



UNIVERSIDAD  
**PABLO DE  
OLAVIDE**  
SEVILLA



REVISTA DE MÉTODOS CUANTITATIVOS PARA  
LA ECONOMÍA Y LA EMPRESA (4). Páginas 56-74.  
Diciembre de 2007. ISSN: 1886-516X. D.L: SE-2927-06.  
URL: <http://www.upo.es/RevMetCuant/art14.pdf>

## Asignación de primas en el seguro del automóvil utilizando el Análisis en Componentes Principales Funcionales

SEGOVIA GONZÁLEZ, M. MANUELA

Departamento de Economía, Métodos Cuantitativos e Historia Económica.

Universidad Pablo de Olavide

Correo electrónico: [mmseggon@upo.es](mailto:mmseggon@upo.es)

GUERRERO CASAS, FLOR M.

Departamento de Economía, Métodos Cuantitativos e Historia Económica.

Universidad Pablo de Olavide

Correo electrónico: [fguecas@upo.es](mailto:fguecas@upo.es)

HERRANZ PEINADO, C. PATRICIA

Departamento de Economía, Métodos Cuantitativos e Historia Económica.

Universidad Pablo de Olavide

Correo electrónico: [pherpei@upo.es](mailto:pherpei@upo.es)

### RESUMEN

El estudio de las primas de riesgo en el seguro del automóvil es de suma importancia, debido a la gran competitividad existente en el mercado asegurador. En este artículo se pretende proponer un sistema de bonificación-penalización de las primas de riesgo utilizando el Análisis Funcional en Componentes Principales.

En concreto, se realiza un estudio empírico con los datos reales de una compañía aseguradora con cobertura nacional. Aplicando el sistema propuesto de bonificación-penalización a los perfiles de comportamiento de los asegurados, se obtienen las primas de riesgo para cada uno de los perfiles considerados y tramos de edad.

**Palabras clave:** primas; seguro del automóvil; análisis funcional en componentes principales; riesgo.

**Clasificación JEL:** C44; C73; D81; G22.

**2000MSC:** 62M99; 91B30; 91B70.

# Premium allocation in the car insurance by using Functional Principal Component Analysis

## ABSTRACT

The study of the premium risk in the car insurance is really important because the insurance market is very competitive. In this article we show a bonus-malus method of risk premiums. To do this, we use the Functional Principal Component Analysis.

More precisely, we explain an empirical study with the real data of an insurance company. By applying the bonus-malus method that we have given, we are able to obtain the premium risk for different profiles and age ranges.

**Keywords:** premium; car insurance; functional principal component analysis; risk.

**JEL classification:** C44; C73; D81; G22.

**2000MSC:** 62M99; 91B30; 91B70.



## 1. Introducción

El seguro del automóvil en España ha venido experimentando grandes cambios en los últimos años, tal y como ocurre en otros países industrializados. La gran competitividad que experimenta el mercado del seguro en general y el ramo de autos en particular, ha provocado entre otras medidas que las compañías aseguradoras cada vez estén más interesadas en establecer métodos eficientes para el control de la evolución de los accidentes y sus costes asociados (Ayuso y Guillén, 1999).

Cuando se realizan estudios sobre la siniestralidad en un determinado ramo del seguro, interesa conocer la realidad lo más ampliamente posible identificando aquellos individuos que son más propensos al riesgo. Cuanto mayor sea la información que proporcione un mejor conocimiento del riesgo, mejor se podrá estudiar el riesgo de cada individuo y, por tanto, mejor se podrá ajustar el precio del seguro. Para ello, es fundamental contar tanto con datos que faciliten detalles de la siniestralidad y de las pólizas suscritas, como modelos estadísticos adecuados con los que se pueda extraer la información que precisa cada Entidad a la hora de tarificar el riesgo. Autores como Boj *et al.* (2005) describen el escenario de bases de datos y estadísticas sectoriales en España, convirtiendo su trabajo en un referente para el estudio a posteriori de los siniestros.

En el negocio del seguro el equilibrio técnico es su prioridad; por un lado, las compañías pretenden ofrecer buenos precios que sean competitivos y adecuados al mercado, y al mismo tiempo deben hacer una política de selección de riesgos que les garantice la solvencia del ramo y que les permita conseguir el mayor beneficio posible. Las normas de contratación de cada compañía es una de las actividades de mayor trascendencia en cuanto a los resultados de su negocio. Existen estudios que describen los comportamientos de fraude teniendo en cuenta las características de los asegurados y de las pólizas (Artis *et al.*, 1999, Ayuso *et al.*, 1999, Artis *et al.*, 2002).

Las compañías aseguradoras que trabajan el ramo de automóvil cuentan con datos tanto externos como datos de su propia cartera de asegurados para su estudio. Uno de los problemas con los que se encuentran, es dar con el sistema de análisis estadístico más adecuado para el cálculo de la prima de riesgo que le proporcione fiabilidad de que las tarifas aplicadas se asemejarán al comportamiento real de la cartera. Tal como recomienda la Unión Española de Entidades Aseguradoras y Reaseguradoras (UNESPA), así como aparece en diversos estudios (Melgar y Guerrero, 2005), se conoce la existencia de una serie de características que a priori son más influyentes a la hora de tener una mayor o menor siniestralidad. Dichas características son: la gama del vehículo, la zona de circulación, el uso del vehículo y las circunstancias personales del conductor, como la edad, el sexo, los años de posesión del carné de conducir, etc. Algunas de ellas son tenidas en cuenta por las aseguradoras para fijar la cuantía de las primas de riesgo; además, para fidelizar a sus buenos conductores, aplican sistemas de bonificación según las normas internas establecidas.

Para la selección de factores de riesgo en la tarificación y el cálculo actuarial podemos encontrar trabajos que realizan análisis multivariante como Boj *et al.* (2004) y Guillén *et al.* (2005).

Nosotros, en este artículo, pretendemos mostrar a las compañías aseguradoras un nuevo sistema de asignación de primas de riesgo de automóviles; para ello nos basaremos en las características descritas como influyentes a la hora de su potencial siniestralidad. Se estudiarán, por perfiles, grupos de individuos que verifiquen una serie de características comunes y trataremos de forma muy específica la edad del conductor habitual. En concreto, definimos un nuevo método que nos servirá para penalizar o bonificar la prima de los individuos según el tramo de edad en el que se encuentren. Este sistema de bonificación-penalización permite discriminar entre los individuos asegurados con un mismo perfil según su tramo de edad y opera a favor de una mayor competitividad de la compañía aseguradora. Además puede resultar útil a las aseguradoras para la búsqueda de clientes óptimos.

En la sección 2, describimos los datos con los que vamos a realizar nuestro trabajo empírico, y establecemos los perfiles objeto de nuestro estudio, así como la base del análisis funcional en componentes principales (AFCP). Hemos elegido esta técnica porque se realiza el análisis de componentes principales teniendo en cuenta el tiempo, edad de los asegurados, que es la variable que consideramos más importante y novedosa para las compañías aseguradoras. Para aplicar este método hemos realizado inicialmente una aproximación funcional<sup>1</sup> utilizando técnicas de suavizado (Ramsay y Silverman, 1997). En el caso que nos ocupa, los datos no se comportan de manera estable, no detectándose ningún tipo de periodicidad en ellos, por tanto la aproximación por medio de series de Fourier no es adecuada. Necesitamos, pues, una gran flexibilidad para poder ajustar de forma apropiada los datos. Por ello, hemos utilizado funciones base del tipo B-spline (Green y Silverman, 1994). Dichos datos son del territorio español y han sido cedidos por una compañía de seguros de ámbito internacional.

En la sección 3, mostramos los resultados obtenidos al aplicar el AFCP que nos determina los tramos de edades que se utilizarán en el resto del trabajo.

En la sección 4, establecemos un sistema de bonificación-penalización para cada uno de los perfiles y tramos de edades detectados. Para ello, se define una primera ponderación en función de las componentes principales funcionales y la edad; seguido de una segunda ponderación según el máximo riesgo de ocurrencia de un siniestro y los tramos de edad establecidos. Con estas dos ponderaciones, calculamos los pesos definitivos que aplicaremos a las primas calculadas por la compañía por el sistema tradicional, obteniendo la prima de riesgo que proponemos más equilibrada.

Finalmente, en la sección 5 se exponen las principales conclusiones y se termina con la bibliografía utilizada en la realización del trabajo.

---

<sup>1</sup>La aplicación detallada se puede ver en Segovia (2005).

## 2. Materiales y métodos

### 2.1. Datos

Hemos de destacar que es difícil obtener información acerca de los asegurados, dadas las restricciones a que obliga la Ley Orgánica de Protección de Datos. No obstante, disponemos de los datos de una compañía de seguros con una amplia implantación a nivel nacional.

En nuestro trabajo, nos restringiremos a los individuos asegurados que conduzcan un vehículo de tipo turismo y cuyo uso sea particular, debido a que si considerásemos otro tipo de vehículos como es el caso de camiones, motos, taxis, etc., se podrían distorsionar los resultados, al hacer un uso distinto del vehículo con un riesgo de siniestro diferente, y por tanto con cálculo de prima distinto. Además, se tendrán en cuenta sólo los siniestros en los que la culpa haya sido directamente del asegurado de la compañía estudiada, o bien, compartida con el asegurado de otra entidad. Si contabilizáramos los siniestros de una póliza en los que el asegurado no tiene la responsabilidad, sino que el culpable ha sido el otro individuo implicado, estaríamos falseando la información al no conocer las características del que comete la falta, ya que la póliza pertenece a otra compañía.

Nuestro trabajo lo hemos desarrollado considerando una serie de perfiles basados en las características del conductor. Para ello, se ha realizado una estratificación de la muestra según el siguiente detalle: sexo, zona geográfica y tipo de turismo. La variable sexo se justifica por los previsibles diferentes comportamientos entre los hombres y las mujeres a la hora de conducir. La zona geográfica, ya que la climatología, la densidad de población y las infraestructuras actúan como componentes importantes en los resultados obtenidos. Se considerarán tres zonas distintas: la zona centro-norte (Castilla-León, Castilla-La Mancha, Madrid, Aragón, La Rioja, Galicia, Asturias, Cantabria, País Vasco y Navarra), la zona mediterránea (Cataluña, Valencia, Murcia y Baleares) y la zona sur (Andalucía, Extremadura y Canarias). Se optó por realizar una agregación territorial debido a que en determinadas zonas no disponíamos de tamaño muestral suficiente para realizar el estudio. Y por último, el tipo de turismo, distinguiendo entre gama baja y media-alta. Un coche de gran tamaño y potencia puede provocar, a iguales circunstancias de accidentes, mayores daños que uno pequeño; de la misma manera, también serán mayores los gastos de reparación de los daños propios del vehículo. Como resultado de esta estratificación, trabajaremos con 12 grupos de individuos.

Una vez aplicadas las restricciones anteriores y depurados los datos nos encontramos con 175.191 asegurados y 30.483 siniestros que serán el objeto de nuestro estudio. En la Tabla 1 se muestran los distintos perfiles con los que vamos a trabajar. De los 175.191 asegurados, 30.987 son mujeres y 144.204 hombres, correspondiendo a un 17,69 % y 82,31 %, respectivamente. De todos ellos, 92.509 (52,8 %) corresponde a coches de gama media-alta y 82.682 (47,2 %) a coches de gama baja. Con respecto a la distribución por zonas geográficas que consideramos, en la zona sur tenemos 95.520 (54,5 %) asegurados; en la zona centro-norte, 49.151 (28,1 %) y en la zona mediterránea, 30.520 (17,4 %).

Perfiles	Sexo	Gama del turismo	Zona geográfica
1	Mujer	Media-Alta	Sur
2	Mujer	Media-Alta	Centro-Norte
3	Mujer	Media-Alta	Mediterránea
4	Mujer	Baja	Sur
5	Mujer	Baja	Centro-Norte
6	Mujer	Baja	Mediterránea
7	Hombre	Media-Alta	Sur
8	Hombre	Media-Alta	Centro-Norte
9	Hombre	Media-Alta	Medierránea
10	Hombre	Baja	Sur
11	Hombre	Baja	Centro-Norte
12	Hombre	Baja	Mediterránea

Tabla 1: Perfiles considerados en el estudio de la siniestralidad.

Estudiaremos el riesgo de ocurrencia de un siniestro en cada uno de los perfiles y en cada una de las edades. Dicho riesgo lo definiremos como el cociente entre el número de siniestros ocurridos en un perfil determinado a una edad determinada y el número total de asegurados que tienen dicho perfil y dicha edad. Las edades van de los 18 a los 88 años; no obstante, al disponer de muy pocos asegurados menores de 25 y mayores de 71 años, hemos decidido estudiar de forma agrupada cada uno de estos grupos de individuos (estas edades se han denotado con los valores  $25^-$  y  $71^+$ ). Esto tiene como desventaja que el comportamiento de los asegurados menores de veinticinco y mayores de setenta y uno no podrán estudiarse minuciosamente, sino que únicamente tendremos el comportamiento global en estos tramos.

## 2.2. Análisis funcional en componentes principales

En los últimos años se han desarrollado técnicas que consideran los datos desde una perspectiva continua. Entre estas técnicas se encuentra el análisis funcional en componentes principales (AFCP), que aparece como una generalización natural del análisis multivariante en componentes principales cuando se trabaja, no con  $p$  variables aleatorias observadas en  $n$  individuos, sino con un conjunto infinito numerable de variables, como ocurre con las series temporales, o incluso cuando hay una cantidad infinita no numerable de variables, como en los procesos estocásticos en tiempo continuo. Una visión general de algunos de los posibles análisis con datos funcionales y, de forma más precisa, acerca del AFCP la podemos encontrar en Rice y Silverman (1991), Silverman (1996), Ramsay y Silverman (1997), entre otros. Cinco años más tarde, los mismos autores publicaron un libro presentando diversas aplicaciones de algunas de las técnicas que habían desarrollado (Ramsay y Silverman, 2002). Sobre los métodos de aproximación del AFCP y su aplicación a modelos predictivos, muchas de las aportaciones se encuentran recogidas en Valderrama *et al.* (2000). En esta

misma línea, podemos citar los trabajos de Aguilera *et al.* (1996a, 1996b, 1996c, 1997a, 1997b, 1999a, 1999b, 1999c) y Escabias *et al.* (2003).

En nuestro estudio utilizaremos los resultados obtenidos al aplicar el AFCP a unas funciones estimadas previamente (Segovia, 2005). Dichas funciones son la función riesgo de ocurrencia de un siniestro para cada uno de los perfiles, que depende de la edad del conductor y varía en el intervalo  $[25^-, 71^+]$ . De forma análoga al caso finito, el objetivo del AFCP es reducir la dimensionalidad del problema. Es decir, pretendemos obtener unas variables ficticias (llamadas componentes principales) que resuman de la forma más precisa posible esta información. Se sabe que la primera componente principal será la que acumule la mayor variabilidad del proceso original; seguidamente, estará la segunda componente principal, la tercera y así sucesivamente. Haciendo uso de la función de correlación entre el proceso y cada una de las componentes obtenidas (Valderrama *et al.*, 2000) se obtienen los tramos de edades con los que están correlacionadas cada una de las componentes principales (Segovia, 2005) y serán los que utilizaremos posteriormente.

### 3. Resultados del análisis funcional en componentes principales

Al realizar el AFCP (Segovia, 2005), se obtuvo que la primera componente principal nos explica el 70,19%; la segunda, el 12,86%; la tercera, el 8,94% y la cuarta, el 4,08%. Luego, con las tres y cuatro primeras componentes explicamos el 91,99% y el 96,07%, respectivamente, de la variabilidad total del proceso, porcentajes elevados en ambos casos. Haciendo uso de la función de correlación se puede observar que la primera componente está muy correlacionada positivamente con el tramo que va de los 31 a los 67 años, ambos inclusive, y con los mayores de 71 años. La información que obtenemos para los mayores de setenta y un año no es muy fiable, pues en nuestro estudio los tuvimos que agrupar. Por tanto, no tendremos en cuenta la significación estadística detectada para la primera componente en dicho tramo de edad. La segunda y tercera componente principal, están muy correlacionadas positivamente con los tramos de edad que van de los 68 a los 70 años y de los 27 a los 30 años, respectivamente. Esto quiere decir que la primera, segunda y tercera componente obtenidas nos explican el comportamiento de los individuos en los tramos de edad  $[31, 67]$ ,  $[68, 70]$  y  $[27, 30]$ , respectivamente. No existe una fuerte correlación de la cuarta componente con ningún tramo de edad. Por ello, a partir de ahora no vamos a tener en cuenta esta componente en los estudios posteriores. Además, si utilizáramos el criterio de contraste de caída o el criterio de porcentaje de la varianza para elegir el número óptimo de componentes, tomaríamos las tres primeras componentes. Con ellas tendremos explicado el 91,99% de la variabilidad del proceso, siendo dicho porcentaje aceptable en estudios similares (Hair *et al.*, 2000).

## 4. Obtención de primas a través del análisis funcional en componentes principales

De manera general, para asignar el importe total de la prima, las compañías aseguradoras consideran una serie de componentes: la prima pura, el recargo de seguridad, los recargos para gastos, los recargos para beneficio o excedente y otros recargos.

Las primas de riesgo puras<sup>2</sup> recibidas, en conjunto, por el asegurador de los tomadores de un determinado seguro le sirven para atender los posibles siniestros o las prestaciones convenidas. La prima pura resulta de aplicar las estadísticas de acaecimiento del riesgo y la intensidad media del mismo.

El recargo de seguridad se destina a cubrir las desviaciones aleatorias desfavorables de la siniestralidad esperada, debiéndose constituir la correspondiente provisión para desviación de la siniestralidad cuando su importe no es consumido en el pago de los siniestros ocurridos.

Los recargos para gastos, que a partir de la entrada en vigor del nuevo Plan Contable para las entidades aseguradoras aprobado por el Real Decreto 2014/1997 de 26 de diciembre, se reclasifican en gastos por destino y se distinguen los gastos de gestión externa e interna. En los gastos de gestión externa denominados gastos de adquisición se incluirán los gastos de mantenimiento del negocio, justificados en función de la organización administrativa y comercial, actual y prevista (comisiones de la red comercial, gastos de cobro y mantenimiento de la cartera, etc.). En cuanto a los gastos de gestión interna son necesarios para administrar el funcionamiento de una entidad aseguradora. Se componen por: los gastos de administración (nóminas, alquileres, suministros, etc.) y otros gastos técnicos que no están vinculados directamente con el negocio, pero son necesarios para el mantenimiento de la empresa (vigilancia, salario del director general, etc.).

Los recargos para beneficio o excedente se destinan a remunerar los recursos financieros e incrementar la solvencia dinámica de la empresa.

Otros recargos atribuibles son impuestos, el recargo al Consorcio de Compensación de Seguros o el recargo a favor de la Comisión Liquidadora de Entidades Aseguradoras.

El resultado de la suma de los anteriores conceptos es la prima total, que es la que realmente satisface el tomador al hacer efectivo el recibo que le es pasado para su cobro por el asegurador. En definitiva, podemos observar que el coste de la prima total depende de la prima de riesgo que se establezca y de otros gastos atribuibles de forma más específica a la política llevada a cabo por la entidad aseguradora en cuestión.

---

<sup>2</sup>De acuerdo al Reglamento de Ordenación y Supervisión de los Seguros Privados en su artículo 77.1.f (BOE, 1998) los gastos de gestión de los siniestros que hasta ahora formaban parte de los gastos de administración pasan a integrarse en la prima de riesgo. Por tanto, la prima de riesgo vendrá determinada por: la prima de riesgo pura correspondiente a los riesgos a cubrir y los gastos de gestión de siniestros. Por consiguiente, en todo nuestro estudio, cuando hablamos de prima de riesgo nos estaremos refiriendo a la prima de riesgo pura.



En este artículo nos centraremos en las primas de riesgo. Con una óptima política de gastos por parte de la compañía, su ventaja competitiva la debe obtener del ajuste de dicha prima; de ahí la importancia que al cálculo de ésta le otorgan las entidades. Plantearemos una nueva forma de calcular las primas de riesgo puras, realizando un tratamiento distinto en lo que a la variable edad del individuo corresponde. A continuación, mostraremos las penalizaciones o bonificaciones según los siguientes tres aspectos: los distintos rangos de edades, las puntuaciones de cada perfil en cada componente principal y la correlación detectada. Teniendo en cuenta todas estas circunstancias, definiremos una función de pesos que toma valores en cada edad,  $t$ , para cada uno de los perfiles considerados y que denotaremos como  $u_i(t)$ . Además, consideraremos el comportamiento de cada uno de los individuos en los tramos detectados, para lo cual definimos una segunda función de ponderación que denotaremos como  $v_i(t)$ . De esta forma, la ponderación que proponemos en nuestro estudio será el resultado de aplicar las dos funciones que se obtienen con anterioridad. Una vez que hayamos calculado los pesos correspondientes, estaremos en disposición de presentar a las compañías aseguradoras una nueva asignación de primas de riesgo puras.

#### 4.1. Asignación de pesos según los distintos tramos de edad

Proponemos un método de bonificación-penalización que será el resultado de aplicar dos ponderaciones distintas. Con la primera de ellas, conseguimos medir el comportamiento de cada perfil con respecto al resto; es decir, tenemos una valoración vertical de las curvas en cada uno de los intervalos obtenidos con la técnica del AFPC. Por su parte, con la segunda ponderación, controlamos la tendencia del riesgo de ocurrencia de un siniestro para cada uno de los perfiles, obteniéndose así una valoración horizontal para cada individuo. El producto de ambas nos dará la ponderación final considerada.

##### 4.1.1. Primera ponderación

Dado un instante de edad cualquiera,  $t$ , consideramos el conjunto  $A_t$  formado por todos los subíndices  $j \in J$  (siendo  $J$  el conjunto formado por los subíndices de las componentes principales obtenidas) tales que la correlación existente entre el proceso riesgo de ocurrencia de un siniestro a la edad  $t$ ,  $X(t)$ , y la componente principal  $\xi_j$  detectada al aplicar el AFPC, verifiquen ser mayor que 0,7 en valor absoluto; es decir,  $A_t = \{j \in J \mid r(X(t), \xi_j) > 0,7\}$ .

Sea  $j_0$  la solución al problema de optimización:

$$\text{Max}_{j \in A_t} |r(X(t), \xi_j)|; \quad (1)$$

es decir, de entre todas las componentes que verifiquen estar fuertemente correlacionadas con el proceso en un determinado instante de tiempo, nos quedamos con aquella en la que la correlación detectada esté más próxima a uno en valor absoluto. En caso de

existir dos componentes principales cuyo subíndice sea solución de (1), nos quedamos con la componente principal que más información nos aporte, esto es, nos quedaremos con la de menor subíndice. De este modo, la función ponderación para el perfil  $i$ -ésimo en el instante de edad  $t$  vendrá dada como:

$$u_i(t) = \begin{cases} 1 + \frac{\xi_{j_0 i}}{K}, & \text{si } r(X(t), \xi_{j_0}) > 0,7; \\ \frac{\sum_{k=1}^K |\xi_{j_0 k}|}{\xi_{j_0 i}}, & \\ 1 - \frac{\xi_{j_0 i}}{K}, & \text{si } r(X(t), \xi_{j_0}) < -0,7; \\ \frac{\sum_{k=1}^K |\xi_{j_0 k}|}{\xi_{j_0 i}}, & \\ 1, & \text{si } A_t = \emptyset; \end{cases}$$

siendo  $j_0$  la solución al problema anterior (1) y  $\xi_{j_0 i}$  la puntuación del perfil  $P_i$  en la componente principal  $\xi_{j_0}$ .

En nuestro estudio, se verifica que  $A_t$  está compuesto por un único valor del conjunto  $J$  en cualquier instante de tiempo. Por tanto, al resolver el problema de optimización,  $j_0$  tomará ese único valor que forma al conjunto  $A_t$ . Los posibles conjuntos  $A_t$  se muestran a continuación:

$$A_t = \begin{cases} \{1\}, & \forall t \in [31, 67]; \\ \{2\}, & \forall t \in [68, 70]; \\ \{3\}, & \forall t \in [27, 30]; \\ \emptyset, & \text{en otro caso.} \end{cases}$$

Una vez que hemos detectado cuales son los posibles  $j_0$ , estamos en condiciones de calcular la función de peso  $u_i$  para el perfil  $i$ -ésimo en el instante de edad  $t$ , donde  $\xi_{j_0 k}$  es la puntuación que toma el perfil  $k$ -ésimo en la componente  $j_0$ -ésima, con  $K = 12$ .

No obstante, la función anterior está definida de forma general y la podremos utilizar en cualquier otra situación que se nos presente. Obsérvese que dicha función está definida a trozos y que se obtiene dependiendo de una serie de consideraciones:

- Que exista o no una correlación estadísticamente significativa de alguna de las componentes con algunos de los instantes de edades.
- En el caso de existir algunos instantes de edad en los que la correlación sea significativa, que ésta sea directa o inversa.
- En el caso de que exista algún instante de edad en el que el proceso esté correlacionado de forma significativa con varias componentes, tomar aquella componente principal en la que la correlación en valor absoluto esté más cercana a la unidad o en el caso de coincidir aquella que mayor información aporte.

En definitiva, para asignar pesos a los distintos perfiles en las diferentes edades, lo que proponemos es considerar las puntuaciones obtenidas en las componentes principales,

así como los tramos de edad correlacionados de forma significativa con alguna de ellas. Si algún tramo de edad no aparece correlacionado de forma significativa con alguna de las componentes calculadas, se le asignará como peso la unidad.

En la Tabla 2 presentamos los perfiles estudiados y cada uno de los pesos asignados en los diferentes rangos de edades considerados. En el tramo de edad que va de los 27 a los 30 los perfiles que mejor se comportan son el tercero y el segundo, para los que planteamos una bonificación del 20,07% y 17,41%, respectivamente. En cuanto a los perfiles que se comportan peor, destacamos el cuarto y el séptimo, a los que se les aplicará un recargo del 13,89% y 13,41%, respectivamente. En el tramo de los 31 a los 67 años las dos bonificaciones y penalizaciones mayores que proponemos se corresponden: en el primer caso a los perfiles 12 y 11 con un 14,68% y 13,01% de descuento; mientras que el segundo caso afecta al primer y segundo perfil, a los cuales le asignaremos un incremento del 23,06% y 11,27%, respectivamente. Por último, de los 68 a los 70 años, los perfiles que mejor se comportan en el tramo de edad que estudiamos son el primer y tercer perfil, para los que planteamos una bonificación del 14,05% y 13,95%, respectivamente. En cuanto a los perfiles que se comportan peor, destacamos el segundo y el séptimo perfil, a los que se les aplicará un recargo del 19,30% y 12,25%, respectivamente. Se puede observar que existen perfiles a los que tendremos que penalizar en cualesquiera de los rangos de edades detectados. Éste sería el caso de los perfiles siete y ocho, es decir, los hombres con coche de gama media-alta de la zona sur o de la zona centro-norte. Por el contrario, aparecen otros perfiles a los que siempre se les tendrá que bonificar. Estaríamos hablando, en este caso, de las mujeres con coche de gama baja que residen en la zona centro-norte (perfil cinco).

Perfiles	Ponderaciones	Ponderaciones	Ponderaciones
	[27, 30]	[31, 67]	[68, 70]
1	1,0375	1,2306	0,8595
2	0,8259	1,1127	1,1930
3	0,7993	1,0017	0,8605
4	1,1389	1,0668	0,9869
5	0,9818	0,9979	0,8760
6	0,9658	0,8826	1,0819
7	1,1341	1,0850	1,1225
8	1,0305	1,0031	1,0834
9	0,9272	0,9522	1,019
10	1,0569	0,9441	0,9988
11	1,0565	0,8699	0,9506
12	1,0456	0,8532	0,9677

Tabla 2: Pesos asignados utilizando las tres primeras componentes principales.

Lo que hemos conseguido hasta ahora es determinar una serie de tramos de edad en los que habría que asignar primas distintas a los distintos perfiles y, por tanto, comparar el comportamiento de cada uno de dichos perfiles con el resto en el tramo de edad que estemos considerando. Con ello, hemos dado un primer sistema de bonificación o penalización dentro de cada uno de los intervalos obtenidos.

No obstante, no queremos decir que todas las curvas en cada uno de los tramos detectados tengan la misma forma, de hecho cada perfil tiene un comportamiento diferente dentro del tramo detectado. En cualquier caso, hemos podido dar una medida del comportamiento de unos perfiles con respecto al de los otros en un tramo de edad concreto. Por ello, proponemos seguidamente un segundo sistema de ponderación.

#### 4.1.2. Segunda ponderación

La ponderación que planteamos a continuación considera una partición en los tramos detectados en el AFCP, de tal forma que se tenga en cuenta el patrón que sigue cada una de las curvas y podamos reflejar la tendencia de cada una de ellas, con lo que tendríamos en cuenta la memoria histórica del individuo. Lo haremos únicamente para el intervalo de edad [31, 67], debido a que los otros dos intervalos tienen muy poca amplitud.

En primer lugar, tomamos una partición de dicho intervalo de cinco en cinco años, salvo el último que será de siete años. De esta forma consideraremos los tramos [31, 35], [36, 40], [41, 45], [46, 50], [51, 55], [56, 60] y [61, 67].

En segundo lugar, calculamos el máximo del riesgo de ocurrencia de un siniestro en cada uno de los rangos de edades especificados anteriormente, así como, la media de todos los máximos.

En tercer lugar, tomamos como ponderación para el intervalo  $[t_i, t_j]$  el cociente entre el máximo detectado en dicho intervalo y la media de todos los máximos. De esta forma, conseguimos que los porcentajes de bonificación o penalización detectados en el intervalo [31, 67] para cada uno de los perfiles se mantengan y a la vez controlamos para cada uno de ellos la tendencia en cuanto al riesgo de ocurrencia de un siniestro en las distintas edades. Se denotará por  $v_i(t)$  al valor de dicha ponderación en el instante  $t$ .

A modo de ejemplo, si consideramos dentro del perfil 1, el tramo de edad [31, 35], el máximo del riesgo de ocurrencia de un siniestro se alcanza en la edad 31,945, con valor 0,217986; por otro lado, la media de los máximos toma el valor 0,26989214. En consecuencia, el valor de  $v_1(t)$  con  $t \in [31, 35]$  será el cociente  $\frac{0,217986}{0,26989214}$ , cuyo valor es 0,80767821. Luego, en este caso, al perfil uno en el tramo de edad [31, 35] le aplicamos una bonificación del 19,23%. En la Tabla 3 exponemos los pesos obtenidos.

#### 4.1.3. Sistema de ponderación propuesto

El sistema que proponemos para el cálculo de los pesos definitivos que asignaremos a cada perfil en cada instante de edad es el resultado de tener en cuenta las dos propuestas

Edades	Perfil 1	Perfil 2	Perfil 3	Perfil 4	Perfil 5	Perfil 6
[31, 35]	0,8077	0,7624	0,7512	1,0275	0,9174	0,9267
[36, 40]	0,9213	0,9484	1,0499	0,8619	0,7572	0,9764
[41, 45]	1,0719	0,9889	1,1433	0,9819	1,0977	0,9124
[46, 50]	1,0732	1,1087	1,0931	1,0242	1,1849	1,1041
[51, 55]	0,9462	1,1599	0,9543	1,0431	1,0859	1,1716
[56, 60]	1,0955	0,9669	0,9482	1,0576	0,9883	1,0888
[61, 67]	1,0841	1,0647	1,0599	1,0037	0,9686	0,8201

Edades	Perfil 7	Perfil 8	Perfil 9	Perfil 10	Perfil 11	Perfil 12
[31, 35]	0,9599	0,9165	0,9893	1,0555	1,1353	1,1787
[36, 40]	0,9546	1,0027	1,1241	0,9577	1,0037	0,9761
[41, 45]	0,9251	1,0053	1,1547	0,9856	0,9724	1,0163
[46, 50]	1,0263	1,0057	0,9353	1,0542	1,0277	0,9851
[51, 55]	1,0905	1,1162	0,9995	1,0471	1,0142	1,0133
[56, 60]	1,0771	1,0645	0,9367	0,9858	0,9354	1,0058
[61, 67]	0,9665	0,8890	0,8604	0,9141	0,9114	0,8248

Tabla 3: Valores de la función  $v_i$  dados por intervalos.

anteriores (Tabla 4). De tal forma que, para el perfil  $i$ -ésimo el sistema de ponderación propuesto será  $w_i(t)$ , donde  $w_i(t) = u_i(t) \cdot v_i(t)$  es el valor de la función ponderación para el perfil  $i$ -ésimo en el instante de edad  $t$ , siendo  $u_i(t)$  y  $v_i(t)$  las funciones de pesos propuestas anteriormente.

En la Tabla 4 podemos observar que entre los asegurados de 27 a 30 años, los individuos penalizados son las mujeres y los hombres de la zona sur, los hombres de la zona centro-norte y los de la zona mediterránea con coche de gama baja. En este tramo de edad solo existen penalizaciones en las mujeres del sur, mientras que las mujeres del resto de zonas tienen bonificaciones; sin embargo los hombres tienen todas penalizaciones, excepto los de la zona mediterránea con vehículos de gama media-alta, siendo las mayores penalizaciones para los hombres del sur.

En cuanto al tramo de edad que va de los 31 a los 35 años, se penalizará a las mujeres del sur con coche de gama baja. Con respecto a los hombres, en ese mismo rango de edad habrá que penalizar a los que tienen un coche de gama media-alta de la zona sur y a los del mediterráneo con coche de gama baja.

En el tramo de edad que va de los 36 a los 40 años, tenemos que, tanto en las mujeres como en los hombres, se penalizarán a aquellos individuos que tengan un coche de gama media-alta, independientemente de la zona de residencia.

Dentro del tramo de edad que comprende de los 41 a los 45 años, se penaliza a todas las mujeres, salvo a las de la zona mediterránea con coche de gama baja. En cuanto a los

Edades	Perfil 1	Perfil 2	Perfil 3	Perfil 4	Perfil 5	Perfil 6
27 <sup>-</sup>	1	1	1	1	1	1
[27, 30]	1,0375	0,8259	0,7993	1,1389	0,9818	0,9658
[31, 35]	0,9939	0,8484	0,7525	1,0961	0,9155	0,8179
[36, 40]	1,1338	1,0553	1,0518	0,9196	0,7557	0,8618
[41, 45]	1,3192	1,1005	1,1453	1,0476	1,0955	0,8053
[46, 50]	1,3207	1,2337	1,0949	1,0927	1,1825	0,9745
[51, 55]	1,1644	1,2907	0,9560	1,1128	1,0837	1,0340
[56, 60]	1,3481	1,0759	0,9499	1,1283	0,9863	0,9609
[61, 67]	1,3340	1,1847	1,0618	1,0708	0,9667	0,7239
[68, 70]	0,8595	1,1930	0,8605	0,9869	0,8760	1,0819
71 <sup>+</sup>	1	1	1	1	1	1

Edades	Perfil 7	Perfil 8	Perfil 9	Perfil 10	Perfil 11	Perfil 12
27 <sup>-</sup>	1	1	1	1	1	1
[27, 30]	1,1341	1,0305	0,9272	1,0569	1,0565	1,0456
[31, 35]	1,0416	0,9194	0,9421	0,9965	0,9876	1,0056
[36, 40]	1,0357	1,0057	1,0704	0,9041	0,8731	0,8327
[41, 45]	1,0038	1,0084	1,0995	0,9305	0,8459	0,8671
[46, 50]	1,1136	1,0088	0,8906	0,9953	0,8940	0,8404
[51, 55]	1,1833	1,1197	0,9517	0,9886	0,8823	0,8645
[56, 60]	1,1686	1,0678	0,8919	0,9306	0,8137	0,8581
[61, 67]	1,0487	0,8917	0,8193	0,8629	0,7929	0,7037
[68, 70]	1,1225	1,0834	1,0192	0,9988	0,9506	0,9677
71 <sup>+</sup>	1	1	1	1	1	1

Tabla 4: Ponderaciones globales por rangos de edades.

hombres se penaliza a los que tienen un coche de gama media-alta.

Pasamos ahora a estudiar el tramo de edad que va de los 46 a los 50 años. En él se penaliza a todas las mujeres, excepto a las de la zona mediterránea con coche de gama baja. Con respecto a los hombres, penalizamos a los que tienen un coche de gama media-alta, excepto a los de la zona mediterránea.

Entre los 51 y los 55 años, son penalizadas todas las mujeres salvo las del mediterráneo con coche de gama media-alta. En cuanto a los hombres penalizamos a los que tienen un coche de gama media-alta de la zona sur y centro-norte.

En el tramo de edad que va de los 56 a los 60 años, se penaliza a las mujeres y hombres con coche de gama media-alta excepto de la zona mediterránea y a las mujeres del sur con coche de gama baja.

Dentro del tramo de edad que va de los 61 a los 67 años, los únicos penalizados son las mujeres con coche de gama media-alta y las del sur con coche de gama baja. En este

tramo de edad se bonificará a todos los hombres salvo a los del sur con coche de gama media-alta.

Consideramos ahora el último tramo de edad que va de los 68 a los 70 años. Se penaliza a las mujeres de la zona centro-norte con coche de gama media-alta, a las de la zona mediterránea con coche de gama baja y a los hombres con coche de gama media-alta.

De manera general, podemos destacar que todos los hombres mayores de 35 años con vehículos de gama baja tendrán bonificaciones. No observamos un comportamiento similar, en el caso de las mujeres, salvo para el tramo de edad entre 31 y 35 años en el que han de aplicarse bonificaciones en todos los casos salvo en las residentes en el sur con vehículo de gama baja. También podemos observar que los conductores, hombres y mujeres, con vehículo de gama media-alta van a tener mayores penalizaciones, siendo esto más acusado en la zona sur.

No obstante, debemos puntualizar que, al asignar estos pesos, no hemos considerado los costes de ellos. A la compañía le interesará conocer no solamente cuándo existe una mayor probabilidad de que ocurra un siniestro; sino, además, la cuantía de dicho siniestro. Una vez que tenemos los pesos para cada uno de los individuos en los distintos tramos de edad, podremos asignar de forma más detallada las primas de riesgo para cada una de las observaciones, según se encuentren en un tramo de edad o en otro.

## 4.2. Determinación de tarifas en función del riesgo del conductor

Clasificaremos, en la presente sección, cada uno de los conductores situándolos en un grupo determinado, de forma que le podamos asignar la tarifa correspondiente.

### 4.2.1. Primas de riesgo propuestas por este estudio en los distintos tramos de edades

Las compañías aseguradoras calculan las primas de riesgo, correspondientes a cada uno de los perfiles que consideran, como el producto del riesgo de ocurrencia de un siniestro para un determinado perfil y el coste medio de los siniestros provocados por los individuos que se encuentran en dicho perfil. Una vez conocidas dichas primas, podremos calcular la recaudación total de la compañía y por perfiles sin más que aplicar las correspondientes tarifas a cada uno de los asegurados.

La recaudación total de la prima de riesgo propuesta por la entidad se puede calcular como:

$$n_1P_{r_1} + n_2P_{r_2} + \dots + n_{12}P_{r_{12}},$$

siendo  $n_i$  el número de asegurados pertenecientes al perfil  $i$ -ésimo y  $P_{r_i}$  es la prima de riesgo asignada a dicho perfil.

En esta sección presentaremos un sistema de recargo o bonificación de las tarifas para cada perfil y cada tramo de edad. Para ello, utilizaremos las primas que propone la compañía estudiada en cada uno de los perfiles, ponderadas éstas por los pesos obtenidos

anteriormente. Por tanto, la prima de riesgo propuesta para el perfil  $i$ -ésimo en la edad  $t$ , que denotaremos por  $P_{r_i}(t)$ , vendrá dada por:

$$P_{r_i}(t) = P_{r_i} w_i(t),$$

donde  $P_{r_i}$  es la prima de riesgo asignada por la compañía para el perfil  $i$ -ésimo y  $w_i(t) = u_i(t) \cdot v_i(t)$  es el valor de la función ponderación para el perfil  $i$ -ésimo en el instante de edad  $t$ , siendo  $u_i(t)$  y  $v_i(t)$  las funciones de pesos propuestas por la primera y segunda ponderación, expuestas ambas anteriormente (Tabla 4).

Queremos dar a las compañías unas indicaciones acerca de la prima de riesgo más adecuada para los clientes que se incorporan a su cartera, de manera que a un posible cliente, dependiendo de la edad que tenga y del perfil en el que se encuentre, se le asigne un valor determinado de prima de riesgo, procurando de esta forma conseguir una mejor selección del riesgo en la compañía. Estamos dando, por tanto, unas pautas para la asignación de las primas base de los nuevos clientes que se van a incorporar a su cartera.

Partiendo del supuesto de que la compañía pretende quedarse en la misma situación en la que está, proponemos las primas de riesgo más adecuadas para estabilizar los ingresos. Por tanto, impondremos que la recaudación total obtenida por la entidad utilizando el sistema original de asignación de primas y el que nosotros proponemos se mantenga con el nuevo método propuesto. Esto podemos conseguirlo de dos formas distintas: simplemente mantener la recaudación final o conseguir que la entidad perciba la misma cantidad en cada uno de los perfiles (con lo que la recaudación global seguiría siendo la misma).

#### 4.2.2. Primas de riesgo asignadas manteniendo la recaudación global

Utilizaremos el método propuesto anteriormente, pero añadiremos una nueva restricción que será que la compañía siga recaudando la misma cantidad.

Al perfil  $i$ -ésimo en el instante  $t$  de edad le asignamos una tarifa que se calcula como:

$$TRG_i(t) = \frac{P_{r_i}(t)}{\int_{t_0}^{t_1} \sum_{k=1}^K n_k(t) P_{r_k}(t) dt} \cdot \sum_{k=1}^K n_k P_{r_k},$$

donde  $P_{r_k}(t)$  y  $n_k(t)$  son, respectivamente, la prima de riesgo propuesta por nosotros y el número de asegurados del perfil  $k$ -ésimo en el instante de edad  $t \in [t_0, t_1]$  y  $n_k$  y  $P_{r_k}$  son, respectivamente, el número de asegurados y la prima de riesgo asignada por la entidad para el perfil  $k$ -ésimo.

#### 4.2.3. Primas de riesgo asignadas manteniendo la recaudación por perfiles

En este apartado pretendemos al igual que en el anterior mantener el beneficio global de la aseguradora. Pero vamos a imponer una restricción más, consiste en pedir que la recaudación se mantenga también por perfiles. En tal caso, la prima propuesta para el



perfil  $i$ -ésimo en el instante de edad  $t$  viene dado por:

$$TRP_i(t) = \frac{P_{r_i}(t)}{\int_{t_0}^{t_1} n_i(t)P_{r_i}(t)dt} \cdot n_i P_{r_i},$$

donde  $n_i(t)$  y  $P_{r_i}(t)$  son, respectivamente, el número de asegurados y la prima de riesgo del perfil  $i$ -ésimo propuesta por nosotros en el instante de edad  $t \in [t_0, t_1]$  y  $n_i$  y  $P_{r_i}$  son, respectivamente, el número de asegurados y la prima de riesgo asignada por la entidad para el perfil  $i$ -ésimo.

## 5. Conclusiones

Nuestro estudio nos permite proponer a las compañías aseguradoras una nueva manera de establecer las primas de riesgo para la nueva adquisición de clientes. La tarificación la hemos desglosado mucho más, de tal modo que no a todos los individuos de un mismo perfil se les aplicará la misma prima. Ahora, además del perfil en el que nos encontremos, la prima dependerá también del instante de edad en el que se encuentre el asegurado. Con esto se consigue que a los individuos que proponemos asignarles una prima de riesgo menor, le podremos ofertar un producto más competitivo. Esto se podría plasmar en una futura entrada de clientes de este tipo en la cartera, mientras que podrían salir de ella aquellos clientes a los que les hemos incrementado su prima. Se podría plantear, como una futura línea de investigación, el estudio de la cartera óptima para la compañía, según las características de los clientes.

Aplicando el sistema de bonificación-penalización que se detalló en la sección 4 a los datos reales de una compañía de seguros, podemos decir que de forma global la gama del coche es una de las características que más influyen a la hora de penalizar o bonificar a los asegurados. Independientemente del sexo y la edad, podemos observar que los individuos con vehículos de gama media-alta son penalizados un mayor número de veces que los de gama baja. No obstante, a las mujeres de la zona sur con coche de gama baja se les penaliza con una frecuencia similar a aquellas que poseen coche de gama media-alta, llegando a estar incluso por encima de las mujeres de la zona mediterránea. Es de destacar que los individuos de la zona mediterránea, a pesar de tener asegurado un coche de gama media-alta, suelen tener bonificaciones en sus primas en casi todos los rangos de edad.

Si tenemos en cuenta el sexo, la edad del asegurado y el tipo de coche podemos destacar que los hombres de más de 35 años de edad y conducen un coche de gama baja tendrán bonificación.

Teniendo en cuenta el sexo y la edad del asegurado se tiene que las penalizaciones en las mujeres se centran en el rango de edad intermedio, de los 41 a los 55 años, mientras que en los hombres el rango de edad más penalizado es el que va de los 27 a los 30 años. Por el contrario, las mujeres con mayores bonificaciones son aquellas cuyas edades están

comprendida entre los 27 y 35 años y entre los 68 y 70 años, y en los hombres en el tramo de edad de los 46 a los 67 años.

En el futuro, pretendemos insistir en la obtención del archivo histórico SINCO, que contiene la información de la mayoría de las entidades aseguradoras, a la que le aplicaremos el AFCP, de esta forma nos podríamos acercar mucho mejor a lo que realmente está ocurriendo en el sector del automóvil.

## Bibliografía

AGUILERA, A.M.; GUTIÉRREZ, R.; VALDERRAMA, M.J. (1996a): Approximations of estimators in the PCA of a stochastic process using B-splines. *Communational Statistics (simulation)* **25** (3), pp. 671 – 690.

AGUILERA, A.M.; OCAÑA, F.A.; VALDERRAMA, M.J. (1996b): Análisis en componentes principales de un proceso estocástico con funciones muestrales escalonadas. *Qüestió* **20** (1), pp. 7 – 28.

AGUILERA, A.M.; OCAÑA, F.A.; VALDERRAMA, M.J. (1996c): On a weighted principal component model to forecast a continuous time series. *Proceed. Computational Statistics: COMPSTAT'96*. (Prat A.,ed.), Physica-Verlag, pp. 169 – 174.

AGUILERA, A.M.; OCAÑA, F.A.; VALDERRAMA, M.J. (1997a): An approximated principal component prediction model for continuous-time stochastic processes. *Applied Stochastic Models and Data Analysis* **13** (1), pp. 61 – 72.

AGUILERA, A.M.; OCAÑA, F.A.; VALDERRAMA, M.J. (1997b): Regresión sobre componentes principales de un proceso estocástico con funciones muestrales escalonadas. *Estadística Española* **39** (142), pp. 5 – 21.

AGUILERA, A.M.; OCAÑA, F.A.; VALDERRAMA, M.J. (1999a): Forecasting time series by functional PCA. Discusión of several weighted approaches. *Computational Statistics* **1** (3), pp. 443 – 467.

AGUILERA, A.M.; OCAÑA, F.A.; VALDERRAMA, M.J. (1999b): Forecasting with unequally spaced data by a functional principal component approach. *Test* **8** (1), pp. 233 – 254.

AGUILERA, A.M.; OCAÑA, F.A.; VALDERRAMA, M.J. (1999c): Stochastic modelling for evolution of stock-prices by means of functional principal component analysis. *Applied Stochastic Models in Bussines and Industry* **15** (4), pp. 227 – 234.

ARTIS, M.; AYUSO, M.; GUILLÉN, M. (1999): Modeling different types of automobile insurance fraud behaviour in Spanish market. *Mathematics and Economics* **824**, 1 – 2, pp. 67 – 81.

ARTIS, M.; AYUSO, M.; GUILLÉN, M. (2002): Detection of automobile insurance fraud with discrete choice models and misclassified claims. *Journal of Risk and Insurance* **69**, 3, pp. 325 – 340.

AYUSO, M.; GUILLÉN, M. (1999): Modelos de detección de fraude en el seguro del automóvil. *Cuadernos actuariales* **8**, pp. 135 – 149.

- AYUSO, M.; GUILLÉN, M.; ARTIS, M. (1999): Técnicas cuantitativas para la detección del fraude en el seguro del automóvil. *Anales del instituto de actuarios españoles* **5**, pp. 51 – 83.
- BOJ, E.; CLARAMUNT, M.M.; FORTIANA, J. (2004): Análisis multivariante aplicado a la selección de factores de riesgo en la tarificación. *Cuadernos de la Fundación Mapfre estudios*. Instituto de Ciencias del Seguro. Madrid.
- BOJ, E.; CLARAMUNT, M.; FORTIANA, J.; VEGAS, A. (2005): Bases de datos y estadísticas del seguro de automóviles en España: influencia en el cálculo de primas. *Estadística Española* **47**, 160, pp. 539 – 566.
- BOE (1998): Reglamento de Ordenación y Supervisión de los Seguros Privados, aprobado por el Real Decreto 2486/1998, de 20 de noviembre. Publicado en el BOE 282/1998 de 25 de noviembre de 1998.
- ESCABIAS, M., AGUILERA, A.M.; VALDERRAMA, M.J. (2003): Una solución al problema de multicolinealidad en regresión logística funcional. *XXVII Congreso Nacional de Estadística e Investigación Operativa. Lleida 8 – 11 Abril de 2003*.
- GREEN, P.J.; SILVERMAN B.W. (1994): *Nonparametric regression and generalized linear models. A roughness penalty approach*. Chapman and Hall/CRC.
- GUILLÉN, M.; AYUSO, M.; BERMÚDEZ, L.; MORILLO, I. (2005): El Seguro de automóviles: estado actual y perspectiva de la técnica actuarial. *Fundación Mapfre estudios*. Instituto de Ciencias del Seguro. Madrid.
- HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L. ; BLACK, W. C. (2000): *Análisis multivariante*. Prentice Hall.
- MELGAR, M.C.; GUERRERO, F.M. (2005): Los siniestros en el seguro del automóvil: un análisis econométrico aplicado. *Estudios de Economía Aplicada (Asepelt)* pp. 355-375.
- RAMSAY, J.O.; SILVERMAN, B.W. (1997): *Functional data analysis*. Springer Series in Statistics.
- RAMSAY, J.O.; SILVERMAN, B.W. (2002): *Applied functional data analysis*. Springer Series in Statistics.
- RICE, J.A.; SILVERMAN, B.W. (1991): Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society Series B* **53**, pp. 233 – 243.
- SEGOVIA, M.M. (2005): *Análisis en componentes principales funcionales: Aplicaciones empresariales*. Tesis Doctoral. Universidad Pablo de Olavide.
- SILVERMAN, B.W. (1996): Smoothed functional principal components analysis by choice of norm. *Annals of Statistics* **24** (1), pp. 1 – 24.
- VALDERRAMA, M.J.; AGUILERA, A.M.; OCAÑA, F.A. (2000): *Predicción dinámica mediante análisis de datos funcionales*. La Muralla-Hespérides.