

CAPÍTULO 6 DISEÑO Y APLICACIONES DE LA METODOLOGÍA *ECONOMÉTRICA NEURONAL*

6.1. Introducción.

Una vez desarrollado en el capítulo 2 los modelos neuronales más habituales y sus posibilidades estadísticas y econométricas en el capítulo 3, ahora abordamos algunas de sus aplicaciones¹ para el sector de entidades de previsión social de Cataluña². Debemos recordar que hasta este momento, poseemos una doble perspectiva de la siniestralidad de seguros de consumo anual no vida del sector: una perspectiva probabilística (véase apartado 5.3.), donde se proponen una familia de modelos de probabilidad para el ratio de siniestralidad³, (y/x), desglosado por tipo de prestación⁴. Y una perspectiva econométrica (véase apartado 5.4.) donde se plantean diferentes especificaciones para el modelo de ratio⁵, con el objetivo de comprobar qué es más conveniente, la especificación del ratio mediante modelos de regresión (entre las componentes que lo integran) o una especificación característica. Los diferentes modelos propuestos incorporan gradualmente una mayor “flexibilidad” en la propia especificación a través del número de parámetros. Uno de ellos, los modelos de regresión con transformaciones Box-Cox (véase apartado 5.4.6.), nos han permitido llegar a la conclusión de que las especificaciones habituales de la forma funcional del modelo de ratio no son concluyentes⁶, existiendo problemas de definición de la misma. Por esta razón consideramos de interés introducir formas con menos restricciones como por ejemplos los modelos econométricos neuronales (MN).

¹ En cada una de ellas solo se utiliza una parte de las herramientas metodológicas comentadas. En futuras líneas de investigación podrán ampliarse.

² Véase el análisis del ratio de siniestralidad en los capítulos 4.1.3.1.1. y 5.2.

³ Véase el apartado 5.3.3.

⁴ Adicionalmente se ha obtenido el margen mínimo de solvencia (MMS) por métodos de Monte-Carlo y a través de los propios modelos de probabilidad del ratio de siniestralidad.

⁵ Véase apartado 1.2. a 1.4.

⁶ Con los datos que se poseen y para el sector analizado.

Dos son los objetivos generales que se abordan en este capítulo. La propuesta de especificaciones alternativas para el modelo de ratio a través de modelos neuronales y el diseño de un “*scoring*” para las compañías que integran dicho sector, a través de una representación neuronal de la siniestralidad del sector.

Para conseguir los objetivos anteriores se han materializado dos aplicaciones, que a continuación detallamos. La *primera* de ellas consiste en ampliar las especificaciones para el modelo de ratio a través de los modelos econométricos neuronales (MN) denominados *Multilayer feed-forward*⁷, y mediante los modelos de regresión neuronal (GRNN)⁸. Ambos poseen muchas similitudes con los modelos definidos en el apartado 5.4. Pero para conseguir este primer objetivo, condición necesaria es comprobar la naturaleza de la forma funcional⁹ existente entre los componentes del ratio de siniestralidad. Los modelos neuronales *Multilayer feed-forward* (MLP) se han utilizado con esta finalidad, escogiendo la arquitectura más idónea que incluye lógicamente la especificación lineal.¹⁰, (véase el apartado 6.2.). La presencia detectada de formas no lineales en el modelo de ratio, nos incita a pensar que pueden existir otras especificaciones cercanas a los modelos de regresión ya especificados, es decir, los modelos de regresión neuronal (GRNN). Para comprobar la bondad de las mismas, se comparan en términos de ajuste los resultados obtenidos, para las dos especificaciones neuronales, junto con los modelos clásicos definidos en el apartado 5.4., (véase apartado 6.3.).

La *segunda aplicación* aborda un aspecto de importancia y es la propia definición de sector o norma. Frente la presencia de formas funcionales desconocidas del modelo de ratio, es lógico pensar que la propia definición de sector debería de incluir una naturaleza mucho más flexible que superase cualquier concepción descriptiva del mismo. Por esta razón, utilizamos la arquitectura de un modelo neuronal *Multilayer feed-forward* (MLP (4:3)) para definir un modelo “*ad-hoc*” generador de ratios del sector, adaptado lógicamente a la morfología del ratio de siniestralidad (4 *inputs*, 3 ratios intermedios en la capa oculta y un

⁷ Utilizamos dichos modelos ya que poseen la propiedad de ser “aproximadores universales”, (véase apartado 2.5.1.).

⁸ Véase su especificación en el apartado 2.5.5.3., siendo GRNN, *Generalized Regresión Neural Network*.

⁹ Que consideramos desconocida del modelo de ratio, a partir de los resultados econométricos obtenidos en el apartado 5.4.

¹⁰ Consideramos que la forma flexible de los modelos neuronales puede sustituir el desconocimiento a priori de la verdadera relación entre los componentes de un ratio.

output que es el ratio de siniestralidad), (véase apartado 6.4.). Para finalizar y a partir del modelo ya definido, *Multilayer feed-forward* estimado (MLP(4:3)), se utiliza la información que contienen sus residuos, para explicar el posicionamiento relativo de las entidades respecto al *benchmark* del sector, representado por dicho modelo neuronal. Esta información se combina para formalizar grupos o cluster de las entidades que constituyen el sector de previsión social en Cataluña, permitiendo obtener puntuaciones relativas de cada una de ellas respecto al sector y manteniendo su naturaleza *Multiprestación*. Los valores “ajustados” del modelo neuronal estimado suministran valores “referencia” para la siniestralidad del sector, estables respecto al tamaño de las compañías, (véase apartado 6.5.).

6.2. Identificación de la forma funcional del modelo de ratio mediante modelos neuronales (MLP($p; q$)).

En este apartado proponemos utilizar la tipología econométrico neuronal MLP($p; q$)¹¹, como un instrumento de aproximación a la verdadera relación entre los componentes que configuran el ratio de siniestralidad, "R₂", comprobando su posible naturaleza no lineal¹², sobre la base de las conclusiones obtenidas, mediante contrastes econométricos, del rechazo de dos hipótesis sobre la forma funcional del modelo de ratio, la primera, *Lineal* y *Proporcional* y la segunda, *Loglineal*. Los modelos neuronales puede ayudarnos en la línea anterior a través de la siguiente especificación neuronal¹³,

$$y = f\left(\beta_0 + \sum_{j=1}^q \beta_j g(x\phi_{ij} + \phi_{0j})\right)$$

$$\left\{ \begin{array}{l} \phi_{ij}, i = 0, 1, \dots, p, j = 1, \dots, q \\ \beta_j, j = 0, 1, \dots, q \end{array} \right\}$$

siendo,

- f : función de salida,
- g : función de activación o transferencia en la capa oculta,
- p : número de *inputs*, en este caso es único, importe ingresado por tipo de prestación i -ésima,
- q : número de neuronas en la capa oculta¹⁴,
- y : *output*, importe pagado por tipo de prestación i -ésima,
- x : *input*, importe ingresado por tipo de prestación i -ésima.

Nos planteamos tres posibles arquitecturas¹⁵ para el modelo MLP($p; q$), que nos permitan representar la posible relación existente entre los dos componentes que configuran el ratio de siniestralidad, (véase ilustración 1.6.2.).

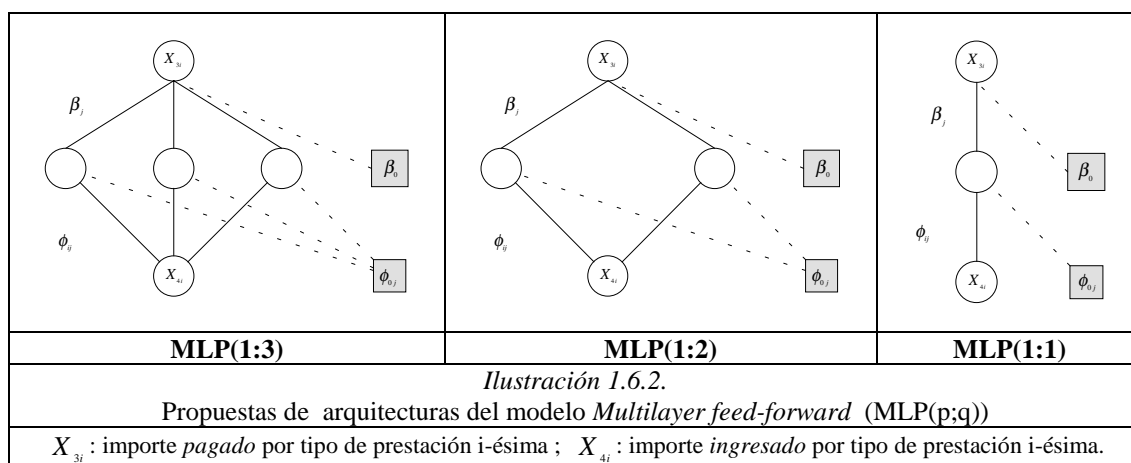
¹¹ Todas las simulaciones de este capítulo se han realizado con el software [NeuroShell@2](#).

¹² Los datos de los diferentes modelos se han transformado para situarlos en el rango $[0;1]$, facilitando de esta forma el aprendizaje de los mismos.

¹³ Véase el capítulo 3.4.3. para la especificación general.

¹⁴ Existen algunas directrices habituales para su determinación: "2n+1" (Lippmann, 1987; Hecht-Nielsen, 1990); "2n" (Wong, 1991); "n" (Tang y Fishwick, 1993); "n/2" (Kang, 1991), donde "n" es el número de inputs, (referenciado en Guoqiang Zhang, B. , Hedí Patuwo, Michael Y. Hu. (1998). **Forecasting with artificial neural networks: The state of the art**. *International Journal of Forecasting*, 14, pp. 35-62).

¹⁵ Las tres arquitecturas propuestas son: **MLP(1;3)**, **MLP(1;2)** y **MLP(1;1)**, utilizando como valor máximo de neuronas en la capa oculta, "2n+1", siendo "n" el número de *inputs* del modelo.



El interés de esta primera aplicación radica en conocer la combinación¹⁶ de funciones de transferencia y de salida que consiga el mejor ajuste en términos de los estadísticos, MAE y RMSE¹⁷, observados en la base de validación o de test¹⁸. A su vez podremos constatar si dicha combinación de funciones es distinta a una combinación de funciones lineales. El proceso de aprendizaje y su validación a través de la base de test consiste en fraccionar la muestra de 801 observaciones en dos submuestras¹⁹. Es decir, se reserva un 20% de los datos para *test* y el resto se escoge de forma aleatoria²⁰. Así la muestra original²¹, queda dividida en 641 observaciones para el aprendizaje (*train*) y 160 para la base de datos de validación (*test*).

¹⁶ Las funciones utilizadas en ambos casos, transferencia y salida, son: *Logística, Lineal, Tangente Hiperbólica, Gaussiana y Sinusoidal*.

¹⁷ RMSE es la raíz del error cuadrático medio y MAE, error absoluto medio. El segundo de estos estadísticos es más resistente a la presencia de *Outliers*.

¹⁸ Como consecuencia de poseer un tamaño de muestra suficiente pero no holgado y que el objetivo final no es la predicción sino comprobar la capacidad de los modelos neuronales de retener estructuras, se ha decidido prescindir de la base de datos de *predicción* y sólo trabajar con la base de *aprendizaje* y de validación o *test*. Además el error en la base de *predicción* será pequeño debido a que es proporcional a la relación entre el número de parámetros (relacionado con el número de neuronas en la capa oculta) y el tamaño de la base de *aprendizaje*, véase Baum, E. (1990). **When are k-nearest neighbor and back propagation accurate for feasible sized sets of examples?**, *Lecture Notes in Computer Science*, 412, Springer. Dicho criterio se ha mantenido para los apartados 6.3. y 6.4.

¹⁹ El número de observaciones definido, 801, supone 36 entidades, por 8 ejercicios económicos (considerados independientes) para cada prestación definida.

²⁰ Existen otras formas de partición de las muestras en función de la tipología de los datos, véase apartado 3.3. En nuestro caso por ser datos de corte transversal que poseen independencia del ejercicio económico se ha considerado más idónea la expuesta en el párrafo.

²¹ Consideramos importante notar que la calidad de los datos puede influenciar tanto en los niveles de error como el grado de ajuste conseguido, véase Klein, B. D.; Rossin, D. F. (1999). **Data quality in neural network models: effect of error rate and magnitude of error on predictive accuracy**, *Omega, Int. J. Mgmt. Sci.*, Vol. 27, No. 5, pp. 569-582.

El proceso de estimación de los parámetros es de carácter iterativo, donde en nuestro caso, la regla utilizada para considerar que el proceso de búsqueda del mínimo (local o global) de la función de error definida es: finalizar el aprendizaje cuando el Error Cuadrático Medio (ECM) cometido en la base de datos de *test*²² no ha variado después de 1000 iteraciones. En este punto se considera que la modificación de los parámetros ya no mejorará la solución obtenida y el aprendizaje se detendrá. Este aspecto ayuda a evitar el posible sobreaprendizaje del modelo²³.

Los resultados de los estadísticos RMSE y MAE obtenidos para la base de datos de *test*²⁴, asociados a la mejor de las tres arquitecturas definidas MLP(1;2), están recogidos en la tabla 1.6.2. Las diferentes combinaciones de funciones de transferencia y salida incorporan lógicamente como caso particular la especificación de funciones lineales [$f : g$].

Tabla 1.6.2. Modelo especificado MLP(1;2) Base Test Número de iteraciones 1000

Capa oculta g	Logística		Lineal		Tangente Hiperbólica		Gaussiana		Sinusoidal	
Capa output f	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Logística	14.2	8.8	15.1	9.2	14.9	9.1	15.0	8.9	14.9	9.3
Lineal	78.6	39.1	12.5	6.2	12.5	5.9	12.5	6.1	12.7	6.6
Tang Hiperb	22.8	17.6	23.7	17.5	24.2	18.2	22.9	18.0	23.9	18.3
Gaussiana	21.8	17.0	22.7	17.2	22.8	17.1	22.3	17.2	22.8	17.1
Sinusoidal	17.2	13.6	18.2	11.1	17.9	11.1	17.1	13.1	18.0	11.0

Nota: Se presenta los resultados del ajuste realizado a partir de la *función de transferencia* utilizada en la capa oculta y de la *función de salida* de la capa output. En sombreado reseñamos la combinación de funciones lineales y en negrita las mejores combinación de funciones.

Los resultados de las tres arquitecturas²⁵ indican persistencia de las mejores combinaciones de funciones, “ f ” y “ g ”, siendo la mejor función de transferencia, g : **Tangente Hiperbólica** y la función de salida, f : **Lineal**. Este aspecto nos induce a pensar en la posibilidad de cierta naturaleza no lineal en la relación entre los dos componentes del ratio, aunque la mejora respecto a la combinación de funciones lineales es pequeña.

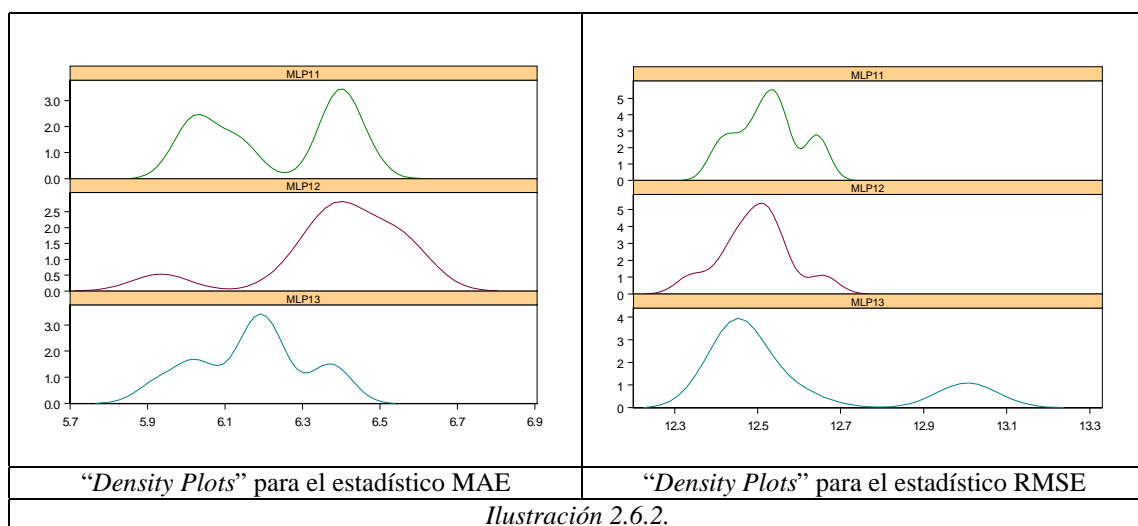
²² La *calibración* para este proceso de validación se ha estipulado cada 200 iteraciones, es decir, cada 200 iteraciones el algoritmo comprueba el error cometido en la base de datos *test*.

²³ El número de iteraciones necesarias para confirmar el proceso de estimación, debe confirmarse de forma gráfica, es decir, a partir del momento del decaimiento de la función de error. En nuestro caso, se ha decidido en 1000 iteraciones, (véase en el anexo 1 de este apartado la evolución del error cometido en cada iteración hasta 1000 iteraciones).

²⁴ Sólo se presentan los resultados estadísticos para la base de datos reservada para la validación o *test*.

²⁵ Véase en el anexo 2 los resultados para las tres arquitecturas.

Las conclusiones son las siguientes. En primer lugar, los resultados no mejoran sustancialmente la especificación lineal de la relación. En segundo lugar, la especificación MLP(1;2) posee cierta ventaja para la misma base de datos utilizada. En tercer lugar, debemos considerar que los resultados están condicionados por la propia naturaleza del ratio de siniestralidad. Éste no es el más idóneo, debido a que los componentes que lo configuran poseen características similares. Una ampliación de la línea de investigación podría consistir en comprobar la bondad de la metodología aplicada, sobre ratios de naturaleza de tipo más económico-financiero, (véase anexo 1 del apartado 7). En último lugar, para comprobar la estabilidad de los resultados obtenidos en términos de RMSE y MAE frente a cambios en las condiciones iniciales, se ha inicializado el proceso de estimación para los tres modelos propuestos, diez veces de forma aleatoria, obteniendo una distribución de los propios estadísticos. Observamos como en términos de MAE la estructura MLP(1:3) posee la distribución más concentrada y por lo tanto con mayor estabilidad, en cambio para el estadístico RMSE, el aumento de la complejidad del modelo empeora su estabilidad, (véase ilustración 2.6.2.).



La presencia de distribuciones con más de una moda es síntoma de la presencia de mínimos locales, una de las posibles soluciones es presentar intervalos de confianza para los estadísticos, con el coste computacional de un mayor número de simulaciones. El proceso habitual es generar diferentes modelos neuronales con inicializaciones distintas, escogidas de forma aleatoria y se selecciona aquel modelo con mejor ajuste en la base de validación o test.

6.3. Propuesta de modelización econométrica neuronal alternativa del modelo de ratio de siniestralidad: Modelo de Regresión Generalizada Neuronal (GRNN).

Una vez desarrollado en el apartado 6.2. diferentes topologías neuronales ($MLP(p : q)$) para representar el modelo de ratio, planteamos ahora una nueva posibilidad a través del modelo *Regresión Generalizada Neuronal*²⁶ (GRNN), tipología muy cercana a las especificaciones clásicas del apartado 5.4. La relación de modelos²⁷ que se comparan en términos de ajuste, donde se incluye la nueva propuesta, es la siguiente, (véase tabla 1.6.3.),

- Modelos de Regresión con Errores Aditivos y Multiplicativos con términos independientes²⁸ (MREA(2);MREM(4)),
- Modelos Característicos con Errores Aditivos y Multiplicativos (MCREA(6); MCREM(7)),
- Modelos de Regresión Box-Cox (FFG(1); FFG(2)),
- Modelos Neuronales Multilayer feed-forward (MN(8); MN(9); MN(10)),
- Modelo de Regresión Generalizada Neuronal (GRNN)(MRN(11)).

Tabla 1.6.3.

Modelos	Especificación General	Nomenclatura Neuronal
MREA(2)	$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$	
MREM(4)	$\ln y_i = \beta_0 + \beta_1 \ln x_i + \varepsilon_i$	
MCREA(6)	$y_i/x_i = \beta_0/x_i + \beta_1 + \varepsilon_i/x_i$	
MCREM(7)	$\ln y_i = \beta_0 + \beta_1 \ln(x_i - \delta_x) + \varepsilon_i^*$	
FFG(1)	$y_i^{(\lambda)} = \beta_0 + \beta_1 x_i^{(\lambda)} + \varepsilon_i$	
FFG(2)	$y_i^{(\lambda_1)} = \beta_0 + \beta_1 x_i^{(\lambda_2)} + \varepsilon_i$	
MN(8)	$y_i = f(\beta_0 + \beta_1 g(\phi_{11} x_i + \phi_{01}))$	MLP(1;1)
MN(9)	$y_i = f\left(\beta_0 + \sum_{j=1}^2 \beta_j g(\phi_{ij} x_i + \phi_j)\right)$	MLP(1;2)
MN(10)	$y_i = f\left(\beta_0 + \sum_{j=1}^3 \beta_j g(\phi_{ij} x_i + \phi_j)\right)$	MLP(1;3)
MRN(11)		GRNN(·)

²⁶ Véase especificación en el apartado 2.5.5.3.

²⁷ Véase para mayor detalle de las especificaciones no neuronales, los apartados 5.4.2. a 5.4.6.

²⁸ El software utilizado no permite el diseño de arquitecturas con ausencia de término “bias”, de ahí, que solo se han realizado comparaciones con modelos clásicos con término independientes. Además el análisis se ha realizado sin distinguir por tipo prestación, sólo para el sector en globalidad. En futuras ampliaciones de la línea de investigación se podría profundizar por tipología de prestaciones.

Todas las especificaciones econométricas²⁹ anteriores se comparan en términos de MAE y RMSE para comprobar su grado de ajuste, pero con anterioridad debemos considerar varios aspectos. Primeramente, para la especificación de los modelos neuronales *Multilayer Perceptron* (MLP) y sus diferentes topologías,

MLP(1:1),MLP(1:2),MLP(1:3),

se ha utilizado la misma combinación de funciones de transferencia y de salida $[f : g]$ ³⁰, permitiendo así una comparación más homogénea entre los diferentes modelos. Si incorporamos dichas funciones a la forma general de los modelos MLP conseguimos unas especificaciones mucho más completas³¹, (véase tabla 2.6.3.).

Tabla 2.6.3.

Modelos	Especificación Inicial	Modelo Neuronal	Especificación Final
MN(8)	$y = f(\beta_0 + \beta_1 g(\phi_{11}x + \phi_{01}))$	MLP(1:1) ➔	$y = \beta_0 + \beta_1 \left(\frac{\exp(\phi_{11}x + \phi_{01}) - \exp(-(\phi_{11}x + \phi_{01}))}{\exp(\phi_{11}x + \phi_{01}) + \exp(-(\phi_{11}x + \phi_{01}))} \right)$
MN(9)	$y = f\left(\beta_0 + \sum_{j=1}^2 \beta_j g(\phi_{ij}x + \phi_{0j})\right)$	MLP(1:2) ➔	$y = \beta_0 + \sum_{j=1}^2 \beta_j \left(\frac{\exp(\phi_{ij}x + \phi_{0j}) - \exp(-(\phi_{ij}x + \phi_{0j}))}{\exp(\phi_{ij}x + \phi_{0j}) + \exp(-(\phi_{ij}x + \phi_{0j}))} \right)$
MN(10)	$y = f\left(\beta_0 + \sum_{j=1}^3 \beta_j g(\phi_{ij}x + \phi_{0j})\right)$	MLP(1:3) ➔	$y = \beta_0 + \sum_{j=1}^3 \beta_j \left(\frac{\exp(\phi_{ij}x + \phi_{0j}) - \exp(-(\phi_{ij}x + \phi_{0j}))}{\exp(\phi_{ij}x + \phi_{0j}) + \exp(-(\phi_{ij}x + \phi_{0j}))} \right)$

Nota: función de transferencia, tangente hiperbólica y como función de salida, lineal

El siguiente aspecto a considerar está relacionado con el diseño del modelo **GRNN**, de forma que debemos tener presente una serie de características. En primer lugar, se ha utilizado la distancia Euclídea³² para cuantificar la similitud entre $[x; y]$. En segundo lugar, el factor de escala³³ “ σ_j ” óptimo, se ha obtenido mediante algoritmos genéticos³⁴. El procedimiento de búsqueda de soluciones mediante algoritmos genéticos es un complemento habitual de los modelos neuronales, que consiste en generar sucesivas generaciones de individuos que muestren un mejor ajuste. Este proceso de reproducción permite seleccionar a los individuos

²⁹ Para poder homogeneizar la nomenclatura en este apartado se ha considerado el término independiente “ α ” como “ β_0 ”.

³⁰ Detectada como la combinación de mejor ajuste, (véase apartado 6.2., $g : \text{Tangente Hiperbólica} / f : \text{Lineal}$).

³¹ En el anexo 1 de este apartado se presentan las matrices de parámetros estimados de los modelos neuronales que permiten su comparativa, en términos de “pesos”, con los respectivos del Modelo de regresión con errores aditivos.

³² Véase especificación del modelo GRNN en el apartado 2.5.5.3.

³³ Equivale al radio de influencia de la neurona en el espacio de los *inputs*.

³⁴ Véase Davis, L. (1991). Handbook of genetic algorithms, Van Nostrand Reinhold, New York.

que muestran mayor compatibilidad y por lo tanto ofrecen la mejor solución al problema de minimización de la función de error en la base de *test*. Y por último, el proceso de aprendizaje se ha detenido en el momento que, después de 20 generaciones de individuos, el error cuadrado medio (ECM) del modelo no ha variado en 1% en la base de *test*.

Los resultados obtenidos en términos de MAE y RMSE para la primera especificación econométrica del modelo de ratio, $y = \beta_0 + \beta_1 x + \varepsilon_i$ (datos sin transformar) y de sus homólogos neuronales están recogidas en la tabla 3.6.3.

Tabla 3.6.3. Diferentes especificaciones del modelo de ratio. Base Test

	<i>Especificación</i>	$R^2 \geq 0.80$	<i>MAE</i>	<i>RMSE</i>	<i>n</i>
MREA(2)	$y = \beta_0 + \beta_1 x + \varepsilon_i$	*	7.0	12.8	160
MN(8)	<i>MLP(1;1)</i>	*	6.2	12.5	160
MN(9)	<i>MLP(1;2)</i>	*	5.9	12.6	160
MN(10)	<i>MLP(1;3)</i>	*	6.4	12.4	160
MRN(11)	<i>GRNN(.)</i> $\sigma = 0.033$	*	9.9	16.6	160
<i>Nota:</i> * valor que cumple restricción $R^2 \geq 0.80$; R^2 es el coeficiente de determinación; <i>MAE</i> error absoluto medio; <i>RMSE</i> , raíz del error cuadrático medio; <i>n</i> : tamaño muestra. En cada uno de los casos de los modelos neuronales se presenta sólo los resultados para la base de datos de validación o test.					
MREA(2): Modelo de regresión errores aditivos con término independiente.					
MN(.): Modelos neuronales <i>Multilayer feed-forward</i> con diferentes topologías.					
MRN(11): Modelo Regresión Neuronal.					

Las conclusiones que se derivan son las siguientes. En primer lugar, el mejor modelo³⁵ es el MLP con la topología MLP(1;2). En segundo lugar, se detecta sobreparametrización en el modelo MLP(1;3) ya que su error cuadrático aumenta con la complejidad del mismo. En tercer lugar, para el modelo GRNN el factor de escala óptimo estimado es $\sigma = 0.033$ obtenido por métodos genéticos. En último lugar, del estudio de los residuos de los diferentes modelos se desprende la presencia de “heterocedasticidad” en todos ellos³⁶. Su persistencia en todos los casos nos incita a considerar que los residuos poseen elementos informativos³⁷ que deberían ser considerados. Adicionalmente y con la finalidad de encontrar similitudes entre los residuos de cada uno de los modelos planteados, se analiza su forma³⁸.

³⁵ La comparación se realiza sobre la misma base de *validación*.

³⁶ Véase anexo 2 de este apartado.

³⁷ En el apartado 6.5. finalmente se ha utilizado los residuos del modelo representativo del ratio de siniestralidad MLP(4;3) por considerarse que es más idóneo para representar las relaciones entre las componentes del mismo.

³⁸ En el anexo 3 de este apartado contiene los “*density plots*” en la misma escala, donde resalta la mayor dispersión de los mismos en el modelo GRNN (MRN(11)).

Finalmente compararemos los resultados de las restantes especificaciones de los modelos clásicos con datos transformados, frente a las alternativas neuronales, (véase tablas 4.6.3. y 5.6.3.). Las conclusiones para los modelos anteriores son las siguientes. Con datos transformados de forma logarítmica el mejor modelo es el MLP(1:3). Para la expresión del propio modelo característico de ratio con término independiente, el mejor modelo es GRNN, con un factor de escala óptimo igual a, $\sigma = 0.056$. Dicha topología neuronal consideramos que puede ser una alternativa válida para una nueva especificación del modelo de ratio, vinculada a una forma funcional del ratio no lineal y desconocida a priori. Y en último lugar, si consideramos el modelo característico de ratio con errores multiplicativos (ratio umbral), los modelos neuronales no consiguen obtener mejores resultados en ninguna de las especificaciones, (véase tabla 4.6.3.).

Tabla 4.6.3.

	Especificaciones	$R^2 \geq 0.80$	MAE	RMSE	n
MREM(4)	$\ln y = \beta_0 + \beta_1 \ln x + \varepsilon_i$	*	0.60	0.83	159
MN(8)	<i>MLP(1;1)</i>	*	0.59	0.86	159
MN(9)	<i>MLP(1;2)</i>	*	0.70	0.86	159
MN(10)	<i>MLP(1;3)</i>	*	0.57	0.83	159
MRN(11)	<i>GRNN(.)</i> $\sigma = 0.15$	*	0.63	0.83	159
<hr/>					
MCREA(6)	$y/x = \beta_0/x + \beta_1 + \varepsilon_i/x$		0.30	0.39	159
MN(8)	<i>MLP(1;1)</i>		0.29	0.38	159
MN(9)	<i>MLP(1;2)</i>		0.28	0.38	159
MN(10)	<i>MLP(1;3)</i>		0.29	0.38	159
MRN(11)	<i>GRNN(.)</i> $\sigma = 0.056$		0.27	0.36	159
<hr/>					
MCREM(7)	$\ln y = \beta_0 + \beta_1 \ln(x - \delta_x) + \varepsilon_i^*$ $\delta_x = -0.083$	*	0.25	0.35	159
MN(8)	<i>MLP(1;1)</i>	*	0.31	0.45	159
MN(9)	<i>MLP(1;2)</i>	*	0.29	0.41	159
MN(10)	<i>MLP(1;3)</i>	*	0.27	0.37	159
MRN(11)	<i>GRNN(.)</i> $\sigma = 0.057$	*	1.15	1.38	159
<p><i>Nota:</i> * valor que cumple restricción $R^2 \geq 0.80$; R^2 es el coeficiente de determinación; MAE error absoluto medio; RMSE, raíz del error cuadrático medio; n: tamaño muestra. En cada uno de los casos de los modelos neuronales se presenta sólo los resultados para la base de datos de validación o test.</p> <p>MREM(4): Modelo de regresión errores multiplicativos con término independiente. MCREA(6): Modelo característico de ratio con errores aditivos. MCREM(7): Modelo característico de ratio con errores multiplicativos (ratio <i>umbral</i> general). MN(.): Modelos neuronales <i>Multilayer feed-forward</i> con diferentes topologías. MRN(11): Modelo Regresión Neuronal.</p>					

Respecto a las restantes especificaciones, véase tabla 5.6.3., las conclusiones son las siguientes. La forma funcional generalizada (FFG(1)) con solo parámetro, “ λ ”, consigue que ningún modelo neuronal genere mejores resultados, neutralizando cualquier posible no linealidad. En cambio si consideramos la forma funcional generalizada (FFG(2)) con dos parámetros, $(\lambda_1; \lambda_2)$, los resultados nos indican que el modelo MLP(1:2) posee un mejor ajuste, es decir, la doble transformación no consigue el mismo efecto que una sola, indicando que la dinámica de las dos variables es muy parecida, (véase las transformaciones en la ilustración 1.6.3.).

Tabla 5.6.3.

	Especificaciones	$R^2 \geq 0.80$	MAE	RMSE	n
FFG(1)	$y_i^{(\lambda)} = \beta_0 + \beta_1 x_i^{(\lambda)} + \varepsilon_i$ $\lambda = 0.3488$	*	0.30	0.40	159
MN(8)	MLP(1;1)	*	0.29	0.39	159
MN(9)	MLP(1;2)	*	0.29	0.40	159
MN(10)	MLP(1;3)	*	0.29	0.40	159
MRN(11)	GRNN(\cdot) $\sigma = 0.037$	*	0.30	0.40	159
FFG(2)	$y_i^{(\lambda_1)} = \beta_0 + \beta_1 x_i^{(\lambda_1)} + \varepsilon_i$ $\lambda_1 = 0.4987; \lambda_2 = 0.5017$	*	0.60	0.86	159
MN(8)	MLP(1;1)	*	0.59	0.84	159
MN(9)	MLP(1;2)	*	0.58	0.82	159
MN(10)	MLP(1;3)	*	0.59	0.85	159
MRN(11)	GRNN(\cdot) $\sigma = 0.056$	*	0.64	0.87	159

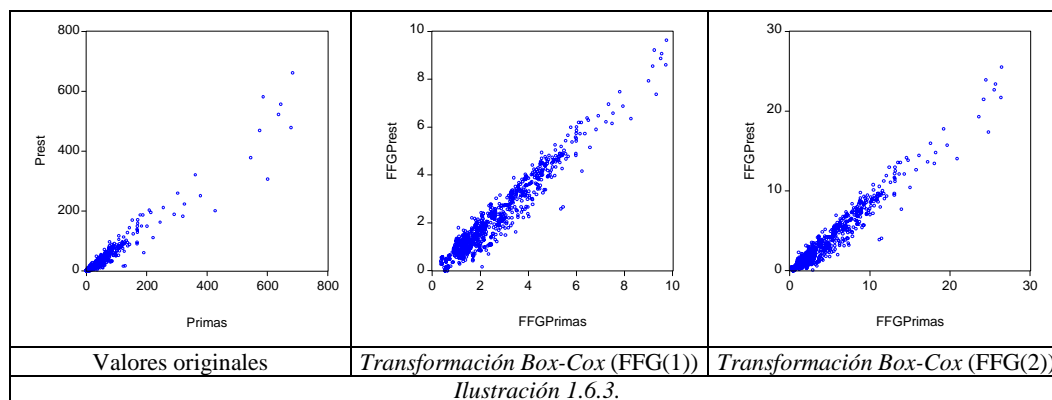
Nota: * valor que cumple restricción $R^2 \geq 0.80$; R^2 es el coeficiente de determinación; MAE error absoluto medio; RMSE, raíz del error cuadrático medio; n : tamaño muestra. En cada uno de los casos de los modelos neuronales se presenta sólo los resultados para la base de datos de validación o test.

FFG(1): Forma Funcional Generalizada ($\lambda_1 = \lambda_2$).

FFG(2): Forma Funcional Generalizada ($\lambda_1 \neq \lambda_2$).

MN(\cdot): Modelos neuronales *Multilayer feed-forward* con diferentes topologías.


MRN(11): Modelo Regresión Neuronal.



6.4. Definición flexible de sector mediante un modelo econométrico neuronal *Multilayer feed-forward* (MLP(4:3)).

La presencia de relaciones entre las variables básicas³⁹, permite obtener ratios de gestión⁴⁰, generando una arquitectura muy parecida a la especificación de un modelo neuronal. Es por ello que en este apartado proponemos una topología que represente y capture el comportamiento del “sector” en términos de siniestralidad, mediante la especificación de un modelo econométrico neuronal⁴¹ *Multilayer feed-forward* (MLP(4:3)), (véase tabla 1.6.4.).

Tabla 1.6.4.

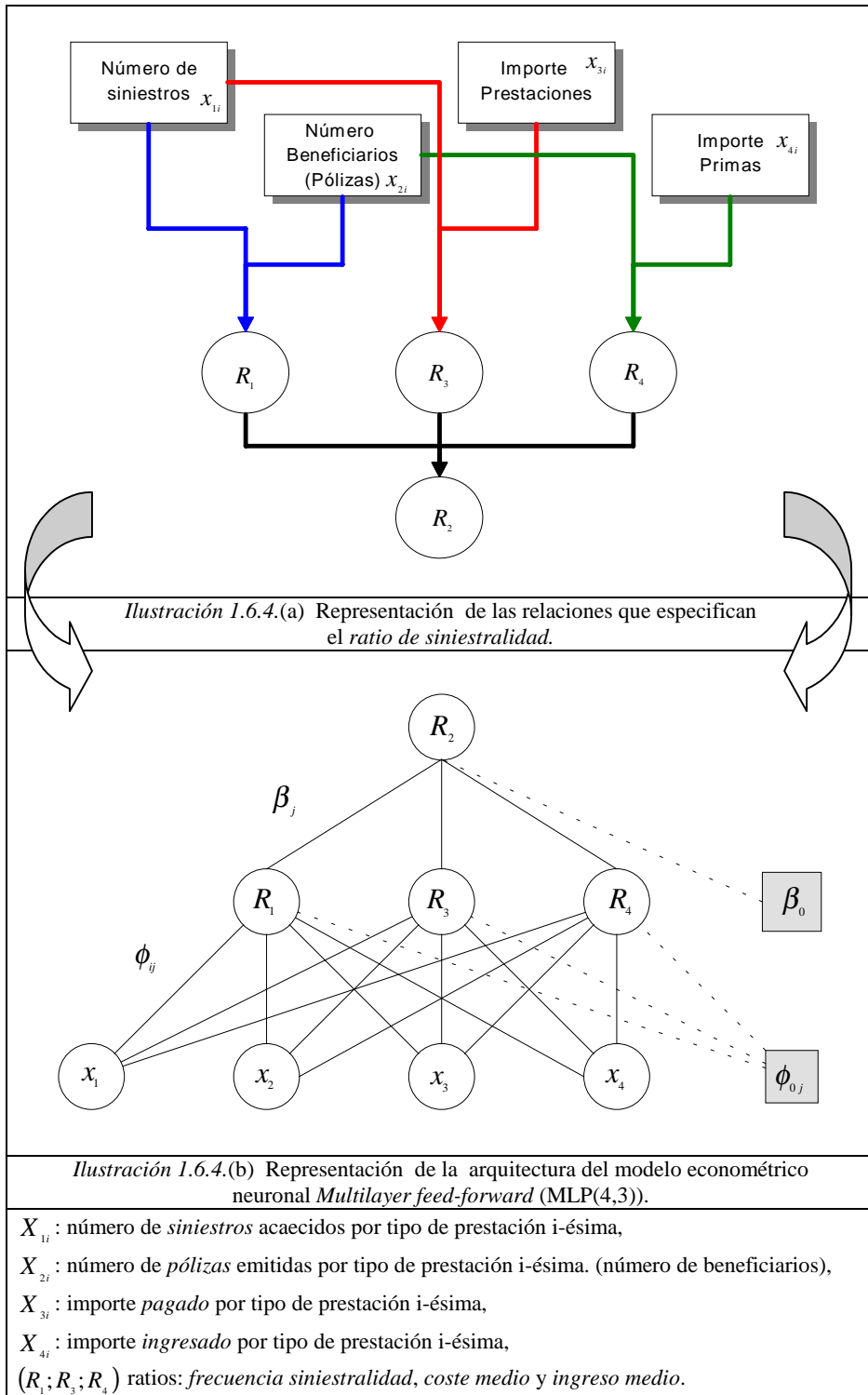
$R_1 = \frac{\text{Número de Siniestros}}{\text{Número de beneficiarios (Pólizas)}} \cdot 1000 = \frac{X_{1i}}{X_{2i}} \cdot 1000$
$R_3 = \frac{\text{Importe satisfecho por Prestaciones}}{\text{Número de Siniestros}} = \frac{X_{3i}}{X_{1i}} = m$
$R_4 = \frac{\text{Importe ingresado por Primas tarifa}}{\text{Número de beneficiarios (pólizas)}} = \frac{X_{4i}}{X_{2i}} = r$

$R_2 = \frac{\text{Gastos Técnicos}}{\text{Ingresos Adquiridos}} = \frac{(\text{Núm de Siniestros}) \cdot (\text{Coste Medio Siniestro})}{(\text{Núm Beneficiarios (Pólizas)}) \cdot (\text{Ingresos por Beneficiarios})} = \frac{X_{1i} \cdot m}{X_{2i} \cdot r}$

La arquitectura de modelo propuesto es la siguiente: cuatro *inputs* ($X_1; X_2; X_3; X_4$) que representan las variables básicas analizadas, tres neuronas en la única capa oculta, equivalentes a los tres ratios de gestión intermedios, ($R_1; R_3; R_4$) y un *output* que es el propio ratio de siniestralidad, “ R_2 ”, calculado a partir de ($R_1; R_3; R_4$), (véase ilustración 1.6.4.(a)). El modelo neuronal *Multilayer feed-forward* (MLP(4:3)) nos permite por lo tanto, centrar nuestra atención en dos aspectos. El primero de ellos consiste en el análisis de la capa oculta, donde teóricamente se generarán los ratios intermedios, ($R_1; R_3; R_4$), que son a su vez utilizados para obtener el propio ratio de siniestralidad, (véase ilustración 1.6.4.(b)). El segundo aspecto, la utilización del modelo estimado como generador de valores “*referencia*” para el ratio de siniestralidad del sector, (véase apartado 6.5.).

³⁹ Las variables básicas son: X_{1i} : número de siniestros acaecidos por tipo de prestación i-ésima, X_{2i} : número de pólizas emitidas por tipo de prestación i-ésima, (número de beneficiarios), X_{3i} : importe pagado por tipo de prestación i-ésima, X_{4i} : importe ingresado por tipo de prestación i-ésima, (véase apartado 5.2.1.).

⁴⁰ Véase apartado 5.2.2.

⁴¹ Dicho planteamiento supone “fijar” a priori la estructura del modelo neuronal en MLP(4:3), para poder representar correctamente las relaciones que generan el ratio de siniestralidad.



La especificación general⁴² del modelo utilizado es la siguiente,

$$R_2 = f \left(\beta_0 + \sum_{j=1}^q \beta_j g \left(\sum_{i=1}^p \phi_{ij} x_i + \phi_{0j} \right) \right)$$

$$X = (x_1, x_2, x_3, x_4)$$

$$\left\{ \phi_{ij}, i = 0, 1, \dots, p, j = 1, \dots, q \right\}$$

$$\left\{ \beta_j, j = 0, 1, \dots, q \right\}$$

siendo, “ f ” la función de salida en la capa *output*. y “ g ” la función de transferencia en la capa oculta. Si particularizamos para este caso [$p = 4; q = 3$], entonces el modelo adquiere la siguiente forma,

$$R_2 = f \left(\beta_0 + \sum_{j=1}^3 \beta_j g \left(\sum_{i=1}^4 \phi_{ij} x_i + \phi_{0j} \right) \right)$$

$$X = (x_1, x_2, x_3, x_4)$$

$$\left\{ \phi_{ij}, i = 0, 1, \dots, 4, j = 1, \dots, 3 \right\}$$

$$\left\{ \beta_j, j = 0, 1, \dots, 3 \right\}$$

Una vez planteado el modelo neuronal, nuestro interés se centra más en la importancia que poseen los valores estimados de los parámetros⁴³ que en el grado de ajuste del mismo. Tal aspecto se justifica debido a que la arquitectura neuronal está fijada de antemano, MLP(4:3) y por lo tanto la reducción del error cometido en el proceso de aprendizaje mediante cambios en su topología no es posible. Además la facilidad de adaptación del propio modelo neuronal quizás nos permita descubrir la presencia de conexiones o parámetros relevantes no previstos en la propia generación del ratio de siniestralidad.

Frente a la restricción de la arquitectura, los grados de libertad del modelo provienen de las funciones [$f; g$]. Sus diferentes combinaciones⁴⁴ nos permitirán buscar aquel modelo que mejor se ajuste a los datos, en términos de MAE y RMSE, para la base de datos de test. De las diferentes posibilidades la mejor combinación de funciones (véase tabla 2.6.4.), es:

Modelo 1: g : Gaussiana / f : Sinusoidal

⁴² Véase el capítulo 3.4.3. para un detalle mayor de las especificaciones.

⁴³ *Conexiones* en el “argot” neuronal.

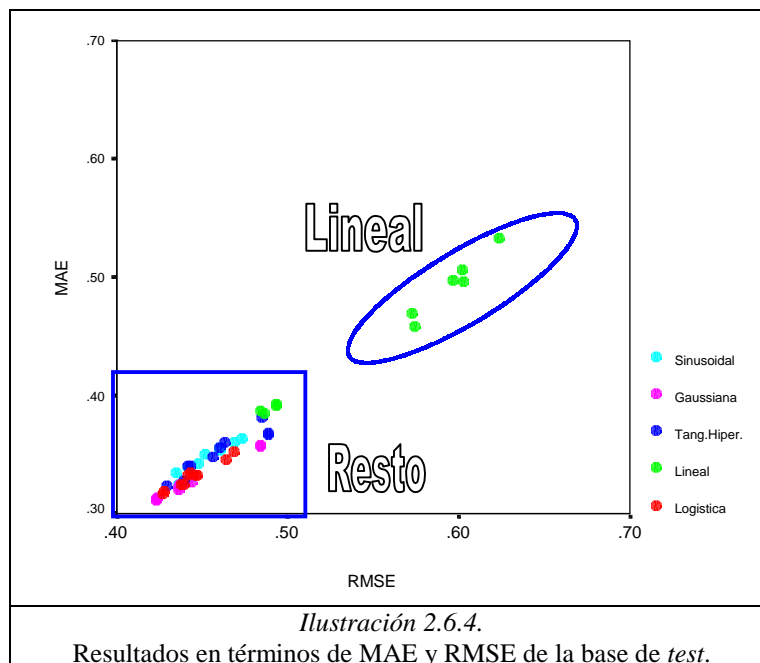
⁴⁴ Recordamos que las funciones disponibles son: *Logística, Lineal, Tangente Hiperbólica, Gaussiana y Sinusoidal*, tanto para la función de “salida” como de “transferencia”.

Tabla 2.6.4. Modelo especificado MLP(4;3) Base Test Número de iteraciones 1000

Capa oculta	Logística		Lineal		Tangente Hiperbólica		Gaussiana		Sinusoidal	
Capa output	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Logística	0.443	0.335	0.602	0.506	0.439	0.328	0.439	0.326	0.460	0.354
Lineal	0.446	0.332	0.603	0.496	0.456	0.348	0.484	0.358	0.473	0.364
Tan Hiperb	0.447	0.332	0.624	0.533	0.485	0.381	0.437	0.324	0.460	0.356
Gaussiana	0.469	0.352	0.493	0.391	0.489	0.368	0.444	0.327	0.445	0.331
Sinusoidal	0.442	0.332	0.493	0.392	0.460	0.356	0.436	0.322	0.448	0.342

Nota: Se presenta los resultados del ajuste realizado a partir de la *función de transferencia* utilizada en la capa oculta y de la *función de salida* de la capa output. En cada uno de los casos de modelos neuronales se presenta sólo los resultados para la base de datos de *validación* o test.

La combinación, por lo tanto con menor error es, *Gaussiana* y *Sinusoidal*, pero debido a los excesivos grados de libertad⁴⁵ que posee la función senoide, se ha optado por considerar como mejor alternativa⁴⁶, la combinación siguiente de menor error, es decir, función de transferencia, *Tangente hiperbólica* y función de salida, *Logística*. Observamos como la combinación de funciones $[f;g]$ de naturaleza lineal es con diferencia la peor especificación, de forma que, posee un 18% más error medido en términos de RMSE y un 33% para el MAE, (véase ilustración 2.6.4.).



⁴⁵ Para evitar problemas de memorización por parte del modelo, (véase Vladimir Vapnik. (2000). **The Nature of Statistical Learning Theory**, Springer, pp. 82.).

⁴⁶ Ésta última combinación será la escogida para representar el modelo de ratio R_2 en el apartado 6.5.

Una vez decididas las dos combinaciones de funciones $[f; g]$ que mejor se ajustan en términos de MAE y RMSE,

Modelo 1: g : **Gaussiana** / f : **Sinusoidal**

Modelo 2: g : **Tangente Hiperbólica** / f : **Logística**

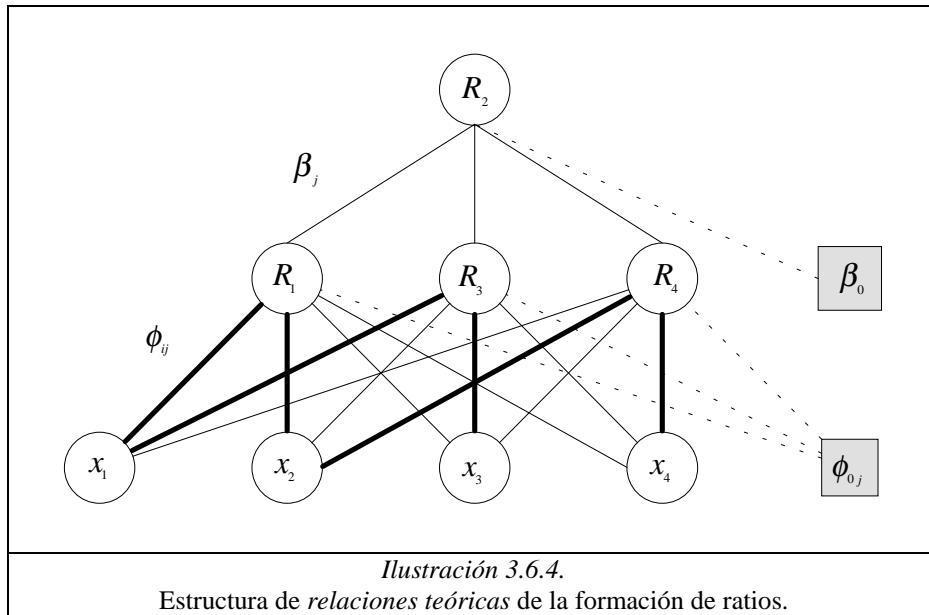
las incorporaremos en la propia especificación del modelo neuronal escogido, (MLP(4:3)), con los objetivos siguientes. En primer lugar, comprobar el grado de aprendizaje⁴⁷, por parte de los modelos, de las relaciones que se forman entre los *inputs* (variables básicas) respecto al ratio de siniestralidad. En segundo lugar, detectar procesos de memorización⁴⁸ al aumentar el número de iteraciones de 1000 a 10000. En tercer lugar, comprobar la naturaleza de los resultados obtenidos en la capa oculta y en último lugar, analizar los residuos de los modelos estimados.

Para poder desarrollar los dos primeros aspectos, debemos definir una tipología teórica de conexiones, que nos represente las relaciones existentes entre las variables básicas y el propio ratio de siniestralidad. El proceso de aprendizaje de los modelos neuronales debería ser capaz de “retener” dicha estructura mediante los parámetros estimados. La ilustración 3.6.4. representa la tipología “modelo” definida, donde las conexiones con trazos gruesos representan las relaciones que realmente deberían de existir entre las variables básicas para generar los *ratios intermedios* (capa oculta) y que posteriormente éstos generan el *ratio de siniestralidad*, R_2 .

Una vez definido el modelo teórico de relaciones, procedemos a estimar el primero de los modelos con las funciones siguientes, g : **Gaussiana** / f : **Sinusoidal**, los resultados de los parámetros o conexiones estimados se presentan en forma matricial en la tabla 3.6.4. El análisis de sus valores nos permitirá conocer la estructura que ha “retenido” el modelo neuronal propuesto.

⁴⁷ Dicho aprendizaje se detecta mediante el valor de los parámetros estimados o conexiones, valores altos de los mismos son indicativo de una mayor importancia o ponderación.

⁴⁸ En el entorno neuronal se entiende por procesos de *memorización*, un excesivo nivel de iteraciones provocando un *sobreaprendizaje* en el modelo. Una forma visual de evitarlo es detener el proceso de aprendizaje cuando a través de la curva de error cometido, observamos que después de un fuerte proceso de reducción del mismo, se estabiliza.



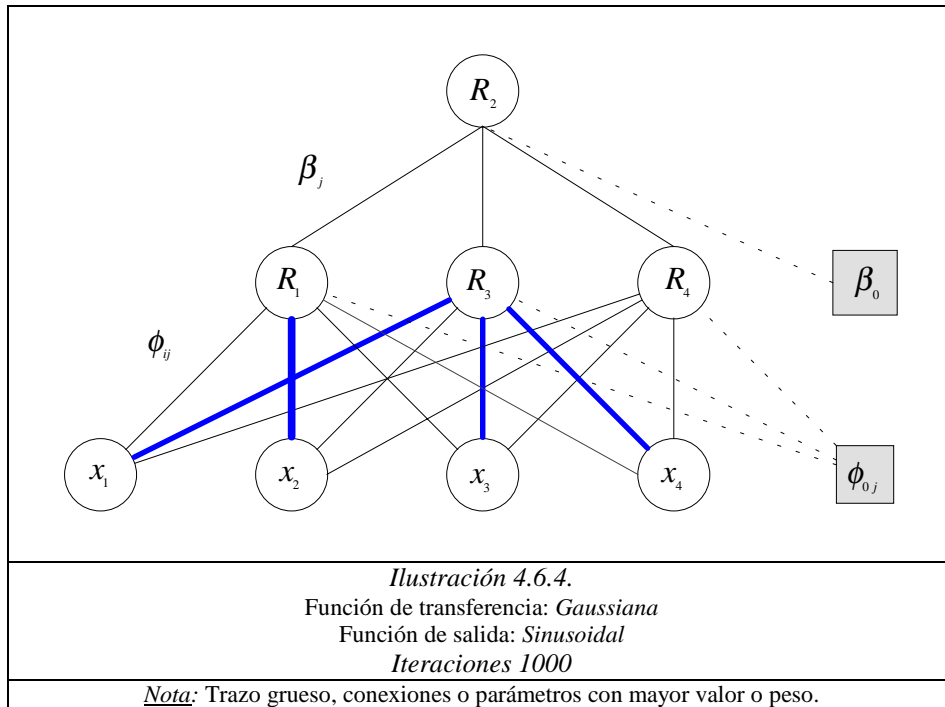
La importancia de estos parámetros estimados nos determinará hasta que punto el modelo replica las relaciones teóricas mediante el estudio de los diferentes casos existentes en la muestra de ratios. Así podemos comprobar que de los parámetros ϕ_{ij} , los de mayor peso⁴⁹ o valor son con diferencia, ϕ_{12} , ϕ_{21} , ϕ_{32} y ϕ_{42} , que debidamente representados con un trazo grueso, no indicará que tipo de relaciones ha retenido o aprendido⁵⁰, (véase ilustración 4.6.4). Si comparamos las dos ilustraciones, 3.6.4. y 4.6.4. podemos observar que de las 6 conexiones relevantes (trazo grueso), el modelo sólo identifica correctamente el 50%, es decir, tres conexiones relevantes coinciden entre ambos.

Tabla 3.6.4.

<i>Modelo MLP (p,q)</i>	
Función de transferencia: <i>Gaussiana</i>	
Función de salida: <i>Sinusoidal</i>	
	<i>MLP (4,3)</i> <i>Base de aprendizaje (n=641)</i> <i>Número de iteraciones 1000</i>
ϕ_{ij}	$\begin{pmatrix} \phi_{01} & \phi_{11} & \phi_{21} & \phi_{31} & \phi_{41} \\ \phi_{02} & \phi_{12} & \phi_{22} & \phi_{32} & \phi_{42} \\ \phi_{03} & \phi_{13} & \phi_{23} & \phi_{33} & \phi_{43} \end{pmatrix} = \begin{pmatrix} -0.40 & 0.08 & 1.15 & 0.05 & 0.11 \\ -0.58 & -0.57 & -0.09 & -0.82 & -0.61 \\ -0.20 & -0.03 & 0.35 & 0.11 & -0.22 \end{pmatrix}$
β_j	$(\beta_0 \ \beta_1 \ \beta_2 \ \beta_3) = (-0.29 \ -0.25 \ -0.08 \ -0.21)$

⁴⁹ Dos aspectos, los pesos se valoran en valor absoluto y no se ha considerado los parámetros del término independiente.

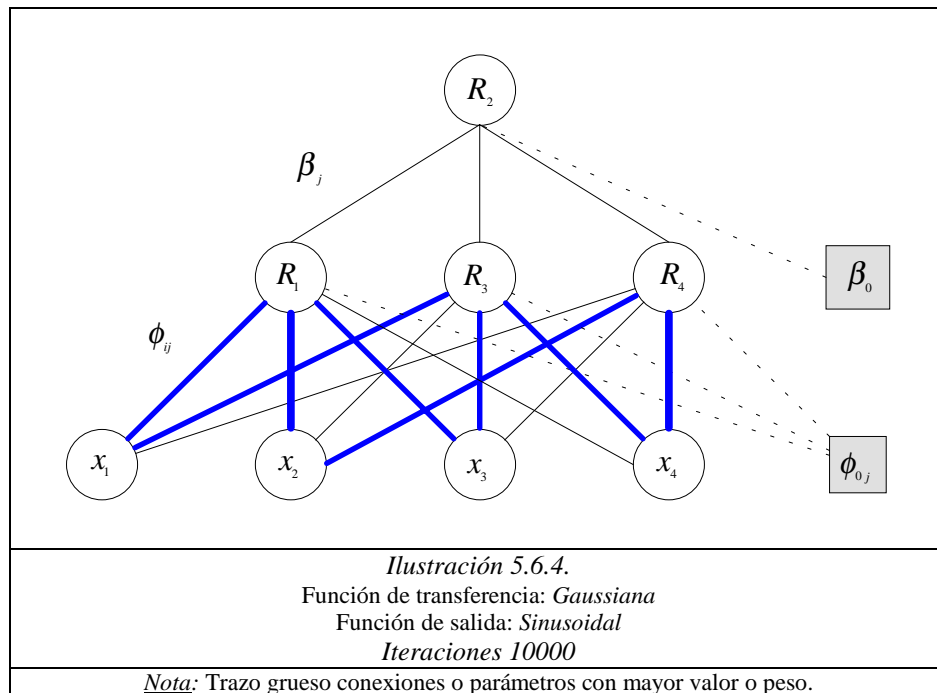
⁵⁰ Utilizamos la habilidad de los modelos neuronales para retener “estructuras”, en nuestro caso de relaciones.



Para poder comprobar el efecto “memorización” versus “generalización”, se ha incrementado el proceso de estimación de 1000 a 10000 iteraciones sobre el modelo anterior. Los resultados de los parámetros estimados son evidentes, (véase tabla 4.6.4.): el número de conexiones con mayor peso se ha incrementado, siendo en este caso los parámetros con mayores valores, $\phi_{11}, \phi_{12}, \phi_{21}, \phi_{23}, \phi_{31}, \phi_{32}, \phi_{42}, \phi_{43}$. El proceso de reconocimiento de relaciones estimado es menos claro, ya que los valores de los parámetros han aumentado en su mayoría, provocando una mayor dificultad para discernir las conexiones más relevantes. Este aspecto esta representado en la ilustración 5.6.4.

Tabla 4.6.4.

Modelo MLP (p,q)	
Función de transferencia: <i>Gaussiana</i>	
Función de salida: <i>Sinusoidal</i>	
	MLP (4,3)
	Base de aprendizaje (n=641)
	Número de iteraciones 10000
ϕ_{ij}	$\begin{pmatrix} \phi_{01} & \phi_{11} & \phi_{21} & \phi_{31} & \phi_{41} \\ \phi_{02} & \phi_{12} & \phi_{22} & \phi_{32} & \phi_{42} \\ \phi_{03} & \phi_{13} & \phi_{23} & \phi_{33} & \phi_{43} \end{pmatrix} = \begin{pmatrix} 2.35 & 2.84 & 8.70 & -30.52 & -2.12 \\ -2.08 & -2.06 & 0.80 & -17.03 & 7.47 \\ 1.73 & 1.90 & -31.14 & 3.08 & -5.21 \end{pmatrix}$
β_j	$(\beta_0 \ \beta_1 \ \beta_2 \ \beta_3) = (-0.25 \ -0.21 \ -38.96 \ -0.38)$



De todos modos después del comentario sobre la función sinusoidal, se ha realizado el mismo proceso de comprobación para el Modelo 2 con las funciones, g : **Tangente Hiperbólica** / f : **Logística**, para 1000 y 10000 iteraciones, (véase tabla 5.6.4.).

Las conclusiones de la comparación entre los dos modelos, son las siguientes. En primer lugar, la combinación de funciones, g : **Tangente Hiperbólica** / f : **Logística** (Modelo 2), no retiene un mayor número de relaciones o conexiones “teóricas” que el Modelo 1, g : **Gaussiana** / f : **Sinusoidal**, pero la matriz estimada gana “nitidez” en la detección de las mismas. En segundo lugar, el Modelo 2 gana en estabilidad ya que al aumentar el número de iteraciones mantiene constante los parámetros relevantes. En último lugar, es posible que la rígida topología impuesta no permita al modelo replicar mejor el modelo teórico. Una solución sería proponer un modelo ampliado.

Tabla 5.6.4.

Modelo MLP (p,q)		
Función de transferencia: <i>tangente hiperbólica</i>		
Función de salida: <i>logística</i>		
	MLP (4,3)	Base de aprendizaje (n=641)
ϕ_{ij}		Número de iteraciones 1000
		$\begin{pmatrix} \phi_{01} & \phi_{11} & \phi_{21} & \phi_{31} & \phi_{41} \\ \phi_{02} & \phi_{12} & \phi_{22} & \phi_{32} & \phi_{42} \\ \phi_{03} & \phi_{13} & \phi_{23} & \phi_{33} & \phi_{43} \end{pmatrix} = \begin{pmatrix} 0.01 & 0.48 & -0.14 & 0.45 & 0.34 \\ 1.06 & 1.07 & 2.03 & 0.83 & 1.03 \\ 0.05 & 0.22 & 0.26 & 0.45 & 0.04 \end{pmatrix}$
β_j		$(\beta_0 \ \beta_1 \ \beta_2 \ \beta_3) = (0.24 \ 0.28 \ -1.88 \ 0.27)$
	MLP (4,3)	Base de aprendizaje (n=641)
ϕ_{ij}		Número de iteraciones 10000
		$\begin{pmatrix} \phi_{01} & \phi_{11} & \phi_{21} & \phi_{31} & \phi_{41} \\ \phi_{02} & \phi_{12} & \phi_{22} & \phi_{32} & \phi_{42} \\ \phi_{03} & \phi_{13} & \phi_{23} & \phi_{33} & \phi_{43} \end{pmatrix} = \begin{pmatrix} 0.04 & 0.53 & -0.22 & 0.49 & 0.39 \\ 0.94 & 0.95 & 3.76 & 0.70 & 0.91 \\ 0.09 & 0.26 & 0.24 & 0.49 & 0.06 \end{pmatrix}$
β_j		$(\beta_0 \ \beta_1 \ \beta_2 \ \beta_3) = (0.36 \ 0.41 \ -2.10 \ 0.27)$

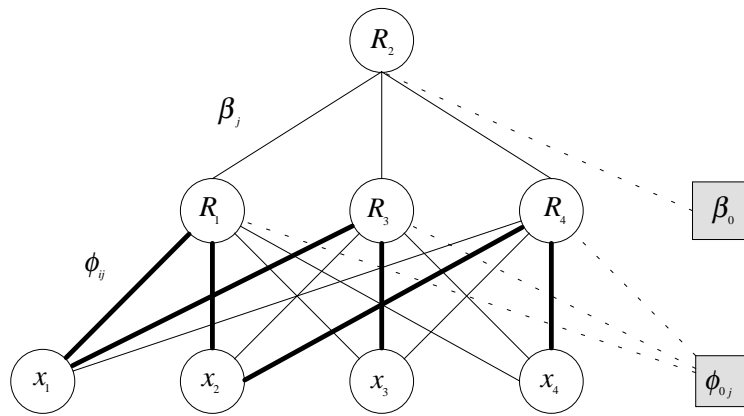


Ilustración 6.6.4.(a)

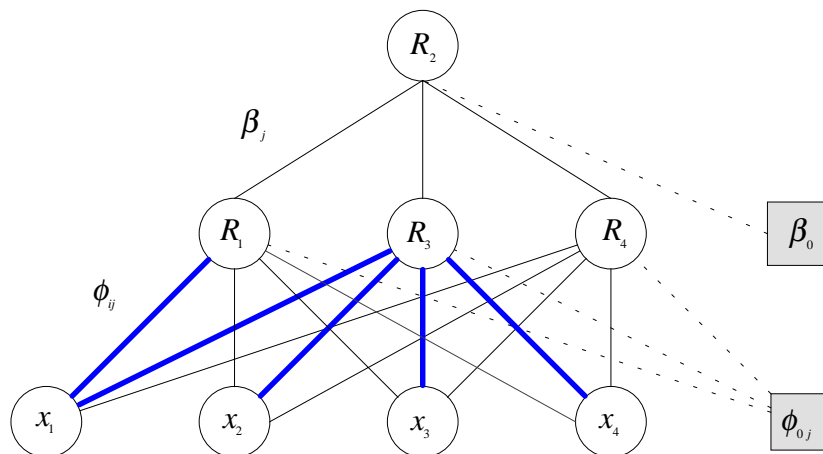


Ilustración 6.6.4.(b) Iteraciones 1000 y 10000.

Nota: La ilustración 6.6.4.(a) representa la relación teórica de los pesos o conexiones en función de la relación de los ratios y la ilustración 6.6.4.(b) esquematiza la *matriz* de pesos estimada por el modelo neuronal MLP(4:3). Trazo grueso conexiones o parámetros con mayor valor o peso.

Una de las posibles soluciones para intentar mejorar los resultados anteriores, es proponer un topología de modelo neuronal con el número de neuronas⁵¹ en la capa oculta aumentado⁵². Es decir, una topología **MLP(4:9)**, donde la mejor combinación de funciones es, g : **Gaussiana** / f : **Logística**, (véase tabla 6.6.4.).

Tabla 6.6.4. Modelo especificado **MLP(4:9)** Base de Test Número de iteraciones 1000

Capa oculta	Logística		Lineal		Tangente Hiperbólica		Gaussiana		Sinusoidal	
Capa output	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Logística	0.446	0.338	0.467	0.356	0.463	0.354	0.435	0.319	0.464	0.353
Lineal	0.452	0.345	0.545	0.445	0.553	0.456	0.526	0.390	0.549	0.450
Tang Hiperb	0.480	0.376	0.492	0.384	0.488	0.381	0.437	0.325	0.490	0.383
Gaussiana	0.454	0.341	0.500	0.398	0.570	0.471	0.442	0.325	0.531	0.429
Sinusoidal	0.477	0.373	0.519	0.400	0.539	0.422	0.449	0.332	0.527	0.409

Nota: Se presenta los resultados del ajuste realizado a partir de la *función de transferencia* utilizada en la capa oculta y de la *función de salida* de la capa output. En cada uno de los casos de modelos neuronales se presenta sólo los resultados para la base de datos de *validación* o test.

Las conclusiones que se obtienen comparando el modelo ampliado, MLP(4:9), respecto el modelo original, MLP(4:3) son las siguientes. Las combinaciones no son estables, la función de salida permanece invariable en todos los casos y en último lugar, la ganancia en términos de RMSE y MAE es mínima. De todos modos debemos constatar si el incremento de la topología ha permitido que el nuevo modelo neuronal sea más receptivo en el proceso de detección de las relaciones teóricas que configuran el ratio de siniestralidad definido. Dos son los aspectos relevantes. En primer lugar los valores altos de los parámetros estimados o conexiones, se concentran alrededor de tres neuronas, 3, 8 y 7. Dichas neuronas representarían a los tres ratios de gestión intermedios ($R_1; R_3; R_4$). En segundo lugar, el nuevo modelo detecta un mayor número de conexiones relevantes comparativamente con el modelo “teórico”, indicador de las fuertes relaciones entre las variables básicas, (véase ilustración 7.6.4. comparativamente con la ilustración 3.6.4.). En último lugar, si observamos en la ilustración 8.6.4. los “*density plots*” de los valores que se han generado en cada nodo de la capa oculta, comprobamos que el nodo 3 representa la gran dispersión que posee el ratio de frecuencia de siniestralidad (R1), el nodo 7 representa los ingresos medios (R4) y el nodo 8 los costes

⁵¹ El número de neuronas en la capa oculta se ha determinado mediante la expresión heurística, “ $2n+1$ ”, siendo “ n ” el número de *inputs*.

⁵² La decisión de aumentar el número de neuronas en la capa oculta esta fundamentada sobre la idea de aumentar los grados de libertad y nos permitirá comprobar si en la capa oculta puede formarse ratios.

medios (R_3). Ambos son muy parecidos en su forma, pero el segundo es algo menor que el primero, como era de esperar.

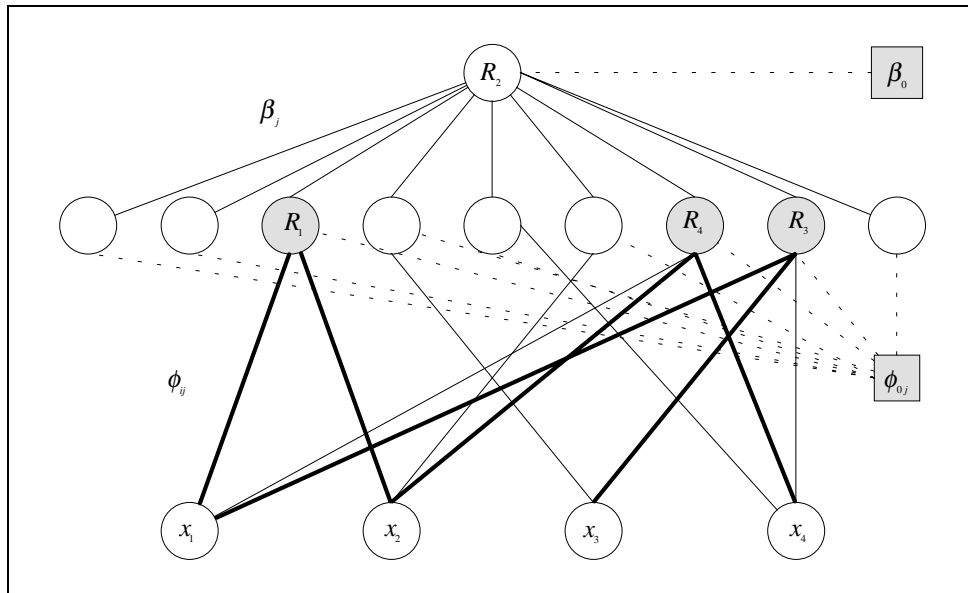


Ilustración 7.6.4.

Función de transferencia: *Gaussiana*, Función de salida: *Logística*, Iteraciones 1000.

Nota: Trazo grueso representan aquellas conexiones detectadas como relevantes por el modelo que además coinciden con las relaciones "teóricas".

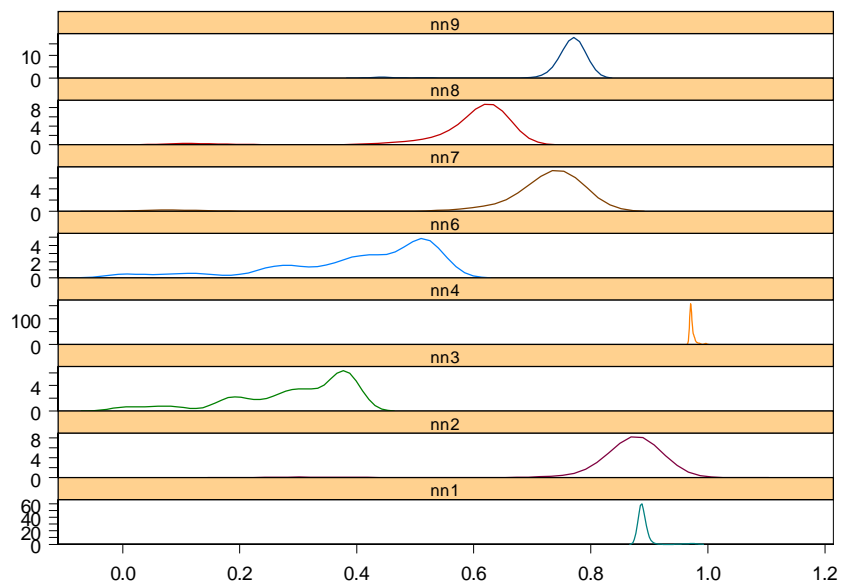


Ilustración 8.6.4.

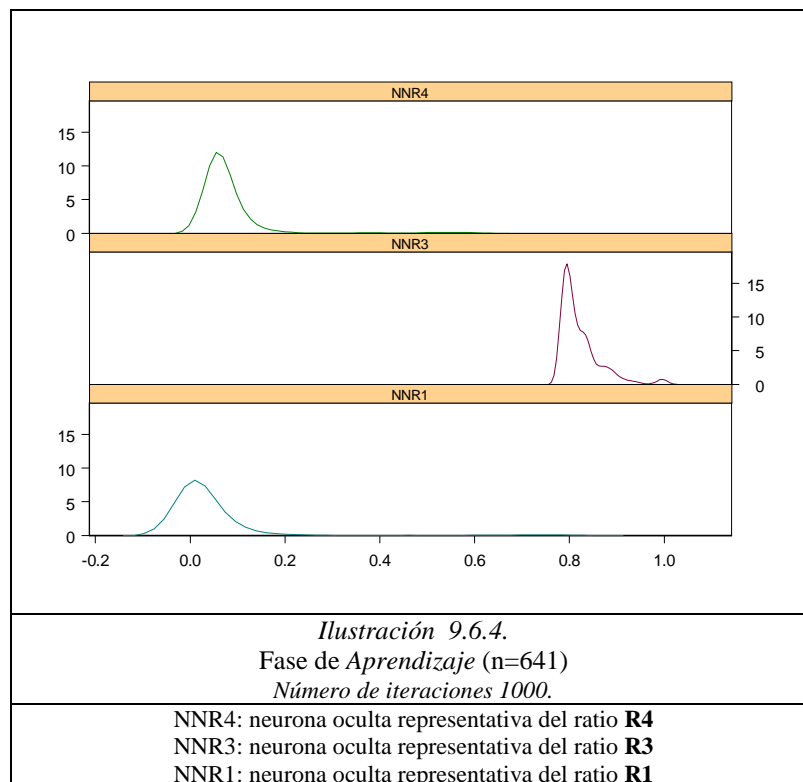
Density Plots de los valores obtenidos en los nodos de la capa oculta.

Fase de Aprendizaje (n=641)

nn3: Ratio de Siniestralidad (R_1); nn7: Ingreso medio (R_4); nn8: Coste medio (R_3)

Nota: No se incorpora el nodo 5 ya que sus valores son muy próximos a cero.

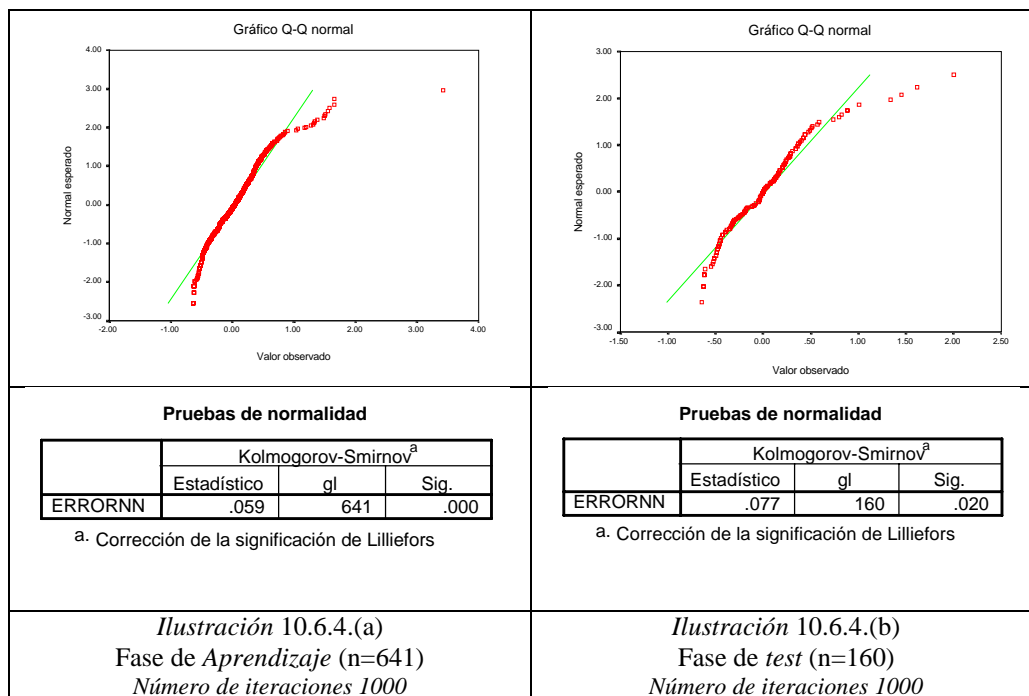
El tercer aspecto considerado, consiste en comprobar la naturaleza de los resultados obtenidos en la capa oculta. Así consideramos de interés desvelar, en parte el comportamiento que se sucede en el interior de las neuronas de la capa oculta de los modelos neuronales, que en nuestro caso supone utilizarlos para generar ratios, a partir del modelo sin ampliar MLP(4:3). Los valores obtenidos coinciden con los ratios de gestión intermedios ($R_1; R_3; R_4$) que de forma indirecta nos permitirán a su vez obtener el propio ratio de siniestralidad. La ilustración 9.6.4. muestra los valores⁵³ que se han generado en el interior de las neuronas de la capa oculta, representativas de los ratios de gestión, para la fase de aprendizaje⁵⁴. Observamos dos aspectos, el primero de ellos es que tanto NNR3 como NNR4, representantes del coste medio por siniestro e ingreso medio por póliza, están en zonas opuestas, de la misma forma que sus valores originales (mayores valores para el coste medio que para el ingreso medio). El segundo de ellos es el comportamiento menos *Gaussiano* del NNR3 frente al NNR4.



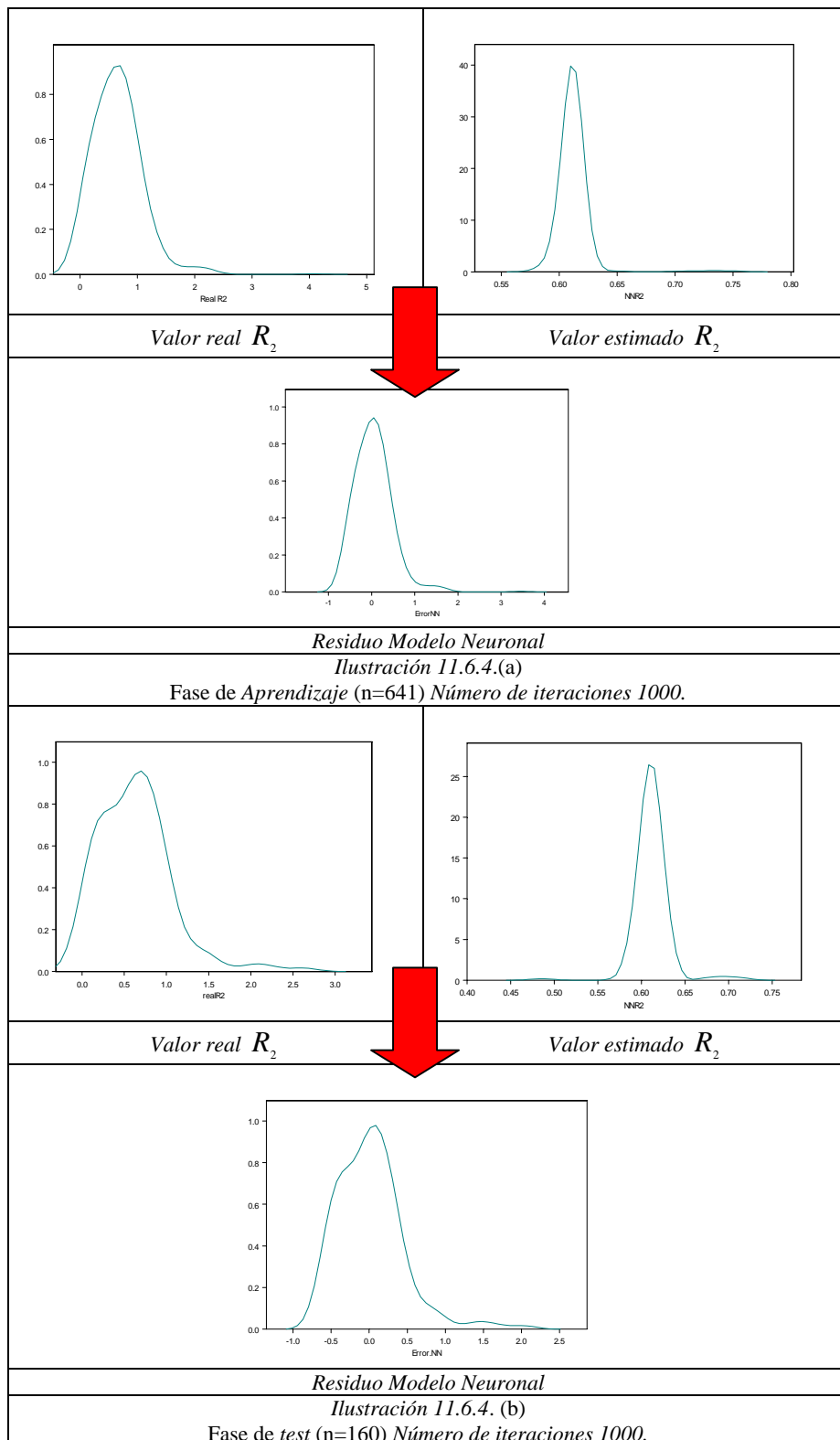
⁵³ Estos valores han sido filtrados por la función *tangente hiperbólica*. Si quisiéramos comprobar su similitud con los valores originales de los ratios deberíamos realizar el proceso inverso de filtrado.

⁵⁴ Los resultados gráficos para la fase de test son muy similares, por esta razón no se incorporan.

En último lugar, no debemos olvidar el análisis de los residuos del modelo MLP(4:3)), de la misma forma que se realiza en los modelos clásicos de regresión. El comportamiento de los residuos será un pieza básica en el apartado 6.5, ya que consideramos que puede poseer información relevante a tener en cuenta. La ilustración 10.6.4. (a y b) muestra los gráficos Normal Q-Q Plot y el resultado del contraste no paramétrico de normalidad Kolmogorov-Smirnov con la adaptación de *Lilliefors*. En ambos casos, aprendizaje y test, rechazamos que el residuo del modelo se ajuste a un modelo normal, lo cual no nos sorprende en el entorno neuronal.



Adicionalmente la ilustración 11.6.4. (a y b) nos muestra: la forma de los valores “ajustados” del modelo, los valores “objetivo” y el “residuo” o error cometido, tanto en la base de *aprendizaje* como en la de validación o *test*. Observamos la naturaleza no Gaussiana de los residuos del modelo y la presencia de colas largas. También se ha realizado un contraste de medias sobre los residuos de ambas bases de datos (*aprendizaje* y validación o *test*) con la finalidad de comprobar la existencia o no de diferencias significativas, en éste caso no existen diferencias (con igualdad de varianzas (*p-value*: 0.506) y sin igualdad (*p-value*: 0.525)). Este aspecto se considerará en el momento que se generen los valores “referencia” de la siniestralidad de las entidades en el apartado 6.5.



Los resultados obtenidos nos invitan a considerar a los modelos neuronales como una metodología econométrica a tener en presente para representar la forma en que se pueden formar los ratios.

6.5. Diseño de un *benchmark sectorial* para la siniestralidad *no vida* mediante un Modelo econométrico neuronal (MLP(4:3)).

6.5.1. Propuesta de *scoring* para el sector de mutualidades de Cataluña.

En los apartados anteriores se ha comprobado que es posible que el residuo posea información relevante para el diagnóstico mediante ratios, tanto para las especificaciones realizadas del modelo de ratio, clásica o neuronal. En este apartado proponemos, en la línea del trabajo de Trigueiros, D.⁵⁵ (1994), utilizar el residuo del modelo **MLP(4:3)** como *input* para generar la posición relativa de las entidades respecto al estándar del mercado⁵⁶ o *benchmark*.

La metodología elaborado por Trigueiros dota a los investigadores de información económico-financiera de un nuevo instrumento de análisis, el gráfico residual rotado⁵⁷ (RRP), el cual consiste en un gráfico de dispersión en donde los ejes (en escala logarítmica) representan la desviación de cada uno de los componentes del ratio estudiado respecto al promedio del sector⁵⁸. Dicho aspecto permitiría una automatización de los análisis financieros más completa, utilizando modelo autoorganizativos tipo SOM⁵⁹. En nuestro caso y debido a la especificidad del ratio estudiado, R_2 , consideramos que una aproximación más sencilla al problema es factible, lógicamente sin desdeñar una ampliación en la línea de Trigueiros en futuras investigaciones con la base de datos del sector estudiado. Nuestra aportación consiste en utilizar como herramienta de posición de las entidades respecto al sector, el residuo generado por el propio modelo **MLP(4:3)** que representa su estructura. De esta forma consideramos que el residuo representa una medida flexible de la desviación de la empresa respecto el sector, definido mediante un modelo neuronal.

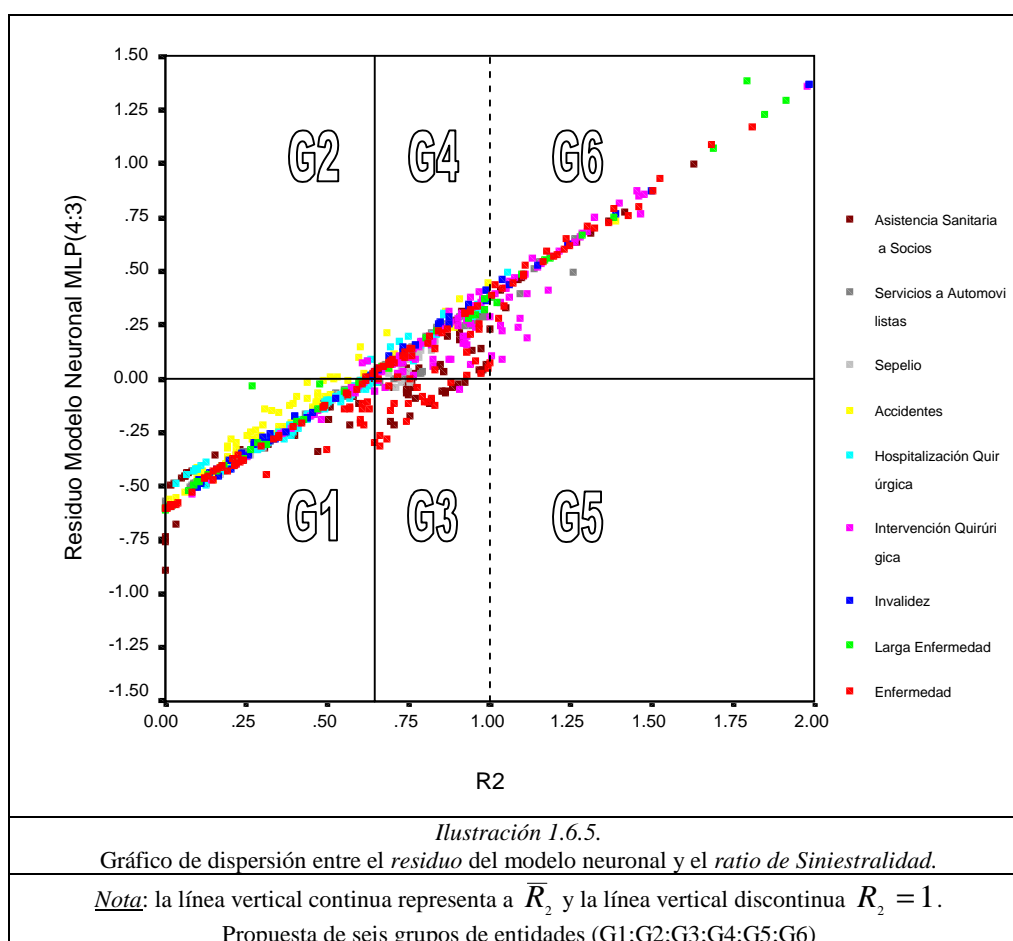
⁵⁵ Véase Trigueiros, D. (1994). **Incorporating Complementary Ratios in The Analysis of Financial Statements**, *Accting., Mgmt. & Info. Tech.*, Vol. 4, No. 3, pp. 149-162. Para aspectos más generales véase Trigueiros, D. y Berry, Taffler, R. (1996). **Neural Networks and Empirical Research in Accounting**, *Accounting and Business Research* Vol. 26, No. 26, pp. 347-355.

⁵⁶ Recordemos que en el apartado 6.3. definimos como posible generador de ratios de siniestralidad, la forma funcional MLP(4:3).

⁵⁷ En inglés, *rotated residual plot* (RRP).

⁵⁸ Generando información sobre la posición estratégica de las compañías. Consideramos que, según los resultados obtenidos en el apartado 6.4., el modelo neuronal **MLP(4:3)** puede representar al sector y por lo tanto sus *outputs* serán tratados como valores estándares. Las desviaciones sobre estos estándares nos indicarán medidas de posición relativa de las entidades.

⁵⁹ Véase en el apartado 2.5.3.



La ilustración 1.6.5. nos muestra un gráfico de dispersión entre el residuo⁶⁰ del modelo neuronal (MLP(4:3)) y el propio ratio de siniestralidad R_2 , desglosado por tipo de prestación. Podemos observar que es posible generar seis grupos o cluster, a partir de los siguientes valores “referencia”: el valor medio del ratio de siniestralidad, el umbral unitario del mismo y el signo del residuo del modelo neuronal. Cada entidad se posicionará en cada uno de estos segmentos y se les asignará un valor ponderador para conseguir una medida de *scoring* compacta, véase tabla 1.6.5. El valor del *scoring* definido para cada entidad tendrá en cuenta la posibilidad de que las mutualidades de previsión social puedan estar en más de un ramo.

⁶⁰ Tal y como se justificó en el apartado anterior, se ha obtenido el “residuo” para la base total de observaciones (aprendizaje y test), mediante la aplicación del modelo validado MLP(4:3) (función de transferencia: tangente hiperbólica y función de salida: logística), permitiendo el cálculo posterior de un *scoring* por entidad y prestación.

Tabla 1.6.5.

Clusters	1 Condición	2 Condición	Ponderación
G1	$R_2 \leq 0.647$	$e_i \leq 0$	1
G2	$R_2 \leq 0.647$	$e_i > 0$	2
G3	$0.647 \leq R_2 \leq 1$	$e_i \leq 0$	3
G4	$0.647 \leq R_2 \leq 1$	$e_i > 0$	4
G5	$R_2 > 1$	$e_i \leq 0$	5
G6	$R_2 > 1$	$e_i > 0$	6

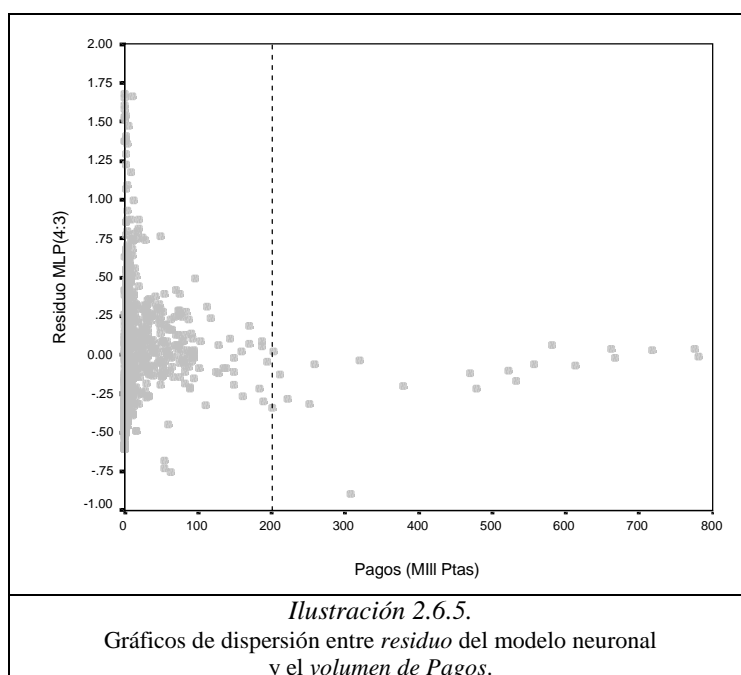
Nota: La referencia para la primera condición puede utilizarse tanto el valor promedio, mediano o cualquier otra medida descriptiva; “e” es el residuo del modelo neuronal.

Las características de cada uno de los cluster es la siguiente:

- Grupo 1: Entidades que poseen, en primer lugar, un ratio de siniestralidad menor que la media del sector. En segundo lugar, el valor de su ratio respecto al sector posee un valor más pequeño del que realmente le correspondería según el modelo neuronal del sector, es decir, su desviación es negativa y además poseen un valor menor que el umbral. A las entidades posicionadas en este cuadrante les asociamos la menor de las penalizaciones, **1**.
- Grupo 2: entidades con un ratio de siniestralidad menor que la media del sector y un valor del mismo superior al que realmente le correspondería según el modelo neuronal del sector, es decir, una desviación es positiva. A las entidades posicionadas en este cuadrante le asociamos la penalización, **2**.
- Grupo 3: Entidades que poseen un ratio de siniestralidad superior a la media del sector pero sin superar el umbral crítico de 1 (el volumen de ingresos se destinaría de forma completa para satisfacer el pago de prestaciones). Además el valor de su ratio respecto al sector posee un valor más pequeño del que realmente le correspondería, es decir, su desviación es negativa. A las entidades posicionadas en este cuadrante le asociamos la penalización, **3**.
- Grupo 4: Entidades con un ratio de siniestralidad superior a la media del sector pero sin superar el umbral crítico de 1 (el volumen de ingresos se destinaría de forma completa para satisfacer el pago de prestaciones). Adicionalmente poseen un valor del ratio respecto al sector superior al que realmente le correspondería según el *output* modelo neuronal, es decir, su desviación es positiva. A las entidades posicionadas en este cuadrante le asociamos la penalización, **4**.

- **Grupo 5:** Entidades que poseen, en primer lugar, un ratio de siniestralidad superior al umbral crítico de 1 (el volumen de ingresos se destinaría de forma completa para satisfacer el pago de prestaciones). En segundo lugar, el valor de su ratio respecto al sector posee un valor más pequeño del que realmente le correspondería según el modelo neuronal del sector, es decir, su desviación es negativa. A las entidades posicionadas en este cuadrante⁶¹ le asociamos la penalización, **5**.
- **Grupo 6:** En último lugar, entidades que poseen, un ratio de siniestralidad superior al umbral crítico de 1 (el volumen de ingresos se destinaría de forma completa para satisfacer el pago de prestaciones), el valor de su ratio respecto al sector es superior al que realmente le correspondería según el modelo propuesto, es decir, su desviación es positiva. A las entidades posicionadas en este cuadrante le asociamos la penalización máxima, **6**.

Adicionalmente y antes de presentar los resultados, consideramos importante plantear la existencia de algún comportamiento sistemático entre el residuo del modelo neuronal y el volumen de riesgo asumido por las entidades, es decir, el volumen anual de pagos, (véase ilustración 2.6.5.).



⁶¹ En este grupo no existe ninguna entidad posicionada.

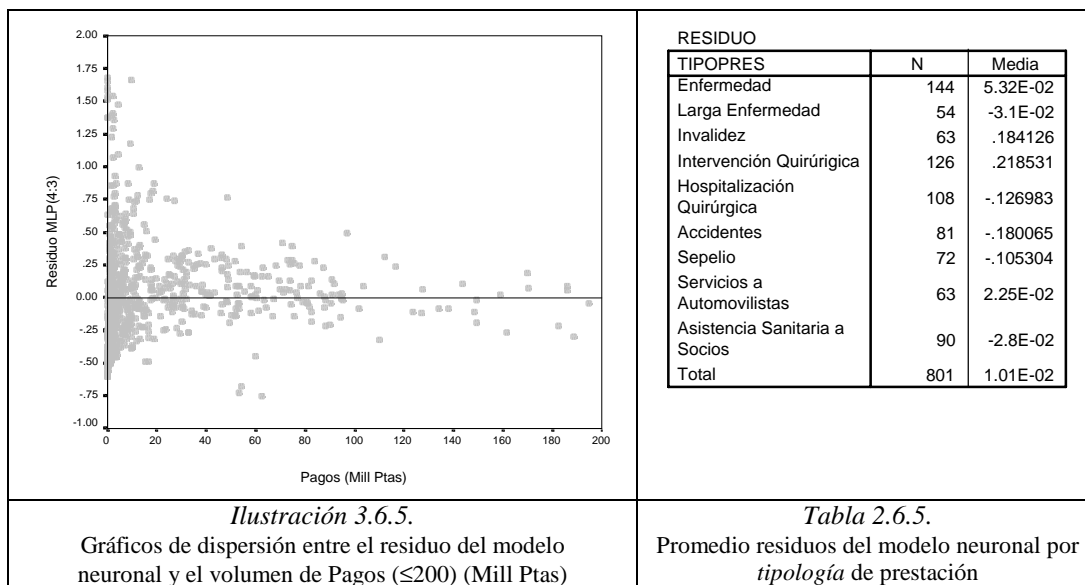


Ilustración 3.6.5.

Gráficos de dispersión entre el residuo del modelo neuronal y el volumen de Pagos (≤ 200) (Mill Ptas)

Tabla 2.6.5.

Promedio residuos del modelo neuronal por tipología de prestación

Las conclusiones obtenidas son las siguientes. El comportamiento diferencial entre entidades, en términos de siniestralidad y en función de su tamaño, se refleja en la evolución del residuo del modelo neuronal. A mayor tamaño su comportamiento converge hacia la “norma” del sector, es decir, el residuo tiende a cero, (véase ilustración 3.6.5.). Las prestaciones que poseen un valor por debajo del sector no vida (residuo negativo) son, Hospitalización Quirúrgica, Accidentes, Sepelio, Asistencia Sanitaria a Socios y Larga Enfermedad. En cambio las prestaciones restantes, Enfermedad, Invalidez, Intervención Quirúrgica, Servicios a Automovilistas, poseen una siniestralidad por encima del “sector” representado por el modelo neuronal (residuo positivo), (véase tabla 2.6.5.).

La metodología anterior de penalización permite obtener para cada una de las entidades estudiadas una puntuación compacta y relativa que tenga en cuenta la multiprestación⁶². Las entidades con los identificadores 17, 22 y 25 son las que poseen una peor situación en términos de *scoring*, es decir, a mayor valor, mayor penalización por un peor posicionamiento, sin distinguir ni por ejercicio económico ni por tipo de prestación, (véase ilustración 4.6.5). Consideramos también de utilidad realizar el mismo estudio por ejercicio económico. Así a partir del ejercicio económico del 1993 aparecen de forma clara dos grupos de entidades que empeoran su comportamiento relativo respecto al resto del sector y posteriormente se mantiene dicho comportamiento relativo, (véase ilustración 5.6.5).

⁶² Recordamos que el número de entidades que poseen información homogénea durante todo el período analizado es de 36.

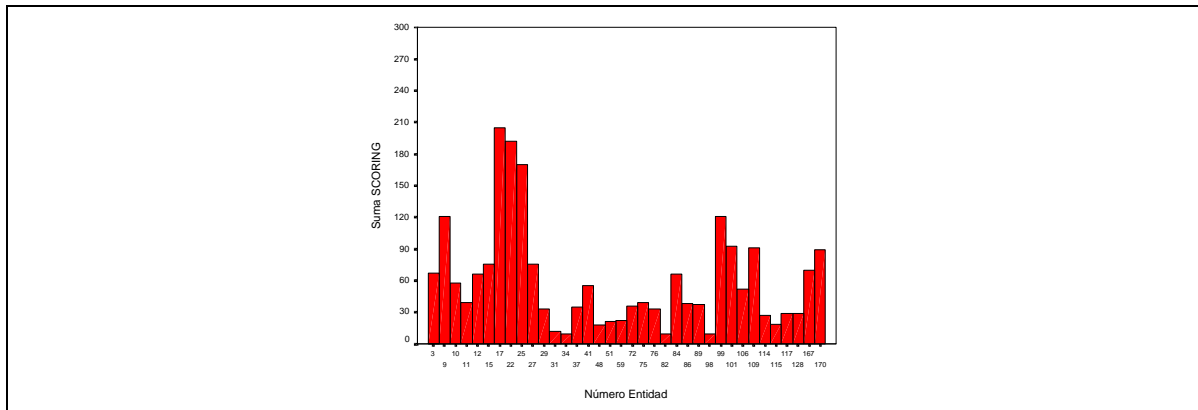


Ilustración 4.6.5.
Puntuaciones agregadas de las diferentes entidades (36)

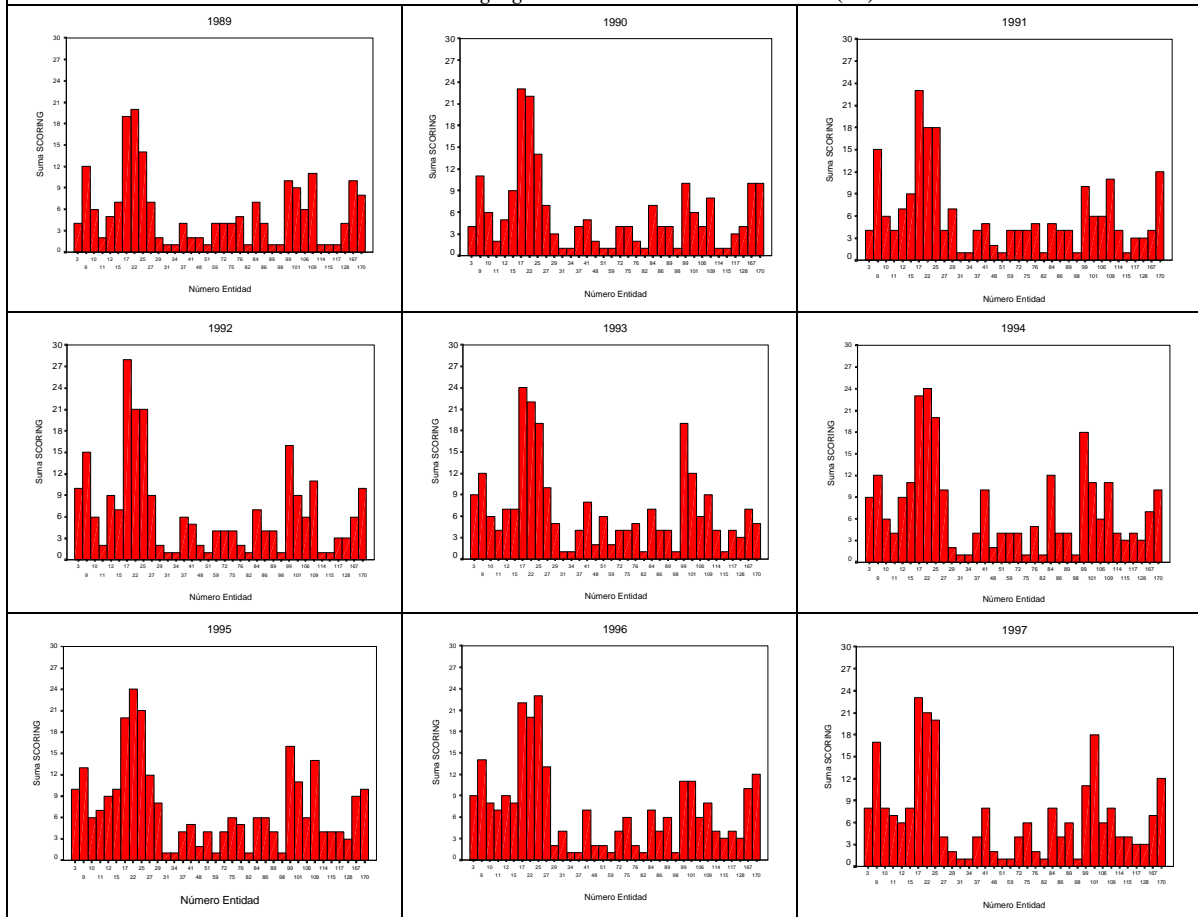
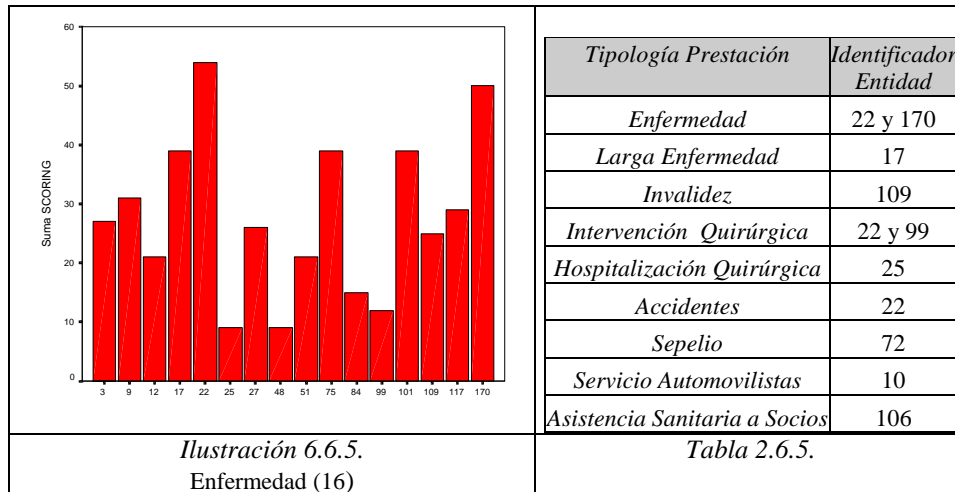


Ilustración 5.6.5.
Puntuaciones de las diferentes entidades (36) desglosadas por ejercicio económico.

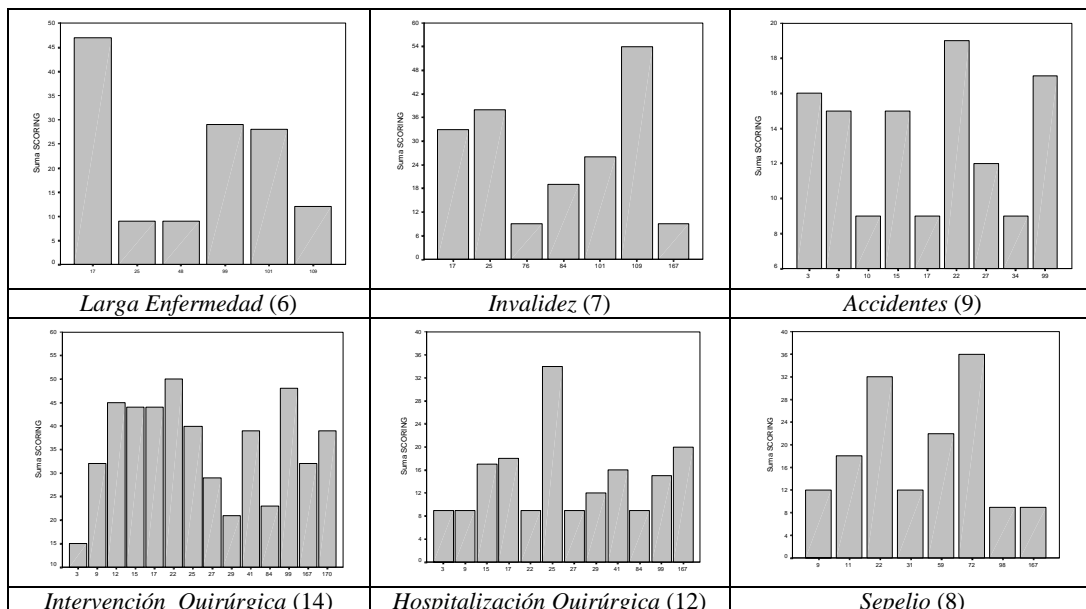
Nota: Valores altos del histograma significa peor posición respecto al sector.

El control que permite dicha metodología debe contemplar el diagnóstico por tipo de prestación. Por ejemplo, la ilustración 6.6.5. presenta los resultados para la prestación de *Enfermedad*, donde observamos que las entidades con un peor *scoring*, indicando un peor

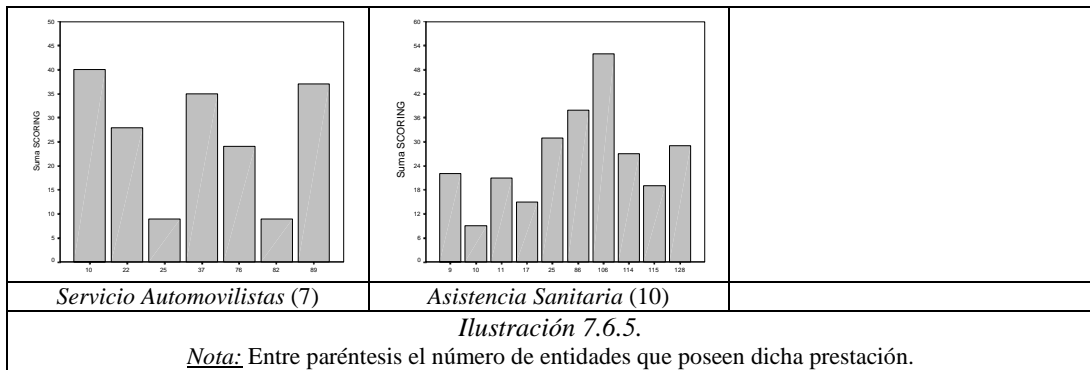
posicionamiento, son las que poseen el identificador 22 y 170 del total de las 16 entidades que ofrecen dicha prestación.



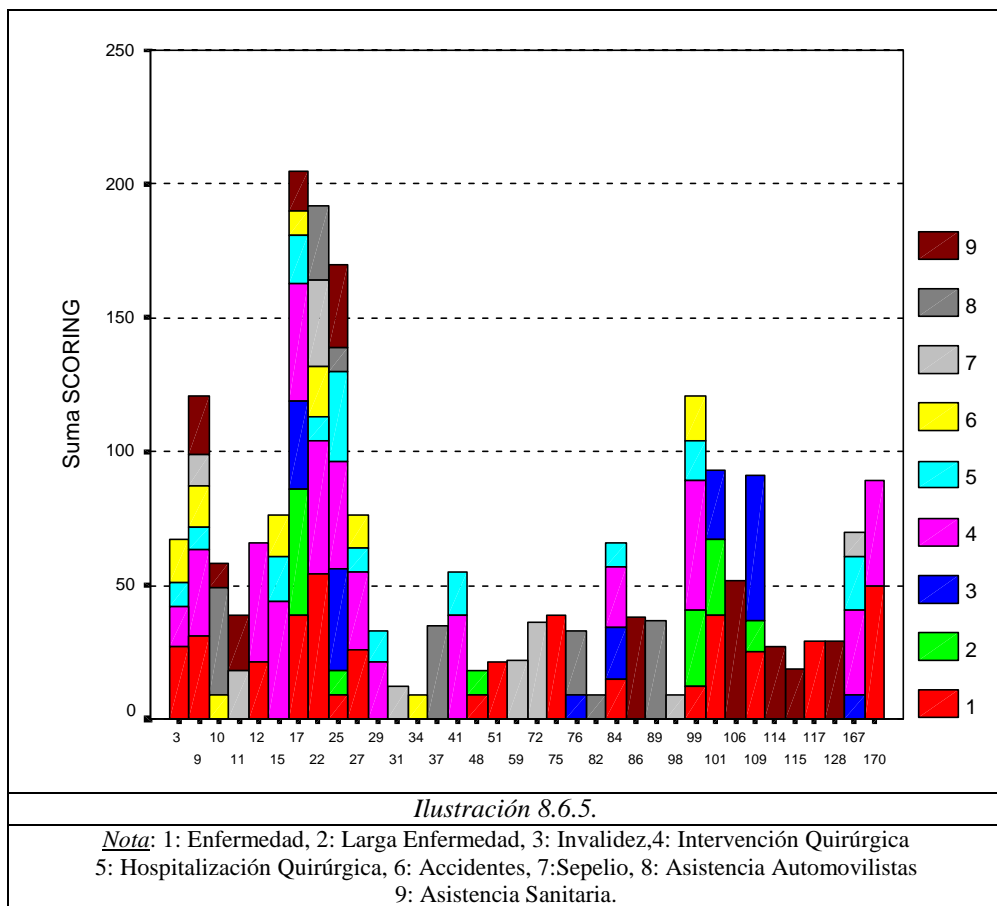
Para las demás prestaciones, la tabla 2.6.5. sintetiza aquellas entidades que poseen prestación por prestación un peor posicionamiento asociado a valores del *scoring* altos⁶³, mediante su identificador. En último lugar, para enriquecer la selección anterior, la ilustración 7.6.5. presenta en formato “histograma” todas las puntuaciones de las restantes prestaciones.



⁶³ Un *scoring* alto significa que su posición relativa respecto al sector no es positiva.



Finalmente y de forma agregada, concluimos que claramente existen tres entidades que poseen un comportamiento diferencial importante en términos de peor posicionamiento relativo y son: 17, 22 y 25, (véase ilustración 8.6.5.).



Para comprobar dichos resultados se presentan los datos estadísticos del propio ratio de siniestralidad para estas tres entidades, las conclusiones son las siguientes. La entidad 17 posee problemas en sus prestaciones de Enfermedad, Larga Enfermedad e Intervención Quirúrgica. La entidad 22, su peor posicionamiento es debido a las prestaciones de

Enfermedad e Intervención quirúrgica. Y la entidad 25, Intervención Quirúrgica y Asistencia Sanitaria a Socios, (véase tabla 3.6.5.).

R2				
TIOPRES	N	Media	Mediana	Desv. típ.
Enfermedad	9	1.020556	.960100	.392927
Larga Enfermedad	9	1.408289	1.381100	.443165
Invalidez	9	.797533	.772900	.489127
Intervención Quirúrgica	9	.930533	.899400	.199649
Hospitalización Quirúrgica	9	.595033	.571300	.188678
Accidentes	9	.231100	.217200	.115907
Asistencia Sanitaria a Socios	9	.568389	.507600	.282846
Total	63	.793062	.722400	.471159

Entidad 17

R2				
TIOPRES	N	Media	Mediana	Desv. típ.
Enfermedad	9	1.389189	1.298400	.327552
Intervención Quirúrgica	9	1.352844	1.237900	.449007
Hospitalización Quirúrgica	9	8.89E-02	9.36E-02	2.73E-02
Accidentes	9	.579922	.519700	.277821
Sepelio	9	.741533	.742800	3.79E-02
Servicios a Automovilistas	9	.703178	.708800	7.78E-02
Total	54	.809254	.749550	.516293

Entidad 22

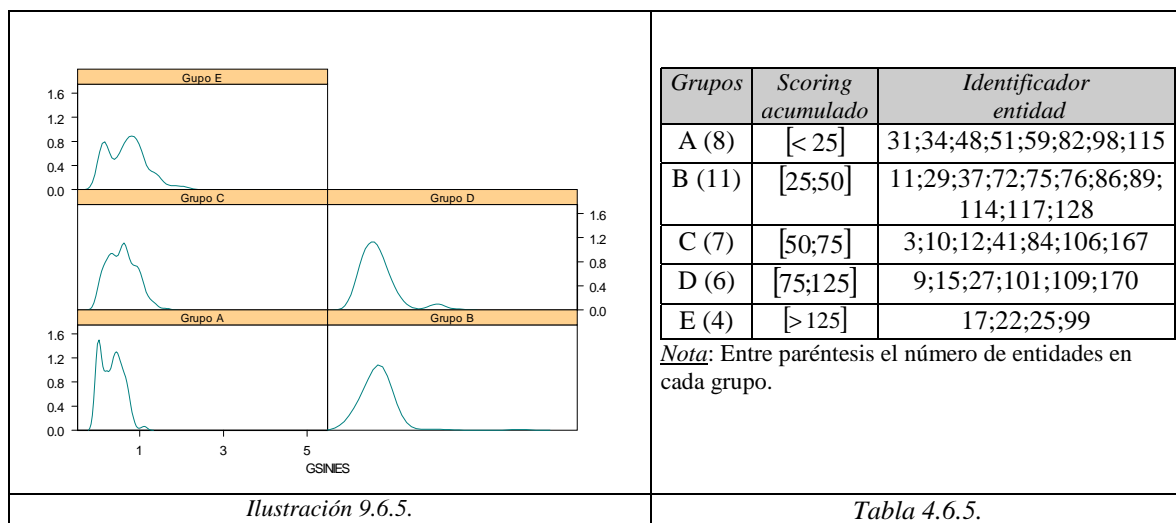
R2				
TIOPRES	N	Media	Mediana	Desv. típ.
Enfermedad	9	.161311	.161200	3.62E-02
Larga Enfermedad	9	.123089	.104000	3.94E-02
Invalidez	9	.881433	.874500	.120021
Intervención Quirúrgica	9	1.035767	.968000	.314706
Hospitalización Quirúrgica	9	.768011	.720700	.136533
Servicios a Automovilistas	9	9.35E-02	8.96E-02	5.55E-02
Asistencia Sanitaria a Socios	9	.920700	.921400	6.12E-02
Total	63	.569110	.686200	.416668

Entidad 25

Tabla 3.6.5.
Entidades con peor *scoring* acumulado.

Consideramos que la propuesta anterior posee las ventajas siguientes. La flexibilidad que poseen los modelos neuronales para definir medidas estándares de sector, la propia definición de modelo de ratio neuronal es robusta frente la presencia de valores atípicos y no es necesaria la transformación previa de los datos, puede ser utilizado como instrumento de control sobre la evolución de las entidades y permite la generación de puntuaciones objetivo.

En la línea de la ventaja anterior y a partir de las puntuaciones acumuladas por entidad, planteamos la generación de grupos⁶⁴ de entidades que nos definan valores referencia de ratio para el sector, (véase tabla 4.6.5.). La ilustración 9.6.5. muestra los “density plot” de los ratios de siniestralidad para cada uno de los grupos generados, donde observamos que la dispersión aumenta conforme la puntuación *acumulada* es mayor y por lo tanto su posicionamiento relativo es peor.



Finalmente los valores “ajustados” del modelo neuronal MLP(4:3) pueden ser considerados como los valores referencia del sector, pero para su utilización eficiente es necesario realizar el análisis desglosado por tipología de prestación, (véase tabla 5.6.5.).

Las conclusiones son las 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, 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, Servicios a Automovilistas. Para poder acercarse a los valores referencia deberían reducir en promedio un 22% su siniestralidad. Para el **grupo D**, las prestaciones son, Enfermedad, Invalidez⁶⁵ e Intervención Quirúrgica, con una reducción necesaria en promedio del orden del 40%.

⁶⁴ Estos grupos se han generado a partir de las puntuaciones acumuladas por entidad.

⁶⁵ Si no se considera esta prestación con valores atípicos, la reducción debería ser del orden del 16%.

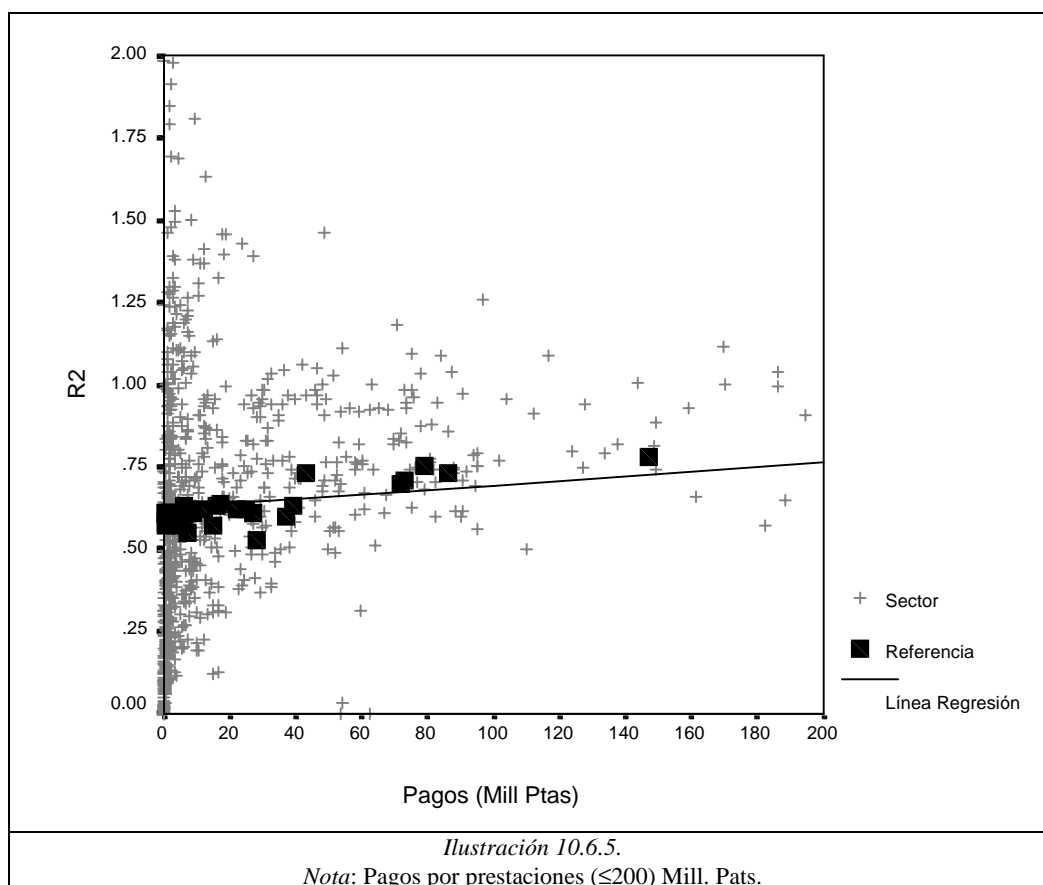
Y en ultimo lugar están las entidades con mayores problemas (4), el **grupo E**, donde las prestaciones peor posicionadas son tres, Larga Enfermedad, Invalidez e Intervención Quirúrgica, que necesitarían una reducción del orden del 25% en promedio. Las prestaciones de Sepelio y Asistencia Sanitaria a Socios, por muy poco, están en línea con los valores referencia.

Tabla 5.6.5.

Grupos Scoring	Prestaciones	Ratio mediano Siniestralidad sector	Ratio mediano Siniestralidad modelo	Tamaño mediano Carteras / sector (Mill Ptas)
A [< 25]	Enfermedad	-----	61%	1
	Larga Enfermedad	-----	60%	-----
	Accidentes	19%	58%	0.5
	Sepelio	49%	61%	2
	Servicios a Automovilistas	39%	60%	6
	Asistencia Sanitaria a Socios	60%	73%	43
B [25;50]	Enfermedad	86%	73%	86
	Invalidez	23%	57%	0.5
	Intervención Quirúrgica	62%	63%	6
	Hospitalización Quirúrgica	44%	60%	2
	Sepelio	73%	64%	17
	Servicios a Automovilistas	84%	61%	9
C [50;75]	Asistencia Sanitaria a Socios	77%	78%	147
	Enfermedad	63%	75%	79
	Invalidez	29%	61%	0.7
	Intervención Quirúrgica	81%	63%	39
	Hospitalización Quirúrgica	43%	61%	9
	Accidentes	37%	53%	28
	Sepelio	36%	61%	7
	Servicios a Automovilistas	96%	70%	72
D [75;125]	Asistencia Sanitaria a Socios	52%	57%	5
	Enfermedad	76%	62%	25
	Larga Enfermedad	49%	61%	0.3
	Invalidez	157%	61%	0.6
	Intervención Quirúrgica	80%	63%	16
	Hospitalización Quirúrgica	45%	62%	12
	Accidentes	39%	57%	15
	Sepelio	49%	61%	1
E [> 125]	Asistencia Sanitaria a Socios	64%	61%	27
	Enfermedad	60%	60%	3
	Larga Enfermedad	69%	62%	8
	Invalidez	86%	60%	2
	Intervención Quirúrgica	104%	62%	22
	Hospitalización Quirúrgica	56%	60%	5
	Accidentes	28%	55%	7
	Sepelio	74%	71%	73
	Servicios a Automovilistas	39%	60%	37
Asistencia Sanitaria a Socios	87%	81%	312	

Nota: En negrita se han señalado, para cada grupo, las prestaciones que poseen un ratio de siniestralidad “mediano” por encima del valor “referencia” mediano del sector obtenido a partir de los valores *ajustados* del modelo neuronal. (----) valores no disponibles.

Podemos observar que en términos generales las prestaciones en cada grupo más alejadas de los valores referencia coinciden con tamaños de carteras pequeñas, en línea con las recomendaciones obtenidas en el apartado 5.2.2.5., (véase tabla 5.6.5.).



Los niveles de siniestralidad medianos de “referencia” obtenidos a partir de los valores ajustados del modelo neuronal utilizado, son relativamente estables respecto al tamaño de la carteras de riesgo. La ilustración 10.6.5. nos muestra donde se sitúan los mismos respecto a la muestra observada para todo el sector sin distinguir por tipo de prestación⁶⁶. Pero si distinguimos por tipología de prestación y obtenemos para cada caso el máximo de los valores medianos de referencia anteriores, conseguimos unos niveles de siniestralidad del sector “teóricos” que pueden ser utilizados como puntos de referencia para las entidades y que lógicamente serán distintos conforme el sector evolucione a lo largo de los ejercicios económicos. Dichos valores son: 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%).

⁶⁶ Véase en el anexo 1 de este apartado, los gráficos de dispersión por tipo de prestación.