

Consideraciones Finales

CAPÍTULO 7 CONSIDERACIONES FINALES

Conclusiones: *Retrospectivas y prospectivas.*

Los resultados del presente trabajo se han estructurado en dos partes. La primera de ellas contiene una revisión de los fundamentos y herramientas metodológicas para el análisis de datos económico-financiero, con una mayor relevancia de los *modelos neuronales* como herramientas de extracción de información. La segunda proporciona el estudio empírico realizado sobre el sector asegurador y de previsión social español, poniendo mayor énfasis en la modelización del comportamiento de la siniestralidad *no vida* del sector de Mutualidades de Previsión Social de Cataluña, incorporando un parte de las herramientas metodológicas propuestas desde las perspectivas: *probabilística, econométrica y neuronal*.

Las conclusiones obtenidas de la **revisión metodológica**, *capítulos 1 a 3*, son las siguientes:

1. La utilización del modelo de ratio, en nuestro caso de *siniestralidad*, está condicionada por un conjunto de puntos débiles de carácter metodológico (Lev, B. (1978); Sunder, S. (1979); Salmi, T. y Martikainen, T. (1994); García-Ayuso Covarsí (1994)), que pueden ser relevantes en el uso habitual de los ratios económicos-financieros por parte de los analistas e investigadores. En este ámbito la verificación de la hipótesis de **estricta proporcionalidad** entre numerador y denominador es un aspecto importante, ya que su cumplimiento permite asegurar la *neutralización* del tamaño de las compañías a través de la utilización del modelo de ratio. Una de las posibles causas de su rechazo es el desconocimiento “a priori” de la forma funcional que relaciona el numerador del ratio con su denominador, o la posible presencia de una **naturaleza no lineal** en la misma. Todo lo anterior recomienda rigor y cuidado en la utilización indiscriminada de los ratios en modelos de clasificación o de predicción.

2. Los esfuerzos por desentrañar el comportamiento del propio ratio en sí mismo, así como el de sus componentes (numerador y denominador), (Whittington (1980); McDonald y Morris (1984; 1985)), ha generado diferentes líneas de trabajo. La primera de ellas consiste en utilizar la **modelización causal** como herramienta de análisis, (Berry, R.H. y Nix, S. (1991)) y la segunda, desde una óptica **probabilística**, consiste en obtener modelos de probabilidad “flexibles” o combinación de ellos que expliquen su comportamiento, (McLeay, S. (1986; 1997; 2000); Barnes, P. (1982;1983;1987), So, J. (1987^a;1994), Lau, H-S.; Lau, A. H-L.; Gribbin, D.W. (1995)). La aportación del presente trabajo consiste en la propuesta de *especificaciones* alternativas para el modelo ratio (véase apartados 5.4, 6.2. y 6.3.), así como, la comprobación de la versatilidad de los modelos de probabilidad habituales para capturar el comportamiento del ratio de siniestralidad, (véase apartado 5.3.3.).
3. Los **nuevos avances** en el diseño del modelo de ratio pueden ser agrupados en tres líneas de actuación. Primeramente, la utilización de procesos estocásticos de carácter continuo para describir la generación de ratios económico-financieros (Tippett M. (1990); Tippett, M. y Whittington, G. (1995); Rhys, H. y Tippett, M. (1993)). En segundo lugar, la aplicación de técnicas econométricas de Co-integración sobre los componentes del propio ratio, buscando comportamientos a largo plazo entre ellos (Whittington, G.; Tippett, M. (1999)). En último lugar están los modelos neuronales (MN) (Rojas, R. (1996); Bishop, C.M. (1995)), en línea con la modelización no paramétrica (Wang, S. (1996)) y las técnicas de Data Mining aplicadas a datos contables y financieros (Trigueiros, D. (1991,1994, 1996 y 1997)), los cuales han sido utilizados como herramientas en el presente trabajo, (véase capítulos 2, 3 y 6).
4. El incremento de la **metodología neuronal** aplicada a la modelización es cada vez más importante, como consecuencia de una mayor interrelación con otras disciplinas, como por ejemplo, los métodos estadísticos (Sarle, W. S. (1994), Kay, J.W. y Titterington, D.M. (1999)), métodos econométricos (Kuan, C. y White, H. (1994)), teoría de la información (Deco, G. y Obradovic, D. (1996)).

5. La **tipología** de los modelos neuronales se divide en: aprendizaje, arquitectura y aplicaciones (Hagan, Martín T. (1996); Chin-Teng Lin, George Lee, C.S. (1996); Patterson, Dan W. (1996)). Respecto a la forma de aprendizaje existen dos grandes métodos, el supervisado (cercano a la modelización causal) y el no supervisado (muy próximo a las técnicas de agrupación). Si consideramos el tipo de arquitectura, encontramos las siguientes tipologías: los modelos *feed-forward*, *feedback* y *recurrentes*. Y en último lugar, las aplicaciones más frecuentes con carácter general son, *predicción* (Refenes, A.P. (1995)) y *clasificación* (Ripley, B.D. (1993); Raudys, S. (2001); Borra. S.; Rocci, R. Vichi, M. Schader, M. (2000)), con muchos paralelismos con los modelos tradicionales estadísticos. En el presente trabajo se han utilizado los modelos *feed-forward*, (véase apartados 6.2. y 6.4.).

6. Como consecuencia de la gran variedad de modelos neuronales existentes es necesario la elaboración de **taxonomías** que faciliten su clasificación. Resaltamos en la misma los modelos híbridos por su número de aplicaciones y porque combinan el aprendizaje no supervisado con el supervisado, (véase apartado 2.5.). En esta tipología neuronal se encuentran los siguientes modelos: *Radial Basis Function* (RBF) (Wasserman, Philip D. (1993), Martín del Brío, B. Sanz Molina, A. (1997)); modelos *probabilísticos* (PNN) que incorporan en su aprendizaje la estadística Bayesiana (Husmeier, D.(1999); Finn V. Jensen (1996)) y modelos generalizados de *regresión neuronal* (GRNN) (Masters, T. (1995); Wasserman, Philip D. (1993)), que serán utilizados como alternativa no paramétrica a los modelos causales tradicionales, (véase el apartado 6.3.). Una mención especial, por su interés, son los modelos neuronales *polinomiales* (PoNN) que utilizan el *algoritmo* GMDH, cuyo proceso iterativo de selección de estructuras les permiten representar cualquier sistema complejo de relaciones (Sung-Kwun Oh; Witold Pedrycz, (2002)). En el presente trabajo se han utilizado los modelos generalizados de *regresión neuronal* (GRNN), (véase apartado 6.3.).

7. Las recientes aportaciones a la disciplina estadística por parte de Cheng, Biang; Titterington, D.M (1994) y Kay, J.W.; Titterington, D.M. (1999) ; Min Qi (1996) y Hastie, T; Tibshirani, R.; Friedman, J. (2001) y al campo econométrico por parte de Kuan, C.; White, H. (1994) y Zaprani y Refenes (1999) desde una óptica neuronal, presagian un terreno fértil en la búsqueda de **paralelismos** con los modelos estadísticos, como pueden ser, los modelos de regresión lineal, modelos de probabilidad Logit / Probit, modelos de variables latentes (MIMIC) y la familia de modelos generalizados o relacionados, (véase el procedimiento adaptativo para la regresión mediante *splines* (MARS) y los modelos de regresión *Projection Pursuit* (PPR)) y los propios modelos neuronales. En última instancia en la mayoría de los casos los modelos anteriores pueden ser especificados como casos particulares de un modelo neuronal *Multilayer Perceptron* (MLP), debido sobre todo, a su mayor versatilidad garantizada por el elevado número de parámetros que poseen: número de capas, número de neuronas en cada capa, funciones de activación o de transferencia y función de salida, (Refenes (1995)).
8. La **interpretación** estadística de los modelos neuronales constituye un problema no resuelto. Su desarrollo futuro precisa de la combinación de los siguientes aspectos: *identificación* correcta de los *inputs* e *outputs* más importantes, elección de la *estructura* más adecuada (número necesario de capas ocultas y número de neuronas para cada una de las capas ocultas) y definición de los criterios de *evaluación* de los modelos estimados. La elección de las variables relevantes depende esencialmente del objetivo principal del estudio y supone un conocimiento profundo del mismo. El diseño de la estructura del modelo posee mayor dificultad, debido a las numerosas posibilidades existentes y al compromiso presente en todo momento entre *generalización* y *memorización*, evitando en la medida de las posibilidades la sobreparametrización del modelo. El mecanismo que permite aumentar la robustez de los resultados consiste en evaluar la capacidad de ajuste del modelo neuronal mediante una base de datos distinta del proceso de *aprendizaje*, de forma iterativa y según una *calibración* predefinida. Éste proceso requiere dividir los datos, siempre que se posean un volumen holgado de ellos, en tres submuestras, *aprendizaje*, *test* y *predicción*. La

base de datos de *aprendizaje* se utiliza para ajustar el modelo, la base de *test* para estimar el error de predicción del modelo seleccionado, y por último, la base de *predicción*, para gestionar el proceso de generalización. Una gestión correcta de este error de predicción supone tratar de forma efectiva el compromiso entre tendencia-varianza, sobre la base de una descomposición del error esperado de predicción y combinado con técnicas de *validación cruzada* (Hastie, T; Tibshirani, R.; Friedman, J. (2001). En última instancia, y desde la óptica de la inferencia estadística, existen varios métodos para interpretar la importancia relativa de cada *input* (variable exógena) sobre cada *output* (variable endógena), algunos de los cuales son: *pseudoponderaciones* (PW), el *sumatorio de los pesos* (parámetros) en valor absoluto de los *inputs* (SW), *análisis de sensibilidad* e *índice de sensibilidad* (Min Qi (1996)).

9. Desde una visión más **econométrica** es esencial, a nuestro entender, la reciente aportación de Zapranis A. y Refenes, A.P. (1999) que proponen una metodología completa de identificación de modelos neuronales cuyas fases son las siguientes: *selección* del modelo, *contraste de significación* de las variables y *contraste de relevancia* del modelo en su conjunto. Incorporan un diseño iterativo de selección de modelos neuronales sobre la base del principio mínimo riesgo de predicción (MPR) y métodos de remuestro (*validación cruzada*, *Bootstrap* y *Jackknife*) para la selección de variables relevantes y contraste de hipótesis en el ámbito neuronal. Consideramos que existen muchas posibilidades en el futuro para configurar un tronco metodológico alrededor del concepto de *econometría neuronal*.

Las conclusiones obtenidas de la **parte empírica**, *capítulos 4 a 6*, pueden dividirse en tres grupos. El primero de ellos contiene los resultados agregados obtenidos para el sector asegurador Español (1991-1997), cuyos datos aún siendo de carácter público han necesitado de un tratamiento específico. También se facilita la misma información para el sector de previsión social desglosado por comunidades con competencias transferidas (Comunidad Autónoma de Cataluña (1991-1997), Comunidad Autónoma del País Vasco (1990-1998)) y aquellas que dependen de la D.G.S. En éste caso los datos han sido facilitados “ad-hoc” con carácter confidencial por parte de las Administraciones Públicas implicadas en el presente

trabajo. En segundo lugar se ha centrado el análisis sobre el comportamiento de la *siniestralidad no vida* para el sector de mutualidades de Cataluña (1989-1997) (véase apartado 5.2.), desglosado por prestaciones y desde un doble enfoque, el probabilístico (véase apartado 5.3.) y el econométrico (véase apartado 5.4.). Y finalmente se han realizado una serie de aplicaciones mediante metodología neuronal, (véase capítulo 6.). A continuación detallamos las conclusiones más relevantes de cada uno de ellos.

Los resultados agregados para el sector asegurador Español (1991-1997), véase el *capítulo 4*, sin considerar las entidades de previsión social, son los siguientes,

1. El crecimiento del seguro directo ha superado aproximadamente un 35% el crecimiento global de la economía Española en el período analizado, aunque si consideramos el negocio neto, es decir, libre de reaseguro, el porcentaje se incrementa hasta niveles de 44%. El peso del seguro directo y del negocio neto sobre el PIB a precios de mercado se han incrementado en casi un punto porcentual situándose en 1997 en 4.85% y 4.62% respectivamente.
2. El número de empresas que operan en el mercado se ha reducido en un 29%. Si consideramos el perfil jurídico de las mismas, se ha producido una reducción de los efectivos en un 24% para las sociedades anónimas, un 89% para las delegaciones extranjeras, un 19% para las mutuas, y un 50% para compañías de reaseguro.
3. Los gastos de administración han aumentado un 33% (5.8% anual acumulativo), las remuneraciones brutas un 18% (3.4% anual acumulativo) y el resultado económico después de impuestos (RE) un 185% (23.3% anual acumulativo). Todo ello asociado a un incremento en el volumen de negocio neto del 76% (11.7% factor anual acumulativo).
4. El importe de las primas consumidas por persona activa y ocupada ha supuesto en los últimos siete años un incremento de 84% (10.7% anual acumulativo), situándose en 1997 en 101.917 Ptas por persona activa y 308.457 Ptas por persona

ocupada, (10.3% anual acumulativo). Además se observa que el consumo de este tipo de productos no es del todo dependiente del ciclo económico representado por la variación del número de ocupados.

5. Si tenemos en cuenta los dos principales segmentos, vida y no vida, se ha producido un incremento de casi diez puntos porcentuales en el peso del segmento vida, en detrimento del no vida, donde la proporción en el último año (1997) es de 47.6% vida y 52.4% no vida.

Comunidad Autónoma de Cataluña (1991-1997):

1. El negocio neto ha pasado de suponer en 1989 un 0.153% del PIB a precios de mercado de Cataluña a un 0.151% nueve años más tarde. Existe un crecimiento paralelo entre la economía catalana y el negocio neto, 76% (7.3% anual acumulativo) y 74% (7.1% anual acumulativo), respectivamente.
2. El mercado de mutualidades en Cataluña ha reducido el número de entidades en un 26%.
3. Los gastos de administración (GA) han aumentado un 140% en diez años (11.5% anual acumulativo), en contraposición con el crecimiento del sector en términos de negocio neto que fue de un 7.1% anual acumulativo. Las remuneraciones brutas (RB) han crecido más que el crecimiento de la propia economía catalana, aproximadamente un 20%, (8.5% anual acumulativo). Finalmente respecto al resultado bruto (RE), su variación ha sido negativa, perdiendo en esta década casi treinta y tres puntos porcentuales, con una tendencia a la baja remarcada.
4. El incremento del consumo anual por persona ocupada de productos de previsión ha supuesto para todo el período analizado un 64% (6% anual acumulativo, sobre el negocio neto), siendo para 1997 de 10041 Ptas.

5. Los resultados anteriores medidos en función del negocio neto generado, supone: reducción de más del 50% en el peso del resultado económico, situándose en 1997 en un 5%. En cambio los gastos de administración han incrementado su peso en un 5% (peso del 24% en 1997) y un 1% las remuneraciones brutas, cuya importancia en 1997, fue del 11%.
6. El empleo total de las entidades de mayor dimensión (modelo *completo*) ha descendido en un 8.5% (-1.3% anual acumulativo), acompañado de un descenso del 26.5% de la cartera de socios (1.5% anual acumulativo). En cambio para las pequeñas entidades la reducción de cartera ha sido algo menos pronunciada, un 22%. Respecto a la cobertura poblacional que el sistema de previsión social en Cataluña ofrece, se observa un descenso del 4% sobre la población total (en 1997 un 19.7%), y del 3.6% sobre ocupados totales (en 1997 un 23.79%). Es decir, casi una cuarta parte de los ocupados en Cataluña, posee cobertura en el ámbito mutual.

Comunidad Autónoma del País Vasco (1990-1998):

1. El peso del seguro directo en el País Vasco ha aumentado considerablemente, ya que en 1991 suponía sobre su propio PIB_{pm} un 1.29% y ocho ejercicios más tarde, supone aproximadamente el doble, 2.40%.
2. El negocio neto se ha incrementado un 190% (16.5% anual acumulativo), con un incremento paralelo del número de entidades en un 27%, durante el período 1991-1998.
3. Incremento de tres puntos porcentuales del peso de los gastos de administración, acompañado de un aumento en el peso del resultado económico sobre el negocio neto de casi 43 puntos, debido sobre todo a una mayor orientación hacia el segmento vida. Adicionalmente el consumo anual por persona ocupada de negocio neto fue en 1998 de 171.423 Ptas, un 182% en ocho ejercicios económicos (15.9% anual acumulativo).

Comunidades autónomas con dependencia directa de la Dirección General de Seguros (D.G.S.) (1992-1997):

1. El peso del seguro directo perteneciente al sector de previsión social con dependencia directa de la *Dirección General de Seguros* ha disminuido, ya que en 1992 suponía sobre el PIB_{pm} un 0.138% y seis ejercicios posteriores, un 0.103%. Adicionalmente el negocio neto ha caído en la misma proporción, con una pérdida del 34% de los efectivos del sector.
2. Respecto a la evolución de los gastos de administración (GA), de las remuneraciones brutas (RB) y el resultado económico (RE) después de impuestos, resalta la caída considerable de las tres partidas, respecto a los gastos de administración un 40%, las remuneraciones un 33% y el resultado, un 35% pero con fuertes oscilaciones en los datos.

Una vez descrito los tres subsectores de previsión social por separado, presentamos a continuación los resultados de forma comparada con el sector asegurador español:

1. El peso del negocio neto de Cataluña respecto al global del sector asegurador (sin incluir las entidades de previsión social) es de 0.59% (1997) con un 25% de las entidades. Para el País Vasco es un 2.68% (1997) con un 39% de las entidades y las entidades dependientes de la D.G.S. representan un 1.67% (1997) con un 22% de las entidades. El total del sector representa un 4.94% (1997) del global del sector asegurador. El reparto del sector de previsión social es el siguiente: Cataluña supone un 12% del total, País Vasco, un 54% y las entidades D.G.S. un 34% en términos de Negocio Neto (NN).
2. Las tasas del crecimiento del negocio neto (homogeneizado para el período común 1992-1997 y en términos de *factor acumulativo anual*) son: Cataluña posee un crecimiento de un 5% anual, País Vasco, 23% anual y para el sector dependiente de la D.G.S. crecimiento cero. Recordamos que el sector asegurador creció en el mismo período un 12% anual acumulativo. La evolución del número de compañías

para el período definido, es la siguiente: Cataluña, una pérdida del 7% de los efectivos, en el País Vasco, un incremento del 24% y para el resto del sector, un decremento del 34% de los efectivos. Para el sector asegurador ha supuesto una pérdida del 29%. Y respecto al peso de los gastos de administración sobre negocio neto, en Cataluña supone un 24% (1997) (reducción de cuatro puntos porcentuales en el período 1992-1997), en el País Vasco, un 4% (1997) (incremento de casi un punto porcentual en el mismo período) y para el resto del sector, 15% (1997) (reducción de diez puntos porcentuales). Comparativamente para el sector asegurador supone un 11%.

Las conclusiones obtenidas sobre el comportamiento de la *siniestralidad no vida* para el sector de mutualidades de Cataluña (1989-1997) (véase apartado 5.2.), desglosado por prestaciones (*enfermedad, larga enfermedad, invalidez, intervención y hospitalización quirúrgica, accidentes, sepelio, servicio a automovilistas y asistencia sanitaria*), y desde un enfoque **probabilístico** (véase apartado 5.3.), son las siguientes:

1. El comportamiento de las prestaciones analizadas es muy distinto en términos de siniestralidad. Los ingresos por primas poseen mayor dispersión que los pagos por prestaciones, lo cual sugiere que la propia variable “ingreso” recoge coberturas muy diferentes. En promedio, la variabilidad de los ingresos es un 80% mayor que la de los pagos, aspecto que afectará al propio *ratio* de siniestralidad.
2. La prestación de Intervención Quirúrgica posee el mayor incremento en términos de *número de siniestros*, en cambio si consideramos el *pago de prestaciones*, la prestación de Accidentes es la de mayor cuantía. Respecto al *número de pólizas*, de nuevo la misma prestación, Accidentes, posee el mayor incremento y en último lugar, resalta la prestación de Invalidez con el mayor incremento desde los *ingresos por primas*.

3. Los resultados en promedio para los *ratios básicos* definidos son: *índice de frecuencia* (R1) (por cada 1000 pólizas emitidas al año), la prestación que con diferencia posee la mayor siniestralidad es asistencia sanitaria a socios, en segundo lugar está servicios a automovilistas, que por cada 1000 pólizas cubiertas genera 146 siniestros y de forma sucesiva hasta la prestación que posee el menor índice de siniestralidad, accidentes, con dos accidentes por cada 1000 pólizas. En términos de *coste medio* (R3), la prestación asistencia sanitaria posee el coste medio por siniestro más bajo, 8461 Ptas; la prestación de servicios a automovilistas, 15115 Ptas, y de la misma forma detectamos que la prestación anterior de accidentes con menor índice de siniestralidad aquí posee un coste medio muy alto, 1083059 Ptas. Finalmente, si tenemos en cuenta los ingresos, los comentarios son similares para las tres prestaciones anteriores, así el *ingreso medio por póliza anual* (R4) para la prestación de asistencia sanitaria es de, 20741 Ptas, servicios a automovilistas, 2404 Ptas y para accidentes, 1781 Ptas.
4. Las prestaciones con mayor “dispersión relativa”, medida por el coeficiente de variación, son: Intervención Quirúrgica, Accidentes y Servicios a Automovilistas, debido a una mayor heterogeneidad en sus coberturas. Dos son los procedimientos utilizados para el cálculo de los momentos de la **siniestralidad anual o daño total** (ξ), (véase apartado 5.3.2.3.): el modelo mixto de Poisson con una Gamma para la mixtura, (véase Latorre (1992)) y el método de Monte-Carlo, así éste último método consigue reducir en términos comparativos la dispersión relativa global en un 3%.
5. Los **modelos teóricos de probabilidad** continuos que mejor representan la siniestralidad anual obtenida mediante simulación de Monte-Carlo, a través de contrastes no paramétricos, son para “modelos no truncados”, el modelo Gaussiano que domina sobre los demás (sí bien el modelo logístico se ajusta mejor para carteras pequeñas [0;10]Mill Ptas), y para “modelos truncados”, el modelo Gamma.

6. Las conclusiones obtenidas del **margen mínimo de solvencia** (MMS) necesario sobre la siniestralidad anual, por tipo de prestación y mediante el método de Monte-Carlo, son las siguientes para los volúmenes de carteras simuladas. En primer lugar, las prestaciones que ocupan una mejor situación por una mejor exigencia son, Asistencia sanitaria a socios y Sepelio con carteras de riesgo superiores a 50 Mill Ptas; Enfermedad con carteras de riesgo superiores a 100 Mill Ptas y Hospitalización Quirúrgica con carteras de riesgo superiores a 20 Mill Ptas. En segundo lugar, Invalidez y Larga Enfermedad necesitan de un mayor volumen de cartera. Consideramos que carteras de volúmenes [10;20]Mill Ptas serían suficientes para alcanzar los niveles oficiales definidos. En tercer lugar, la prestación de Intervención Quirúrgica posee comparativamente con Hospitalización Quirúrgica, un coste medio por siniestro superior en un 38% y un 100% en número de siniestros. Recomendamos políticas de eficiencia para reducir el coste medio. En cuarto lugar, Servicio a Automovilistas posee un coste medio de 19246 Ptas por siniestro y para una cartera de 50 Mill Ptas de riesgo, un número demasiado elevado de siniestros, 2598. La recomendación en este caso es incrementar el volumen de cartera hasta niveles de [150;200]Mill Ptas, ya que las variables estocásticas generadoras de la siniestralidad son de difícil control. Como podemos observar la *concentración* del sector es necesaria para reducir la necesidad de requerimientos de recursos propios.
7. El **comportamiento empírico** promedio del propio ratio de siniestralidad (**R2**), para la muestra de entidades analizada y por prestación es: enfermedad (71%), larga enfermedad (57%), invalidez (79%), intervención quirúrgica (87%) y hospitalización quirúrgica (48%), accidentes (38%), sepelio (53%), servicio a automovilistas (65%) y asistencia sanitaria (69%). Resaltamos que el sector en promedio posee una siniestralidad del 64%, sin distinguir por tipo de prestación, (apartado 5.3.3.1.).

8. El **comportamiento probabilístico** del ratio de siniestralidad (**R2**), está representado por los modelos teóricos siguientes, (apartado 5.3.3.2.). El modelo Gamma es el que posee una mayor flexibilidad para representar el ratio de siniestralidad del sector sin distinguir por prestaciones. Si consideramos cada una de ellas, el modelo Rayleigh es el más idóneo para las prestaciones de *enfermedad* y *servicios a automovilistas*; el modelo Weibull para *invalidez, hospitalización quirúrgica* y *asistencia sanitaria a socios*; el modelo Erlang para *sepelio* y *larga enfermedad* y finalmente el modelo Gamma para las prestaciones de *accidentes* e *intervención quirúrgica*. De todos modos, los gráficos de dispersión de los momentos superiores (asimetría y curtosis) para los ratios (Lau, H-S.; Lau, A. H-L.; Gribbin, D.W. (1995), McLeavy, S. (1997)) nos indican que para algunas de las prestaciones, los modelos de probabilidad estimados no son muy representativos y por lo tanto nos plantea dudas sobre su fiabilidad.
9. Para resolver las dudas anteriores, planteamos la utilización de **distribuciones complementarias Log-Log**, (Mittnik, S.; Paoella, M.S. (1999); Höpfner, R.; Rüschemdorf, L. (1999)). El exponente característico del modelo *Pareto estable* $S_{\alpha}(\beta, \mu, c)$ permite identificar si la distribución posee momentos finitos o no, como consecuencia del tamaño o “peso” de su cola. Los resultados nos indican que todas las prestaciones poseen momentos finitos ya que su valor es superior a 2, garantizando por lo tanto que son estables al aumentar el tamaño de la muestra, (apartado 5.3.3.3.).

La segunda línea de investigación, consistente en atajar la vertiente econométrica del modelo de ratio, está justificada por los aspectos metodológicos del capítulo 1. Así la posibilidad de la existencia de una relación entre los componentes que generan el propio ratio de siniestralidad invita a plantear un conjunto de **especificaciones** para el modelo de ratio. De todos modos somos conscientes de las dudas que pueden aflorar sobre la forma en puede afectar a dichas especificaciones el hecho de que, una de sus componentes, la siniestralidad, posea claramente un comportamiento asimétrico del tipo *Gamma*, sobre todo en términos del control sobre el comportamiento del residuo del modelo.

Las especificaciones son:

- Modelos de Regresión con Errores Aditivos (MREA(1) y (2)) (McDonald y Morris (1984, 1985) y García-Ayuso Covarsí (1994)) (apartado 5.4.2.),
- Modelos de Regresión con Errores Multiplicativos (MREM(3) y (4)), (García-Ayuso Covarsí (1994)) y Trigueiros, D. (1997))(apartado 5.4.3.),
- Modelo Característico de Ratio con Errores Aditivos (MCREA(5) y (6)) (apartado 5.4.4.),
- Modelo Característico de Ratio Errores Multiplicativos (MCREM(7)) (apartado 5.4.5.), (*ratio umbral* (Trigueiros, D (1997))),
- Formas Generalizadas sobre la base de la transformación de Box-Cox (Box y Cox (1964)) (**FFG(1) y FFG(2)**) Sundarsanam y Taffler (1995) (apartado 5.4.6.),
- Modelos Econométricos Neuronales (MN) (capítulo 6).

Para todas las especificaciones econométricas anteriores se han realizado las comprobaciones siguientes: significación del término independiente (si está definido); el grado de ajuste a través del modelo (coeficiente de determinación, R^2 y estadísticos MAE, RMSE); detección de heterocedasticidad en los residuos (mediante el contraste de White combinado posteriormente con el test de Glesjer); normalidad de las perturbaciones (contraste de Jarque-Bera), presencia de autocorrelación espacial (contraste de Durbin-Watson) y finalmente, el contraste Reset propuesto por Ramsey (1969) que permite detectar un error en la especificación de la forma funcional. Los resultados obtenidos se detallan a continuación.

1. Los resultados obtenidos para la especificación aditiva del modelo de ratio mediante el análisis de regresión, (MREA(1) y (2)) son: mayoritariamente el esquema de heterocedasticidad detectado por el contraste de Glesjer es del tipo, $\varepsilon_i \rightarrow N(0; \sigma^2 x_i)$. Este aspecto implica que el modelo de ratio no neutralizará del todo el “efecto tamaño” de las empresas, además la presencia de términos independientes significativos en cinco de las nueve prestaciones es un síntoma de que existen otros factores a neutralizar a parte del propio tamaño empresarial. Finalmente, cinco de las nueve prestaciones poseen un error en la especificación funcional, siendo este aspecto una posible causa de los problemas econométricos detectados.
2. Si consideramos ahora las especificaciones con errores multiplicativos (MREM(1) y (2)) tenemos: sólo tres de las nueve prestaciones no poseen error en la forma funcional (no existe mejora alguna respecto a los modelos de regresión con errores aditivos). Respecto a la heterocedasticidad, sigue existiendo una mayoría de prestaciones con este problema, seis de nueve prestaciones. El esquema detectado más frecuente es $u_i \rightarrow N(0; \sigma^2 x_i)$. El problema de la no normalidad en los residuos continúa, si bien en este caso existen tres de nueve prestaciones que han aceptado el test de Jarque-Bera. La autocorrelación espacial es presente y no mejora entre los dos grupos de modelos estudiados. En último lugar, la hipótesis de **proporcionalidad estricta** es rechazada en ocho de las nueve prestaciones, pero debe tomarse con cautela debido a los diferentes problemas detectados de carácter econométrico. En síntesis no parece que exista una clara mejoría en la especificación multiplicativa frente a la aditiva, consideramos que la naturaleza especial de los componentes del ratio de siniestralidad ha condicionado los resultados.

3. Para la especificación característica del modelo de ratio con errores aditivos (MCREA(5) y (6)) detectamos los siguientes aspectos: cinco de las nueve prestaciones poseen normalidad en los residuos; los problemas de heterocedasticidad se han reducido en el 30% de los casos; continúan los problemas de error en la especificación funcional, en cinco de los nueve casos; el término independiente es significativo en ocho de las nueve prestaciones disponibles, generando dudas sobre la eficiencia del ratio con deflactor de la dimensión empresarial.
4. Interesante es la aportación realizada por Trigueiros (1997), para la especificación característica con errores multiplicativos (MCREM(7)), consistente en un modelo de **ratio umbral** (bajo la hipótesis de comportamiento lognormal del ratio y de sus componentes). Uno de los tres parámetros que lo constituyen, “ δ ”, captura el desplazamiento de la distribución del ratio causada por los efectos fijos, que se considera una posible causa de la presencia de no proporcionalidad. En nuestro caso no añade ninguna mejora econométrica y sólo en cuatro de las prestaciones es significativo dicho efecto.
5. En definitiva, la **especificación característica del ratio** es “relativamente” mejor que el modelo de regresión por las razones siguientes: reduce en un 25% (en promedio) el error en la forma funcional y los problemas de heterocedasticidad; disminuye casi en un 50% los problemas de no normalidad de los residuos; persiste la significación del término independiente (una de las posibles causas de la no proporcionalidad) y la existencia de autocorrelación espacial. El reducido tamaño de la muestra en algunos casos y la posibilidad de relaciones no lineales de carácter desconocido que vinculen los dos componentes estudiados, pueden ser causas de los problemas econométricos detectados, en especial el error en la forma funcional.

6. Como consecuencia de las conclusiones anteriores, se propone realizar un contraste de la forma funcional mediante la utilización de la **transformación de Box-Cox** (Sundarsanam y Taffler (1995)(FFG(1) y FFG(2)). Así en primer lugar especificamos un modelo de tres parámetros, $[\alpha; \beta; \lambda]$, que incluye como casos particulares, el modelo Lineal y el modelo Loglineal. Posteriormente se podrá realizar su contraste econométrico y la comprobación de la hipótesis de proporcionalidad estricta a través del planteamiento siguiente: $H_0 : \alpha = 0$ y $H_0 : \beta = 1$. Además la especificación de un modelo recíproco (caso particular de Box-Cox), permite contrastar la validez de la forma funcional del modelo característico de ratio ($H_0 : \lambda = 1$), (véase apartado 5.4.6.1.). En segundo lugar, presentamos un modelo de regresión de Box-Cox más general con cuatro parámetros, $[\alpha; \beta; \lambda_1; \lambda_2]$, donde λ_1 y λ_2 suponen distintas transformaciones sobre las variables, $[y; x]$. A través de este modelo podemos plantar un paralelismo entre un modelo recíproco y el propio modelo característico de ratio con término independiente y realizar su contraste, $H_0 : \lambda_1 = 1; \lambda_2 = -1$, (véase apartado 5.4.6.2.).
7. El análisis del planteamiento anterior de tres parámetros nos aporta las conclusiones siguientes. Se rechazan las hipótesis de forma funcional *Lineal* y de proporcionalidad y *Loglineal*, tanto para el sector como para cada una de las prestaciones. Si aplicamos dichas transformaciones a los datos, las prestaciones con error en la especificación funcional se mantienen, los problemas de heterocedasticidad se reducen un 50%, se solucionan los problemas de no normalidad en los residuos en un 70% de los casos, permanecen los problemas de autocorrelación espacial y finalmente, sólo para la prestación de Intervención Quirúrgica se acepta el modelo característico de ratio con término independiente, es decir, en este caso sería metodológicamente correcto utilizar la especificación característica de ratio a efectos comparativos.

8. Si consideramos ahora la especificación con cuatro parámetros, los resultados no mejoran, de forma que, continúa el rechazo al planteamiento sobre la forma funcional *Lineal* y de proporcionalidad y *Loglineal*. La doble transformación ha permitido anular los problemas de heterocedasticidad en casi su totalidad pero persisten los problemas de no normalidad en los residuos y de error en la especificación funcional. Y sólo para aquellas prestaciones que ha sido posible, se rechaza la forma funcional de un modelo recíproco idéntico al modelo característico de ratio con término independiente. En definitiva, los modelos de regresión de Box-Cox sólo reducen claramente los problemas de heterocedasticidad.

A la vista de los anteriores resultados econométricos consideramos que una nueva vía de investigación puede ser la utilización de modelos econométricos neuronales (MN). El *capítulo 6* contiene varias aplicaciones en esta línea metodológica, desarrollada en los capítulos 2 y 3. Las conclusiones obtenidas son las siguientes:

1. La *identificación* de la forma funcional del modelo de ratio mediante modelos neuronales, se ha realizado a partir de tres estructuras diferentes del modelo *Multilayer feed-forward (MLP(1:3), MLP(1:2) y MLP(1:1))*. Los resultados no mejoran sustancialmente la especificación lineal de la relación a comprobar, si bien, una de las posibles causas es la propia naturaleza del ratio objeto del estudio, donde los componentes que lo configuran poseen una naturaleza de “flujo” muy similar, (véase apartado 6.2.).
2. Mediante la comparación de los resultados obtenidos a través de las diferentes especificaciones del modelo de ratio y la nueva propuesta, modelo de regresión generalizado neuronal (GRNN), las conclusiones son las siguientes: con datos sin transformar el mejor modelo es el MLP(1:2); con datos transformados de forma logarítmica, MLP(1:3); para la expresión del propio modelo característico de ratio con término independiente, el modelo GRNN y frente la especificación de Box-

Cox, el modelo con mejor ajuste es el MLP(1:2). Concluimos que la modelización neuronal *Multilayer feed-forward* es superior a la propuesta alternativa, GRNN, si bien consideramos que esta pequeña ventaja proviene de la propia naturaleza del ratio de siniestralidad, (véase apartado 6.3.).

3. Como consecuencia de las dificultades detectadas para especificar la forma del modelo de ratio, consideramos que una posible ampliación del análisis es utilizar los modelos neuronales como “**generadores**” de ratios partir de sus variables básicas, de forma que, conforma una arquitectura muy parecida a un modelo econométrico neuronal *Multilayer feed-forward* **MLP(4:3)**, (véase apartado 5.4.1.). La detección de **relaciones relevantes** (pesos o conexiones con mayor valor absoluto) por parte de modelo neuronal estimado es de un 50%, es decir, retiene sólo la mitad de las conexiones que vinculan las variables básicas con los ratios intermedios. Una de las posibles soluciones para intentar mejorar los resultados anteriores, es una nueva topología *ampliada* **MLP(4:9)**. Para esta arquitectura la combinación de funciones es parcialmente idéntica al modelo anterior MLP(4:3): función de transferencia, modelo *Gaussiano* y función de salida, modelo *logístico*. La ganancia en términos de RMSE y MAE es mínima, si bien podemos comprobar como el modelo ampliado es capaz de identificar los ratios intermedios ($R_1; R_3; R_4$) en su capa oculta, reteniendo la totalidad de las relaciones relevantes, (véase apartado 6.4.).
4. Una de las dificultades comentadas en el capítulo 1 en el análisis de ratios es la definición de un *benchmark sectorial*. En este ámbito y reafirmado por las dificultades detectadas para encontrar una especificación clara del modelo de ratio, se propone utilizar el residuo del modelo MLP(4:3) como medida complementaria del posicionamiento de las entidades, cuantificando de forma flexible la desviación de la entidad respecto del sector, representado por el propio modelo neuronal. Se han generado seis clusters, mediante los siguientes puntos de referencia: valor medio del ratio de siniestralidad, umbral unitario del mismo, y signo del residuo. Cada entidad se posiciona en cada uno de los grupos obteniendo un *scoring*, que tendrá en cuenta la posibilidad de que las mutualidades de previsión social puedan

estar en más de un ramo. Con carácter agregado y para cada una de las entidades, se han detectado tres de ellas con un comportamiento diferencial importante, aquellos con los identificadores siguientes: 17, 22, 25. Dicho procedimiento permite idear un ranking de entidades para su mayor control por parte de la Administración Pública competente.

5. Si acumulamos las puntuaciones obtenidas anteriormente para cada entidad en función del posicionamiento de su siniestralidad, podemos generar cinco grupos, (Grupo A [< 25]; Grupo B [$25;50$]; Grupo C [$50;75$]; Grupo D [$75;125$] y Grupo E [> 125]), de cuyo análisis se derivan las conclusiones siguientes. En primer lugar, para el **grupo A**, en ninguna de las prestaciones de las 8 entidades que lo forman, su valor mediano es superior al valor mediano referencia del sector. Las entidades que configuran el **grupo B**, formado por 11 entidades, poseen tres prestaciones cuyos ratios son superiores al sector, Enfermedad, Sepelio y Servicios a Automovilistas. En promedio deberían reducir un 15% su siniestralidad para acercarse a los valores referencia. Para el **grupo C**, siete entidades, las prestaciones en peor situación son, Intervención Quirúrgica y Servicios a Automovilistas. Para poder acercarse a los valores referencia deberían reducir en promedio un 22% su siniestralidad. El **grupo D** posee las prestaciones siguientes: Enfermedad, Invalidez e Intervención Quirúrgica con un mayor número de problemas. En este caso es necesario en promedio una reducción de un 40% en su siniestralidad. En ultimo lugar están las entidades con mayores problemas (4), el **grupo E**, donde las prestaciones peor posicionadas son tres, Larga Enfermedad, Invalidez e Intervención Quirúrgica, que necesitarían una reducción del orden del 25% en promedio. Las prestaciones de Sepelio y Asistencia Sanitaria a Socios, por muy poco, están en línea con los valores referencia.

6. Finalmente se propone escoger como **referencia** para el sector, aquellos valores medianos obtenidos de los generados por el modelo neuronal que representa al sector no vida. Así pueden ser considerados como límites máximos a tener en cuenta por las entidades, de forma que, si una de ellas se sitúa por encima del límite establecido nos indica que su posición no es la correcta respecto al global de un sector en constante evolución, siendo relativamente estables respecto al tamaño de las carteras de riesgo simuladas. Los valores son los siguientes: Enfermedad (75%), Larga Enfermedad (62%), Invalidez (61%), Intervención Quirúrgica (63%), Hospitalización Quirúrgica (62%), Accidentes (58%), Sepelio (71%), Servicios a Automovilistas (70%) y Asistencia Sanitaria a Socios (81%).

Algunas de las futuras líneas de investigación que pueden surgir del presente trabajo son las siguientes:

1. El estudio podría ampliarse para otros muchos ratios, más de carácter económico-financiero, permitiendo realizar un modelo explicativo de la posición estimada por cada entidad a partir de los mismos y generando *alertas*, (véase anexo 1 de este apartado).
2. Profundizar en los resultados econométricos neuronales mediante la utilización de medidas de sensibilidad que permitan una mayor interpretación de los mismos.
3. Proponer un sistema de evaluación de entidades no vida en la línea del sistema híbrido N.E.W.S. propuesto Kramer, B. (1997), ampliado a un mayor número de ratios.
4. La ampliación de los resultados obtenidos para todo el sector de previsión social. Por ejemplo sería de mucho interés para la Comunidad Autónoma del País Vasco y las entidades dependientes de la Dirección General de Seguros.
5. Comprobar si el modelo Pareto-Estable (SP) es adecuado para modelizar los ratios del propio sector asegurador, cuya especificidad es notoria.
6. Consideramos que en el ámbito docente y en cursos superiores, sería interesante introducir gradualmente la metodología neuronal como una herramienta más de modelización, además de tener una mayor presencia en los cursos de doctorado de métodos cuantitativos.

7. Propuesta de un nuevo modelo oficial de recogida de datos que permita retener con mayor claridad el comportamiento de las prestaciones que este momento las entidades de previsión social facilitan a sus socios, desde la óptica actuarial. Consideramos que los datos de este modelo oficial son lo suficientemente importantes como para exigir mayor rigor al suministrar la información, (véase anexo 2 de este apartado).
8. Finalmente y desde el punto de vista metodológico, consideramos que los modelos neuronales pueden llegar a ser muy útiles para el ámbito de la modelización económica, ya sea empresarial o de carácter general.
9. Profundización en la utilización de “mixturas” de distribuciones para modelar ratios desde la óptica probabilística.

