Cuadernos del CIMBAGE N°8 (2006) 1-36

CÁLCULO DE LAS PROVISIONES PARA SINIESTROS PENDIENTES DE DECLARACIÓN CON REGRESIÓN BORROSA

Jorge de Andrés Sánchez
Departamento de Gestión de Empresas
Facultad de Ciencias Económicas y Empresariales
Universidad Rovira i Virgili
Avenida de la Universidad 1
43204 – Reus - España
jorge.deandres@urv.net

Recibido 24 de noviembre 2004, aceptado 3 de febrero 2006

Resumen

La correcta determinación de la provisión para siniestros pendientes de declaración es fundamental para la estabilidad de las compañías de seguros, sobre todo en ramos con un notable diferimiento de muchas de las reclamaciones como el del automóvil, el de responsabilidad civil, etc. Por tanto, la literatura actuarial ha propuesto una gran variedad de metodologías para su cálculo, basadas esencialmente en el empleo de la estadística. No obstante, dadas las condiciones cambiantes del entorno en que se mueve la actividad aseguradora, no es aconsejable partir de una experiencia excesivamente amplia en el cálculo de dichas provisiones, lo que conlleva una pérdida de fiabilidad de los métodos estocásticos. Por estas razones, entendemos que la utilización de instrumentos de la lógica borrosa puede ser muy adecuada. Concretamente, en este trabajo proponemos una metodología que aplica la regresión borrosa y el método propuesto por Sherman (1984).

Palabras clave: provisiones para siniestros pendientes de declaración (IBNR), triángulo de siniestralidad, link-ratio, lógica borrosa, números borrosos, regresión borrosa.

ADJUSTING INCURRED BUT NOT REPORTED CLAIMS RESERVE WITH FUZZY REGRESSION

Jorge de Andrés Sánchez
Departamento de Gestión de Empresas
Facultad de Ciencias Económicas y Empresariales
Universidad Rovira i Virgili
Avenida de la Universidad 1
43204 – Reus - España
jorge.deandres@urv.net

Received 24 November 2004, accepted 3 February 2006

Abstract

Determining the correct value for the Incurred but not reported claims reserves is crucial for the insurer's financial stability. That's why actuarial literature has proposed a great number of methods for calculating those reserves. These methods are usually built up from statistical concepts. However, the mutant and uncertain behaviour of insurance environment makes not advisable to use a wide data-base when calculating these reserves. So, on one hand, that fact implies a great loose of reliability of statistical methods but, on the other, it makes very attractive the use of fuzzy logic instruments. Concretely, in this paper we propose a method that combines fuzzy regression and the scheme proposed by Sherman (1984).

*Keyw*ords: Incurred but not reported claims reserves (IBNR), Claims triangle, Link-ratio, Fuzzy logic, Fuzzy numbers, Fuzzy regression.

---

1. INTRODUCCIÓN

Una de las provisiones más importantes en las compañías de no-vida es la provisión para siniestros pendientes de declaración (IBNR[1]), por lo cual no es extraño que su determinación sea objeto de un amplio tratamiento por la literatura actuarial, generalmente, desde una perspectiva estadística. En este contexto, podemos diferenciar dos grandes tipos de enfoques. El primer enfoque los podemos denominar como "clásico" y adopta una perspectiva determinística. El segundo tipo es comúnmente calificado de "estocástico" y busca predicciones de las IBNR más completas que los métodos clásicos.

Entre los métodos clásicos son especialmente conocidos el basado en el coste medio del siniestro, el basado en el período medio de liquidación; y dentro de los que parten de la información organizada a través del triángulo de siniestralidad, podemos mencionar el *chain ladder* y los basados en el *link ratio*. El denominador común de todas estas técnicas es que únicamente ofrecen predicciones puntuales de las IBNR, es decir, se enmarcan en un ambiente determinístico. En Van Eeghen (1981), Taylor (1986), el Institute of Actuaries (1989) puede encontrarse una amplia recopilación de este tipo de métodos.

El segundo enfoque, denominado como "estocástico", es más sofisticado, y su origen se sitúa, según Taylor *et al.* (2003), a mediados de los años 70 del pasado siglo. En última instancia, este tipo de enfoque supone aceptar que la evolución de la siniestralidad a lo largo del tiempo para cualquier año de ocurrencia es aleatoria, tal como se apunta en Kass *et al.* (2001). Así, los métodos englobados en este grupo, tal como afirman England y Verrall (2002), buscan obtener

---

[1] Se trata del acrónimo del nombre de estas provisiones en inglés: *Incurred But Not Reported Reserves*.

estimaciones tanto del valor como de la variabilidad de las IBNR, siendo la situación ideal, la descripción de las IBNR a través de funciones de distribución de probabilidad completas. Dentro de los métodos estocásticos que se han ido proponiendo en la literatura, tal como afirman England y Verrall (2002), muchos de éstos parten de la formulación pura del método *chain ladder* (o de ligeras modificaciones el mismo) y realizan refinamientos estadísticos sobre éste. En este contexto podemos encontrar los trabajos de Mack (1993), England y Verrall (1999) o Verrall (2000). Otro enfoque habitual consiste en considerar el incremento de la siniestralidad de un año de ocurrencia durante los diversos años de desarrollo; y más concretamente, su logaritmo neperiano, a través de una función de distribución de predefinida. En este contexto, lo más común es considerar, tal como hacen Kremer (1982) o Verrall (1989), distribuciones de probabilidad logarítmico-normales para el incremento de la siniestralidad en los años de desarrollo (es decir, distribuciones de probabilidad normales para cuantificar su logaritmo neperiano).

Para que el empleo de métodos estocásticos sea lo más efectivo posible, es necesario partir de una amplia experiencia pero, desafortunadamente, ello no es aconsejable en este problema. Straub (1997) apunta que tomar en consideración observaciones muy alejadas del momento en que deben calcularse las IBNR puede llevarnos a obtener estimaciones poco realistas para éstas. Por ejemplo, si las reclamaciones están relacionadas con daños corporales, el valor final de las indemnizaciones dependerá de variables como la inflación, los cambios en las prácticas judiciales y en la legislación, etc. Por tanto, dado que para la determinación de las IBNR debe utilizarse una experiencia reducida y próxima al momento de su cálculo, los métodos estadísticos pierden una buena parte de su potencia. Cuando es necesario utilizar una información bastante escasa respecto al fenómeno de estudio, un instrumento muy adecuado es la regresión borrosa (RB), basada en la lógica borrosa (LB). Un ejemplo en este sentido es Tseng *et al.* (2001) en el análisis de series temporales univariantes ARIMA. Así, en este trabajo, se adapta la metodología de determinación de las IBNR propuesta en Sherman (1984), basada en la utilización de Mínimos Cuadrados Ordinarios (MCO) para la determinación de la tendencia temporal de la evolución de la siniestralidad, a la utilización de la RB.

Debemos remarcar que la utilización de instrumentos de RB puede reportarnos otras ventajas. En primer lugar, las estimaciones que obtenemos con la RB tras ajustar los coeficientes no serán variables aleatorias de difícil manipulación aritmética, sino números borrosos (NB), cuya operativa es mucho más sencilla. Cuando partimos de magnitudes estimadas como variables aleatorias, éstas, con el fin de facilitar su manipulación, son reducidas a su esperanza matemática (que puede ser corregida o no por su varianza). En cambio, esta pérdida de información no es necesaria cuando operamos con números borrosos.

También queremos reseñar que en el campo actuarial, varios de los instrumentos derivados de la LB han sido utilizados para modelizar problemas que requieren de la utilización de juicios subjetivos por parte del actuario o bien en los cuales, como es nuestro caso, la información de partida es lo suficientemente escasa o vaga como para aceptar la total fiabilidad de los métodos estadísticos. Una visión panorámica sobre las numerosas utilizadas que la LB puede tener en la ciencia actuarial es ofrecida en Ostaszewski (1993), Derrig y

Ostaszewski (1998) y de Andrés y Terceño (2003a). En cualquier caso, podemos remarcar como aportaciones más relevantes las siguientes:

- En aspectos de la toma de decisiones del asegurador como la determinación de riesgo asegurable o la fijación de una política de reaseguro: de Wit (1982), Jablonowski (1991), Lemaire (1990) y Caro (1997).
- Evaluación de la bondad de diversos métodos econométricos en la predicción de la siniestralidad, utilizando el concepto de decisión en ambiente borroso de Bellman y Zadeh (1970): Cummins y Derrig (1993).
- Tarificación y agrupación riesgos: Derrig y Ostaszewski (1995) y Young (1996).
- Valoración financiera de seguros de vida y no vida con tipos de interés borrosos: Lemaire (1990) y Ostaszewki (1993) en el campo de vida y Cummins y Derrig (1997) y Derrig y Ostaszewki (1997) en un contexto de no-vida.
- Determinación de las provisiones para siniestralidad con lógica borrosa y el esquema de Benjamin y Eagles (1986), en de Andrés y Terceño (2003).

En el siguiente epígrafe expondremos los instrumentos de lógica borrosa que utilizaremos en el trabajo para, posteriormente, presentar el esquema de Sherman (1984) de calcula de las IBNR. Posteriormente desarrollamos nuestra metodología, que es una adaptación de la de Sherman (1984) a la utilización de LB. Finalizamos extrayendo las conclusiones más relevantes acerca del trabajo desarrollado.

2. ARITMÉTICA BORROSA Y REGRESIÓN BORROSA

2.1. NÚMEROS BORROSOS TRIANGULARES SIMÉTRICOS Y SU ARITMÉTICA

Un número borroso (NB) $\tilde{A}$, es un subconjunto borroso definido sobre los números reales y es el principal instrumento de la lógica borrosa (LB) para cuantificar magnitudes inciertas. Adicionalmente, los NBs cumplen dos condiciones adicionales. En primer lugar, son subconjuntos borrosos normales, es decir, existe al menos algún elemento cuyo grado de pertenencia es total (es decir, 1). En segundo lugar, deben ser un subconjunto borroso convexo, es decir, sus α-cortes (o conjuntos de nivel) son conjuntos convexos.

En aplicaciones prácticas, y sobre todo en un contexto de regresión borrosa, son muy utilizados los *números borrosos triangulares simétricos* (NBTSs), ya que son fáciles de manipular aritméticamente y presentan una interpretación muy intuitiva. Nosotros denotaremos a un NBTS $\tilde{A}$ como $\tilde{A}=(a, r_a)$ donde a es el centro y r_a el radio. Obsérvese que una estimación sobre la rentabilidad que se obtendrá con una determinada inversión que viniera dada por la sentencia "*el* 4%*, no esperándose desviaciones sobre este valor del* 1%", puede ser cuantificada de forma inmediata por el NBTS $\tilde{A}$=(4%, 1%). Un NBTS queda caracterizado por su función de pertenencia, $\mu_A(x)$, o alternativamente, por sus α-cortes, A_α, como:

$$\alpha = \mu_A(x) = \begin{cases} 1 - \dfrac{|a-x|}{r_a} & a - r_a \leq x \leq a + r_a \\ 0 & \text{en otro caso} \end{cases} \tag{1a}$$

$$A_\alpha = \{x \mid \mu_A(x) \geq \alpha\} = \left[\underline{A}(\alpha), \overline{A}(\alpha)\right] = \left[a - r_a(1-\alpha), a + r_a(1-\alpha)\right] \tag{1b}$$

La combinación lineal de NBTS es un nuevo NBTS[2]. Si partimos de los NBTS $\tilde{A}_i = \left(a_i, r_{a_i}\right)$, i=1,2,...,n, el resultado de la combinación lineal $\tilde{B} = \sum_i k_i \tilde{A}_i$, $k_i \in \Re$, i=1,2,...,n, es $\tilde{B} = \left(b, r_b\right)$, que se obtiene como:

$$\tilde{B} = \left(b, r_b\right) = \left(\sum_i k_i a_i, \sum_i |k_i| r_{a_i}\right) \tag{2}$$

Por otra parte, si evaluamos una función real definida sobre $\Re^n$: $y=f(x_1, x_2,...,x_n)$ que sea contínua, donde sus argumentos $x_1, x_2,...,x_n$ son los NBs $\tilde{A}_1, \tilde{A}_2,..., \tilde{A}_n$, se obtiene el NB $\tilde{B} = f\left(\tilde{A}_1, \tilde{A}_2,..., \tilde{A}_n\right)$. No obstante, aunque $\tilde{A}_1, \tilde{A}_2,..., \tilde{A}_n$ sean NBTS, si $f(\cdot)$ no es lineal, $\tilde{B}$ no será un NBTS. No obstante, a partir de Dubois y Prade (1993) es fácil comprobar que si $f(\cdot)$ es creciente respecto a las m primeras variables, donde $m \leq n$, y decreciente respecto al resto, $\tilde{B}$ se aproxima razonablemente bien con $\tilde{B}' = (b, r_b)$, donde:

$$b = f(\vec{a}),\ r_b = \sum_{i=1}^{m} \frac{\partial f(\vec{a})}{\partial x_i} r_{a_i} - \sum_{i=m+1}^{n} \frac{\partial f(\vec{a})}{\partial x_i} r_{a_i}\ ,\ \text{con}\ \vec{a} = (a_1,...,a_m,a_{m+1},...,a_n) \tag{3}$$

En muchas ocasiones, aunque estimemos las variables que describen un problema mediante números borrosos, será necesario cuantificar las magnitudes que finalmente proyectemos mediante un valor cierto. En nuestro caso, esto ocurrirá cuando debamos determinar el valor definitivo de las IBNR. Este problema es lo que se conoce como "desfuzzyficar" un NB. Nosotros utilizaremos el concepto de valor esperado de Campos y González (1989). El valor esperado de un NB $\tilde{A}$,

---

[2] Ello es consecuencia de que la manipulación aritmética de subconjunto borrosos se realiza mediante el uso de convoluciones max-min, en lugar de convoluciones del tipo suma-producto, que son las utilizadas si operamos con variables aleatorias.

que denotaremos como $EV[\tilde{A},\beta]$, se obtiene a partir de (1b) introduciendo la aversión al riesgo del decisor con un parámetro $0\leq\beta\leq 1$ como:

$$EV[\tilde{A},\beta]=(1-\beta)\int_0^1 \underline{A}(\alpha)d\alpha+\beta\int_0^1 \overline{A}(\alpha)d\alpha \tag{4}$$

y si $\tilde{A}$ es el NBTS (a, r_a), entonces:

$$EV[\tilde{A},\beta]=a+\left(\beta-\frac{1}{2}\right)r_a \tag{5}$$

2.2. REGRESIÓN BORROSA CON NÚMEROS BORROSOS TRIANGULARES

El modelo de regresión borrosa que a continuación expondremos es el de mayor utilización en aplicaciones empíricas[3], y se encuentra desarrollado en los trabajos de Tanaka (1987) y Tanaka e Ishibuchi (1992). En cualquier caso, una panorámica más amplia sobre los diferentes modelos de regresión borrosa existentes puede encontrarse en Chang y Ayyub (2001). Como toda técnica de regresión, el objetivo de la regresión borrosa es determinar una relación funcional entre una variable dependiente con las variables explicativas. Como en las técnicas de regresión más clásicas, partimos de la hipótesis de que la relación entre la variable explicada, y las explicativas es lineal. Dicha relación se infiere partiendo de un conjunto de n observaciones $\{(Y_1,X_1), (Y_2, X_2),\ldots,(Y_j,X_j),\ldots,(Y_n,X_n)\}$, donde:

---

[3] En la economía, la utilización de instrumentos de regresión borrosa es sugerida por ejemplo, en Fedrizzi, M. *et al* (1993). Aplicaciones económicas del modelo que exponemos pueden encontrarse, por ejemplo, en Ramenazi y Duckstein (1992), Watada (1992), Profillidis *et al.* (1999), Tseng *et al.* (2001), Lee y Chen (2001).

X_j: Es la observación j-ésima con j=1,2,…,n, de las variables explicativas. X_j es un vector de $\Re^{m+1}$: X_j = $\left(X_{0j}, X_{1j}, ..., X_{ij}, ..., X_{mj}\right)$, donde X_{0j} = 1 $\forall j$, y X_{ij} es el valor observado sobre la variable explicativa i-ésima en el j-ésimo componente de la muestra.

Y_j: Es la observación j–ésima de la variable dependiente, j=1,2,…,n, que se supone que puede ser un valor crisp o un intervalo de confianza. Así, queda cuantificada a través de su centro y su radio como Y_j = <y_j' $r_{y'_j}$ >, donde y_j' es el centro y $r_{y'_j}$ el radio.

Suponemos que la j-ésima observación de la variable dependiente es un α^*-corte del número borroso del que proviene, donde α^* viene prefijado por el decisor de forma arbitraria. El NB que cuantifica el valor de la j-ésima observación sobre la variable dependiente es, asimismo, un NBTS que notamos como $\tilde{Y}_j = \left(y_j, r_{y_j}\right)$. De esta forma, como el α^*-corte de $\tilde{Y}_j$, $Y_{j\alpha^*}$, es:

$$Y_{j\alpha^*} = Y_j = \left\langle y'_j, r_{y'_j} \right\rangle = \left[y_j - r_{y_j}\left(1-\alpha^*\right), y_j + r_{y_j}\left(1-\alpha^*\right)\right], \; j=1,2,\ldots,n \quad (6)$$

podemos inferir el centro y el radio de $\tilde{Y}_j$ a través de su α^*-corte como:

$$y_j = y'_j \qquad \text{y} \qquad r_{y_j}\left(1-\alpha^*\right) = r_{y'_j} \Rightarrow r_{y_j} = \frac{r_{y'_j}}{1-\alpha^*} \quad (7)$$

Por tanto, la función lineal borrosa que debemos estimar puede expresarse, según las hipótesis realizadas, como:

$$\tilde{Y}_j = \tilde{A}_0 + \tilde{A}_1 X_{1j} + \ldots + \tilde{A}_m X_{mj} \quad (8)$$

Así, en la RB, el término de error no queda introducido como sumando aleatorio, que es lo que ocurre con la regresión convencional, sino que se incorpora en los coeficientes $\tilde{A}_i$, i=0,1,...,m, al asumirse que son números borrosos. Por supuesto, el objetivo final debe ser ajustar los números borrosos $\hat{\tilde{A}}_i$ que estiman a $\tilde{A}_i$ con la muestra disponible. Dada la naturaleza de $\tilde{Y}_j$, los parámetros $\tilde{A}_i$, i=0,1,2,....,m también serán NBTS y, por tanto, pueden ser expresados como $\tilde{A}_i = (a_i, r_{ai})$, i=0,1,2,....,m. Las estimaciones que se obtienen con la muestra de estos coeficientes son también los NBTS $\hat{\tilde{A}}_i = (\hat{a}_i, r_{\hat{a}i})$, i=0,1,2,....,m. Tras determinar $\hat{\tilde{A}}_i$, obtendremos una predicción de las observaciones de variables explicadas, que denotamos como $\hat{\tilde{Y}}_j = \left(\hat{y}_j, r_{\hat{y}_j}\right)$ y que, por analogía con (8), se hallan como:

$$\hat{\tilde{Y}}_j = \hat{\tilde{A}}_0 + \hat{\tilde{A}}_1 X_{1j} + \ldots + \hat{\tilde{A}}_m X_{mj} \tag{9}$$

siendo por tanto $\hat{\tilde{Y}}_j$, a partir de (2):

$$\hat{\tilde{Y}}_j = \left(\hat{y}_j, r_{\hat{y}_j}\right) = \sum_{i=0}^{m} \left(\hat{a}_C,\ r_{\hat{a}_i}\right) X_{ij} = \left(\sum_{i=0}^{m} \hat{a}_i X_{ij}, \sum_{i=0}^{m} r_{\hat{a}_i} \left|X_{ij}\right|\right) \tag{10}$$

con X_{0j}=1, j=1,2,...,n.

Los parámetros $\hat{a}_i$ y $r_{\hat{a}i}$, deben minimizar la incertidumbre de $\hat{\tilde{Y}}_j$, es decir, su radio; y simultáneamente, maximizar su congruencia con la observación que predicen, $\tilde{Y}_j$, que medimos como en grado en que $\tilde{Y}_j$ está incluida en $\hat{\tilde{Y}}_j$, y que denotamos como $\mu\left(\tilde{Y}_j \subseteq \hat{\tilde{Y}}_j\right)$. Para ello, debemos resolver el programa multiobjetivo no lineal:

$$\text{Minimizar } z=\sum_{j=1}^{n} r_{\hat{y}_j} = \sum_{j=1}^{n}\sum_{i=0}^{m} r_{\hat{a}_i}\left|X_{ij}\right|,\ \text{Maximizar } \alpha \tag{11a}$$

sujeto a:

$$\mu\left(\tilde{Y}_j \subseteq \hat{\tilde{Y}}_j\right) \geq \alpha \quad j=1,2,\ldots,n, \tag{11b}$$

$$r_{\hat{a}i} \geq 0 \qquad i=0,1,\ldots,m, \quad \alpha\in[0,1] \tag{11c}$$

Si exigimos para la segunda función objetivo un cumplimiento mínimo de α^*, que es el nivel de presunción al cual el decisor considera que se obtiene $<y_j'\ r_{y_j'}>$, $j=1,2,\ldots,n$, (11a)-(11c) queda transformado en el siguiente programa lineal:

$$\text{Minimizar } z= \sum_{j=1}^{n}\sum_{i=0}^{m} r_{\hat{a}_i}\left|X_{ij}\right| \tag{12a}$$

sujeto a:

$$\sum_{i=0}^{m}\hat{a}_i X_{ij} - \sum_{i=0}^{m} r_{\hat{a}_i}\left|X_{ij}\right|\left(1-\alpha^*\right) \leq y'_j - r_{y'_j} \quad j=1,2,\ldots,n \tag{12b}$$

$$\sum_{i=0}^{m}\hat{a}_i X_{ij} + \sum_{i=0}^{m} r_{\hat{a}_i}\left|X_{ij}\right|\left(1-\alpha^*\right) \geq y'_j + r_{y'_j} \quad j=1,2,\ldots,n \tag{12c}$$

$$r_{\hat{a}i} \geq 0,\ i=0,1,\ldots,m \tag{12d}$$

Las restricciones (12b)-(12c) son consecuencia de la exigencia $\mu\left(\tilde{Y}_j \subseteq \hat{\tilde{Y}}_j\right) \geq \alpha^*$, mientras que con el último bloque aseguramos que los radios sean no negativos. En este contexto, debemos remarcar que, mientras que en el modelo de regresión de Tanaka e Ishibuchi (1992), en el programa delimitado por (12a)-(12d), las variables decisión son tanto los centros como los radios de los parámetros $\hat{\tilde{A}}_i$, i=0,1,...,n, otros autores como Savic y Pedrycz (1992) proponen ajustar previamente los centros a través de mínimos cuadrados, quedando como parámetros a determinar con el programa (12a)-(12d), únicamente los radios.

Creemos que la regresión lineal borrosa ofrece ciertas ventajas sobre la tradicional mínimo-cuadrática. En primer lugar, porque las estimaciones que obtengamos después de ajustar los coeficientes borrosos no serán variables aleatorias; que, en muchas ocasiones, son de difícil tratamiento numérico, sino NBs, cuyo tratamiento es más sencillo. Por otra parte, si el fenómeno de estudio es de carácter económico o social, las observaciones que del mismo se obtienen son consecuencia de la interacción entre las creencias, expectativas, etc. de los agentes involucrados en dicho fenómeno, con un alto componente subjetivo; siendo por tanto un instrumento más adecuado en su tratamiento la Teoría de los Subconjuntos Borrosos. Creemos que es más realista modelizar el sesgo que puede darse entre las realizaciones de la variable dependiente y el valor que, teóricamente, éstas pueden tomar asumiendo que la relación entre variable dependiente y variables explicativas es borrosa, que si damos una naturaleza aleatoria a dicho sesgo, ya que estamos asumiendo, como mínimo, el fuerte componente subjetivo que implica su determinación. Por otra parte, en muchas ocasiones las observaciones obtenidas no vienen dadas por un número cierto, sino por un intervalo de confianza (por ejemplo, son valores extraídos de diferentes fuentes estadísticas). Para utilizar la regresión tradicional debe cuantificarse las observaciones de la variable explicada y de las explicativas a través de un único valor representativo. Es evidente que este proceder implica una pérdida de información. Para implementar los métodos de regresión borrosa no hace falta reducir el valor de las variables observadas a un número real, pudiendo ser utilizados todos los valores observados.

3. CÁLCULO DE LAS IBNR AJUSTANDO DE LA TENDENCIA TEMPORAL DE LA SINIESTRALIDAD CON REGRESIÓN

La determinación de las IBNR, con independencia del método de estimación que se utilice, parte de la información sobre la experiencia de la siniestralidad ordenada en un triángulo denominado como *run-off* o de *siniestralidad*, al estilo del de la tabla 1. En la tabla 2 se ofrece un ejemplo numérico de triángulo de siniestralidad, que utilizamos en nuestras aplicaciones numéricas.

		Año de desarrollo						
		0	**1**	**…**	***j***	**…**	***n*-1**	***n***
Año de origen u ocurrencia	**0**	$Z_{0,0}$	$Z_{0,1}$	…	$Z_{0,j}$	…	$Z_{0,n-1}$	$Z_{0,n}$
	1	$Z_{1,0}$	$Z_{1,1}$	…	$Z_{1,j}$	…	$Z_{1.n-1}$	
	⋮	⋮	⋮	⋮	⋮	⋮		
	i	$Z_{i,0}$	$Z_{i,1}$	…	$Z_{i,j}$	…		
	⋮	⋮	⋮	⋮				
	***n*-1**	$Z_{n-1,0}$	$Z_{n-1,1}$	…				
	n	$Z_{n,0}$		…				

Tabla 1. Triángulo *run-off*

En la tabla 1, $Z_{i,j}$ es el coste acumulado de la siniestralidad ocurrida en el año i (nótese que disponemos de una historia de n años de siniestros) al final del año de desarrollo j (obsérvese que la siniestralidad de las pólizas suscritas en un ejercicio se desarrolla en n años), donde j=0 denota el año de ocurrencia más lejano y j=n el último utilizable. El año en que debemos calcular las IBNR es el año n (en la tabla 2, el año 4 ó 2004). Por tanto, la siniestralidad acumulada conocida para el año de origen i finaliza en el año de desarrollo j, cumpliéndose $i+j=n$, lo que implica que $j=n-i$. En el triángulo de siniestralidad de la tabla 2 se observa que, para el año de origen de origen 3 se conoce la siniestralidad acumulada sólo hasta el año de desarrollo 1 (3+1=4=n).

La diferencia entre unos métodos u otros de estimación de las IBNR estriba en como obtener el valor de la siniestralidad acumulada en todos los años de origen vigentes al final de sus n años de desarrollo; es decir, los valores $Z_{i,n}$, i=0,1,...,n. A continuación, la estimación de las IBNR es inmediata. Para el año de origen i, el valor acumulado de la siniestralidad llega al año de desarrollo $j=n-i$; de tal forma que la IBNR parcial de este año de origen i es $IBNR_i= Z_{i,n}- Z_{i,n-i}$; es decir, la diferencia entre la totalidad de la siniestralidad acumulada para el año de origen i-ésimo, que debe ser inferida y el último valor de su siniestralidad acumulada conocido. Así, la provisión total ($IBNR$) para todos los años de origen en vigor, se halla sumando las IBNR parciales:

$$IBNR = \sum_{i=0}^{n} IBNR_i \qquad (13)$$

			Año de desarrollo				
			0	1	2	3	4
Año de origen u ocurrencia	2000	0	1120	2090	2610	2920	3130
	2001	1	1030	1920	2370	2710	
	2002	2	1090	2140	2610		
	2003	3	1300	2650			
	2004	4	1420				

Tabla 2. Triángulo *run-off* con el que desarrollaremos nuestras aplicaciones numéricas

A continuación explicamos los pasos que seguiremos en la determinación de las IBNR con el método de Sherman (1984).

Paso 1: Determinación del triángulo de link-ratios

El primer paso consiste en determinar el denominado como "triángulo de link ratios". El link ratio para un año de ocurrencia i y de desarrollo j, que denotaremos como $r_{i,j}$, es la razón de la siniestralidad acumulada en dicho año de ocurrencia "i" entre el año de desarrollo j+1 y el inmediatamente anterior (j). Así:

$$r_{i,j} = \frac{Z_{i,j+1}}{Z_{i,j}} \tag{14}$$

El triangulo de link ratios que se deduce del de siniestralidad inicial (la tabla 1), viene dado en la tabla 3. En cambio, el triangulo de link ratios derivado de la tabla 2, viene dado en la tabla 4.

		Año de desarrollo						
		0	1	…	j	…	n-1	n
Año de origen u ocurrencia	0	$r_{0,0}=Z_{0,1}/Z_{0,0}$	$r_{0,1}=Z_{0,2}/Z_{0,1}$	…	$r_{0,j}=Z_{0.\,j+1}/Z_{0,j}$	…	$r_{0,n-1}=Z_{0,\,n}/Z_{0,n-1}$	----
	1	$r_{1,0}=Z_{1,1}/Z_{1,0}$	$r_{1,1}=Z_{1,2}/Z_{0,1}$	…	$r_{1,j}=Z_{1.\,j+1}/Z_{1,j}$	…	----	----
	⋮	⋮	⋮	⋮	⋮	⋮		
	I	$r_{i,0}=Z_{i,1}/Z_{i,0}$	$r_{i,1}=Z_{i,2}/Z_{i,1}$	…	$r_{i,j}=Z_{i,\,j+1}/Z_{i,j}$	…	----	----
	⋮	⋮	⋮	⋮				
	n-1	$r_{n-1,0}=Z_{n-1,1}/Z_{n-1,0}$						
	N	----	----	…	----	----	----	----

Tabla 3. Triángulo general de link ratios

			Año de desarrollo				
			0	1	2	3	4
Año de origen u ocurrencia	2000	0	1.866	1.249	1.119	1.072	----
	2001	1	1.864	1.234	1.143	----	----
	2002	2	1.963	1.220	----	----	----
	2003	3	2.038	----	----	----	----
	2004	4	----	----	----	----	----

Tabla 4. Triángulo de link-ratios de nuestro ejemplo numérico

Paso 2: Ajuste de la tendencia en el tiempo del desarrollo de la siniestralidad

A partir del triángulo link-ratios, debemos determinar para la transición entre dos años de desarrollo dados, un link-ratio único con independencia del año de origen. El link ratio representativo para la evolución de la siniestralidad entre el año de desarrollo j y el siguiente, el j+1, lo denominaremos como r_j y debe determinarse a partir de los link-ratios particulares conocidos para dicho año de desarrollo: $\{r_{0,j}, r_{1,j},\ldots,r_{n-j-1,j}\}$. En nuestro ejemplo numérico, el link ratio r_0, que relaciona el año de desarrollo 0 y el 1, deberá ser inferido a partir de los link-ratios $\{r_{0,0}, r_{1,0}, r_{2,0}, r_{3,0}\}$=$\{1.866, 1.864, 1.963, 2.038\}$. Las alternativas que nos encontramos en este contexto son diversas, y se encuentran glosadas de forma más extensa, por ejemplo, en Institute of Actuaries (1989). Las más comunes consisten en tomar el máximo link ratio o bien calcular las medias de éstos, simples o ponderadas.

Tal como apunta Sherman (1984), esta forma de reducir los link ratios puede suponer, si los datos de partida son escasos, como suele ser el caso, un problema de "sobreparametrización". Así, la capacidad de generalización de las estimaciones sobre la evolución de la siniestralidad puede no ser muy buena. Para obtener una mayor capacidad de generalización en la estimación de la siniestralidad acumulada, deberíamos intentar reducir, en la medida de lo posible, el número de parámetros que debemos ajustar. Obsérvese que en nuestra aplicación numérica, el número de parámetros que debemos ajustar para representar la evolución de la siniestralidad acumulada a lo largo del tiempo son 4: r_0, r_1, r_2, r_3. La alternativa a este problema propuesta por Sherman (1984), consiste en ajustar los link-ratios mediante la expresión común:

$$r_j = 1+a\cdot(j+1)^b \tag{15}$$

donde los parámetros a estimar son a y b. Obsérvese que, en nuestro ejemplo, esto significaría pasar a tener que estimar 2 parámetros (a y b), en lugar de tener que estimar individualmente r_0, r_1, r_2, r_3 (4 parámetros). Obsérvese que (15), aunque no es lineal, es fácilmente linealizable tomando logaritmos. Concretamente, el modelo lineal equivalente a (15), al que se le añade un término de error aleatorio, ε_j, es:

$$R_j=c + b\cdot\ln(j+1)+ \varepsilon_j \tag{16}$$

Con R_j =ln (r_j – 1) y c= ln (a). En la recta de regresión (16) puede observarse que los parámetros a ajustar son c y b, mientras que el link-ratio (15) puede ser también escrito como:

$$r_j = 1+e^c(j+1)^b \tag{17a}$$

Obsérvese que r_j es una función de c y b. Posteriormente nos será de utilidad conocer el valor de sus derivadas parciales:

$$\frac{\partial r_j}{\partial c} = e^c(j+1)^b \quad \text{y} \quad \frac{\partial r_j}{\partial b} = e^c(j+1)^b \ln(j+1) \tag{17b}$$

Con los datos de la tabla 4, el modelo lineal (16) que queda finalmente ajustado es ln (r_j – 1)= -0.096-1.835·ln(j+1), siendo el coeficiente de determinación de la regresión el 99.03%. Por tanto, la expresión final de los link-ratios, r_j, en (17a) es: r_j= $1+e^{-0.096}\times(j+1)^{-1.835}$. La tabla 5 muestra, para el caso práctico que estamos exponiendo, el valor de los link ratios estimados en cada año de desarrollo.

	Link-ratio			
	r_0	r_1	r_2	r_3
Valor	1.908	1.255	1.121	1.071

Tabla 5. Link-ratios finales correspondientes a los 4 años de desarrollo con la metodología de Sherman (1984) en nuestro ejemplo

Paso 3: Determinación de los factores de proyección

Tras haberse obtenido los link-r-atios para cada año de desarrollo "comunes" a cualquier año de origen, debe proyectarse la siniestralidad acumulada en el año de desarrollo n, que para cada el año de origen i simbolizamos como $Z_{i,n}$. Ésta debe proyectarse a partir de la última conocida, $Z_{i,n-i}$. Para ello debemos determinar los *factores de proyección.* Para una siniestralidad acumulada en el año de desarrollo j, del año de origen i, $Z_{i,j}$, su factor de proyección hasta el final de todos los años de desarrollo $f_{j,n}$, es:

$$f_{j,n} = r_j \times r_{j+1} \times \ldots \times r_{n-1} \qquad (18a)$$

Así, la siniestralidad acumulada en el año de desarrollo n es:

$$Z_{i,n} = Z_{i,j} \times f_{j,n} \qquad (18b)$$

Por supuesto, si queremos proyectar la siniestralidad acumulada hasta un año de desarrollo intermedio s, anterior a n, y donde $s>j$, el factor de proyección que debemos utilizar, $f_{j,s}$, se halla como:

$$f_{j,s} = r_j \times r_{j+1} \times \ldots \times r_{s-1} \qquad (19a)$$

Así, la siniestralidad acumulada del año de ocurrencia i en el año de desarrollo s, a partir de la conocida en el año de desarrollo j, es:

$$Z_{i,s}= Z_{i,j}\times f_{j,s} \tag{19b}$$

Obsérvese que $f_{j,s}$ es una función de los link ratios r_j, r_{j+1}, ...,r_{s-1} y su derivada parcial respecto al link-ratio r_h, $j\leq h<s$ es:

$$\frac{\partial f_{j,s}}{\partial r_h} = \prod_{\substack{h=j \\ h\neq k}}^{s-1} r_h \tag{19c}$$

En la tabla 6 se obtienen los factores de proyección $f_{j,4}$, j=0,1,2,3.

	Link-ratio			
	$f_{0,4}$	$f_{1,4}$	$f_{2,4}$	$f_{3,4}$
Valor	2.875	1.507	1.201	1.071

Tabla 6. Factores de proyección de la siniestralidad hasta el año de desarrollo 4

Paso 4: Completar el triángulo run-off y determinar las IBNR

A partir de las expresiones (18a) y (18b) y de las más generales, (19a) y (19b), podemos completar la parte inferior del triángulo de siniestralidad de la tabla 1, lo que se realizaría siguiendo la tabla 7. Asimismo, en la tabla 8, completamos el triángulo de siniestralidad de nuestro ejemplo. Tanto en la tabla 7 como en la tabla 8 destacamos con letra más gruesa los valores de la siniestralidad acumulada cuando estos no han sido observados, sino proyectados.

		Año de desarrollo						
		0	**1**	**…**	$\boldsymbol{j}$	**…**	$\boldsymbol{n}$**-1**	$\boldsymbol{n}$
Año de origen u ocurrencia	**0**	$Z_{0,0}$	$Z_{0,1}$	…	$Z_{0,j}$	…	$Z_{0,n-1}$	$Z_{0,n}$
	1	$Z_{1,0}$	$Z_{1,1}$	…	$Z_{1,j}$	…	$Z_{1,n-1}$	$\boldsymbol{Z_{1,n}}= Z_{1,n-1}\times f_{n-1,n}$
	⋮	⋮	⋮	⋮	⋮	⋮		
	$\boldsymbol{i}$	$Z_{i,0}$	$Z_{i,1}$	…	$Z_{i,j}$	…	$\boldsymbol{Z_{i,n-1}}=Z_{i,j}\times f_{i,n-1}$	$\boldsymbol{Z_{i,n}}=Z_{i,j}\times f_{i,n}$
	⋮	⋮	⋮	⋮				
	$\boldsymbol{n-1}$	$Z_{n-1,0}$	$Z_{n-1,1}$	…	$\boldsymbol{Z_{n-1,j}}= Z_{n-1,1}\times f_{1,j}$	…	$\boldsymbol{Z_{n-1,n-1}}=Z_{n-1,1}\times f_{1,n-1}$	$\boldsymbol{Z_{n-1.n}}=Z_{n-1,1}\times f_{1,n}$
	$\boldsymbol{n}$	$Z_{n,0}$	$\boldsymbol{Z_{n,1}}= Z_{n,0}\times f_{0,1}$	**…**	$\boldsymbol{Z_{n,j}}=Z_{n,0}\times f_{0,j}$	**…**	$\boldsymbol{Z_{n,n-1}}=Z_{n,0}\times f_{0,n-1}$	$\boldsymbol{Z_{n,n}}=Z_{n,0}\times f_{0,n}$

Tabla 7. Triángulo *run-off* tras predecir la siniestralidad acumulada con los factores de proyección

			Año de desarrollo				
			0	1	2	3	4
Año de origen u ocurrencia	2000	0	1120	2090	2610	2920	3130
	2001	1	1030	1920	2370	2710	**2902.41**
	2002	2	1090	2140	2610	**2925.81**	**3133.54**
	2003	3	1300	2650	**3325.75**	**3728.17**	**3992.87**
	2004	4	1420	**2709.36**	**3400.25**	**3811.68**	**4082.31**

Tabla 8. Triángulo *run-off* con el que desarrollamos nuestras aplicaciones numéricas completado

A continuación, es fácil estimar las provisiones a dotar para cada año de origen "i", que se hallan como la diferencia entre la siniestralidad total proyectada al cabo de los n años de desarrollo, y la siniestralidad acumulada ya conocida, que abarca n-i años de desarrollo. En nuestro ejemplo, puede observarse que las IBNR parciales para cada año de origen son:

Año 0: $IBNR_0$= 3130-3130 =0

Año 1: $IBNR_1$= 2902.41-2710 =192.41

Año 2: $IBNR_2$= 3133.54-2610 =523.54

Año 3: $IBNR_3$= 3992.87-2650 =1342.87

Año 4: $IBNR_4$= 4082.31-1420 =2662.31

La provisión final se obtiene sumando las IBNR parciales como (13):

$IBNR$=0+192.41+523.54+1342.87+2662.31 = 4721.12

4. DETERMINACIÓN DE LAS IBNR UTILIZANDO REGRESIÓN BORROSA

Ya apuntamos en la introducción, que la utilización de la metodología expuesta en el epígrafe anterior, basado en el uso de regresión convencional, presenta diversos problemas. En primer lugar, la utilización de MCO es bastante fiable cuando se dispone de una muestra amplia, pero ello no es aconsejable en nuestro problema. Obsérvese que en el ejemplo desarrollado, en el mejor de los casos, hemos partido de 10 observaciones. Por otra parte, la utilización de toda la información disponible en el triángulo IBNR requiere de la estimación del montante de la siniestralidad acumulada final a través de variables aleatorias o al menos, de intervalos de confianza probabilísticos, dada la existencia de término de error aleatorio en (16), lo cual implica un alto coste en términos computacionales, ya que su aproximación requerirá de simulación estocástica. Así, en este epígrafe desarrollamos los 4 pasos que nos llevan a obtener las IBNR expuestos en el epígrafe 3, incidiendo, lógicamente, en las cuestiones diferenciales derivadas de la utilización de una metodología borrosa.

Paso 1: Determinación del triángulo de link ratios

A partir del triángulo de link-ratios de la tabla 3, podemos obtener para el crecimiento de la siniestralidad acumulada entre el año de desarrollo

j y j+1, un valor superior, $\overline{R}_j$ y uno inferior, $\underline{R}_j$, de $R_{i,j}$=ln($r_{i,j}$-1), i=0,1,...,n que determinamos como:

$$\overline{R}_j = \ln(\mathrm{Max}\{r_{0,j}, r_{1,j},\ldots,r_{n-j-1,j}\}-1) \quad (20a)$$

$$\underline{R}_j = \ln(\mathrm{Min}\{r_{0,j}, r_{1,j},\ldots,r_{n-j-1,j}\}-1) \quad (20b)$$

Así, los valores $R_{i,j}$, i=0,1,...,n-j-1 están contenidos en un intervalo de confianza $R_j = < \rho'_j, r_{\rho'_j} >$, donde:

$$\rho'_j = \frac{\overline{R}_j + \underline{R}_j}{2} \quad (21a)$$

$$s_{\rho'_j} = \frac{\overline{R}_j - \underline{R}_j}{2} \quad (21b)$$

Siguiendo (6), suponemos que el intervalo R_j es un α-corte tomado a un nivel predefinido por el decisor α^* del NBTS $\tilde{R}_j = (\rho_j, s_{\rho j})$. Así, ρ_j, $s_{\rho j}$ se obtienen de forma inmediata a partir de ρ'_j, $s_{\rho'_j}$ y α^* con (7). En la tabla 9 determinamos en el ejemplo que estamos desarrollando los valores de R_j y $\tilde{R}_j$, j=0,1,2,3, para un nivel de presunción α^*=0.5.

j	R_j	$\tilde{R}_j$
0	<-0.054, 0.092>	(-0,054, 0.184)
1	<-1.453, 0.062>	(-1.453, 0.125)
2	<-2.036, 0.094>	(-2.036, 0.189)
3	<-2.632, 0.000>	(-2.632, 0.000)

Tabla 9. Valores de R_j, y $\tilde{R}_j$ en el ejemplo que estamos desarrollando

Paso 2: Ajuste de la tendencia en el tiempo del desarrollo de la siniestralidad

A continuación, debemos adaptar el modelo lineal (16) al hecho de que estemos trabajando con regresión borrosa. Así, el modelo que debemos estimar es:

$$\tilde{R}_j = \tilde{c} + \tilde{b}\ln(j+1) \tag{22a}$$

y como estamos trabajando con NBTS, siguiendo la notación establecida en el epígrafe 2.1., el modelo (22a) quedará expresado, utilizando (2), como:

$$\left(\rho_j, s_{\rho_j}\right) = (c, s_c) + (b, s_b)\ln(j+1) = \left(c + b\ln(j+1), s_c + s_b\ln(j+1)\right) \tag{22b}$$

Así, debemos ajustar las estimaciones de $\tilde{c}$ y $\tilde{b}$, $\hat{\tilde{c}} = (\hat{c}, s_{\hat{c}})$ y $\hat{\tilde{b}} = \left(\hat{b}, s_{\hat{b}}\right)$ con (12a)-(12d), debiéndose previamente haber fijado el nivel de presunción para el que se obtienen las observaciones (α^*). Por tanto, planteamos:

$$\text{Minimizar } z = ns_{\hat{c}} + \left[\sum_{j=0}^{n-1}\ln(j+1)\right]s_{\hat{b}} \tag{23a}$$

sujeto a:

$$\hat{c} + \hat{b}\ln(j+1) - \left[s_{\hat{c}} + s_{\hat{b}}\ln(j+1)\right]\left(1-\alpha^*\right) \le \rho'_j - s_{\rho'_j}\ ,\ j=0,1,\ldots,n\text{-}1 \tag{23b}$$

$$\hat{c} + \hat{b}\ln(j+1) + \left[s_{\hat{c}} + s_{\hat{b}}\ln(j+1)\right]\left(1-\alpha^*\right) \ge \rho'_j + s_{\rho'_j}\ ,\ j=0,1,\ldots,n\text{-}1 \tag{23c}$$

$$s_{\hat{c}}, s_{\hat{b}} \ge 0 \tag{23d}$$

Puede observarse que, en definitiva, buscamos determinar con regresión borrosa la tendencia a lo largo de los años de una serie temporal (en este caso, de link-ratios), lo cual ya ha sido propuesto en trabajos anteriores como Watada (1992).

En nuestro ejemplo, a partir de los datos de la tabla 9, y suponiendo que la regresión se prefija para α*=0.5, la obtención de $\tilde{\hat{c}}$ y $\tilde{\hat{b}}$ supone resolver el programa matemático:

Minimizar z= $4s_{\hat{c}} + [\ln(1) + \ln(2) + \ln(3) + \ln(4)]s_{\hat{b}}$

sujeto a:

$$\hat{c} + \hat{b}\ln(1) - [s_{\hat{c}} + s_{\hat{b}}\ln(1)]\,(1-0.5) \leq \text{-0.054- 0.092}$$

$$\hat{c} + \hat{b}\ln(2) - [s_{\hat{c}} + s_{\hat{b}}\ln(2)](1-0.5) \leq \text{-1.453 - 0.062}$$

$$\hat{c} + \hat{b}\ln(3) - [s_{\hat{c}} + s_{\hat{b}}\ln(3)](1-0.5) \leq \text{-2.036 - 0.094}$$

$$\hat{c} + \hat{b}\ln(4) - [s_{\hat{c}} + s_{\hat{b}}\ln(4)](1-0.5) \leq 2.632$$

$$\hat{c} + \hat{b}\ln(1) + [s_{\hat{c}} + s_{\hat{b}}\ln(1)](1-0.5) \geq \text{-0.054+ 0.092}$$

$$\hat{c} + \hat{b}\ln(2) + [s_{\hat{c}} + s_{\hat{b}}\ln(2)](1-0.5) \geq \text{-1.453 + 0.062}$$

$$\hat{c} + \hat{b}\ln(3) + [s_{\hat{c}} + s_{\hat{b}}\ln(3)](1-0.5) \geq \text{-2.036 + 0.094}$$

$$\hat{c} + \hat{b}\ln(4) + [s_{\hat{c}} + s_{\hat{b}}\ln(4)](1-0.5) \geq 2.632$$

$$s_{\hat{c}}, s_{\hat{b}} \geq 0$$

Si aplicamos la metodología de Tanaka e Ishibuchi (1992) se obtiene: $\tilde{\hat{c}}$ = (-0.114, 0.304) y $\tilde{\hat{b}}$ =(-1.802, 0). Con la de Savic y Pedrycz (1992), debemos prefijar previamente los centros, $\hat{c}$ y $\hat{b}$, con MCO. Su valor ya se obtuvo en el apartado 3, siendo $\hat{c}$=-0.096 y $\hat{b}$=-1.835. A continuación obtenemos los radios con el programa matemático anterior, siendo $s_{\hat{c}}$ = 0.269) y $s_{\hat{b}}$ = 0.064.

Posteriormente deberemos obtener los link-ratios comunes evaluando (17a) con NBs; de tal forma que los link-ratios, que denominamos como $\tilde{r}_j$ serán también NBs. Concretamente, debemos determinar:

$$\tilde{r}_j = 1 + e^{\tilde{\hat{c}}}(j+1)^{\tilde{\hat{b}}} \tag{24}$$

Obviamente, aunque $\tilde{\hat{c}}$ y $\tilde{\hat{b}}$, son NBTS, $\tilde{r}_j$ no lo será. No obstante, utilizando (3), y teniendo en cuenta la relación de los link-ratios con los parámetros c y b apuntada en (17b), puede obtenerse su aproximación triangular. Es decir, $\tilde{r}_j \approx (r_j,\ s_{r_j})$, donde:

$$r_j = 1 + e^{\hat{c}}(j+1)^{\hat{b}} \tag{25a}$$

$$s_{r_j} = e^{\hat{c}}(j+1)^{\hat{b}} s_{\hat{c}} + e^{\hat{c}}(j+1)^{\hat{b}} \ln(j+1) s_{\hat{b}} = e^{\hat{c}}(j+1)^{\hat{b}}\left(s_{\hat{c}} + s_{\hat{b}} \ln(j+1)\right) \tag{25b}$$

En la tabla 10 mostramos los valores de los link-ratios finalmente estimados con (25a)-(25b) en para nuestra aplicación numérica.

j	Regresión de Tanaka e Ishibuchi	Regresión de Savic y Pedrycz
0	$\tilde{r}_0$=(1.892, 0.271)	$\tilde{r}_0$=(1.908, 0.244)
1	$\tilde{r}_1$=(1.256, 0.078)	$\tilde{r}_1$=(1.255, 0.080)
2	$\tilde{r}_2$=(1.123, 0.037)	$\tilde{r}_2$=(1.121, 0.041)
3	$\tilde{r}_3$=(1.073, 0.022)	$\tilde{r}_3$=(1.071, 0.026)

Tabla 10. *Link-ratios* borrosos de nuestro ejemplo numérico

Paso 3: Determinación de los factores de proyección

A continuación deberemos obtener los factores de proyección de la siniestralidad desde el año de desarrollo j hasta el s, donde $j<s\leq n$. La expresión general cuando los *link-ratios* eran crisp venía dada en (19a). En el caso en que estos sean NBs, el factor de proyección desde el año de desarrollo j hasta el s será el NB $\tilde{f}_{j,s}$, que se obtiene como:

$$\tilde{f}_{j,s} = \tilde{r}_j \times \tilde{r}_{j+1} \times \ldots \times \tilde{r}_{s-1} \tag{26}$$

Dado que el factor de proyección no es una combinación lineal de los link-ratios, aunque éstos vengan aproximados por NBTSs, $\tilde{f}_{j,s}$ no será un NBTS. No obstante, utilizando (3), obtenemos una aproximación triangular simétrica a $\tilde{f}_{j,s}$, $\tilde{f}_{j,s} \approx \left(f_{j,s}, s_{f_{j,s}}\right)$. Para obtener esta aproximación, deberemos evaluar (19a) en los centros de los link-ratios y utilizar las derivadas parciales de los factores de proyección respecto a los link-ratios que lo componen, apuntada en (19b), tal como indica (3):

$$f_{j,s} = r_j \times r_{j+1} \times \ldots \times r_{s-1} \tag{27a}$$

$$s_{f_{j,s}} = \sum_{k=j}^{s-1} \left(\prod_{\substack{h=j \\ h\neq k}}^{s-1} r_h \right) s_{r_k} \tag{27b}$$

La tabla 11 muestra el valor de las aproximaciones triangulares simétricas de los factores de proyección que deberíamos utilizar para completar el triángulo *run-off* de la tabla 2.

j	s	Regresión de Tanaka e Ishibuchi	Regresión de Savic y Pedrycz
0	1	(1.892, 0.271)	(1.908, 0.244)
0	2	(2.377, 0.487)	(2.394, 0.458)
0	3	(2.669, 0.637)	(2.684, 0.612)
0	4	(2.865, 0.743)	(2.875, 0.724)
1	2	(1.256, 0.078)	(1.255, 0.080)
1	3	(1.411, 0.134)	(1.406, 0.141)
1	4	(1.514, 0.176)	(1.507, 0.187)
2	3	(1.123, 0.037)	(1.121, 0.041)
2	4	(1.206, 0.065)	(1.201, 0.073)
3	4	(1.073, 0.022)	(1.071, 0.026)

Tabla 11. Aproximaciones con NBTS a los factores de proyección de nuestra aplicación numérica

La siniestralidad acumulada en el año de desarrollo *s* para el año de ocurrencia *i*, se obtiene como el NBTS $\tilde{Z}_{i,s} = \left(Z_{i,s}, s_{Z_{i,s}}\right)$, adaptando (19c) al hecho a que los factores de proyección son borrosos con (2):

$$\tilde{Z}_{i,s} = \left(Z_{i,s}, s_{Z_{i,s}}\right) = Z_{i,j}\tilde{f}_{j,s} = Z_{i,j}\left(f_{j,s}, s_{f_{j,s}}\right) = \left(Z_{i,j} f_{j,s}, Z_{i,j} s_{f_{j,s}}\right) \quad (28)$$

Paso 4: Completar el triángulo *run-off* y determinar las IBNR

Completar el triángulo *run-off* es inmediato a partir de (28), de tal forma que la parte inferior del triángulo queda cuantificada con NBTS. En la tabla 12 completamos el triángulo *run-off* de la tabla 2, para lo que utilizamos los factores de proyección obtenidos a partir de la metodología de regresión de Savic y Pedrycz (1992) de la tabla 11.

			Año de desarrollo				
			0	1	2	3	4
Año de origen	2000	0	1120	2090	2610	2920	3130
	2001	1	1030	1920	2370	2710	**(2902.41, 69.26)**
	2002	2	1090	2140	2610	**(2925.81, 107.22)**	**(3133.54, 189.66)**
	2003	3	1300	2650	**(3325.75, 211.41)**	**(3728.17, 373.58)**	**(3992.87, 495.52)**
	2004	4	1420	**(2709.36, 346.40)**	**(3400.25, 650.77)**	**(3811.68, 869.19)**	**(4082.31, 1028.66)**

Tabla 12. Triángulo *run-off* final con el que desarrollamos nuestras aplicaciones numéricas

Una vez completado el triángulo *run-off*, las IBNR quedan cuantificadas como el NBTS que denotamos como $\widetilde{IBNR}$. Para ello, en primer lugar, deberemos obtener la correspondiente a cada año de ocurrencia:

$$\widetilde{IBNR}_i = \tilde{Z}_{i,n} - Z_{i,n-i} = \left(Z_{i,n} - Z_{i,n-i}, s_{Z_{i,n}}\right),\ i=0,1,\ldots,n \qquad (29)$$

La IBNR borrosa de todos los años de ocurrencia se obtiene con (2) y (13) como:

$$\widetilde{IBNR} = \sum_{i=0}^{n} \widetilde{IBNR}_i = \sum_{i=0}^{n} \left(Z_{i,n} - Z_{i,n-i}, s_{Z_{i,n}}\right) = \left(\sum_{i=0}^{n} \left(Z_{i,n} - Z_{i,n-i}\right), \sum_{i=0}^{n} s_{Z_{i,n}}\right) \qquad (30)$$

Para determinar la cuantía final de la IBNR a efectos contables, deberemos reducir $\widetilde{IBNR}$ a un número cierto $IBNR^*$. Para ello, proponemos utilizar el concepto de valor esperado de un número borroso, cuya expresión para NBTS viene dada en (5). En este caso, β debe fijarse a partir de la necesaria prudencia que debe tener el actuario, es decir, debería cumplirse que $\beta > 0.5$. Así, el valor final cierto de las IBNR es, a partir de (30) y (5):

$$IBNR^* = EV\left[\widetilde{IBNR}, \beta\right] = \sum_{i=0}^{n} \left(Z_{i,n} - Z_{i,n-i}\right) + \left(\beta - \frac{1}{2}\right) \sum_{i=0}^{n} s_{Z_{i,n}} \qquad (31)$$

Con los datos de la tabla 12 y (29), es inmediato determinar las IBNR asociadas a cada uno de los años de ocurrencia en vigor. Concretamente, se obtiene:

Año 0: $IB\tilde{N}R_0$ = (3130, 0)-3130 =(0, 0)

Año 1: $IB\tilde{N}R_1$ = (2902.41, 69.26) -2710 =(192.41, 69.26)

Año 2: $IB\tilde{N}R_2$ = (3133.54, 189.66) -2610 =(523.54, 189.66)

Año 3: $IB\tilde{N}R_3$ = (3992.87, 495.52) -2650 =(1342.87, 495.52)

Año 4: $IB\tilde{N}R_4$ = (4082.31, 1028.66) -1420 =(2662.31, 1028.66)

La provisión final se obtiene sumando las IBNR parciales para cada año de origen:

$$IB\tilde{N}R = (0,0)+(192.41,69.26)+(523.54,189.66)+(1342.87,495.52)+ \\ +(2662.31,1028.66)=(4721.12,\ 1783.10)$$

Así, la IBNR total ha sido ajustada por $IB\tilde{N}R$ =(4721.12, 1783.10). Es decir, se considera que, a partir de la experiencia pasada, el valor más verosímil para las IBNR es 4721.12, pero éste valor esperado de las IBNR podría desviarse hasta 1783.10 unidades monetarias. Así, para un nivel de aversión al riesgo β=1, la IBNR final ($IBNR^*$) a reflejar en los estados contables se calcula con (31) como:

$$IBNR^* = EV[IB\tilde{N}R,\ 1] = 4721.12+(1-\frac{1}{2})\ 1783.10 = 5612.67$$

5. CONCLUSIONES

Dada la relativamente escasa información que debe utilizarse en la determinación de las IBNR, entendemos que la aplicación de los métodos estadísticos no es del todo adecuada. Por esta razón, en nuestro trabajo proponemos el uso de instrumentos de lógica borrosa, concretamente, la regresión borrosa y la aritmética de números borrosos. Como se ha podido comprobar, su empleo reporta una ventaja adicional: es más sencillo trabajar con la incertidumbre asociada a las estimaciones, que si partimos de las estimaciones de una regresión convencional, ya que en este último caso hubiéramos tenido que recurrir a la utilización de la simulación estocástica.

Para determinar el montante final de las IBNR con nuestro método, será necesario reducir su estimación borrosa a un valor cierto. Nuestra propuesta consiste en la utilización del concepto de valor esperado, que permite introducir de forma sencilla e intuitiva la prudencia del actuario en la determinación de las IBNR.

REFERENCIAS

[1] Bellman, R.E.; L.A. Zadeh (1970). "Decision Making in a fuzzy environment". *Management Sciences.* 17, pp.B141-B164.

[2] Benjamin, S.; L.M. Eagles (1986). "Reserves in Lloyd's and the London market". *Journal of the Institute of Actuaries.* 113, 2, pp.197-257.

[3] Campos, L.M.; A. González (1989). "A subjective approach for ranking fuzzy numbers". *Fuzzy Sets and Systems.* 29, pp.145-153.

[4] Caro, R. (1997). "Una aproximación a una operación de seguro de autos mediante la técnica fuzzy". *Anales del Instituto Español de Actuarios.* 3, 3° época.

[5] Chang, Y-H.O.; Ayyub, B.M. (2001). "Fuzzy regression methods-a comparative assessment". *Fuzzy Sets and Systems.* 119, pp.187-203.

[6] Cummins, J.D.; R.A. Derrig (1993). "Fuzzy trends in property-liability insurance claim costs". *Journal of Risk and Insurance.* 60, 3, pp.429-465.

[7] Cummins, J.D.; R.A. Derrig (1997). "Fuzzy financial pricing of property-liability insurance". *North American Actuarial Journal.* 1, 4, pp.21-44.

[8] De Andrés, J.; A. Terceño (2003a). "Aplicaciones actuariales de la teoría de los subconjuntos borrosos"*. Cuadernos del Instituto de Investigaciones de Estadística y Matemática Actuarial de la Universidad de Buenos Aires* 5, pp.1-40.

[9] De Andrés, J.; A. Terceño (2003b). "Applications of Fuzzy Regression in Actuarial Analysis". *Journal of Risk and Insurance.* 70, 4, pp.665-699.

[10] De Andrés, J.; A. Terceño (2004). "Estimating a fuzzy term structure of interest rates using fuzzy regression methods". *European Journal of Operational Research.* 154, pp.804-818.

[11] De Wit, G.W. (1982). "Underwriting and uncertainty". *Insurance: Mathematics and Economics.* 1, pp.277-285.

[12] Derrig, R.A.; K Ostaszewski (1998). "Fuzzy sets methodologies in Actuarial Science". En: Zimmermann, H.-J. (ed.): *Practical applications of fuzzy technologies.* Kluwer Academic Publishers, pp.531-556.

[13] Derrig, R.A.; K. Ostaszewski (1995). "Fuzzy techniques of pattern recognition in risk and claim classification". *Journal of Risk and Insurance.* 62, pp.447-482.

[14] Derrig, R.A.; K. Ostaszewski (1997). "Managing the tax liability of a property liability insurance company". *Journal of Risk and Insurance.* 64, pp.695-711.

[15] Dubois, D.; Prade, H. (1993). "Fuzzy numbers: an overview". En Dubois, D., Prade, H.; Yager, R.R. (eds.). *Fuzzy sets for intelligent systems.* San Mateo (California), Morgan Kaufmann Publishers, pp.113-148.

[16] England, P.D.; R.J. Verrall (1999). "Analytic and bootstrap estimates of prediction errors in claim reserving". *Insurance: Mathematics and Economics.* 25, pp.281-293.

[17] England, P.D.; R.J. Verrall (2002). *Stochastic claims reserving in general insurance.* London, Institute of Actuaries. Disponible en http://www.actuaries.org.uk/ sessional/sm0201.pdf

[18] Fedrizzi, M.; Fedrizzi, M.; Ostasiewicz, W. (1993). "Towards fuzzy modelling in economics". *Fuzzy Sets and Systems.* 54, pp.259-268.

[19] Institute Of Actuaries (1989). *Claims reserving manual.* London, Institute of Actuaries.

[20] Jablonowski, M. (1991). "Fuzzy logic and insurance decisions". *CPCU Journal.* September, pp.181-187.

[21] Kass, R.; Goovaerts, M.; Dhaene, J.; Denuit, M. (2001). *Modern Actuarial Risk Theory.* Kluwer Academic Publishers (Dordetch).

[22] Kremer, E. (1982). "IBNR claims and the two-way of ANOVA". *Scandinavian Actuarial Journal,* pp.47-55.

[23] Lee, H.T; Chen, S.H. (2001). “Fuzzy regression model with fuzzy input and output data for manpower forecasting”. *Fuzzy Sets and Systems.* 119, pp.205-213.

[24] Lemaire, J. (1982). “Claims Provisions in Liability Insurance”. *Journal of Forecasting.* 1, pp.303-318.

[25] Lemaire, J. (1990). “Fuzzy insurance”. *Astin Bulletin.* 20, pp.33-55.

[26] Mack, T. (1993). "Distribution-free calculation of the standard error of chain-ladder reserve estimates". *Astin Bulletin.* 23, pp.213-223.

[27] Ostaszewski, K. (1993). *An investigation into possible applications of fuzzy sets methods in actuarial science.* Schaumburg (USA) Society of Actuaries.

[28] Profillidis, V.A.; Papadopoulos, B.K.; Botzoris, G.N. (1999). “Similarities in fuzzy regression models and application on transportation”. *Fuzzy Economic Review.* 4, 1, pp.83-98.

[29] Ramenazi, R.; Duckstein, L. (1992). "Fuzzy regression analysis of the effect of university research on regional technologies". In Kacprzyk, J.; Fedrizzi M. (eds.). *Fuzzy regression analysis.* Physica-Verlag. Heildelberg. Pp.237-263.

[30] Savic, D.; Predrycz, W. (1992). “Fuzzy linear models: construction and evaluation”. En: J. Kacprzyk, M. Fedrizzi (Eds.), *Fuzzy Regression Analysis*. Heildelberg, Physica-Verlag, pp.91-100.

[31] Sherman, R.E. (1984). "Extrapolating, Smoothing and Interpolating development factors". *Proceedings of the Casualty Actuarial Society.* 71, pp.122-123.

[32] Straub, E. (1997). *Non-life insurance mathematics.* Berlin, Springer.

[33] Tanaka, H. (1987). "Fuzzy data analysis by possibilistic linear models", *Fuzzy Sets and Systems.* 24, pp.363-375.

[34] Tanaka, H.; Ishibuchi H. (1992). "A possibilistic regression analysis based on linear programming". In Kacprzyk, J.; Fedrizzi M. (eds.). *Fuzzy regression analysis.* Heildelberg, Physica-Verlag, pp.47-60.

[35] Taylor, G.; G. McGuire; A. Greenfield (2003). *Loss reserving: past, present and future.* Melbourne, Centre for Actuarial Studies of the University of Melbourne.

[36] Taylor, G.C. (1986). *Claims reserving in non-life insurance.* Amsterdam, North-Holland.

[37] Tseng, F.-M.; G.-H. Tzeng; H.-C. Yu; B. J.-C. Yuan (2001). "Fuzzy ARIMA model for forecasting the foreign exchange market". *Fuzzy Sets and Systems.* 118, pp.9-19.

[38] Van Eeghen, J. (1981). "*Loss Reserving methods, Surveys on actuarial studies 1*." Rotterdam, Nationale-Nederlanden.

[39] Verrall, R.J. (2000). "An investigation into stochastic claims reserving models and the chain-ladder technique". *Insurance: Mathematics and Economics.* 19, pp.31-43.

[40] Watada, J. (1992). "Fuzzy time-series analysis and forecasting of sales volume". En Kacprzyk, J.; Fedrizzi M. (eds.). *Fuzzy regression analysis.* Heildelberg, Physica-Verlag, pp.211-227.

[41] Young, V.R. (1996). "Insurance rate changing: a fuzzy logic approach". *Journal of Risk and Insurance.* 63,3, pp.461-484.