

SISTEMAS INFORMÁTICOS APLICADOS A LA MATEMÁTICA ACTUARIAL NO VIDA. UNA PROPUESTA CON R

M^a MERCEDES CLARAMUNT BIELSA

mmclaramunt@ub.edu

*Departamento de Matemática Económica, Financiera y Actuarial, Universidad de Barcelona
Avenida Diagonal 690, 08034_Barcelona.*

EVA BOJ DEL VAL

evaboj@ub.edu

*Departamento de Matemática Económica, Financiera y Actuarial, Universidad de Barcelona
Avenida Diagonal 690, 08034_Barcelona.*

TERESA COSTA COR

tcosta@ub.edu

*Departamento de Matemática Económica, Financiera y Actuarial, Universidad de Barcelona
Avenida Diagonal 690, 08034_Barcelona.*

M^a TERESA MÁRMOL JIMÉNEZ

mmarmol@ub.edu

*Departamento de Matemática Económica, Financiera y Actuarial, Universidad de Barcelona
Avenida Diagonal 690, 08034_Barcelona.*

Recibido 11/11/2011

Revisado 02/02/2012

Aceptado 15/03/2012

RESUMEN: El presente trabajo es el resultado de la investigación realizada por las autoras sobre los sistemas informáticos aplicados a la Matemática Actuarial No Vida (MANV). El principal objetivo es presentar y explicar la utilización del lenguaje de programación R a partir de la experiencia iniciada el curso 2009-2010 en la Universidad de Barcelona (UB). La aportación del trabajo consiste en la recopilación y aplicación de librerías realizadas por distintos autores sobre las temáticas incluidas en MANV y su complementación con funciones de elaboración propia. El artículo incluye, para algunos de los bloques que se explican dentro de MANV en la UB, un pequeño recordatorio teórico y diversos ejemplos, calculados con R y con datos reales de carteras de seguros, extraídos de artículos de referencia en la ciencia actuarial.

Palabras claves: Software libre, Matemática actuarial no vida, Número de siniestros, Coste total, Credibilidad, Provisiones técnicas.

ABSTRACT: The present work is the result of the investigation realized by the authoresses on programming languages applied to Non Life Actuarial Mathematics (NLAM). The main objective of this paper is to present and to expose the use of the R programming language in NLAM looking at the experience initiated at the University of Barcelona (UB) during 2009-2010. The contribution of the paper consists of the summary and application of libraries realized by different authors on the topics included in NLAM and its improvement with functions of own production. The article includes, for some of the different blocks that are treated inside NLAM at the UB, a small theoretical summary and numerical examples with R, and with real data of insurance portfolios that can be found in basic papers of actuarial science.

Keywords: Free software, Non life actuarial mathematics, Number of claims, Total cost, Credibility, Technical provisions.

1. Introducción

En la Universidad de Barcelona, en la Facultad de Economía y Empresa, en el primer semestre del segundo curso del segundo ciclo de la licenciatura de Ciencias Actariales y Financieras, se imparte la asignatura MANV que consta de 9 créditos, de los cuales 6 son de teoría y 3 son de práctica. Los contenidos teóricos de la asignatura incluyen temas diversos: ajuste de distribuciones estadísticas (tanto al número de siniestros, como a la cuantía, como al coste total), modelos de credibilidad aplicados a la tarificación o cálculo de provisiones técnicas de prestaciones pendientes. Evidentemente, todos ellos son temas que necesitan de la aplicación de *software* informático que nos permita y facilite la realización de cálculos numéricos.

Durante el curso 2009-2010 el grupo de profesoras encargadas de la docencia de la asignatura iniciamos la introducción del lenguaje de programación R en la parte práctica de la asignatura. Los detalles del proceso de cambio en la metodología docente que se llevaron a cabo pueden encontrarse en los trabajos Claramunt *et al.* (2010) y Claramunt *et al.* (2012).

Previamente el *software* utilizado había sido el APL2 que no es de libre acceso, motivo por el cuál generaba problemas a los alumnos. Así, consideramos necesario trabajar con un *software* de libre acceso, más teniendo en cuenta el aumento de trabajo autónomo del alumno y la disminución de clases presenciales en el nuevo contexto del Espacio Europeo de Educación Superior (Costa *et al.*, 2009).

Con tal de conocer la situación de la utilización de *software* en las universidades españolas que imparten estudios en ciencias actuariales y financieras, enviamos una encuesta a los responsables de las asignaturas. La conclusión (Claramunt *et al.*, 2012) fue que todas ellas utilizan *software* informático como herramienta en las clases prácticas realizadas en aulas de informática dentro de las facultades. El programa informático varía, pasando por *Excel* (macros y *Visual Basic*), *Maple* e incluso *Matlab*, que es el programa de pago equivalente al de libre uso *Octave*.

El lenguaje de programación R, que consta de un sistema base y de librerías adicionales, cubre las necesidades descritas como herramienta de cálculo gratuita y de libre acceso, siendo además un *software* muy dinámico, ya que las funciones y las librerías se van ampliando gracias a las aportaciones de usuarios que pasan un estricto proceso de elaboración y contrastación. Estamos hablando, además, de un *software* ampliamente utilizado en el campo actuarial, siendo constante la aparición de librerías especializadas en temas actuariales, que son de gran utilidad para MANV (Kaas *et al.*, 2008).

Para el profesorado implicado en la aplicación de R, el nuevo lenguaje ha sido de gran utilidad para explicar y aplicar conceptos teóricos por ser un lenguaje dinámico, fácil de utilizar y gratuito. De cara a la transformación de la licenciatura en Máster Universitario será necesario tener presente todas estas consideraciones en el correcto diseño de dicho máster y de los planes docentes de las distintas asignaturas. De hecho, en tres de las trece universidades españolas que impartían dicha licenciatura ya se ha iniciado durante el curso 2010-2011 un máster adaptado al EEES (Universidad Complutense de Madrid, Universidad Carlos III y Universidad de Valencia).

Por otro lado, la implementación del *software* R en el curso 2009-2010 ha sido bien recibida por los alumnos. Para conocer la opinión de los estudiantes pasamos una encuesta al inicio y al final de curso, donde éstos expresaron que el cambio era positivo y que la adaptación al nuevo *software* no les representó ningún problema. Remarcan como aspectos positivos del lenguaje R en su aplicación a MANV que es *software* libre, que hay mucha información disponible en la *web* (ayuda, información de funciones, ...), la disponibilidad de muchas librerías y funciones adaptadas a temas actuariales, que es un *software* muy dinámico o que es muy fácil de aprender y muy lógico en el modo de operar (Claramunt *et al.*, 2012).

El trabajo que se presenta se estructura siguiendo algunos de los temas que se explican en la asignatura, incluyendo para cada tema una breve introducción y la resolución de ejemplos con uso de R. En algunos temas, además de usar librerías y funciones ya existentes en R, ha sido necesaria la programación de funciones específicas para la resolución de determinados aspectos. En concreto, hemos utilizado dos librerías específicas en temas actuariales:

- *actuar* (Dutang *et al.* 2008, Goulet and Pouliot, 2009)

- *ChainLadder* (Gesmann, 2009),
- y hemos elaborado las siguientes funciones propias, cuyo código adjuntamos en el Anexo:
- *conchi*. Contraste de bondad de ajuste.
 - *bonusma*. Modelo de credibilidad exacta Gamma-Poisson-Gamma.
 - *ibnrchl*, *ibnrchlvar1*, *ibnrchlvar2*, *ibnrvylder*. Métodos deterministas de cálculo de las provisiones técnicas.

2. Distribución del número de siniestros

Simbolizamos por N a la variable aleatoria (v.a.) número de siniestros en un determinado período. N puede tomar valores naturales positivos, usualmente 0, 1, 2,.... Las distribuciones de probabilidad discretas que mejor se adaptan a esta variable son: Poisson, Binomial Negativa y Poisson-Inversa Gaussiana.

Para la estimación de los parámetros de la distribución elegida usaremos, por su facilidad y, en general, buenos resultados, el método de los momentos (para más detalles sobre este método de estimación así como sobre las principales características de las distribuciones referidas, ver Panjer and Willmot, 1992).

Se dispone, en la Tabla 1, de la siguiente distribución empírica de datos sobre el número de siniestros por póliza y número de pólizas para un determinado periodo en Suiza (Bühlmann, 1970):

Tabla 1. Datos sobre el número de siniestros por póliza y número de pólizas para un determinado periodo en Suiza, extraídos de Bühlmann (1970).

Número siniestros	Número pólizas
0	103704
1	14075
2	1766
3	255
4	45
5	6
6	2
Total	119853

Ajustaremos a los datos observados las tres distribuciones anteriormente citadas, es decir, calcularemos las frecuencias absolutas (número de pólizas) teóricas que se obtendrían si la probabilidad de que ocurra 0,1,2,... siniestros se distribuyese según el modelo Poisson, Binomial Negativa o Poisson-Inversa Gaussiana.

Estimación de las frecuencias absolutas teóricas según Poisson

Consideramos que $N \sim Po(\lambda)$, con $\hat{\lambda} = \bar{n}$, siendo \bar{n} la media de la muestra de datos sobre el número de siniestros. A continuación mostramos el cálculo en R de la media muestral, que utilizaremos como parámetro $\hat{\lambda}$, así como las probabilidades de que el número de siniestros tome valores 0, 1, 2, ... 6 según la distribución de Poisson:

```
> valn<-0:6
> valf<-c(103704,14075,1766,255,45,6,2)
> medn<-weighted.mean(valn,valf);medn
[1] 0.1551400
> prpois<-dpois(valn,medn);prpois
[1] 8.562952e-01 1.328457e-01 1.030484e-02 5.328979e-04 2.066845e-05
[6] 6.413009e-07 1.658191e-08
> sum(prpois)
[1] 1
```

En este ejemplo, la suma de las probabilidades teóricas correspondientes a los siete primeros valores de N es prácticamente 1. De hecho, las probabilidades teóricas para $N = 7, 8$ y 9 son igual a $3.675018e-10$, $7.12678e-12$ y $1.228498e-13$ respectivamente. Teniendo en cuenta que el número total de pólizas es 119.853, con el objetivo de contrastar el ajuste del modelo teórico de Poisson, podemos trabajar con una probabilidad teórica acumulada hasta $N = 6$ de 1. De este modo, los valores obtenidos de las frecuencias absolutas teóricas o esperadas según la distribución de Poisson son:

```
> fteopois<-rpois*sum(valf);fteopois
[1] 1.026296e+05 1.592195e+04 1.235066e+03 6.386942e+01
[5] 2.477176e+00 7.686184e-02 1.987392e-03
```

Estimación de las frecuencias absolutas teóricas según Binomial Negativa

Supongamos que $N \square BN(a, b)$, es decir, la distribución Binomial Negativa, que depende de dos parámetros, a y b , cuyos estimadores se obtienen a partir de la media (\bar{n}) y varianza (S_N^2) de los datos muestrales:

$$\hat{a} = \frac{\bar{n}^2}{S_N^2 - \bar{n}} \quad \hat{b} = \frac{S_N^2 - \bar{n}}{\bar{n}}. \quad (1)$$

Igual que en el caso anterior, estimaremos los parámetros de la distribución y calcularemos las probabilidades teóricas a partir de la distribución Binomial Negativa:

```
> varn<-sum((valn-medn)^2*valf)/sum(valf);varn
[1] 0.1793140
> a<-medn^2/(varn-medn);a
[1] 0.9956332
> b<-(varn-medn)/medn;b
[1] 0.1558205
> prnegbin<-dnbinom(valn,a,1/(1+b));prnegbin
[1] 8.657335e-01 1.162031e-01 1.563157e-02 2.104283e-03 2.833766e-04
[6] 3.816969e-05 5.142054e-06
> sum(prnegbin)
[1] 0.9999992
```

En este ejemplo, a diferencia del ajuste mediante la distribución de Poisson, en el caso de la distribución Binomial Negativa, la suma de las probabilidades teóricas hasta $N = 6$ es significativamente distinta de 1. Por lo tanto, calcularemos la probabilidad de que $N > 6$, así como la frecuencia absoluta teórica correspondiente:

```
> prnegbin<-c(prnegbin,1-sum(prnegbin));prnegbin
[1] 8.657335e-01 1.162031e-01 1.563157e-02 2.104283e-03 2.833766e-04
[6] 3.816969e-05 5.142054e-06 8.006705e-07
> fteonegbin<-prnegbin*sum(valf);fteonegbin
[1] 1.037608e+05 1.392729e+04 1.873491e+03 2.522047e+02
[5] 3.396353e+01 4.574752e+00 6.162906e-01 9.596276e-02
```

Estimación de las frecuencias absolutas teóricas según Poisson-Inversa Gaussiana

Se considera que $N \square IG(\mu, \beta)$, con parámetros μ y β , que pueden estimarse a partir de las siguientes expresiones:

$$\hat{\mu} = \bar{n} \quad \hat{\beta} = \frac{S_N^2}{\bar{n}} - 1. \quad (2)$$

Para calcular las probabilidades teóricas a partir de la distribución Poisson-Inversa Gaussiana, utilizaremos la función *dPIG* de la librería *gamlss.dist*. Se muestra, a continuación, la ejecución de dicha función y los resultados obtenidos:

```
>library(gamlss.dist)
> mu<-medn;mu
[1] 0.1551400
> beta<-varn/medn-1;beta
[1] 0.1558205
> prpoig<-dPIG(valn,mu,beta/mu);prpoig
[1] 8.653384e-01 1.172202e-01 1.490219e-02 2.128848e-03 3.389167e-04
[6] 5.832093e-05 1.059993e-05
> sum(prpoig)
[1] 0.9999975
> prpoig<-c(prpoig,1-sum(prpoig));prpoig
[1] 8.653384e-01 1.172202e-01 1.490219e-02 2.128848e-03 3.389167e-04
[6] 5.832093e-05 1.059993e-05 2.492419e-06
> fteopoig<-prpoig*sum(valf);fteopoig
[1]1.037134e+05 1.404919e+04 1.786072e+03 2.551488e+02
[5]4.062018e+01 6.989939e+00 1.270433e+00 2.987239e-01
```

En la Tabla 2 se muestran los resultados de los ajustes para las tres distribuciones de probabilidad aplicadas utilizando los datos de la Tabla 1:

Tabla 2. Ajuste de las frecuencias teóricas de las distribuciones Poisson, Binomial Negativa y Poisson-Inversa Gaussiana con los datos de la Tabla 1.

Número siniestros	Frecuencias reales	Frecuencias teóricas Poisson	Frecuencias teóricas Binomial Negativa	Frecuencias teóricas Poisson-Inversa Gaussiana
0	103704	102629.5543	103760.7619	103713.4079
1	14075	15921.9538	13927.2921	14049.1917
2	1766	1235.0663	1873.4908	1786.0723
3	255	63.8694	252.2047	255.1488
4	45	2.4772	33.9635	40.6202
5	6	0.0769	4.5748	6.9899
6	2	0.002	0.6163	1.2704
>6	0	0	0.096	0.2987

Si disponemos de los datos completos de la muestra, como en este caso, podremos comprobar el ajuste del modelo propuesto, es decir, de la distribución elegida. Entre los distintos procedimientos y tests estadísticos, utilizaremos, por ser uno de los más usuales en el caso de las variables discretas, el de la Bondad del Ajuste (para más detalles sobre este método, ver Panjer and Willmot, 1992).

El estadístico de contraste es:

$$D = \sum_{i=0}^{k-1} \frac{(O_i - n \cdot p_i)^2}{n \cdot p_i} \quad (3)$$

siendo:

p_i : probabilidad de que la variable aleatoria tome valores dentro del intervalo $[a_i, a_{i+1})$, según la distribución elegida,

O_i : número de observaciones reales,

$n = \sum_{i=1}^k O_i$: tamaño de la muestra.

Los intervalos deben elegirse de manera que $n \cdot p_i \geq 5$ en todos ellos; si no es así, se reagrupan. En el caso de las distribuciones de N , teniendo en cuenta que el número de valores distintos de la variable (k) normalmente es pequeño, los intervalos se eligen conteniendo cada uno de ellos (excepto el último) un número natural.

Simbolizando por r el número de parámetros que se han estimado según la distribución elegida, si el tamaño de la muestra es grande ($n \rightarrow \infty$), $D \approx \chi^2_{(k-r-1)}$; siendo $k-r-1$ los grados de libertad. En general, a mejor ajuste, menor es el valor de D . Sin embargo, el valor de D por si solo no permite rechazar o no la hipótesis de que la distribución real del número de siniestros es la considerada. Para ello calcularemos la probabilidad:

$$p = P\left[\chi^2_{(k-r-1)} > D\right], \quad (4)$$

y a mayor valor de p (p -value o valor de significación), mejor ajuste. Pueden utilizarse también otros estadísticos como la ratio de verosimilitud o el estadístico de Kolmogorov-Smirnov, que pueden consultarse en Panjer and Willmot (1992).

Para poder realizar el contraste de bondad de ajuste con los datos obtenidos según las distribuciones Poisson, Binomial Negativa y Poisson-Inversa Gaussiana, ejecutaremos la función que se ha programado en R para dicho fin. Esta función, que se denomina *conchi*, utiliza la función *pchisq* de R, de manera que para su ejecución es necesario disponer de los vectores con los valores de las frecuencias absolutas reales y teóricas, convenientemente agrupados para que se cumpla que $n \cdot p_i \geq 5$, además del número de parámetros que se han estimado en cada distribución. Recordemos que la distribución de Poisson sólo depende de un parámetro, mientras que las distribuciones Binomial-Negativa y Poisson-Inversa Gaussiana dependen de dos parámetros.

Los cálculos a realizar en el contraste de bondad de ajuste son los siguientes (ver Anexo para el código de la función *conchi*):

```
> fteopoiscont<-c(fteopois[1:3],sum(fteopois[4:7]));fteopoiscont
[1] 102629.55434 15921.95384 1235.06633 66.42544
> valfpoiscont<-c(valf[1:3],sum(valf[4:6]));valfpoiscont
[1] 103704 14075 1766 306
> conchi(valfpoiscont,fteopoiscont,1)
[1] 1317.801
[1] 0
> fteonegbincont<-c(fteonegbin[1:5],sum(fteonegbin[6:8]));fteonegbincont
[1]1.037608e+05 1.392729e+04 1.873491e+03 2.522047e+02
[5] 3.396353e+01 5.287005e+00
> conchi(valfnegbincont,fteonegbincont,2)
[1] 12.77428
[1] 0.005151069
> valfnegbincont<-c(valf[1:5],sum(valf[6:7]));valfnegbincont
[1] 103704 14075 1766 255 45 8
> fteopoigcont<-c(fteopoig[1:5],sum(fteopoig[6:8]));fteopoigcont
[1]1.037134e+05 1.404919e+04 1.786072e+03 2.551488e+02
[5]4.062018e+01 8.559096e+00
> valfpoigcont<-c(valf[1:5],sum(valf[6:7]));valfpoigcont
[1] 103704 14075 1766 255 45 8
> conchi(valfpoigcont,fteopoigcont,2)
[1] 0.782696
[1] 0.8536013
```

En resumen, tenemos en la Tabla3, los siguientes resultados de la bondad del ajuste:

Tabla 3. Resultado de la bondad del ajuste de las distribuciones Poisson, Binomial Negativa y Poisson-Inversa Gaussiana para los datos de la Tabla 1 y con las frecuencias teóricas obtenidas en la Tabla 2.

	Poisson	Binomial Negativa	Poisson-Inversa Gaussiana
D	1317.801	12.77428	0.782696
p-value	0	0.005151069	0.8536013

Con estos resultados, puede afirmarse que la distribución Poisson-Inversa Gaussiana es la que mejor se ajusta a los datos.

3. Distribución del coste total en un período

El coste total, S , de todos los siniestros que ha tenido una póliza en un determinado período (usualmente un año), es:

$$S = \begin{cases} 0 & \text{si } N = 0 \\ X_1 + X_2 + \dots + X_N & \text{si } N \neq 0 \end{cases} \quad (5)$$

con N : v.a. número de siniestros en un período, y X_i : v.a. que indica el importe del siniestro i -ésimo. Si no hay siniestros, entonces $N = 0$ y $S = 0$. Por tanto, aunque los importes de los siniestros sean v.a. continuas, S será una v.a. mixta pues tiene un punto de probabilidad finita, el 0.

Se realizan, en la Teoría Colectiva del Riesgo, dos hipótesis básicas:

- (1) Dado un valor concreto, k , para la variable aleatoria N , número de siniestros, las variables aleatorias X_1, X_2, \dots, X_k , que nos indican la cuantía de cada uno de los siniestros, son mutuamente independientes e idénticamente distribuidas. Por ello, podemos hablar de la variable aleatoria X , cuantía de un siniestro.
- (2) Esta distribución común para la cuantía de un siniestro no depende del número de siniestros.
- (3) Con las dos hipótesis anteriores, el coste total en un período sigue una distribución compuesta. El objetivo es encontrar una expresión para la función de distribución de la variable aleatoria S coste total por período, $F_S(c) = P(S \leq c)$.

Incluiremos en este apartado:

- Cálculo exacto mediante convoluciones,
- Método de recurrencia de Panjer.

El concepto estadístico de convolución puede encontrarse en diversos libros de estadística o matemática actuarial (ver, por ejemplo, Bowers *et al.*, 1997). Dada la función de distribución $F_X(x)$ de una variable aleatoria X , se denomina convolución k -ésima de la misma a la función $F_X^{*(k)}(x)$, que da la probabilidad de que la suma $X_1 + X_2 + \dots + X_k$ sea no superior a x , siendo X_i independientes y equidistribuidas. Así, $F_X^{*(n)}(x) = P(X_1 + X_2 + \dots + X_n \leq x)$.

Si X es discreta, $P(i) = P[X = i]$, $P(X_1 + X_2 + \dots + X_n = i) = P^{*(n)}(i)$.

Para el cálculo exacto de la función de distribución del coste total mediante convoluciones, aplicando el teorema de la probabilidad total, se tiene:

$$F_S(s) = \sum_{k=0}^{\infty} P(N = k) \cdot F_X^{*(k)}(s) \quad \text{siendo } F_X^{*(0)}(s) = 1 \forall s. \quad (6)$$

Veamos un ejemplo práctico, que se puede encontrar en Bowers *et al.* (1997), en el que la cuantía de un siniestro es una variable discreta que puede tomar valores 1, 2 o 3, con probabilidades 0.25, 0.375 y 0.375, respectivamente. El número de siniestros sigue una distribución de Poisson, con $\lambda = 0.8$. Se pide calcular la función de distribución $F_S(s)$ para $s = 0, 1, \dots, 6$.

En primer lugar, calculamos las probabilidades $P(N=k)$ mediante la función de R correspondiente:

```
> pn<-dpois(0:6,0.8);pn
[1]0.4493289641 0.3594631713 0.1437852685 0.0383427383 0.0076685477
[6] 0.0012269676 0.0001635957
```

Se pueden obtener los cálculos de las convoluciones utilizando la función *aggregateDist* de la librería *actuar*. Como parámetros de la función, además de las probabilidades para el número de siniestros, necesitamos las probabilidades para la cuantía de un siniestro (incluyendo la probabilidad de que sea 0).

```

>library(actuar)
> fx<-c(0,0.25,0.375,0.375)
> Fs <- aggregateDist("convolution", model.freq=fn,model.sev=fx)
> Fs(0:6)
[1] 0.4493290 0.5391948 0.6829800 0.8453376 0.8952430 0.9426035
[7] 0.9735264

```

El método de recurrencia de Panjer (en Kaas *et al.*, 2008 se puede consultar con detalle) sirve para distribuciones del número de siniestros que cumplen la siguiente ecuación de recurrencia en el cálculo de sus probabilidades $p_k = P[N = k]$:

$$\frac{p_{k+1}}{p_k} = c + \frac{d}{k+1} \quad k = 0, 1, 2, \dots \quad (7)$$

Las únicas distribuciones que cumplen la recurrencia anterior son: Poisson, Binomial, Geométrica y Binomial Negativa.

Supondremos que la cuantía de los siniestros sólo puede ser positiva. En este caso la probabilidad de que el coste total sea 0 coincide con la probabilidad de que el número de siniestros sea también cero, $P[S = 0] = P[N = 0] = p_0$. El método recurrente que pasamos a desarrollar sirve entonces únicamente para valores del coste total positivos.

Si X es discreta, el coste total también lo será por lo que, en lugar de calcular su función de distribución, podemos calcular directamente sus probabilidades, es decir, su función de cuantía.

La fórmula recurrente de Panjer es:

$$P[S = s] = \sum_{y=1}^s \left(c + \frac{d \cdot y}{s} \right) \cdot P[X = y] \cdot P[S = s - y] \quad s = 1, 2, \dots \quad (8)$$

Si la cuantía de un siniestro X tiene un valor máximo r , la recurrencia anterior se simplifica y se reduce a:

$$P[S = s] = \sum_{y=1}^j \left(c + \frac{d \cdot y}{s} \right) \cdot P[X = y] \cdot P[S = s - y] \quad \text{siendo } j = \min\{s, r\}. \quad (9)$$

A título de ejemplo, vamos a recalculer la función de distribución del coste total para el ejemplo anterior. Se utiliza la misma función de la librería *actuar*, aunque los parámetros varían, en este caso se indica la distribución que sigue el número de siniestros y sus parámetros, así como las probabilidades para la cuantía de un siniestro. Hay que tener en cuenta que los resultados nos muestran la función de distribución y no la probabilidad para el coste total.

```

>Fs<-aggregateDist("recursive",model.freq="poisson",model.sev=fx,lambda=0.8)
> Fs(0:6)
[1] 0.4493290 0.5391948 0.6829800 0.8453376 0.8952430 0.9426035
[7] 0.9735264

```

Se puede observar que los resultados coinciden con los obtenidos mediante convoluciones.

4. Sistemas de tarificación

La tarificación es el proceso de determinación de las primas de tal forma que se correspondan a los siniestros a pagar. En los procesos de tarificación *a priori* se determinan una serie de factores de riesgo y niveles de los mismos con el objetivo de clasificar a los individuos en grupos más homogéneos respecto al riesgo. Así, se cobra a todos los individuos del grupo la misma prima común. Sin embargo, dentro de cada grupo no se ha eliminado totalmente la heterogeneidad, no son totalmente homogéneos, debido a varios factores (factores de riesgo insuficientes, factores inobservables, etc...) y, por tanto, si aplicamos a

todos los individuos del grupo la tasa media de siniestralidad común, habrán algunos beneficiados y otros perjudicados, provocando a largo plazo una antiselección.

Una manera de conseguir que cada asegurado pague una prima de acuerdo con su riesgo es introducir modificaciones *a posteriori* según la experiencia que vaya teniendo. Dentro del esquema general de “experience-rating” o tarificación *a posteriori* podemos destacar, entre otros, los modelos de credibilidad exacta y de credibilidad de distribución libre.

Más detalles sobre los modelos de credibilidad aplicados a la tarificación pueden consultarse, por ejemplo, en Goovaerts and Hoogstad (1987) y Kaas *et al.* (2008).

Nos centraremos, en este apartado, en un caso particular dentro de los modelos de credibilidad exacta, en el que la modificación se produce según el número de siniestros que tenga una póliza determinada y en los modelos básicos de credibilidad de distribución libre.

Modelo de credibilidad exacta: Gamma-Poisson-Gamma en tarificación a posteriori

Vamos a analizar el modelo más sencillo, conocido como Gamma-Poisson-Gamma, que nos proporcionará un sistema de *Bonus-Malus* basado en el número de siniestros. Un análisis detallado de los sistemas *Bonus-Malus* puede encontrarse en Lemaire (1985).

Variables:

- N : número de siniestros de una póliza en un período.
- N_t : número de siniestros de una póliza en t períodos.
- λ : número medio de siniestros de una póliza en un período.

Hipótesis:

- (1) La v.a. N , número de siniestros de una póliza en un periodo, conocido λ , sigue una distribución Poisson, $N|\lambda \sim Po(\lambda)$
- (2) El valor de λ de cada individuo es desconocido y asumimos que se distribuye según una Gamma , es decir, $\lambda \sim Ga(a = n \cdot h, b = h)$.

A partir de las hipótesis anteriores es fácil demostrar que tanto N como N_t siguen distribuciones Binomial Negativa.

La prima a priori es $E[\lambda] = n$, y la prima a posteriori, es decir, la esperanza a posteriori del número medio de siniestros condicionado a una cierta experiencia de siniestralidad, viene dada por:

$$E[\lambda|N_t = k] = \frac{t}{t+h} \cdot \frac{k}{t} + \frac{h}{t+h} \cdot n \tag{10}$$

siendo $z = \frac{t}{t+h}$ el factor de credibilidad de la experiencia individual. El porcentaje de aumento sobre el número medio aplicado inicialmente, es:

$$\frac{E[\lambda|N_t = k] - E[\lambda]}{E[\lambda]} = \frac{t}{t+h} \cdot \frac{k - n \cdot t}{n \cdot t} \tag{11}$$

Consideremos el siguiente ejemplo con los datos de la Tabla 4. Se ha observado una cartera de seguros de automóviles de una compañía belga durante un año y los datos sobre el número de siniestros y número de pólizas son (Lemaire, 1985):

Tabla 4. Datos sobre número de siniestros y número de pólizas extraídos de Lemaire (1985).

Número siniestros	Número pólizas
0	96978
1	9240
2	704
3	43
4	9
>4	0

Suponiendo que el primer año de incorporación de una nueva póliza se le aplica la tasa media de siniestralidad del grupo y después se establece un sistema *Bonus-Malus* que tiene en cuenta su

experiencia individual, vamos a obtener un cuadro que nos indique la esperanza del número medio de siniestros condicionado a una cierta experiencia individual dada por:

- número de años de experiencia de 1 a 7
- número total de siniestros en dichos años, de 0 a 4.

y los correspondientes porcentajes de reducción/aumento sobre la prima común de la clase.

Para ello, previamente (tal y como se ha comentado en el apartado del número de siniestros), hemos comprobado que el número de siniestros se puede modelizar con una distribución Binomial Negativa de parámetros $a = 1.604935$ y $b = 0.062298114$, siendo, a su vez, $n = a \cdot b = 0.1010806$ y

$h = \frac{1}{b} = 15.87777$, los parámetros que indican el número medio de siniestros y el coeficiente de

heterogeneidad, respectivamente. De esta forma, podemos aplicar el modelo Gamma-Poisson-Gamma.

El número medio de siniestros a aplicar en el cálculo de la prima del próximo período, teniendo en cuenta los años de experiencia individual, es:

$$E[\lambda/N_t = k] = \frac{t}{t+15.87777} \cdot \frac{k}{t} + \frac{15.87777}{t+15.87777} \cdot 0.1010806, \quad (12)$$

y el porcentaje de aumento sobre el número medio aplicado inicialmente,

$$\frac{E[\lambda/N_t = k] - E[\lambda]}{E[\lambda]} = \frac{t}{t+15.87777} \cdot \frac{k - 0.1010806 \cdot t}{0.1010806 \cdot t} \quad (13)$$

con $t = 1, \dots, 7$ y $k = 0, \dots, 8$.

Para calcular cómodamente los valores correspondientes a todas las combinaciones entre t y k , utilizaremos la función $bonusma(a,b,t,k)$ que se ha programado en R (ver Anexo para el código de la función $bonusma$):

```
> t <- 1:7; t
[1] 1 2 3 4 5 6 7
> k <- 0:4; k
[1] 0 1 2 3 4
> res <- bonusma(a,b,t,k); res
$h
[1] 15.87777
$numesp
  1  2  3  4  5  6  7
0 0.09509166 0.08977267 0.0850172 0.0807402 0.07687292 0.07335917 0.0701526
1 0.15434119 0.14570806 0.1379896 0.1310477 0.12477075 0.11906767 0.1138632
2 0.21359073 0.20164345 0.1909619 0.1813551 0.17266859 0.16477617 0.1575737
3 0.27284027 0.25757884 0.2439343 0.2316626 0.22056643 0.21048467 0.2012843
4 0.33208981 0.31351423 0.2969066 0.2819700 0.26846427 0.25619317 0.2449948
$porcen
  1  2  3  4  5  6  7
0 94.07505 88.81292 84.10829 79.87702 76.05108 72.5749 69.40261
1 152.69116 144.15032 136.51433 129.64665 123.43685 117.7947 112.64586
2 211.30727 199.48771 188.92038 179.41627 170.82262 163.0146 155.88912
3 269.92338 254.82511 241.32642 229.18590 218.20839 208.2344 199.13237
4 328.53949 310.16250 293.73246 278.95553 265.59416 253.4543 242.37563
```

Por ejemplo:

- El factor de credibilidad que se da a la experiencia individual obtenida durante 4 años, es:

$$\frac{t}{t+h} = \frac{4}{4+15.87777} = 0.2012298. \quad (14)$$

- El número medio de siniestros para calcular la prima del próximo período, teniendo en cuenta un individuo que ha tenido 3 siniestros en 4 años asciende a:

$$E[\lambda/N_t = k] = \frac{4}{4+15.87777} \cdot \frac{3}{4} + \frac{15.87777}{4+15.87777} \cdot 0.1010806 = 0.2316626. \quad (15)$$

- El porcentaje de aumento de este individuo, sobre el número medio aplicado inicialmente, es igual a:

$$\frac{t}{t+h} \cdot \frac{k-n \cdot t}{n \cdot t} = 0.2012298 \cdot \frac{3-0,1010806 \cdot 4}{0.1010806 \cdot 4} = 1.291859. \quad (16)$$

Cálculo de primas con los modelos de credibilidad de distribución libre

Mientras que en los modelos de credibilidad exacta es necesario conocer la distribución a priori del parámetro de riesgo, en la fórmula de credibilidad de distribución libre desarrollada por Bühlmann (1967) no se hacen supuestos acerca del tipo de función de distribución del parámetro de riesgo individual, ni es necesario introducir ninguna distribución a priori del mismo.

El **modelo de Bühlmann** puede ser descrito a través de las siguientes variables:

θ_j : Parámetro de riesgo para la póliza j -ésima, con $j = 1, 2, \dots, k$ siendo k el número total de pólizas de la cartera.

X_{js} : Variable que indica la experiencia de reclamaciones para la póliza j -ésima en el periodo s , donde $j = 1, 2, \dots, k$ y $s = 1, 2, \dots, t$, siendo t el número de periodos observados para cada póliza. Se considera que X_{js} es una variable aleatoria pero con realizaciones observables x_{js} . Puede interpretarse X_{js} como el importe medio de las indemnizaciones por siniestro o también como frecuencia de siniestralidad.

Así, la póliza j -ésima vendrá descrita por el siguiente vector:

$$(\theta_j, X_{j1}, X_{j2}, \dots, X_{jt}) = (\theta_j, \bar{X}_j). \quad (17)$$

Hipótesis:

(1) Las pólizas $j = 1, 2, \dots, k$, es decir, los pares $(\theta_1, \bar{X}_1), (\theta_2, \bar{X}_2), \dots, (\theta_k, \bar{X}_k)$ son independientes y están idénticamente distribuidos.

(2) Para cada póliza $j = 1, 2, \dots, k$ y para un θ_j dado, las variables condicionadas $X_{j1} | \theta_j, X_{j2} | \theta_j, \dots, X_{jt} | \theta_j$ son independientes y están idénticamente distribuidas.

La primera hipótesis es la de independencia entre las pólizas y la segunda hipótesis refleja la independencia dentro de una póliza y la homogeneidad en el tiempo, es decir, las v.a. X_{js} conservan la media y la varianza en el tiempo.

Bühlmann estableció que lo que le interesaba era que las primas fuesen lineales. De acuerdo con el criterio de los mínimos cuadrados, se denomina estimador de credibilidad, $\hat{\mu}(\theta_j)$:

$$\hat{\mu}(\theta_j) = (1-z) \cdot m + z \cdot \bar{X}_j \text{ donde } z = \frac{a \cdot t}{S^2 + a \cdot t}, \bar{X}_j = \sum_{s=1}^t \frac{X_{js}}{t}. \quad (18)$$

En la fórmula del estimador ajustado por credibilidad aparece el parámetro z , denominado factor de credibilidad, que toma valores comprendidos entre 0 y 1, y es único para toda la cartera.

El **modelo de Bühlmann-Straub** es una ampliación del modelo de Bühlmann, en el cual se mantiene la homogeneidad en el tiempo pero se incorpora un peso o ponderación a cada observación, W_{js} , que son números positivos conocidos, con $j = 1, 2, \dots, k$ y $s = 1, 2, \dots, t$.

Las observaciones esperadas siguen siendo homogéneas en el tiempo, pero la varianza deja de serlo, pasando a depender del período considerado a través de los pesos o ponderaciones. Se obtendrá, además, un factor de credibilidad distinto para cada póliza, a diferencia del modelo de Bühlmann, en el cual era único para toda la cartera. Así, el estimador ajustado por credibilidad es:

$$\hat{\mu}(\theta_j) = (1-z_j) \cdot m + z_j \cdot X_{jw}, \text{ con } X_{jw} = \sum_{s=1}^t \frac{X_{js} \cdot W_{js}}{W_{j\bullet}}, z_j = \frac{a \cdot W_{j\bullet}}{S^2 + a \cdot W_{j\bullet}} \text{ y } W_{j\bullet} = \sum_{s=1}^t W_{js}. \quad (19)$$

A continuación mostramos un ejemplo, que fue desarrollado por Bühlmann and Straub (1970) (Una traducción de este artículo a inglés se encuentra en *Actuarial Research Clearing House*, 1972, nº 2, con título *Credibility for loss ratios*).

La información sobre la cuantía de los siniestros de un año para una cartera de pólizas, divididas en 7 clases y durante los últimos 5 años es la siguiente:

Tabla 5. Datos sobre cuantías de los siniestros de un año para una cartera de pólizas, divididas en 7 clases y durante los últimos 5 años.

	1		2		3		4		5		6		7	
	P	X	P	X	P	X	P	X	P	X	P	X	P	X
1	5	0	14	11.3	18	8	20	5.4	21	9.7	43	9.7	70	9
2	6	0	14	25	20	1.9	22	5.9	24	8.9	47	14.5	77	9.6
3	8	4.2	13	18.5	23	7	25	7.1	28	6.7	53	10.8	85	8.7
4	10	0	11	14.3	25	3.1	29	7.2	34	10.3	61	12	92	11.7
5	12	7.7	10	30	27	5.2	35	8.3	42	11.1	70	13.1	100	7
	41	3.1	62	19.5	111	5	131	7	149	9.5	274	1.2.1	424	9.2

Siendo P el volumen de primas y X la cuantía de los siniestros, expresada como tanto por ciento de los capitales en riesgo.

Vamos a estimar el valor de X para el próximo período, en cada uno de los grupos. Dicho valor nos indicará la prima a cobrar en función del capital asegurado. Para ello, aplicaremos el método de Bühlmann y de Bühlmann-Straub. Las funciones que nos proporcionan los resultados para ambos modelos se hallan incluidas en la librería *actuar* de R.

Modelo de Bühlmann

```

> library(actuar)
> state <- c(1,2,3,4,5,6,7)
> ratio.1 <- c(0,11.3,8,5.4,9.7,9.7,9)
> ratio.2 <- c(0,25,1.9,5.9,8.9,14.5,9.6)
> ratio.3 <- c(4.2,18.5,7,7.1,6.7,10.8,8.7)
> ratio.4 <- c(0,14.3,3,1,7.2,10.3,12,11.7)
> ratio.5 <- c(7.7,30,5.2,8.3,11.1,13.1,7)
> weight.1 <- c(5,14,18,20,21,43,70)
> weight.2 <- c(6,14,20,22,24,47,77)
> weight.3 <- c(8,13,23,25,28,53,85)
> weight.4 <- c(10,11,25,29,34,61,92)
> weight.5 <- c(12,10,27,35,42,70,100)
> dataBS <- cbind(state,ratio.1,ratio.2,ratio.3,ratio.4,ratio.5,
+ weight.1,weight.2,weight.3,weight.4,weight.5); dataBS
  state ratio.1 ratio.2 ratio.3 ratio.4 ratio.5 weight.1 weight.2 weight.3 weight.4 weight.5
[1,]  1  0.0   0.0   4.2  0.0  7.7    5    6    8   10   12
[2,]  2 11.3  25.0  18.5 14.3 30.0   14   14   13   11   10
[3,]  3  8.0   1.9   7.0  3.1  5.2   18   20   23   25   27
[4,]  4  5.4   5.9   7.1  7.2  8.3   20   22   25   29   35
[5,]  5  9.7   8.9   6.7 10.3 11.1   21   24   28   34   42
[6,]  6  9.7  14.5  10.8 12.0 13.1   43   47   53   61   70
> B <- cm(~state, dataBS, ratios = ratio.1:ratio.5)
> summary(B)
Call:
cm(formula = ~state, data = dataBS, ratios = ratio.1:ratio.5)
Structure Parameters Estimators
Collective premium: 9.225714
Between state variance: 29.22030
Within state variance: 12.587
Detailed premiums
Level: state

```

```

state Indiv. mean Weight Cred. factor Cred. premium
1 2.38 5 0.920681 2.922995
2 19.82 5 0.920681 18.979673
3 5.04 5 0.920681 5.372006
4 6.78 5 0.920681 6.973991
5 9.34 5 0.920681 9.330935
6 12.02 5 0.920681 11.798360
7 9.20 5 0.920681 9.202040
> predict(B)
[1] 2.922995 18.979673 5.372006 6.973991 9.330935 11.798360
[7] 9.202040

```

Según estos resultados, para cada uno de los grupos, el valor de X para el período 6 será:

$$\hat{\mu}(\theta_1) = 2.922995, \hat{\mu}(\theta_2) = 18.979673, \hat{\mu}(\theta_3) = 5.372006,$$

$$\hat{\mu}(\theta_4) = 6.973991, \hat{\mu}(\theta_5) = 9.330935, \hat{\mu}(\theta_6) = 11.798360, \hat{\mu}(\theta_7) = 9.202040.$$

Estos valores se han obtenido con un factor de credibilidad, igual para todos los grupos, cuyo valor es $z = 0.920681$.

Modelo de Bühlmann-Straub

```

> BS <- cm(~state,dataBS,ratios=ratio.1:ratio.5,weights=weight.1:weight.5)
> summary(BS)
Call:
cm(formula = ~state, data = dataBS, ratios = ratio.1:ratio.5,
weights = weight.1:weight.5)
Structure Parameters Estimators
Collective premium: 9.379879
Between state variance: 12.45453
Within state variance: 216.0749
Detailed premiums
Level: state
state Indiv. mean Weight Cred. factor Cred. premium
1 3.073171 41 0.7026672 4.948362
2 19.451613 62 0.7813573 17.249502
3 4.963717 113 0.8669028 5.551496
4 6.981679 131 0.8830522 7.262144
5 9.538926 149 0.8957067 9.522339
6 12.116788 274 0.9404525 11.953812
7 9.162972 424 0.9606908 9.171498
> predict(BS)
[1]4.948362 17.249502 5.551496 7.262144 9.522339 11.953812 9.171498

```

Se ha obtenido que para cada grupo, para el período 6, el valor de X es:

$$\hat{\mu}(\theta_1) = 4.948362, \hat{\mu}(\theta_2) = 17.249502, \hat{\mu}(\theta_3) = 5.551496, \hat{\mu}(\theta_4) = 7.262144,$$

$$\hat{\mu}(\theta_5) = 9.522339, \hat{\mu}(\theta_6) = 11.953812, \hat{\mu}(\theta_7) = 9.171498.$$

A diferencia del modelo de Bühlmann, como ya hemos indicado, a cada grupo le corresponde un factor de credibilidad distinto, que viene indicado como z_j . Los valores obtenidos para estos factores han sido:

$$z_1 = 0.7026672, \quad z_2 = 0.7813573, \quad z_3 = 0.8669028, \quad z_4 = 0.8830522,$$

$$z_5 = 0.8957067, \quad z_6 = 0.9404525, \quad z_7 = 0.9606908.$$

5. Cálculo de la provisión técnica de prestaciones pendientes

La provisión de prestaciones pendientes es el importe total de las obligaciones pendientes del asegurador derivadas de los siniestros ocurridos con anterioridad a la fecha de cierre del ejercicio.

Aplicaremos en este apartado, en primer lugar, diferentes métodos estadísticos de los denominados deterministas, es decir, aquellos que utilizan información sobre los siniestros pagados en el pasado para estimar los valores futuros, sin la intervención de variables aleatorias. Los datos básicos necesarios son:

- c_{ij} : cuantía pagada en el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i
- C_{ij} : cuantía acumulada pagada hasta (e incluido) el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i : $i \in I, j \in J, \{i, j / i \in I, j \in J\} = D$,

siendo D el dominio (años de origen, años de desarrollo para los que tenemos datos). Si $i = 0, 1, \dots, k$; $j = 0, 1, \dots, k$, el dominio D tiene forma triangular, denominándose triángulo de desarrollo.

Por F representamos el conjunto de combinaciones de i, j que permiten completar el cuadrado. A continuación resumimos las características más importantes de un grupo de métodos estadísticos que consideramos como clásicos, por ser de los primeros planteados, y que por su sencillez son de los más utilizados en la práctica. Estos métodos pueden encontrarse desarrollados en Taylor (1986).

Tabla 6. Características más importantes de los métodos de Chain-ladder y sus variantes, y del método de mínimos cuadrados de de Vylder.

Método	Datos	Estimación parámetros	\hat{C}_{ij} $i, j \in F$
Chain-ladder	C_{ij}	$\hat{m}_h = \frac{\sum_{i=0}^{k-h-1} C_{i,h+1}}{\sum_{i=0}^{k-h-1} C_{i,h}}$	$\hat{C}_{i,j} = C_{i,k-i} \cdot \prod_{h=k-i}^{j-1} \hat{m}_h$
Variantes de Chain-ladder	$d_{ij} = \frac{C_{i,j+1}}{C_{ij}}$ $i, j = 0, 1, \dots, k-1$	Ajuste tendencias lineales	$\hat{C}_{i,j} = C_{i,k-i} \cdot \prod_{h=k-i}^{j-1} \hat{d}_h$
		$\hat{d}_h = \frac{\sum_{i=0}^{k-h-1} w_{ih} \cdot d_{ih}}{\sum_{i=0}^{k-h-1} w_{ih}}$	
Mínimos Cuadrados De Vylder	c_{ij} Hipótesis: $c_{ij} = x_i \cdot p_j$ $\sum_{j=0}^k p_j = 1$	$\text{Min} \sum_{\forall i, j \in D} (c_{ij} - x_i \cdot p_j)^2$ $\hat{x}_i = \frac{\sum_{j \in J_i} c_{ij} \cdot p_j}{\sum_{j \in J_i} p_j^2}$ $\hat{p}_j = \frac{\sum_{\forall i \in I_j} c_{ij} \cdot x_i}{\sum_{\forall i \in I_j} x_i^2}$	$\hat{c}_{ij} = \hat{x}_i \cdot \hat{p}_j$

Para poder aplicar estos métodos de cálculo utilizaremos los datos de la Tabla 7 que hacen referencia a una cartera de pólizas de seguros de automóviles de Singapur entre 1997 y 2001, extraídos de Frees (2009), relativos a los pagos anuales, en miles, por daños a terceros:

Tabla 7. Datos sobre una cartera de pólizas de seguros de automóviles de Singapur entre 1997 y 2001, extraídos de Frees (2009), relativos a pagos anuales, en miles, por daños a terceros.

	Año de origen <i>i</i> / Año desarrollo <i>j</i>				
	0	1	2	3	4
1997	1.188.675	2.257.909	695.237	166.812	92.129
1998	1.235.402	3.250.013	649.928	211.344	
1999	2.209.850	3.718.695	818.367		
2000	2.662.546	3.487.034			
2001	2.457.265				

Se han elaborado distintos programas para poder realizar los cálculos en cada método, nos remitimos al Anexo para ver el detalle de las funciones *ibnrchl*, *ibnrchlvar1*, *ibnrchlvar2*, *ibnrvylder*). En cada caso, el resultado que muestra el programa se corresponde al importe total de la provisión y al importe de la misma desglosada por año futuro de pago.

Método de Chain-ladder

```
> c0<-c(1188675,2257909,695237,166812,92129)
> c1<-c(1235402,3250013,649928,211344)
> c2<-c(2209850,3718695,818367)
> c3<-c(2662546,3487034)
> c4<-2457265
> ibnrchl(c(c0,c1,c2,c3,c4))
[1] "triángulo cuantías"
  [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 2257909 695237 166812 92129
[2,] 1235402 3250013 649928 211344 0
[3,] 2209850 3718695 818367 0 0
[4,] 2662546 3487034 0 0 0
[5,] 2457265 0 0 0 0
[1] "triángulo cuantías acumuladas"
  [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 3446584 4141821 4308633 4400762
[2,] 1235402 4485415 5135343 5346687 5346687
[3,] 2209850 5928545 6746912 6746912 6746912
[4,] 2662546 6149580 6149580 6149580 6149580
[5,] 2457265 2457265 2457265 2457265 2457265
[1] "factores de desarrollo"
[1] 2.742438 1.156093 1.040762 1.021382
[1] "rectángulo con cuantías acumuladas"
  [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 3446584 4141821 4308633 4400762
[2,] 1235402 4485415 5135343 5346687 5461012
[3,] 2209850 5928545 6746912 7021930 7172076
[4,] 2662546 6149580 7109486 7399283 7557497
[5,] 2457265 6738897 7790790 8108359 8281735
[1] "rectángulo con cuantías desacumuladas"
  [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 2257909 695237.0 166812.0 92129.0
```

```
[2,] 1235402 3250013 649928.0 211344.0 114325.1
[3,] 2209850 3718695 818367.0 275017.8 150145.9
[4,] 2662546 3487034 959905.6 289797.0 158214.6
[5,] 2457265 4281632 1051893.7 317568.4 173376.3
[1] "provisiones (suma de pagos futuros)"
[1] 7771876
[1] "vector pagos futuros"
[1] 5630880.1 1491836.6 475783.0 173376.3
```

Método de Chain-ladder, variante I

```
> ibnrchlvar1(c(c0,c1,c2,c3,c4))
[1] "triángulo cuantías"
      [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 2257909 695237 166812 92129
[2,] 1235402 3250013 649928 211344 0
[3,] 2209850 3718695 818367 0 0
[4,] 2662546 3487034 0 0 0
[5,] 2457265 0 0 0 0
[1] "triángulo cuantías acumuladas"
      [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 3446584 4141821 4308633 4400762
[2,] 1235402 4485415 5135343 5346687 5346687
[3,] 2209850 5928545 6746912 6746912 6746912
[4,] 2662546 6149580 6149580 6149580 6149580
[5,] 2457265 2457265 2457265 2457265 2457265
[1] "matriz factores desarrollo"
      [,1] [,2] [,3] [,4]
[1,] 2.899518 1.201718 1.040275 1.021382
[2,] 3.630733 1.144898 1.041155 1.000000
[3,] 2.682782 1.138038 1.000000 1.000000
[4,] 2.309662 1.000000 1.000000 1.000000
[5,] 1.000000 1.000000 1.000000 1.000000
[1] "matriz factores desarrollo completa"
      [,1] [,2] [,3] [,4]
[1,] 2.899518 1.201718 1.040275 1.021382
[2,] 3.630733 1.144898 1.041155 1.021382
[3,] 2.682782 1.138038 1.040715 1.021382
[4,] 2.309662 1.097872 1.040715 1.021382
[5,] 2.201294 1.066032 1.040715 1.021382
[1] "rectángulo con cuantías acumuladas"
      [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 3446584 4141821 4308633 4400762
[2,] 1235402 4485415 5135343 5346687 5461012
[3,] 2209850 5928545 6746912 7021612 7171751
[4,] 2662546 6149580 6751452 7026337 7176577
[5,] 2457265 5409162 5766342 6001118 6129437
[1] "rectángulo con cuantías desacumuladas"
      [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 2257909 695237.0 166812.0 92129.0
[2,] 1235402 3250013 649928.0 211344.0 114325.1
```

```
[3,] 2209850 3718695 818367.0 274700.0 150139.1
[4,] 2662546 3487034 601872.4 274884.8 150240.1
[5,] 2457265 2951897 357180.4 234776.1 128318.4
[1] "provisiones (suma de pagos futuros)"
[1] 5238333
[1] "vector pagos futuros"
[1] 3942794.1 782204.3 385016.2 128318.4
```

Método de Chain-ladder, variante II

En este caso, los parámetros del modelo se estiman a partir de medias ponderadas, es decir, debemos elegir qué pesos o ponderaciones, w_{ij} , se aplican. El programa creado permite seleccionar entre distintas opciones de pesos, en este ejemplo consideraremos que $w_{ij} = i + j + 1$.

```
> ibnrchlvar2(c(c0,c1,c2,c3,c4))
[1] "triángulo cuantías"
      [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 2257909 695237 166812 92129
[2,] 1235402 3250013 649928 211344 0
[3,] 2209850 3718695 818367 0 0
[4,] 2662546 3487034 0 0 0
[5,] 2457265 0 0 0 0
[1] "triángulo cuantías acumuladas"
      [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 3446584 4141821 4308633 4400762
[2,] 1235402 4485415 5135343 5346687 5346687
[3,] 2209850 5928545 6746912 6746912 6746912
[4,] 2662546 6149580 6149580 6149580 6149580
[5,] 2457265 2457265 2457265 2457265 2457265
[1] "matriz factores desarrollo"
      [,1] [,2] [,3] [,4]
[1,] 2.899518 1.201718 1.040275 1.021382
[2,] 3.630733 1.144898 1.041155 1.000000
[3,] 2.682782 1.138038 1.000000 1.000000
[4,] 2.309662 1.000000 1.000000 1.000000
[5,] 1.000000 1.000000 1.000000 1.000000
[1] "opciones de ponderaciones: 1: wij=1 2: wij=i+j+1 3: wij=(i+j+1)^2 4: wij=2^(i+j+1)"
[1] "entrar opción: 1,2,3 o 4"
1: 2
2:
Read 1 item
[1] "matriz estimaciones de dij"
      [,1] [,2] [,3] [,4]
[1,] 2.744797 1.154476 1.040778 1.021382
[2,] 2.744797 1.154476 1.040778 1.021382
[3,] 2.744797 1.154476 1.040778 1.021382
[4,] 2.744797 1.154476 1.040778 1.021382
[5,] 2.744797 1.154476 1.040778 1.021382
[1] "rectángulo con cuantías acumuladas"
      [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 3446584 4141821 4308633 4400762
[2,] 1235402 4485415 5135343 5346687 5461012
```

```
[3,] 2209850 5928545 6746912 7022036 7172184
[4,] 2662546 6149580 7099542 7389045 7547041
[5,] 2457265 6744695 7786588 8104107 8277393
[1] "rectángulo con cuantías desacumuladas"
     [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 2257909 695237 166812.0 92129.0
[2,] 1235402 3250013 649928 211344.0 114325.1
[3,] 2209850 3718695 818367 275123.9 150148.1
[4,] 2662546 3487034 949962 289503.4 157995.7
[5,] 2457265 4287430 1041893 317519.6 173285.4
[1] "provisiones (suma de pagos futuros)"
[1] 7757186
[1] "vector pagos futuros"
[1] 5626840.9 1481544.5 475515.3 173285.4
```

Método de mínimos cuadrados de De Vylder

```
> ibnrvylder(c(c0,c1,c2,c3,c4))
[1] "triángulo cuantías"
     [,1] [,2] [,3] [,4] [,5]
[1,] 1188675 2257909 695237 166812 92129
[2,] 1235402 3250013 649928 211344 0
[3,] 2209850 3718695 818367 0 0
[4,] 2662546 3487034 0 0 0
[5,] 2457265 0 0 0 0
[1] "Resultados intermedios"
[1] "xi"
[1] 4319575 5713616 7212040 7284444 8085469
[1] "pj"
[1] 0.30391124 0.51504421 0.12213462 0.03758167 0.02132825
[1] "rectángulo con cuantías totales"
     [,1] [,2] [,3] [,4] [,5]
[1,] 1312767 2224772 527569.7 162336.9 92129.0
[2,] 1736432 2942765 697830.3 214727.3 121861.5
[3,] 2191820 3714519 880839.8 271040.5 153820.2
[4,] 2213824 3751810 889682.7 273761.6 155364.5
[5,] 2457265 4164374 987515.7 303865.5 172448.9
[1] "provisiones (suma de pagos futuros)"
[1] 7493735
[1] "vector pagos futuros"
[1] 5446958.8 1415097.5 459229.9 172448.9
```

A continuación aplicaremos la versión estocástica del modelo de Chain-ladder siguiendo a Mack (1993), en el cual, el factor de cambio de columna \hat{m}_h de Chain-ladder se corresponde con el valor esperado del factor de desarrollo supuesto conocido el importe total pagado hasta el período $h-1$ (incluido). El modelo de Mack tiene la ventaja de ser un modelo de distribución libre, de forma que no hace hipótesis sobre la distribución de las variables aleatorias incluidas, sino tan sólo hipótesis generales, válidas para cualquier distribución.

Tabla 8. Características más importantes del método de Chain-ladder estocástico.

Método	Datos	Estimación Parámetros	\hat{C}_{ij} $i, j \in F$
Chain-ladder estocástico	C_{ij} Hipótesis: $E \left[\frac{C_{i,h+1}}{C_{i,h}} \middle C_{i,h} \right] = m_h$ $V \left[\frac{C_{i,h+1}}{C_{i,h}} \middle C_{i,h} \right] = \sigma^2(h)$ c_{i_1, j_1} y c_{i_2, j_2} son v.a independientes para $i_1 \neq i_2$	$\hat{m}_h = \frac{\sum_{i=0}^{k-h-1} C_{i,h+1}}{\sum_{i=0}^{k-h-1} C_{i,h}}$ $\sigma^2(h) = \frac{1}{k-h} \sum_{i=0}^{k-h-1} C_{i,h} \left(\frac{C_{i,h+1}}{C_{i,h}} - \hat{m}_h \right)^2$ $h = 0, \dots, k-2$ (para k-1 Mack, 1993) propone diversos métodos)	$\hat{C}_{i,j} = C_{i,k-i} \cdot \prod_{h=k-i}^{j-1} \hat{m}_h$

En este modelo estocástico, a partir de las hipótesis, se obtienen también estimadores para el error cuadrático medio de las reservas estimadas, por año de origen y del total. Vamos a aplicar este método a los datos anteriores de Frees (2009) y utilizaremos las funciones de la librería *ChainLadder*. Los datos deben introducirse en una matriz, que contenga los pagos acumulados por año de origen en cada fila y con NA en el caso en que no se haya realizado aún el pago correspondiente.

```

> C0<-cumsum(c0)
> C1<-c(cumsum(c1),NA)
> C2<-c(cumsum(c2),NA,NA)
> C3<-c(cumsum(c3),NA,NA,NA)
> C4<-c(c4,NA,NA,NA,NA)
> C<-rbind(C0,C1,C2,C3,C4);C
  [,1] [,2] [,3] [,4] [,5]
C0 1188675 3446584 4141821 4308633 4400762
C1 1235402 4485415 5135343 5346687 NA
C2 2209850 5928545 6746912 NA NA
C3 2662546 6149580 NA NA NA
C4 2457265 NA NA NA NA
> mch<-MackChainLadder(C)
Warning message:
In Mack.S.E(CL[["Models"]], FullTriangle, est.sigma = est.sigma, :
'loglinear' model to estimate sigma_n doesn't appear appropriate.
p-value > 5.
est.sigma will be overwritten to 'Mack'.
Mack's estimation method will be used instead.
> mch
MackChainLadder(Triangle = C)

Latest Dev.To.Date Ultimate IBNR Mack.S.E CV(IBNR)
C0 4,400,762 1.000 4,400,762 0 0.00e+00 NaN
C1 5,346,687 0.979 5,461,012 114,325 8.83e+01 0.000773
C2 6,746,912 0.941 7,172,076 425,164 4.65e+03 0.010928
C3 6,149,580 0.814 7,557,497 1,407,917 2.20e+05 0.156460
C4 2,457,265 0.297 8,281,735 5,824,470 1.60e+06 0.274325

Totals
Latest: 25,101,206.00
Dev: 0.76
Ultimate: 32,873,081.95
IBNR: 7,771,875.95
Mack S.E.: 1,623,032.41
CV(IBNR): 0.208834060636341
    
```

6. Conclusiones

El presente trabajo recoge los frutos de la investigación realizada sobre los sistemas informáticos aplicados a MANV. Dentro de la diversidad de lenguajes posibles, las autoras se decantaron por el R. Entre las razones del cambio destacamos el hecho de que se empezaron a encontrar libros y artículos internacionales que incluían ejemplos actuariales solucionados con R o que presentaban nuevas librerías especializadas en dicha temática, pensadas tanto para la práctica profesional como para la enseñanza y la investigación.

La apuesta por este lenguaje de programación ha sido arriesgada (Claramunt *et al.*, 2012), pues ha implicado trabajo extra en la preparación de la asignatura pero el resultado ha sido muy satisfactorio tanto para los estudiantes como para el profesorado. El nuevo lenguaje de programación no ha representado un problema para los alumnos en el seguimiento de la asignatura, los cuales han considerado, en general, que el conocimiento de un software libre como R les ha proporcionado más autonomía que el conocimiento del lenguaje que antes se utilizaba, el APL2, y les ha aportado un plus para la entrada en el mercado laboral, debido a que R es un lenguaje utilizado actualmente en las Compañías de Seguros a diferencia del APL2.

En este trabajo hemos presentado las principales cuestiones de MANV resueltas a través de ejemplos con datos reales, seleccionados de la literatura actuarial, mediante la aplicación de las librerías actuariales existentes hasta el momento junto con funciones de elaboración propia. Por lo tanto, una de las conclusiones de nuestro trabajo es que R puede ser un buen lenguaje de programación a utilizar en los cálculos actuariales de no vida y en su enseñanza.

La aparición de libros de matemática actuarial en los que la parte de cálculo se realiza en R es cada vez más habitual. Así, la incorporación en la docencia de la asignatura de MANV de paquetes diseñados por investigadores/docentes de universidades no españolas implica que la enseñanza que realizamos de MANV en la Universidad de Barcelona es de alto nivel y se encuentra en sintonía con la del resto de países europeos, lo que está en concordancia con la tendencia de los colegios profesionales.

Agradecimientos

Trabajo financiado por el Instituto de Ciencias de la Educación (ICE) de la Universidad de Barcelona con el proyecto A0801-06; por el Programa de Mejora de la Innovación Universitaria (PMID) de la Universidad de Barcelona con el proyecto 2009PID-UB/32; y por el Ministerio de Educación y Ciencia (MEC) mediante los proyectos MTM2010-17323 y ECO2010-22065-C03-03; y finalmente, por la Agencia de Gestión de Ayudas Universitarias y de Investigación (AGAUR) de la Generalitat de Catalunya con el proyecto 2009SGR970.

Las autoras agradecen los comentarios realizados por el editor adjunto y el evaluador anónimo, que han contribuido a la mejora del trabajo.

Referencias bibliográficas

1. M. M. Claramunt, E. Boj, T. Costa y M. Mármol, Enseñanza práctica de la Matemática Actuarial No Vida con R: Análisis de una experiencia innovadora. Ponencia presentada en el *VI Congreso Internacional de Docencia Universitaria e Innovación CIDUI2010*, celebrado en Barcelona, el 30 de junio, 1 y 2 de julio de 2010.
2. M. M. Claramunt, A. Alegre, E. Boj, T. Costa, M. Mármol y I. Morillo, Enseñanza práctica de la Matemática Actuarial No-Vida con R: Experiencia innovadora en la Universidad de Barcelona, *REIRE. Revista d'Innovació i Recerca en Educació* **5**, 1 (2012) 123–139.

3. T. Costa, A. Alegre, E. Boj, M. M. Claramunt, M. Mármol y I. Morillo, Una reflexió sobre la relació entre el grau de presencialitat i el procés d'aprenentatge-avaluació, *RIDU: Revista d'Innovació Docent Universitària* **1** (2009) 16–26.
4. R. Kaas, M. Goovaerts, J. Dhaene and M. Denuit, *Modern Actuarial Risk Theory Using R. Second Edition* (Springer, Heidelberg, 2008).
5. Ch. Dutang, V. Goulet and M. Pigeon, *actuar: An R Package for Actuarial Science*, *Journal of Statistical Software* **25**, 7 (2008) 1–37.
6. V. Goulet and L. P. Pouliot, Simulation of compound hierarchical models in R. *North American Actuarial Journal* **12**, 4 (2009) 401–412.
7. M. Gesmann, *Package 'ChainLadder'. Mack-, Bootstrap and Munich-chain-ladder methods for insurance claims reserving* (2009). <http://code.google.com/p/chainladder/>
8. H. H. Panjer and G. E. Willmot, *Insurance risk models* (The Society of Actuaries, 1992).
9. H. Bühlmann, *Mathematical Models in Risk Theory* (Springer-Berlag, New York, 1970).
10. N. L. Bowers, H. U. Gerber, J. C. Hickman, D. A. Jones and C. J. Nesbitt, *Actuarial Mathematics* (The Society of Actuaries, 1997).
11. M. J. Goovaerts and W. J. Hoogstad, *Credibility Theory*, *Surveys of Actuarial Studies* **4**, (Nationale-Nederlanden N. V., 1987).
12. J. Lemaire, *Automobile Insurance: Actuarial Models* (Kluwer-Nijhof Publishing, Boston, MA, 1985).
13. H. Bühlmann, *Experience rating and credibility*, *ASTIN Bulletin* **4** (1967) 119–207.
14. H. Bühlmann and E. Straub, Glaubwürdigkeit für Schadensätze, *Bulletin of the Association of Swiss Actuaries* **70**, 1 (1970) 111–113.
15. G.C. Taylor, *Claim reserving in non-life insurance*, (North-Holland, 1986).
16. E.W. Frees, *Regression modeling with actuarial and financial applications*, (Cambridge University Press, 2009).
17. T. Mack, *Distribution-free calculation of the standard error of chain-ladder reserve estimates*, *Astin Bulletin* **23** (1993) 213-225.

Anexo. Listado de las funciones de elaboración propia.

conchi

```
conchi<-function(valfcont,fteocont,par) {
D<- sum((valfcont-fteocont)^2/fteocont)
grad<-length(valfcont)-par-1
pvalue<-1-pchisq(D,grad)
print(D)
print(pvalue)
}
```

bonusma

```
bonusma<-function(a,b,t,k) {
h <- 1/b; h
n <- a*b; n
kmax <- max(k)
tmax <- max(t)
kmax1 <- kmax+1
numesp <- matrix(0,nrow=kmax+1,ncol=tmax)
porcen <- matrix(0,nrow=kmax+1,ncol=tmax)
for (i in 1:kmax1){
for (j in 1:tmax){
numesp[i,j] <- (j/(j+h))*((i-1)/j)+(h/(j+h))*n
porcen[i,j] <- (j/(j+h))*((i-1-n*j)/(n*j))
}
}
porcen <- (1+porcen)*100
colnames(numesp) <- t
rownames(numesp) <- k
colnames(porcen) <- t
rownames(porcen) <- k
return(list(h=h,numesp=numesp,porcen=porcen))
}
```

ibnrchl

```
ibnrchl<-function(Xij) {
TT<-trunc(sqrt(2*length(Xij)))
i<-rep(1:TT,TT:1)
j<-sequence(TT:1)
Xij.mat.cum<-Xij.mat<-matrix(0,nrow=TT,ncol=TT)
for(k in 1:length(Xij)) Xij.mat[i[k],j[k]]<-Xij[k]
for(k in 1:TT) Xij.mat.cum[k,]<-cumsum(Xij.mat[k,])
print("triángulo cuantías")
print(Xij.mat)
print("triángulo cuantías acumuladas")
print(Xij.mat.cum)
m<-rep(0,TT-1)
for (i in 1:(TT-1))
m[TT-i]<-sum(Xij.mat.cum[1:i,TT+1-i])/sum(Xij.mat.cum[1:i,TT-i])
print("factores de desarrollo")
print(m)
for (i in 1:(TT-1))
Xij.mat.cum[TT:(TT-i+1),i+1]<-Xij.mat.cum[TT:(TT-i+1),i]*m[i]
print("rectángulo con cuantías acumuladas")
print(Xij.mat.cum)
Xij.mat[,1]<-Xij.mat.cum[,1]
for (i in 1:TT)
Xij.mat[i,2:TT]<-diff(c(Xij.mat.cum[i,1],Xij.mat.cum[i,2:TT]))
print("rectángulo con cuantías desacumuladas")
print(Xij.mat)
}
```

```

prov<-sum(Xij.mat.cum[,TT])-sum(Xij)
print("provisiones (suma de pagos futuros)")
print(prov)
pagofut<-rep(0,TT-1)
for (k in 1:TT-1) {
future <- row(Xij.mat) + col(Xij.mat) - 1 == TT+k
pagofut[k] <- sum(Xij.mat[future])
}
print("vector pagos futuros")
print(pagofut)
}

```

ibnrchlvar1

```

ibnrchlvar1<-function(Xij) {
TT<-trunc(sqrt(2*length(Xij)))
i<-rep(1:TT,TT:1)
j<-sequence(TT:1)
Xij.mat.cum<-Xij.mat<-matrix(0,nrow=TT,ncol=TT)
for(k in 1:length(Xij)) Xij.mat[i[k],j[k]]<-Xij[k]
for(k in 1:TT) Xij.mat.cum[k,]<-cumsum(Xij.mat[k,])
print("triángulo cuantías")
print (Xij.mat)
print("triángulo cuantías acumuladas")
print(Xij.mat.cum)
dij<-matrix(0,nrow=TT,ncol=TT-1)
for (i in 1:(TT-1))
dij[,i]<-Xij.mat.cum[,i+1]/Xij.mat.cum[,i]
print("matriz factores desarrollo")
print(dij)
dijest<-matrix(0,nrow=TT,ncol=TT-1)
dijest[,TT-1]<-dij[1,TT-1]
dijest[,TT-2]<-c(dij[1:2,TT-2],rep(mean(dij[1:2,TT-2]),TT-2))
for (i in 1:(TT-3)) {
y<-dij[1:(TT-i),i]
t<-1:(TT-i)
medy<-mean(y)
medt<-mean(t)
covyt<-(sum(y*t)/length(y))-medy*medt
vart<-sum(t^2)/length(t)-medt^2
b<-covyt/vart
a<-medy-b*medt
dijest[,i]<-c(y,a+b*(length(y)+(1:i)))}
print("matriz factores desarrollo completa")
print(dijest)
for (i in 1:(TT-1))
Xij.mat.cum[,i+1]<-Xij.mat.cum[,i]*dijest[,i]
print("rectángulo con cuantías acumuladas")
print (Xij.mat.cum)
Xij.mat[,1]<-Xij.mat.cum[,1]
for (i in 1:TT)
Xij.mat[i,2:TT]<-diff(c(Xij.mat.cum[i,1],Xij.mat.cum[i,2:TT]))
print("rectángulo con cuantías desacumuladas")
print(Xij.mat)
prov<-sum(Xij.mat.cum[,TT])-sum(Xij)
print("provisiones (suma de pagos futuros)")
print(prov)
pagofut<-rep(0,TT-1)
for (k in 1:TT-1) {
future <- row(Xij.mat) + col(Xij.mat) - 1 == TT+k
pagofut[k] <- sum(Xij.mat[future])
}
}

```

```
print("vector pagos futuros")
print(pagofut)
}
```

ibnrchlvar2

```
ibnrchlvar2<-function(Xij) {
TT<-trunc(sqrt(2*length(Xij)))
i<-rep(1:TT,TT:1)
j<-sequence(TT:1)
Xij.mat.cum<-Xij.mat<-matrix(0,nrow=TT,ncol=TT)
for(k in 1:length(Xij)) Xij.mat[i[k],j[k]]<-Xij[k]
for(k in 1:TT) Xij.mat.cum[k,]<-cumsum(Xij.mat[k,])
print("triángulo cuantías")
print (Xij.mat)
print("triángulo cuantías acumuladas")
print(Xij.mat.cum)
dij<-matrix(0,nrow=TT,ncol=TT-1)
for (i in 1:(TT-1))
dij[,i]<-Xij.mat.cum[,i+1]/Xij.mat.cum[,i]
print("matriz factores desarrollo")
print(dij)
print("opciones de ponderaciones: 1: wij=1 2: wij=i+j+1 3: wij=(i+j+1)^2 4: wij=2^(i+j+1)")
print("entrar opción: 1,2,3 o 4")
opc<-scan()
wij<-matrix(0,nrow=TT,ncol=TT-1)
if (opc==1)
{ wij<-matrix(1,nrow=TT,ncol=TT-1) }
else{
if (opc==2)
{ wi<-matrix(0:(TT-1),nrow=TT,ncol=TT-1)
wj<-matrix(0:(TT-2),nrow=TT,ncol=TT-1,byrow=T)
w1<-matrix(1,nrow=TT,ncol=TT-1)
wij<-wi+wj+w1 }
else
{
if (opc==3)
{ wi<-matrix(0:(TT-1),nrow=TT,ncol=TT-1)
wj<-matrix(0:(TT-2),nrow=TT,ncol=TT-1,byrow=T)
w1<-matrix(1,nrow=TT,ncol=TT-1)
wij<-(wi+wj+w1)^2
}
else{
wi<-matrix(0:(TT-1),nrow=TT,ncol=TT-1)
wj<-matrix(0:(TT-2),nrow=TT,ncol=TT-1,byrow=T)
w1<-matrix(1,nrow=TT,ncol=TT-1)
w2<-matrix(2,nrow=TT,ncol=TT-1)
wij<-w2^(wi+wj+w1)
}
}
}
}
}
dijest<-rep(0,TT-1)
for (i in 1:(TT-1))
dijest[i]<-weighted.mean(c(dij[1:(TT-i),i]),c(wij[1:(TT-i),i]))
dijest.mat<-matrix(dijest,nrow=TT,ncol=TT-1,byrow=T)
print("matriz estimaciones de dij")
print(dijest.mat)
for (i in 1:(TT-1))
Xij.mat.cum[TT:(TT-i+1),i+1]<-Xij.mat.cum[TT:(TT-i+1),i]*dijest[i]
print( "rectángulo con cuantías acumuladas")
print (Xij.mat.cum)Xij.mat[,1]<-Xij.mat.cum[,1]
for (i in 1:TT)
```

```
Xij.mat[i,2:TT]<-diff(c(Xij.mat.cum[i,1],Xij.mat.cum[i,2:TT]))
print("rectángulo con cuantías desacumuladas")
print(Xij.mat)
prov<-sum(Xij.mat.cum[,TT])-sum(Xij)
print("provisiones (suma de pagos futuros)")
print(prov)
pagofut<-rep(0,TT-1)
for (k in 1:TT-1) {
future <- row(Xij.mat) + col(Xij.mat) - 1 == TT+k
pagofut[k] <- sum(Xij.mat[future])
}
print("vector pagos futuros")
print(pagofut)
}
```

ibnrvylder

```
ibnrvylder<-function(Xij) {
TT<-trunc(sqrt(2*length(Xij)))
i<-rep(1:TT,TT:1)
j<-sequence(TT:1)
Xij.mat.cum<-Xij.mat<-matrix(0,nrow=TT,ncol=TT)
for(k in 1:length(Xij)) Xij.mat[i[k],j[k]]<-Xij[k]
for(k in 1:TT) Xij.mat.cum[k,]<-cumsum(Xij.mat[k,])
print("triángulo cuantías")
print (Xij.mat)
pj<-rep(0,TT)
xi<-rep(0,TT)
pit<-rep(0,TT);xit<-rep(0,TT)
for (i in 1:TT)
pj[i]<-Xij.mat[1,i]/Xij.mat.cum[1,TT]
pj2<-pj^2
for (i in TT:1)
xi[i]<-sum(Xij.mat[i,1:(TT-i+1)]*pj[1:(TT-i+1)])/sum(pj2[1:(TT-i+1)])
xi2<-xi^2
it<-1
for (i in TT:2)
pit[i]<-sum(Xij.mat[1:(TT-i+1),i]*xi[1:(TT-i+1)])/sum(xi2[1:(TT-i+1)])
pit[1]<-1-sum(pit[2:TT])
pit2<-pit^2
for (i in TT:1)
xit[i]<-sum(Xij.mat[i,1:(TT-i+1)]*pit[1:(TT-i+1)])/sum(pit2[1:(TT-i+1)])
xit2<-xit^2
while(it<50) {
for(i in TT:2)
pit[i]<-sum(Xij.mat[1:(TT-i+1),i]*xit[1:(TT-i+1)])/sum(xit2[1:(TT-i+1)])
pit[1]<-1-sum(pit[2:TT])
pit2<-pit^2
for (i in TT:1)
xit[i]<-sum(Xij.mat[i,1:(TT-i+1)]*pit[1:(TT-i+1)])/sum(pit2[1:(TT-i+1)])
xit2<-xit^2
it<-it+1
}
print("Resultados intermedios")
print("xi")
print(xit)
print("pj")
print(pit)
xi.mat<-matrix(xit,TT,TT,byrow=T)
pj.mat<-matrix(pit,TT,TT)
xij.est.mat<-xi.mat*pj.mat
print("rectángulo con cuantías totales")
}
```

```
print(t(xij.est.mat))
pagofut<-rep(0,TT-1)
for (k in 1:TT-1) {
future <- row(xij.est.mat) + col(xij.est.mat) - 1 == TT+k
pagofut[k] <- sum(xij.est.mat[future])
}
prov<-sum(pagofut)
print("provisiones (suma de pagos futuros)")
print(prov)
print("vector pagos futuros")
print(pagofut)
}
```