

A MODELIZACIÓN DOS CAMBIOS NA LONXEVIDADE DA POBLACIÓN DO PAÍS BASCO E A SÚA ESTIMACIÓN FUTURA

Amaia Jone BETZUEN ÁLVAREZ

amaiajone.betzuen@ehu.es

Amancio BETZUEN ZALBIDEGOITIA

amancio.betzuen@ehu.es

Universidade do País Basco

RESUMO

O obxectivo deste traballo é estimar a tendencia futura da mortalidade da poboación do País Basco a través da mellora da mortalidade para o período 1991-2013, para o cal o estudo é validable. A investigación realízase a través dun modelo estocástico bietápico ao que realizamos dúas achegas. Por unha banda, unha información inicial supervisada que posibilitará prognosticar mellor o avellentamento da poboación do País Basco baixo os criterios de Solvencia II. E por outra, na estimación inclúese unha variante na extensión do parámetro do tempo de calendario mediante unha regresión logarítmica lineal no canto da coñecida "random walk with drift".

CLASIFICACIÓN JEL: C22, C53, G17 e G22

PALABRAS CHAVE: Medida da mellora da mortalidade, Estimación futura do avellentamento, Risco de supervivencia, Poboación do País Basco e EPSVs.

ABSTRACT

The aim of this paper is to estimate the future trend of the Basque population's mortality focusing on its improvement for the period 1991-2013 for which the study is validable. The research involves the use of a two-stage stochastic model graduation, for which we add two contributions. On the one hand, supervised initial information will help in forecasting the ageing of Basque Country's population under Solvency II guidelines. On the other hand, when estimating, a variant is included in the extension of the calendar time's parameter using a logarithmic-linear regression, rather than the known "random walk with drift".

JEL CLASSIFICATION: C22, C53, G17 and G22

KEY WORDS: Longevity improvements measure, Forecasting aging, Survivor risk, Basque Country Population, and Voluntary Social Welfare Entities (EPSVs).

1. INTRODUCCIÓN

A estimación futura da lonxevidade sempre foi un tema prioritario entre os investigadores do campo demográfico e actuarial de todo o mundo, polo grande impacto que supón o custo das pensións públicas, as pensións privadas, os coidados de longa duración, a asistencia sanitaria e, no caso que nos atangue, o tema das pensións vitalicias que as Entidades de Previsión Social Voluntaria do País Basco deben garantir.

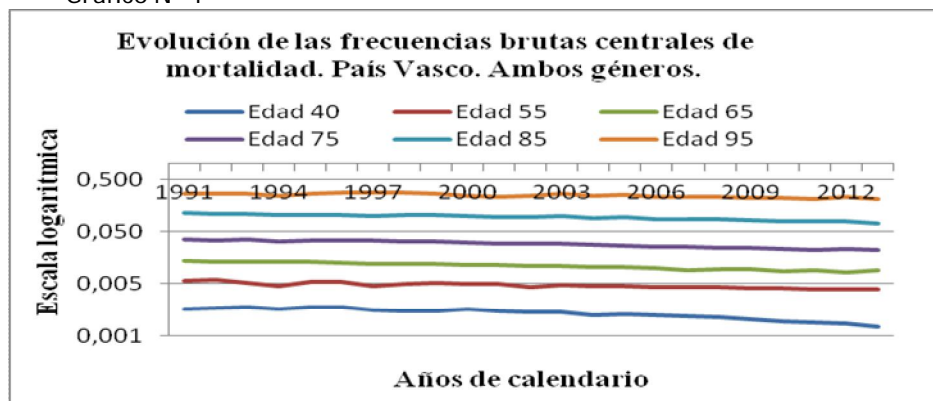
Este traballo estruturouse da seguinte forma: en primeiro lugar, seleccionouse o período de anos de calendario óptimo para obter os datos a incluír no modelo matemático de graduación e elixiuse o modelo a utilizar para a estimación de parámetros. A continuación, procedeuse á estimación dos parámetros mediante o proceso de graduación de Lee-Carter (1992) introducindo unha mellora na estimación do parámetro que depende do tempo de calendario. Posteriormente, realizouse unha análise dos resultados da estimación, así como da súa volatilidade e das táboas de supervivencia, e, finalmente, ofrécense varias conclusións con respecto á investigación realizada.

2. ESCOLLA DO TRAMO DE ANOS DE CALENDARIO A ESTUDO E DO MODELO DE ESTIMACIÓN INICIAL DOS PARÁMETROS DE GRADUACIÓN

Un dos puntos importantes no estudo dun colectivo concreto está representado polo intervalo de captación de información. Neste caso, a capacidade de escolla era limitada, dado que para o colectivo da poboación do País Vasco, a súa diferenciación por xénero supoñía que para certas persoas de idade moi nova, como son as comprendidas entre os tres e os dez anos, o número de falecidos non chegaba a cinco persoas, número que se estima necesario para proceder a unha graduación cun mínimo de garantías. Así mesmo, o tramo de idades seleccionado para a investigación será sempre superior aos 39 anos, dado que as prestacións por pensións teñen vixencia para idades superiores e ademais, a partir desta idade evitamos a incidencia das turbulencias que se producen para idades comprendidas entre os vinte e os trinta anos, debido aos riscos propios desas idades, como accidentes de diversa índole, que perturban a evolución parsimoniosa da propensión á mortalidade.

Do tratamento dos datos, obtivéronse as frecuencias que se amosan a continuación no seguinte gráfico:

Gráfico N° 1



Fonte: Elaboración propia.

No gráfico preséntanse as frecuencias centrais brutas a intervalos basicamente de 10 anos de calendario.

A tendencia da mellora na mortalidade é clara como unha diminución continuada e regular. Co obxecto de aplicar o máis fielmente posible a base de datos referente ao colectivo a estudo é necesario inspeccionala previamente. Por iso, debe ser un experto na materia quen elixa o modelo, e o que seleccione o período de tempo de calendario a incluír no mesmo.

En todo caso, e tendo en conta as directrices sinaladas en Solvencia II¹, este tramo debe corresponder aos anos de calendario máis recentes, evitando incluír períodos nos que, por algunha circunstancia exógena ao devir normal dos acontecementos, produciranse alteracións nos resultados de forma anómala. Por iso, eliximos un modelo estocástico dos factores.

¹ É unha Directiva que incorpora unhas normas Europeas do Seguro para protexer esta industria, garantir a súa solvencia e propiciar produtos aseguradores en termos de obxectividade, transparencia e equidade para os cidadáns.

Neste traballo de investigación usaremos o modelo de Lee-Carter (1992) polas súas características de sinxeleza e estimación conxunta dos parámetros, que dependen da idade das persoas e do tempo de calendario para a estimación da mortalidade cara ao futuro. Este modelo estocástico, recolle a aleatoriedade dos resultados e o ruído que a base de datos contén, como consecuencia do deterioro na captación da información, os erros típicos da mostra, etc.

Unha forma de representación do modelo é por medio da seguinte ecuación:

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

onde,

$m_{x,t}$ é o tanto bruto central de mortalidade á idade x e no ano t .

a_x é o parámetro que representa a evolución dos tantos promedio centrais en función da idade.

b_x é o parámetro que representa a variación dos tantos promedio centrais en función da idade ao longo do tempo.

k_t é o parámetro que representa a evolución da mortalidade a través do tempo de calendario.

$\varepsilon_{x,t}$ é o termo erro de estimación.

No modelo a utilizar, constrúese unha matriz de datos de dobre entrada, por idade e por tempo de calendario. Deste xeito, captúrase a variación da frecuencia bruta da mortalidade que se inclúe nesta matriz, principalmente a través dos anos de calendario que son a base para a nosa estimación futura.

Sobre a base desta información, o proceso de graduación proporciona uns resultados sobre a tendencia futura da mortalidade, para os anos de calendario sucesivos, e no noso caso, para idades superiores aos 40 anos. O obxectivo é estimar o máis fielmente posible a evolución da supervivencia futura das personas contratantes de operacións, que como a hipoteca inversa ou plans de pensións complementarios, requiren dunha valoración o máis xusta posible.

A través do "*Singular Value Decomposition (SVD)*" da matriz constituída, obtéñense os valores de b_x e k_t . O procedemento consiste en minimizar a suma dos cadrados dos erros obtidos calculando as súas derivadas primeiras respecto dos tres parámetros, da seguinte forma, como ilustra o traballo de Betzuen, A. (2010):

$$\hat{a}_x = \frac{\sum_t (\ln m_{x,t} - b_x * k_t)}{T} \quad \hat{b}_x = \frac{\sum_t k_t (\ln m_{x,t} - \hat{a}_x)}{\sum_t k_t^2} \quad \hat{k}_t = \frac{\sum_x b_x (\ln m_{x,t} - \hat{a}_x)}{\sum_x b_x^2}$$

Estas estimacións proporcionan un conxunto de ecuacións normais para cada parámetro².

² Co obxecto de garantir unha única solución ao modelo débense establecer certas condicións de contorna que no noso caso foron:

$$\sum_x \hat{b}_x^2 = 0 \quad \sum_t \hat{k}_t = 1$$

Aínda que existen outras restricións para a total identidade do modelo, entendemos que estas son as máis sinxelas.

3. ESTIMACIÓN DOS PARÁMETROS E ALGUNHAS CORRECCIÓN S AO MODELO PARA A PROXECCIÓN FUTURA DO AVELLAMENTO

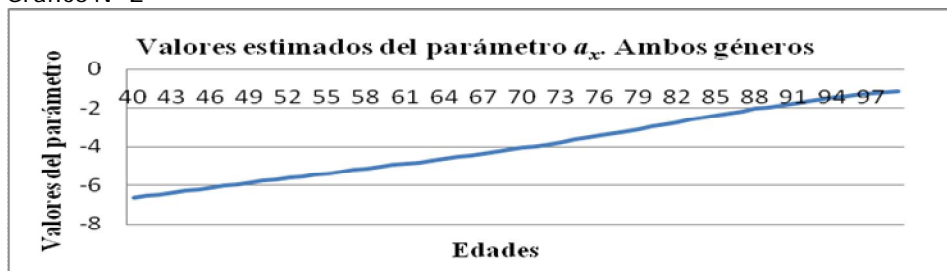
Nunha primeira aproximación, obtemos o valor do parámetro promedio a_x que depende unicamente da idade das persoas.

Pódese obter directamente a partir da consideración dos valores formados por:

$$a_x = \frac{\sum_{t=1991}^{2011} m_{x,t}}{T}$$

A estimación obtida preséntase no seguinte gráfico:

Gráfico N° 2



Fonte: Elaboración propia.

No gráfico observamos os valores do parámetro a_x que representa a evolución dos tantos centrais da mortalidade para todos os anos de calendario elixidos.

Subliñamos o feito de que para o caso do País Basco os valores do parámetro, no decorrer dos anos de calendario elixidos, son os normais ao comparalos cos países da contorna europea.

Para a estimación do segundo parámetro que captura información da mortalidade para cada idade, construímos unha colección de datos usando a seguinte ecuación:

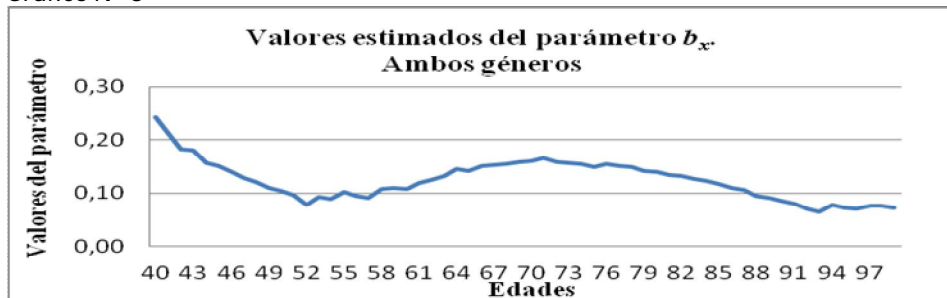
$$\ln m_{x,t} - a_x$$

onde os valores do parámetro a_x corresponden aos valores xa estimados no paso anterior. Trátase de minimizar a suma dos cadrados dos erros dados por:

$$\sum_{x,t} [(\ln m_{x,t} - a_x) - b_x * k_t]^2 \quad (2)$$

O valor do parámetro b_x para cada idade x represéntase no seguinte gráfico:

Gráfico N° 3



Fonte: Elaboración propia.

Os valores do parámetro combinado b_x representan a mellora dos tantos centrais de mortalidade, para cada idade incluída no modelo.

O seguinte gráfico recolle a información da mellora da mortalidade de forma lonxitudinal que vén representada polo parámetro k_t . Nunha primeira aproximación os valores da estimación son os seguintes:

Gráfico N° 4



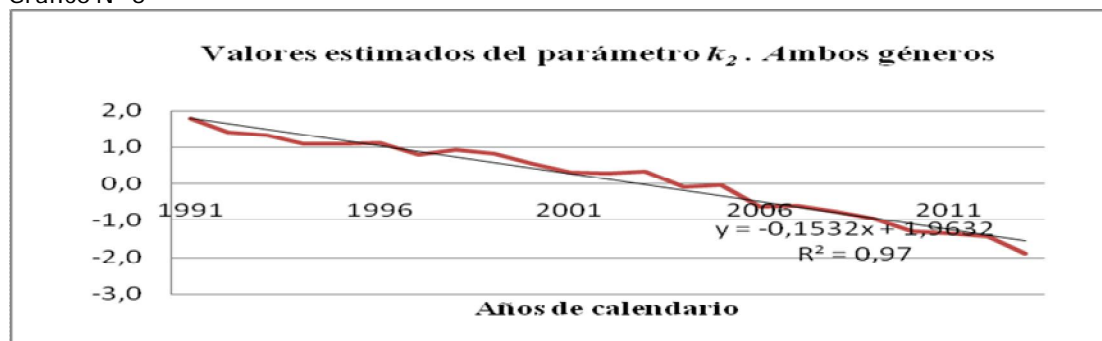
Fonte: Elaboración propia.

Os valores do parámetro combinado k_t representan a evolución no tempo de calendario dos tantos centrais de mortalidade para todo o período incluído no modelo. Obsérvase unha tendencia casi lineal na evolución dos valores do parámetro co decorrer do tempo de calendario.

Unha vez estimados os valores dos parámetros, realizamos un primeiro contraste con algúns dos valores reais obtidos inicialmente sobre o colectivo real. No mundo científico, resulta habitual comprobar a distribución do número de falecidos para cada ano de calendario. Por iso, realizamos unha re-estimación do parámetro dependente do tempo de calendario, incluíndo *a posteriori* esta restrición, como recolle Wilmoth (1993). Desta forma, corríxese en certa medida, á súa vez, a estimación da esperanza matemática de vida, do colectivo actual e para os anos futuros.

Introducimos unha aproximación por medio dunha regresión lineal. Convén destacar que dada a aceptable regularidade dos valores de k_t , aínda que unha regresión de tipo cadrático proporciona un R^2 mellor (0,988) que o R^2 da regresión lineal (0,975), despois de estudar a tendencia da esperanza matemática de vida, entendemos que con esta última, a esperanza de vida segue unha tendencia máis racional en base á experiencia e en base á análise contrastada, coas tendencias dos países da contorna cuxa mellora na mortalidade está máis próxima ao noso, como é caso do Reino Unido, Alemaña, Suecia, Holanda, Francia, etc. Visualmente pódese percibir a regresión aos valores do parámetro k_t mediante o seguinte gráfico.

Gráfico N° 5

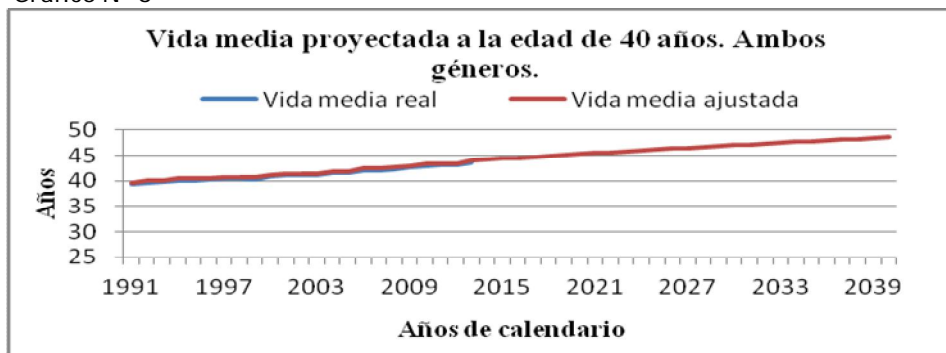


Fonte: Elaboración propia.

Os novos valores do parámetro k_t proporcionan unha tendencia un chisco máis suave que se se fixera uso dunha regresión cadrática. Por outra banda, aproximan notablemente mellor os valores da esperanza matemática de vida.

Pódese apreciar que os valores estimados están, para os anos intermedios, como para o período 2000-2005, aproximadamente, un chisco por enriba dos valores reais publicados por institucións oficiais como o INE. A tendencia da esperanza matemática de vida cara ao futuro é claramente regular. Conforme este modelo, ao que se incorpora a nosa corrección, a tendencia preséntase crecente cara ao futuro, pero este crecemento é suave e regular cunha lixeira diminución no crecemento da tendencia como se amosa no seguinte gráfico.

Gráfico N° 6



Fonte: Elaboración propia.

Os valores da esperanza matemática de vida, para unha persoa de 40 anos, seguen os valores brutos da mesma, con grande exactitude.

Os valores estimados da esperanza matemática de vida para unha persoa de 40 anos preséntanse, para poder ser utilizados de forma numérica na seguinte táboa.

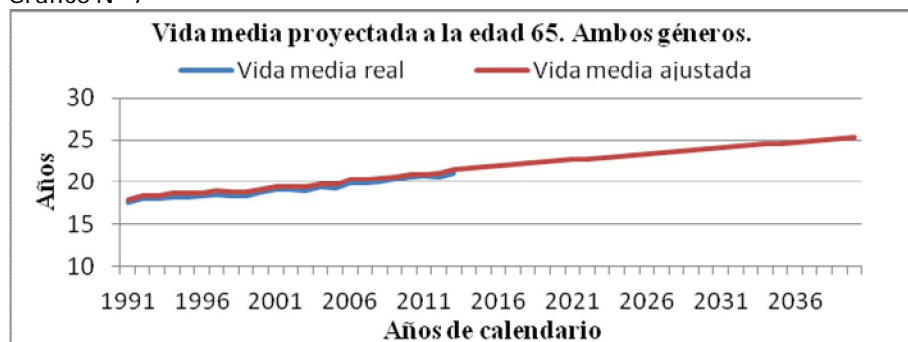
Táboa N° 1

Esperanza matemática de vida					
Idade 40 anos					
Ano	$e_{(x)}$	Ano	$e_{(x)}$	Ano	$e_{(x)}$
1991	39,53	2008	42,72	2025	45,63
1992	39,73	2009	42,90	2026	45,79
1993	39,92	2010	43,08	2027	45,95
1994	40,11	2011	43,25	2028	46,11
1995	40,30	2012	43,43	2029	46,27
1996	40,49	2013	43,61	2030	46,43
1997	40,68	2014	43,78	2031	46,58
1998	40,87	2015	43,95	2032	46,74
1999	41,06	2016	44,13	2033	46,89
2000	41,25	2017	44,30	2034	47,05
2001	41,44	2018	44,47	2035	47,20
2002	41,62	2019	44,64	2036	47,35
2003	41,81	2020	44,81	2037	47,50
2004	41,99	2021	44,97	2038	47,64
2005	42,17	2022	45,14	2039	47,79
2006	42,36	2023	45,31	2040	47,94
2007	42,54	2024	45,47		

Fonte: Elaboración propia

A continuación presentamos os valores estimados proxectados da esperanza matemática de vida para persoas de 65 anos ou máis que é o grupo máis significativo para a cobertura de pensións e seguros, como tamén recolle Byrne (2005) no seu traballo.

Gráfico N° 7



Fonte: Elaboración propia.

Os valores da esperanza matemática de vida para unha persona de 65 anos, seguen os valores brutos da mesma con grande exactitude. Os valores estimados da esperanza matemática de vida para unha persoa de 65 anos preséntase de forma numérica na seguinte táboa.

Táboa N° 2

Esperanza matemática de vida					
Idade 65 anos					
Ano	Ex	Ano	ex	Ano	ex
1991	17,83	2008	20,40	2025	22,78
1992	17,98	2009	20,54	2026	22,92
1993	18,14	2010	20,69	2027	23,05
1994	18,29	2011	20,83	2028	23,18
1995	18,45	2012	20,98	2029	23,31
1996	18,60	2013	21,12	2030	23,44
1997	18,75	2014	21,26	2031	23,57
1998	18,90	2015	21,41	2032	23,70
1999	19,05	2016	21,55	2033	23,83
2000	19,21	2017	21,69	2034	23,96
2001	19,36	2018	21,83	2035	24,08
2002	19,51	2019	21,97	2036	24,21
2003	19,66	2020	22,11	2037	24,33
2004	19,81	2021	22,24	2038	24,45
2005	19,95	2022	22,38	2039	24,58
2006	20,10	2023	22,52	2040	24,70
2007	20,25	2024	22,65		

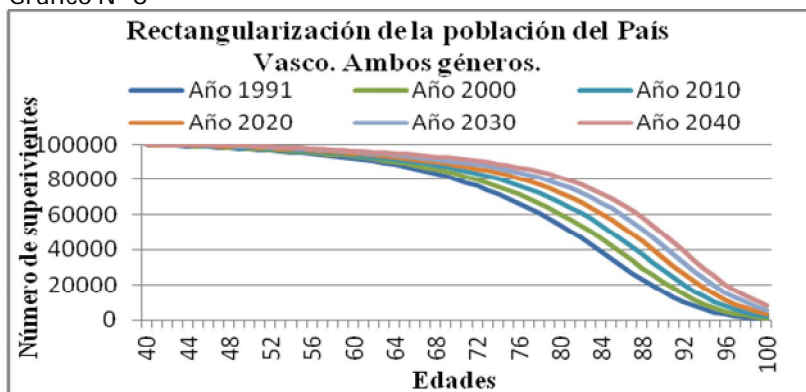
Fonte: Elaboración propia.

A tendencia cara ao futuro, como podemos observar, é crecente pero a unha intensidade un chisco inferior á resultante para o tramo de datos histórico 1991-2013.

Nos últimos anos están xurdindo novos mercados de aplicación de índices de mortalidade. Referímonos ao mercado de bonos a emitir para a cobertura dos riscos de supervivencia,

swaps de mortalidade, titulizacións, etc, onde cremos que os investidores estarían interesados en investir sempre que houbera unha medida que reflectise coa maior exactitude posible a tendencia do risco de mortalidade, como ocorre co noso traballo.

Gráfico N° 8



Fonte: Elaboración propia.

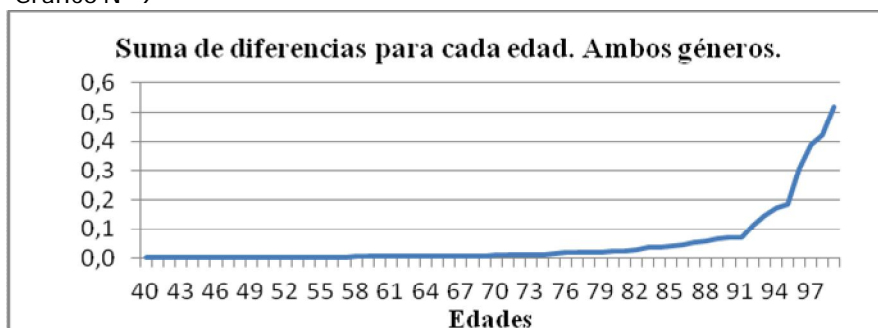
Os valores correspondentes ao número de superviventes, a partir dos 40 anos (no noso caso a estudo), tanto para o período de datos histórico como para o período de anos de calendario proxectados, amosan unha tendencia regular cara á mellora da supervivencia. Obsérvase no desprazamento cara á dereita da curva de supervivencia.

4. MOSTRAS DA BONDADE DO MODELO

A continuación presentamos algunhas evidencias que amosan a bondade dos resultados obtidos polo modelo, co fin de indagar sobre a fiabilidade dos mesmos. Existen múltiples procedementos, máis ou menos ortodoxos, para averiguar a posible verosimilitude dos resultados obtidos dun axuste estático pero non dunha proxección futura, como é o noso caso.

A continuación amosamos graficamente a suma das diferenzas absolutas en sentido horizontal, isto é, para cada idade, sumando por tanto, as diferenzas ao longo de todo o período de calendario considerado no estudo.

Gráfico N° 9



Fonte: Elaboración propia.

Corresponden aos valores da suma de diferenzas absolutas tomadas de forma transversal, para cada unha das diferentes idades incluídas no modelo. A evidencia é clara. As desviacións son inapreciables se se ten en conta que corresponden ás sumas absolutas, de todo o tramo de anos de calendario (1991-2013) das desviacións para cada idade. O

aumento das desviacións cara as idades máis altas é un comportamento razoable en todos os países da nosa cotorna, como consecuencia dun menor número de datos dispoñibles e dunha peor precisión na toma de información.

No seguinte gráfico, amósase o resultado de sumar as diferenzas absolutas por anos de calendario, e por tanto, corresponden á suma para todas as idades (40-100) co mesmo criterio que no parágrafo anterior.

Gráfico N° 10



Fonte: Elaboración propia.

Os valores seguen sendo insignificantes. Dado que observamos que o comportamento das frecuencias brutas presenta certos altibaixos, procedemos á análise da volatilidade da mortalidade do colectivo. Os resultados numéricos obtidos para todo o horizonte temporal foron:

Táboa N° 3

Idade	Volatilidade en termos porcentuais
40	2,14
55	5,17
65	1,88
75	1,59
85	2,11
95	4,74

Fonte: Elaboración propia.

Chama a atención o resultado obtido para a idade de 55 anos. Porén, todos os valores están en consonancia cos resultados observables para outros países da contorna europea.

5. CONCLUSIÓNS

As melloras na mortalidade da poboación do País Basco observadas ao longo deste traballo, lonxe de diminuír, seguirán aumentando cara ao futuro, a un ritmo similar ao que observamos nos últimos 25 anos de datos históricos.

Para a estimación futura da mellora na mortalidade do País Basco utilizamos basicamente o modelo de Lee-Carter, pero introducimos unha pequena variación ao utilizar como proxección do parámetro dependente do tempo de calendario, unha regresión lineal, nunha contorna logarítmica, no canto dunha estimación "random walk with drift". A mellora non é significativa pero suaviza a tendencia da mellora cara ao futuro máis acorde coa evolución da supervivencia nos países da nosa contorna.

Os valores presentados nos cadros correspondentes ás esperanzas matemáticas de vida construíronse conforme a unha métrica de supervivencia coas máximas garantías para

seren utilizadas non só na medida futura da supervivencia das persoas, sobre todo maiores, senón que tamén sería un indicativo solvente a utilizar nun mercado que se podería construír para negociar produtos de cobertura nos riscos de supervivencia, tales como bonos de supervivencia, swaps de mortalidade, etc, como podemos observar no traballo de Dowd (2006).

A calibración dos resultados da graduación futura do risco de mortalidade dun colectivo como o que aquí se estuda, é imprescindible no marco da Directiva europea Solvencia II, que entra en vigor o 1 de xaneiro do ano que vén e afecta directamente á solvencia das Entidades de Previsión Social Voluntaria do País Basco, en canto que deben garantir a súa solvencia ante as obrigas no pago das prestacións en forma de pensión vitalicia e, por tanto, a moi longo prazo.

BIBLIOGRAFÍA

Andreev, K.F. (2004) "A method for estimating size of population aged 90 and over with application to the 2000 USA census data". Demographic Research 11. pp. 235-262.

Betzuen, A. (2010). "Un análisis sobre las posibilidades de predicción de la mortalidad futura aplicando el modelo Lee-Carter". Anales Instituto Actuarios Españoles. PP. 111-138.

Byrne, A. and Harrison, D. (2005). "Longevity risk weighs on trustees". Pension's management. Disponible en www.pensions-management.co.uk.

Dowd, K. Blake, D. Cairns, A. and Blake, D. (2006) "Mortality Dependent financial risk measures". Insurance: Mathematics and Economics, 38. PP. 427-440.

Instituto Nacional de Estadística. Información demográfica y de la población. Disponible en www.ine.es.

Lee, R. and Carter, L.R. (1992). "Modeling and Forecasting US mortality". Journal of the American Statistical Association, 87, pp. 659-671.

Solvencia II. Disponible en DGSFP.mineco.es.

Wilmoth, J.R. (1993) "Computational Methods for Fitting and Extrapolating the Lee-Carter Model of Mortality Change". National Institute on Aging. USA.