor número de hombres en esta categoría. Como se verá más adelante, una situación opuesta ocurrirá con el PAS, especialmente en el grupo E donde el número de mujeres es casi 2.5 veces el de hombres. La Figura 8.25 presenta la evolución de

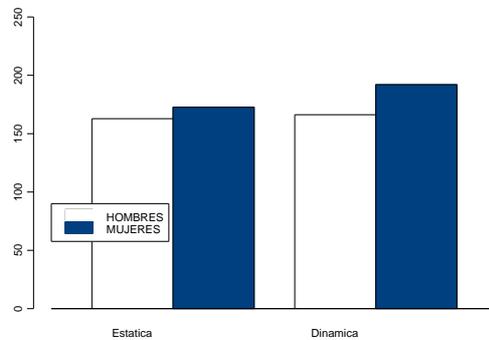


Figura 8.25: Aportaciones para Catedráticos de Escuela

las aportaciones de la categoría de Catedráticos de Escuela Universitaria. El crecimiento de las aportaciones por el efecto de las tablas es mucho mayor en el caso de las mujeres (10.26%) que en el de los hombres (1.69%). Mientras que si analizamos en la diferencia para una misma tabla entre hombres y mujeres es prácticamente inapreciable para tablas estáticas (0.92%), pero sí vemos que es relevante en el caso de las tablas dinámicas (9.43%). La Figura 8.26 presenta los resultados para la categoría de los Titulares de Universidad. La

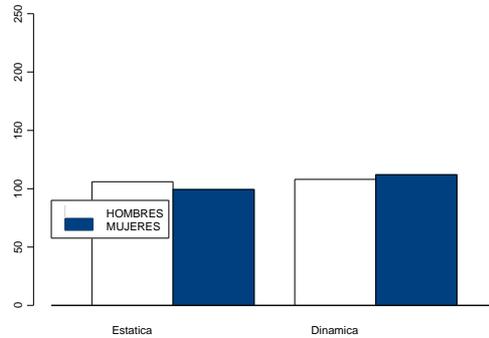


Figura 8.26: Aportaciones para Titulares de Universidad

situación en este colectivo es similar a los anteriores, pero con unas diferencias menores en el caso de los hombres (1.97%) y mayores para las mujeres (12.77%). La comparación entre hombres y mujeres para las tablas homogéneas presenta diferencias ya que las aportaciones de las mujeres son inferiores a las de los hombres con las tablas estáticas, pero mayores para las tablas dinámicas como consecuencia de que hay un 75% más de hombres que de mujeres en ese colectivo. La evolución de las aportaciones para los Titulares de Escuela

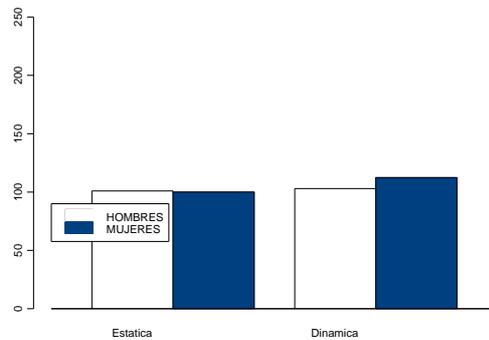


Figura 8.27: Aportaciones para Titulares de Escuela Universitaria

Universitaria aparece en la Figura 8.27. El comportamiento de este colectivo tiene muchas similitudes con los Titulares de Universidad. Las diferencias para los hombres según el tipo de tablas es baja (2.0%) mientras que para las mujeres es alta (12.35%). Si se analiza las diferencias entre hombres y mujeres para el mismo tipo de tabla, las aportaciones de hombres son ligeramente mayores a las de mujeres (0.95%) para las tablas estáticas,

mientras que la aportación de las mujeres es un 9.10% mayor que la de los hombres para las tablas dinámicas. La razón estriba en que el número de hombres es ligeramente superior (4%) al de mujeres dentro de los TEU's.

Como se ha señalado anteriormente el colectivo del Personal de Administración y Servicios (PAS) presenta una estructura (por sexos) muy diferente al del PDI.

Las figuras siguientes presentan la evolución de las aportaciones de cada uno de los grupos dentro del PAS, según el tipo de tablas utilizadas y el sexo del partícipe. Un rasgo a destacar a todas ellas es que las aportaciones realizadas con tablas dinámicas en el colectivo de hombres varían entre 1.58% (grupo A) y 1.72% (grupo E), lo cual denota una gran estabilidad.

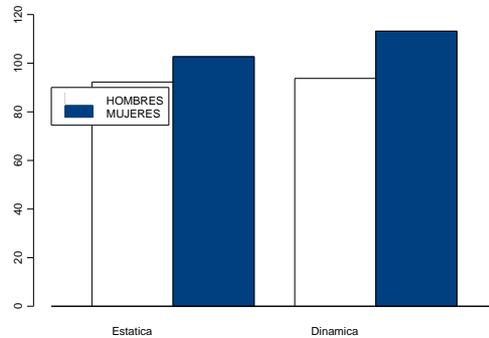


Figura 8.28: Aportaciones para el grupo A

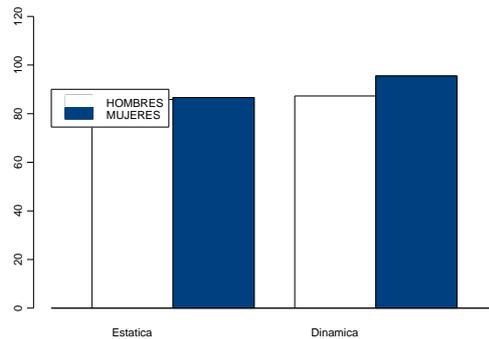


Figura 8.29: Aportaciones para el grupo B

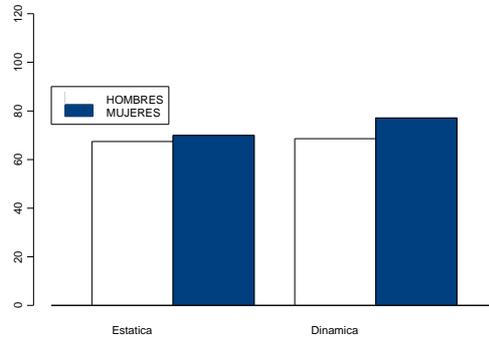


Figura 8.30: Aportaciones para el grupo C

En el caso de las mujeres el uso de unas tablas u otras sí que tiene importancia, ya que los cambios van desde el 9.18 % del grupo E hasta el 10.35 % del grupo D.

Las comparaciones entre las aportaciones entre hombres y mujeres con tablas estáticas son más divergences pasando del 0.92 por ciento (grupo B) superior de las aportaciones de mujeres con relación a los hombres hasta el 33.39 % en el grupo E.

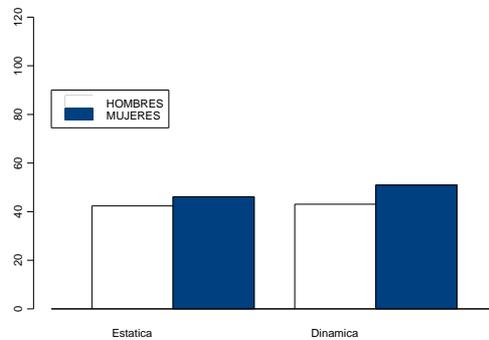


Figura 8.31: Aportaciones para el grupo D

Una situación similar, ocurre con las aportaciones para tablas dinámicas, siendo la mayor discrepancia, al igual que en el caso anterior, la del grupo E. La explicación reside en la estructura del colectivo E, formado casi exclusivamente por mujeres de edad próxima a la jubilación. Esto es así porque el grupo se formó a partir de mujeres que fueron contratadas por la Universidad, para las tareas de limpieza de locales (años 70 y 80) que fueron asimiladas a PAS cuando posteriormente se contrató este servicio con empresas

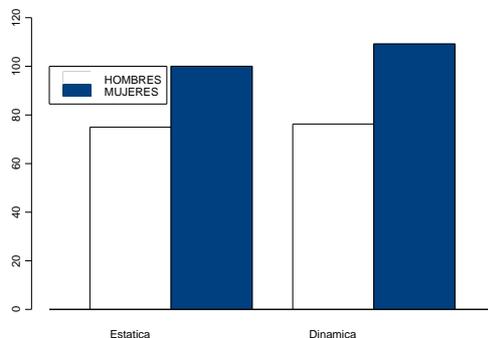


Figura 8.32: Aportaciones para el grupo E

externas. Se trata de un grupo “a extinguir” que por sus especiales características tiene aportaciones más elevadas de lo que realmente correspondería.

Si se hubiera planteado la construcción de un plan de pensiones con aportación “patronal” o de la empresa, este sería un caso en donde la aportación del promotor debería aplicarse para poder “nivelar” las aportaciones de los diferentes grupos de forma que fuera lo más homogénea y ajustada posible a sus niveles salariales.

8.4. Conclusiones

En este epígrafe resumiremos brevemente las principales conclusiones que se han señalado en la páginas anteriores, tanto en lo referente a las operaciones individuales (seguros y rentas) como a las colectivas (aportaciones a un hipotético plan de pensiones) como a la influencia que sobre ellas tienen los cambios en los elementos básicos de las tablas de mortalidad: tipo de interés y probabilidades de supervivencia.

Los cambios en el tipo de interés técnico supone unos cambios “previsibles” en los valores actuales de las rentas y de las primas únicas de los seguros.

Sin embargo, las variaciones debidas a los cambios en el uso de los diferentes tipos de tablas de mortalidad suponen unas modificaciones significativas en el caso de las rentas y los seguros, proporcionando variaciones importantes.

Un caso más especial es el referido a las aportaciones con capitalización colectiva, donde el efecto de la composición del colectivo al que se va a aplicar hace que las diferencias se vean amortiguadas o amplificadas. Esto se ha puesto de manifiesto especialmente en el caso del colectivo de PAS entre los grupos B y E. En el grupo B el porcentaje de mujeres excede al de hombres en un 48 %, mientras que en el grupo E hay un 145 % más de mujeres que hombres, es decir, un factor de amplificación de 3.02 lo que se ha traducido en el incremento de las aportaciones usando tablas dinámicas entre hombre y mujeres (las diferencias van del 9.43 % al 43.18 %). El análisis de la capitalización colectiva, como se ha

señalado queda condicionado de forma importante por la estructura de edad y el número de hombres y mujeres del colectivo, pero en aras de asegurar una mejor y mayor cobertura de las prestaciones es de desear el uso de las tablas dinámicas frente a las estáticas, ya que se adaptan mejor a la evolución de la mortalidad.

Pero una vez establecido el plan conviene ir adecuando las reservas generadas de acuerdo con las variaciones de la mortalidad y con unas tablas ajustadas a cada uno de los colectivos, ya que la no corrección de estos desajuste supondría un problema importante a la hora de materializar las prestaciones que se han generado de forma colectiva.

Bibliografía

- Andersen, P., Borgan, O., Gill, D., y Keiding, N. (1994). *Statistical Models Based in Counting Proceses*. Springer, New York.
- Azzalati, A. y Bowman, A. (1993). On the use of nonparametric regresion for checking relationships. *Journal of the Royal Statistical Society B*, 55:549–557.
- Benjamin, B. y Pollard, J. (1992). *The Analysis of Mortality and Other Actuarial Statistics*. Butterworth-Heinemann, London, 6^a edición.
- Benjamin, B. y Soliman, A. (1995). *Mortality on the Move*. Actuarial Education Service, Oxfort.
- Betzuen, A. (1989). *Planes y Fondos de Pensiones. Su Cálculo y Valoración*. Ediciones Deusto, Bilbao.
- Betzuen, A. (1995). La problematica de los tanto interanuales, la asignación del rango de la edad y del intervalo de exposición al riesgo en un colectivo de activos. *Anales del Instituto de Actuarios*, 3^a Epoca(1):51–97.
- Betzuen, A. (1997). An aproach to people´s in employment mortality through smoothing splines. Edimburgo. Heriot-Watt University.
- Bloomfield, D. y Haberman, S. (1987). Graduation: Some experiments with kernel methods. *Journal of the Institute of Actuaries*, 114:339–369.
- Bureau, C. M. I. (1990). *Continuous Mortality Investigation Reports CMIR 10*. The Institute of Actuaries and Faculty of Actuaries, UK.
- Bureau, C. M. I. (1999). *Continuous Mortality Investigation Reports CMIR 17*. The Institute of Actuaries and Faculty of Actuaries, UK.
- Carter, L. y Lee, R. (1992). Modeling and forecasting u.s. sex differentials in mortality. *International Journal of Forecasting*, 8(3):393–411.
- Chan, F., Chan, L., Falkenberg, J., y Yu, M. (1986). Applications of linear and quadratic programmings to some cases of the whittaker-henderson graduation method. *Scandinavian Actuarial Journal*, (1):141–153.

- Cleveland, W. (1979). Robust locally-weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74:829–836.
- Cleveland, W., Devlin, S., y Groose, E. (1988). Regression by local fitting. *Journal of Econometrics*, 37:87–114.
- Congdon, P. (1993). Statistical graduation in local demographic analysis and projection. *Journal of the Royal Statistical Society*, 156(2):237–270.
- Copas, J. y Haberman, S. (1983). Non parametric graduation using kernel methods. *Journal of the Institute of Actuaries*, 110:135–156.
- Dellaportas, P., Smith, A., y Stavropoulos, P. (2001). Bayesian analysis of mortality data. *Journal of the Royal Statistical Society A*, 164(2):275–291.
- Eubank, R. y Thomas, W. (1993). Detecting heterocedasticity in nonparametric regression. *Journal of the Royal Statistical Society B*, 55:145–155.
- Fan, J. y Gijbels, I. (1996). *Local Polinomial Modelling and its Applications*. Chapman and Hall, London.
- Felipe, A., Guillen, M., y Nielsen, J. (2001). Longevity studies based on kernel hazard estimation. *Insurance: Mathematics and Economics*, 28:191–204.
- Felipe, M. y Guillén, M. (1999). *Evolución y Predicción de las Tablas de Mortalidad Dinámicas para la Población Española*. Cuadernos de la Fundación, Fundación Mapfre Estudios.
- Forfar, D., McCutcheon, J., y Wilkie, A. (1988). On graduation by mathematical formula. *Journal of the Institute of Actuaries*, 115 part I(459):1–149.
- Gavin, J., Haberman, S., y Verrall, R. (1993). Moving weighted average graduation using kernel estimation. *Mathematics and Economics*, 12(2):113–126.
- Gavin, J., Haberman, S., y Verrall, R. (1994). On the choice of bandwidth for kernel graduation. *Journal of the Institute of Actuaries*, 121:119–134.
- Gavin, J., Haberman, S., y Verrall, R. (1995). Graduation by kernel and adaptive kernel methods with a boundary correction. *Transactions. Society of Actuaries*, XLVII:173–209.
- Gerber, H. (1997). *Life Insurance Mathematics*. Springer-Verlag, Berlín.
- Gompertz, B. (1825). On the nature of the function of the law of human mortality and on a new mode of determining the value of life contingencies. *Transactions of The Royal Society*, 115:513–585.
- Haberman, S. y Renshaw, A. (1996). Generalized linear models and actuarial science. *The Statistician*, 45(4):407–436.

- Haberman, S. y Renshaw, A. (1999). A simple graphical method for the comparison of two mortality experiences. *Applied Stochastic Models in Business*, 15(4):333–352.
- Hall, P. y Wehrly, T. (1991). A geometrical method for removing edge effects from kernel-type nonparametric regression estimates. *Journal of the American Statistical Association*, 86:665–672.
- Hamilton, J. (1994). *Time Series Análisis*. Princeton University Press, Princeton, New Jersey.
- Hartmann, M. (1983). Past and recent experiments in modeling mortality at all ages. *A: Stockholm Research Reports in Demography*, (107).
- Hastie, T. J. y Tibshirani, R. (1990). *Generalized Additive Models*. Chapman-Hall, London.
- Heligman, L. y Pollard, J. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, 107:49–80.
- Hoem, J. M. y Linnemann, P. (1988). The tails in moving average graduation. *Scandinavian Actuarial Journal*, (1):193–229.
- INE (1997). *Evolución de la población de España entre los censos de 1981 y 1991*. INE.
- Lay, D. C. (2001). *Algebra Lineal y sus Aplicaciones*. Prentice-Hall, Mexico, 2ª edición.
- Lecina, J. (1989). *Los Planes de Previsión. Un Tratamiento Actuarial*. Caja de Cataluña, Barcelona.
- Lee, R. (2000). The lee-carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal*, páginas 80–91.
- Lee, R. y Carter, L. (1992). Modelling and forecasting u.s. mortality. *Journal of the American Statistical Association*, 87(419):659–671.
- London, D. (1985). *Graduation: the Revision of Estimates*. Actex Publication, Winsted, Connecticut.
- Macdonald, A. (1996a). An actuarial survey of statistical models for decrement and transition data i: Multiple state, poisson and binomial models. *British Actuarial Journal*, 2(1):129–155.
- Macdonald, A. (1996b). An actuarial survey of statistical models for decrement and transition data ii: Competing risks, non-parametric and regression models. *British Actuarial Journal*, 2(2):429–448.
- Macdonald, A. (1996c). An actuarial survey of statistical models for decrement and transition data.iii: counting process models. *British Actuarial Journal*, 2(3):703–726.
- Makeham, W. (1860). On the law of mortality. *Journal of the Institute of Actuaries*, 13:325–358.

- McCullagh, P. y Nelder, J. (1989). *Generalized Linear Models*. Chapman and Hall, London.
- McNown, R. y Rogers, A. (1988). Time series analysis and forecasting of parametrized model schedules: mortality. *A: IBS Population Program Working Paper*, 8(88).
- McNown, R. y Rogers, A. (1989). Forecasting mortality: A parametrized time series approach. *Demography*, 26(4):645–660.
- McNown, R. y Rogers, A. (1992). Forecasting cause-specific mortality using time series methods. *International Journal of Forecasting*, 8:413–432.
- Navarro, E. (1991). *Tablas de Mortalidad de la Población Española 1982. Metodología y Fuentes*. Mapfre, Madrid.
- Navarro, E., Ferrer, R., Gonzalez, C., y Nave, J. (1995). *Tablas de Mortalidad de la Comunidad Valenciana 1990-91. Censos de Población i Habitatges*, volumen I. IVE, Valencia.
- Nayadara, E. (1964). On estimating regression. *Theory of Probability and its Applications*, 9:141–142.
- Nielsen, J. (1992). A transformation approach to bias correction in kernel hazard estimation. Technical Report 115, University of Copenhagen: Laboratory of Actuarial Mathematics.
- Nieto, U. y Vegas, J. (1993). *Matemática Actuarial*. Mapfre, Madrid.
- Olivieri, A. (2001). Uncertainty in mortality projections: an actuarial perspective. *Insurance: Mathematics and Economics*, 29(4):231–245.
- Palacios, H. (1996). *Introducción al Cálculo Actuarial*. Mapfre, Madrid.
- Perks, W. (1825). On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, 63:12–40.
- Prieto, E. y Fernández, M. (1994). *Tablas de mortalidad de la población española de 1950 a 1990: tabla proyectada del año 2000: tablas con margen y sin margen de seguridad*. Aseguradora, Madrid.
- Ramlau-Hansen, H. (1983). The choice of a kernel function in the graduation of counting process intensities. *Scandinavian Actuarial Journal*, páginas 165–182.
- Renshaw, A. y Haberman, S. (1997). Dual modelling and select mortality. *Insurance: Mathematics and Economics*, 19(2):105–126.
- Renshaw, A. y Haberman, S. (2000). Modelling for mortality reduction factors. *Actuarial Research Report*, 127.
- Renshaw, A. y Haberman, S. (2001). On the forecasting of mortality reduction factors. *Actuarial Research Report*, 135.

-
- Renshaw, A. y Haberman, S. (2002). Lee-carter mortality forecasting with age specific enhancement. *IME 2002*.
- Renshaw, A. y Haberman, S. (2003). Lee-carter mortality forecasting: a parallel generalized linear modelling approach for england and wales mortality projections. *Journal of the Royal Statistical Society C*, 52:119–137.
- Renshaw, A., Haberman, S., y Hatzopoulos, P. (1996). The modelling of recent mortality in united kingdom male assured lives. *British Actuarial Journal*, 2(II):449–477.
- Renshaw, A., Haberman, S., y Hatzopoulos, P. (1997). On the duality of assumptions underpinning the construction of life tables. *Astin Bulletin*, 27(I):5–22.
- Renshaw, E. (1991). Actuarial graduation practice and generalised linear a non-linear models. *Journal of the Institute of Actuaries*, 118(II):295–312.
- Renshaw, W. y Hatzopoulos, P. (1996). On the graduation of amounts. *British Actuarial Journal*, 2(I):185–205.
- Rue, M. (1992). *Les Lleis de mortalitar: Un ajust parametric per a Catalunya y España*. Tesis doctoral, Facultad de Ciencias Economicas y Empresariales. Departamento de Matematica economica, financiera y actuarial.
- Sithole, T., Haberman, S., y Verrall, R. J. (2000). An investigation into parametric models for mortality projections, with applications to immediate annuitants' and life office pensioners' data. *Insurance: Mathematics and Economics*, 27(4):285–312.
- Thiele, P. (1972). On a mathematical formula to express the rate of mortality throughout the whole of life. *Journal of the Institute of Actuaries*, 16:313.
- Vegas, A. (1982). *Estadística, Aplicaciones Económicas y Actuariales*. Piramide, Madrid.
- Venables, W. y Ripley, B. D. (1994). *Modern Applied Statistics with S-PLUS*. Springer, Nueva York.
- Verrall, R. (1993a). A state space formulation of whittaker graduation, with extensions. *Mathematics and Economics*, 13:7–14.
- Verrall, R. (1993b). Graduation by dynamic regression methods. *Journal of the Institute of Actuaries*, 120:153–170.
- Verrall, R. (1996). A unified framework for graduation. *Actuarial Research Paper*, 91:2–25.
- Villalón, J. G. (1994). *Manual de Matemáticas Financiero-Actuariales*. Fernández Ciudad, S L, Madrid.
- Villalón, J. G. (1997). *Operaciones de seguros clásicas y modernas*. Pirámide, Madrid.
- Watson, G. (1964). Smooth regression analysis. *Sankhya A*, 26:359–372.

Wilmoth, J. (1993). Computational methods for fitting and extrapolating the lee-carter model of mortality change. *Technical Report, Department of Demography, University of California, Berkeley.*