

TRABAJO FINAL DE MÁSTER

Título: Estimación Bootstrap del Valor en Riesgo para el Cálculo de Reservas de Seguros No Vida

Autoría: Pau Bernat Morillo

Tutoría: Eva Boj y Teresa Costa

Curso académico: 2023-2024



UNIVERSITAT DE
BARCELONA

Facultat d'Economia
i Empresa

Màster
de Ciències
Actuarials
i Financeres

Facultad de Economía y Empresa

Universidad de Barcelona

Trabajo Final de Máster

Máster en Ciencias Actuariales y Financieras

Estimación Bootstrap del Valor en Riesgo para el Cálculo de Reservas de Seguros No Vida

Pau Bernat Morillo

Eva Boj y Teresa Costa

“El contenido de este documento es de exclusiva responsabilidad del autor, quien declara que no ha incurrido en plagio y que la totalidad de referencias a otros autores han sido expresadas en el texto”

“The content of this document is the sole responsibility of the author, who declares that he/she has not incurred plagiarism and that all references to other authors have been expressed in the text”

Resumen:

Este trabajo de fin de máster se centra en la estimación *bootstrap* para el cálculo de reservas en seguros de no vida. El objetivo es analizar las estimaciones por año de calendario y calcular el error de predicción como herramienta para las entidades aseguradoras. La estimación se realiza utilizando un método estocástico, el Modelo Lineal Generalizado (GLM), empleando dos métodos *bootstrap*, los denominados *residual* y *pairs*. Se calcula el Valor en Riesgo (VaR), el valor actual de las reservas y el valor actual del VaR. Todos los cálculos y estimaciones se realizan utilizando el lenguaje de programación R.

Palabras clave: *Bootstrap residuals, bootstrap pairs, Chain-Ladder, provisiones, VaR.*

Abstract:

This master's thesis focuses on bootstrap estimation for the calculation of reserves in non-life insurance. The objective is to analyze yearly calendar estimates and compute the prediction error as a tool for insurance entities. The estimation is carried out using a stochastic method, the Generalized Linear Model (GLM), employing two bootstrap methods known as residual and pairs. We calculate the Value at Risk (VaR), the present value of reserves, and the present value of VaR. All calculations and estimations are performed using the R programming language.

Key words: Bootstrap residuals, bootstrap pairs, Chain-Ladder, reserves, VaR.

Índice

1.- Introducción y Motivación	7
1.1.- Objetivos.....	8
1.2.- Estructura del trabajo.....	9
2.- Marco Teórico	9
2.1.- Datos	9
2.2.- Modelos deterministas	10
2.3.- Modelos estocásticos	12
2.4.- <i>Bootstrap</i>	15
2.5 Aplicación de <i>bootstrap</i> en el cálculo de provisiones no vida	15
2.5.1 Errores de predicción	17
2.5.2 Distribuciones predictivas y VaR.....	17
2.6.- Provisiones: valor actual de los pagos futuros	18
3.- Caso práctico	19
3.1.- Análisis código de R.....	19
3.2.- Datos	22
4.- Análisis resultados	23
5.- Conclusiones	41
6.- Bibliografía	43
7.- Anexos.....	45

1.- Introducción y Motivación

La aplicación de la técnica *bootstrap* en la estimación del VaR para el cálculo de reservas es de interés en seguros no vida.

Esta técnica, mediante remuestreo, genera muestras de datos simuladas para estimar la distribución del VaR. Por ello, facilita la toma de decisiones informada por parte de las compañías de seguros al comprender mejor los riesgos asociados con las reservas.

Resulta una herramienta esencial para una gestión efectiva del riesgo y fortalece la posición financiera de las aseguradoras en un sector complejo y cambiante.

El trabajo consiste en la investigación de las estimaciones *bootstrap* del VaR para el cálculo de reservas de seguros no vida. Se usará un modelo estocástico para realizar las estimaciones y las comparaciones pertinentes con otros modelos clásicos, como Chain-Ladder. Estos modelos están descritos en Boj *et al.* (2021).

La metodología *bootstrap* en cálculo de provisiones se aplica en diversos trabajos teóricos y prácticos, obteniéndose las distribuciones predictivas de los pagos por siniestros desde el enfoque del año de ocurrencia de los siniestros, como en Paule-Vianez *et al.* (2018, 2020). Algunos paquetes de R (R Development Core Team, 2024), como el paquete ChainLadder (Gesmann *et al.*, 2023), también abordan este tipo de análisis. En relación al paquete ChainLadder encontramos su código adaptado en un aplicativo de Shiny elaborado por Hinkley (2017) donde se pueden visualizar los resultados numéricos con diferentes bases de datos.

En este trabajo se adaptará la metodología según el enfoque del año de calendario en el que se pagan los importes de los siniestros, siguiendo las formulaciones de Boj *et al.* (2014) y Boj y Costa (2018).

Además, en la aplicación de la técnica *bootstrap* se van a tener en cuenta dos metodologías distintas, las denominadas *bootstrap residuals* y *bootstrap pairs*, descritas en los trabajos de Pinheiro *et al.* (2003) y Björkwall *et al.* (2009).

En el campo de los seguros no vida, el cálculo de reservas es crucial para la estabilidad financiera de las compañías aseguradoras. Una medida clave en este proceso es el VaR, que estima la pérdida máxima esperada en un horizonte de tiempo específico con un nivel

de confianza específico. La estimación *bootstrap* del VaR es una técnica fiable y flexible que aborda la incertidumbre asociada con las reservas de seguros no vida.

Este método se basa en el remuestreo de datos para generar múltiples muestras y estimar la distribución subyacente del VaR. En el contexto del cálculo de reservas, el VaR determina el capital necesario para cubrir pérdidas inesperadas. Sin embargo, calcular el VaR de forma precisa es un desafío debido a la complejidad de los datos y la incertidumbre.

La estimación *bootstrap* del VaR supone una metodología capaz de evaluar una amplia gama de escenarios posibles, capturando la variabilidad de los datos y permitiendo una evaluación más completa de los riesgos. Al iterar sobre muestras simuladas, este método ofrece una distribución empírica del VaR, lo que mejora la comprensión de las estimaciones y brinda mayor confianza a las compañías aseguradoras.

Por otro lado, el cálculo de provisiones no vida por año de origen implica evaluar el riesgo asociado con cada período en el que se originaron los siniestros. El método *bootstrap* es una herramienta valiosa en este proceso, ya que simula múltiples escenarios futuros, considerando la variabilidad en los datos históricos de siniestralidad. Esto permite una estimación más precisa de la incertidumbre asociada con las provisiones calculadas.

En este trabajo se aplicará esta metodología para los pagos futuros a los que debe hacer frente el asegurador según el año de calendario en el que se realice dicho pago. Al emplear el *bootstrap* para el cálculo de los pagos por año de calendario, las entidades aseguradoras pueden comprender mejor la variabilidad de sus estimaciones y evaluar los riesgos financieros de manera más precisa. Esto facilita la toma de decisiones respecto a la gestión de riesgos y la asignación de capital.

1.1.- Objetivos

Este trabajo tiene como objetivo aplicar un método estocástico, el GLM, para el cálculo de provisiones en seguros de no vida, empleando metodologías *bootstrap*. Se calcularán los errores de predicción para los pagos por año de calendario y se determinará el VaR a partir de cada una de las distribuciones predictivas obtenidas. Estas metodologías se aplicarán a nivel práctico utilizando un conjunto de datos ampliamente utilizado en la literatura actuarial, con referencia específica al trabajo de Taylor y Ashe (1983).

Finalmente, se analizarán y compararán los resultados obtenidos para evaluar la eficacia y precisión de las estimaciones realizadas.

1.2.- Estructura del trabajo

En el trabajo, después de la introducción y motivación, en primer lugar, se describe el marco teórico, para explicar la terminología y las fórmulas que se van a aplicar a lo largo del trabajo.

Se hace referencia brevemente al modelo determinista de Chain-Ladder y se desarrolla con más detalle el GLM, que va a ser el modelo estocástico a aplicar. También se explica la metodología *bootstrap* y su aplicación en el cálculo de provisiones no vida.

En la parte práctica del trabajo, se aplica la metodología teórica a unos datos numéricos utilizando el código de R, siguiendo los pasos que se describen. Se analizan los resultados obtenidos con detalle y se comparan tanto las metodologías *bootstrap* aplicadas como las distribuciones que se asumen para los datos en el GLM.

Finalmente, el último punto son las conclusiones.

Al final del trabajo se adjunta la bibliografía junto a los anexos del trabajo. En los anexos se incluyen los resultados de las estimaciones y los errores de predicción de las provisiones por año de origen, así como algunos gráficos relativos a los pagos futuros por años de calendario.

2.- Marco Teórico

2.1.- Datos

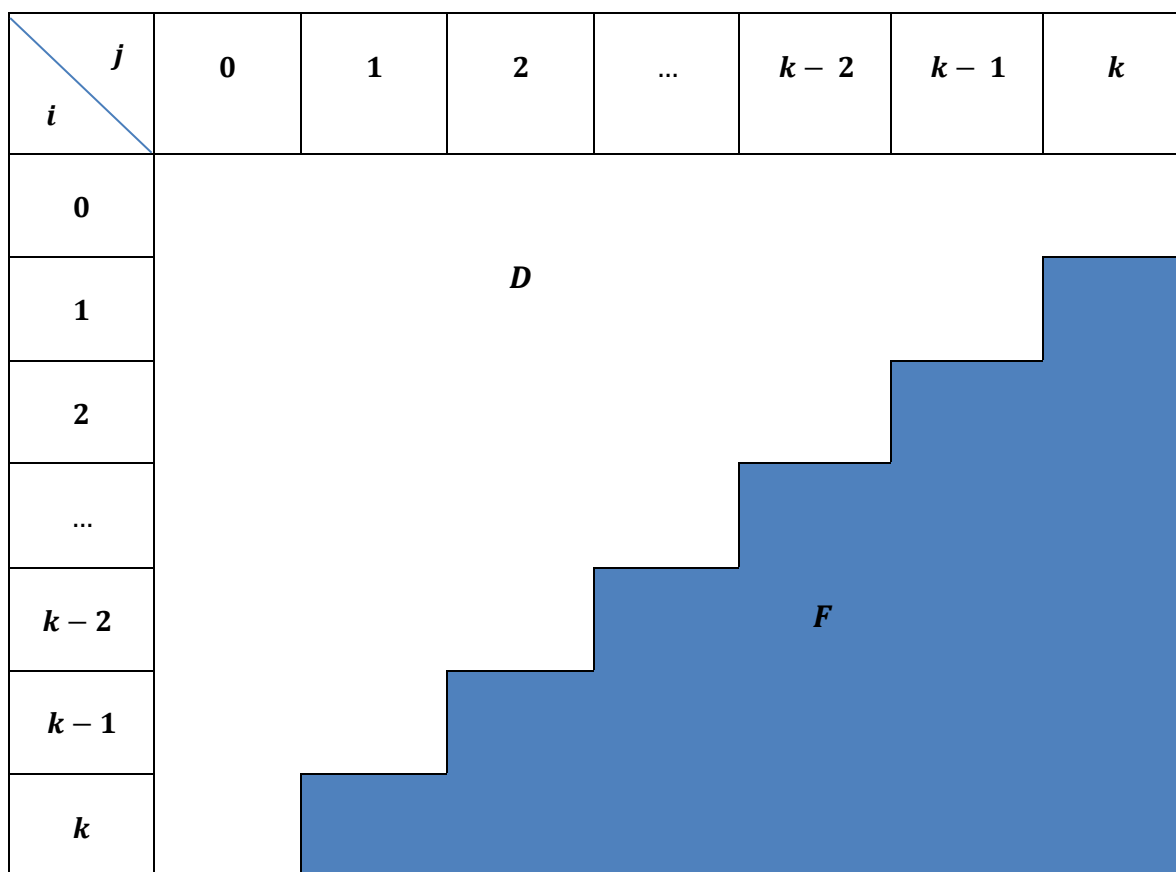
La terminología de los datos de los modelos que se van a reproducir en el trabajo es la siguiente, a partir de Boj *et al.* (2021):

- 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 .
- $D = \{i, j | i \in I, j \in J\}$ el dominio (años de origen, años de desarrollo para los que tenemos datos).

- Si $i = 0, \dots, k$ y $j = 0, 1, \dots, k$, el dominio D tiene forma triangular, denominándose triángulo de desarrollo.

Gráficamente:

Figura 1: Rectángulo de cuantías



Fuente: Elaboración propia.

F representa el conjunto de combinaciones de i, j que permiten completar el triángulo con los pagos futuros que deberá realizar el asegurador y en las contradiagonales están los importes que corresponden al mismo año de calendario.

2.2.- Modelos deterministas

Los modelos deterministas para el cálculo de las provisiones en los seguros de no vida se usan de forma más común debido a su simplicidad. Estos modelos suponen que todas las variables relevantes están completamente especificadas y que el resultado futuro está completamente determinado por las condiciones iniciales y las reglas del modelo, es decir, no intervienen variables aleatorias en el modelo.

Dentro de los métodos deterministas, el método Chain-Ladder representa una técnica altamente empleada en la industria aseguradora para la estimación de reservas de siniestros pendientes. Las primeras ideas del método clásico fueron inicialmente apuntadas en Tarbell (1934). Su utilización se remonta a los años 70 y algunas de las referencias en que fue utilizado son Kramreiter y Straub (1973), Skurnick (1973) y Clarke y Harland (1974). Este enfoque se basa en la extrapolación de los desarrollos históricos de los siniestros, estableciendo una secuencia de relaciones entre datos pasados y presentes para anticipar los pagos futuros. Su simplicidad y eficacia lo convierten en una opción atractiva para empresas aseguradoras de diversos sectores.

La ejecución del método Chain-Ladder consta de varias etapas. Inicialmente, se dispone un conjunto de pagos acumulados por siniestros en un triángulo de desarrollo, donde las filas representan los años de ocurrencia y las columnas representan los períodos de desarrollo.

Figura 2: Rectángulo de cuantías acumuladas pagadas

$j \backslash i$	0	1	2	...	$k-2$	$k-1$	k
0	$C_{0,0}$	$C_{0,1}$	$C_{0,2}$	...	$C_{0,(k-2)}$	$C_{0,(k-1)}$	$C_{0,k}$
1	$C_{1,0}$	$C_{1,1}$	$C_{1,2}$	...	$C_{1,(k-2)}$	$C_{1,(k-1)}$	$\hat{C}_{1,k}$
2	$C_{2,0}$	$C_{2,1}$	$C_{2,2}$	...	$C_{2,(k-2)}$	$\hat{C}_{2,(k-1)}$	$\hat{C}_{2,k}$
...	...	...	...	...	...	...	...
$k-2$	$C_{(k-2),0}$	$C_{(k-2),1}$	$C_{(k-2),2}$	...	$\hat{C}_{(k-2),(k-2)}$	$\hat{C}_{(k-2),(k-1)}$	$\hat{C}_{(k-2),k}$
$k-1$	$C_{(k-1),0}$	$C_{(k-1),1}$	$\hat{C}_{(k-1),2}$	...	$\hat{C}_{(k-1),(k-2)}$	$\hat{C}_{(k-1),(k-1)}$	$\hat{C}_{(k-1),k}$
k	$C_{k,0}$	$\hat{C}_{k,1}$	$\hat{C}_{k,2}$	...	$\hat{C}_{k,(k-2)}$	$\hat{C}_{k,(k-1)}$	$\hat{C}_{k,k}$

Fuente: elaboración propia

Posteriormente, se calculan factores de desarrollo que describen la evolución de los pagos de los siniestros en el tiempo. La hipótesis que considera el método Chain-Ladder es que las columnas del triángulo de cuantías acumuladas son proporcionales, es decir, la relación entre una columna y la siguiente es el estimador del cambio, éste se calcula con la división del sumatorio de todas las cuantías acumuladas pagadas del dominio de un año de desarrollo entre las cuantías acumuladas pagadas del dominio del año de desarrollo anterior.

El factor de desarrollo se obtiene de la siguiente manera:

$$\hat{m}_h = \frac{\sum_{i=0}^{k-h-1} C_{i,h+1}}{\sum_{i=0}^{k-h-1} C_{i,h}}.$$

Estos factores se aplican a los datos de desarrollo más recientes para proyectar los pagos futuros, iterando el proceso hasta alcanzar estabilidad en las estimaciones:

$$\hat{C}_{i,j} = C_{i,k-i} \cdot \prod_{h=k-i}^{j-1} \hat{m}_h.$$

Con el tiempo, se han desarrollado diversas extensiones y variaciones del método Chain-Ladder para adaptarse a diferentes situaciones y mejorar la precisión de las estimaciones. Estas incluyen ajustes para considerar la volatilidad en los datos, la inclusión de factores estacionales y el desarrollo de modelos más sofisticados que incorporan información adicional.

2.3.- Modelos estocásticos

En este caso, los siniestros se consideran que son variables aleatorias.

En las últimas décadas, se han propuesto métodos estocásticos, como el GLM, con el objetivo de medir el error de predicción en la estimación de las reservas. Los errores de predicción con GLM pueden calcularse mediante formulaciones analíticas o mediante la metodología de remuestreo *bootstrap* (Boj *et al.*, 2014).

Dentro del GLM, podemos seleccionar la distribución del error, la función de enlace y el predictor lineal. Dependiendo de las suposiciones que hagamos, podemos obtener algunos métodos deterministas como casos particulares. Por ejemplo, el método Chain-Ladder puede derivarse de estas hipótesis.

A continuación, se van a comentar las fórmulas que se han usado en el trabajo y que son del modelo GLM, extraídas de Boj *et al.* (2021).

Los datos c_{ij} de las cuantías están disponibles para D triangular:

$$D = \{i, j | 0 \leq i \leq k, 0 \leq j \leq k, 0 \leq i + j \leq k\}.$$

Se asumen las siguientes hipótesis:

$$\mu_{ij} = E[c_{ij}],$$

$$V(\mu_{ij}) = \mu_{ij}^\theta,$$

con lo que:

$$\mu_{ij} = E[c_{ij}]$$

$$Var[c_{ij}] = (\phi/w_{ij})V(\mu_{ij}) = (\phi/w_{ij})\mu_{ij}^\theta.$$

Donde ϕ es el parámetro de dispersión y w_{ij} es el peso a priori de los datos, asumiendo que $w_{ij} = 1$, para las cuantías de siniestros de un triángulo de desarrollo.

Considerando la función de enlace logarítmica:

$$\log \mu_{ij} = \eta_{ij},$$

el predictor lineal en el GLM es de la forma:

$$\eta_{ij} = c_0 + \alpha_i + \beta_j.$$

Siendo:

- α_i el factor correspondiente a los años de origen $i = 1, \dots, k$,
- β_j el factor correspondiente a los años de desarrollo $j = 1, \dots, k$,
- c_0 el término que se correspondería al año de origen y desarrollo 0.

A partir de aquí podemos realizar las estimaciones de las cuantías para completar el cuadro mediante la expresión:

$$\hat{c}_{ij} = \exp(c_0 + \alpha_i + \beta_j).$$

A continuación, se describe el cálculo de las provisiones a partir del trabajo de Boj y Costa (2018). Las provisiones para los diferentes años de origen $i = 1, \dots, k$ se obtienen sumando las cuantías futuras estimadas en la fila correspondiente del cuadrado:

$$P_i = \sum_{j=k-i+1}^k \hat{c}_{ij}.$$

La provisión total se calcula sumando todas las pérdidas incrementales futuras pronosticadas en la parte inferior derecha del triángulo de desarrollo:

$$P = \sum_{i=1}^k \sum_{j=k-i+1}^k \hat{c}_{ij}.$$

Y los pagos futuros para los diferentes años calendarios $t = k + 1, \dots, 2k$ se obtienen sumando las cuantías que se pagarán en los años de calendario futuros, es decir, los valores de la misma contra diagonal t .

$$P_t = \sum_{j=t-k}^k \hat{c}_{t-j,j}.$$

Podemos calcular el error cometido en la predicción de las reservas, en primer lugar, utilizando la fórmula analítica del GLM, teniendo en cuenta la distribución supuesta en los datos.

En Boj *et al.* (2014) y Boj y Costa (2018) se describen las expresiones del error de predicción con relación a cada cuantía estimada, a las reservas por años de origen y total y a los pagos futuros por año de calendario.

Error cuadrático medio para cada estimación de las cuantías:

$$MSE(c_{ij}) \cong \phi \mu_{ij}^\theta + \mu_{ij}^2 Var[\eta_{ij}] \quad i, j = 1, \dots, k.$$

Error cuadrático medio para las reservas por años de origen:

$$MSE(P_i) = E[(P_i - \hat{P}_i)^2] \approx \sum_{\substack{j=1, \dots, k \\ i+j > k}} \phi \mu_{ij}^\theta + \mu_i^T Var[\eta_i] \mu_i \quad i = 1, \dots, k.$$

Error cuadrático medio para la estimación de la reserva total:

$$MSE(P) = E[(P - \hat{P})^2] \approx \sum_{\substack{j=1, \dots, k \\ i+j > k}} \phi \mu_{ij}^\theta + \mu^T Var[\eta] \mu.$$

Error cuadrático medio para los pagos futuros por años de calendario:

$$MSE(P_t) = E[(P_t - \hat{P}_t)^2] \approx \sum_{\substack{j=1, \dots, k \\ i+j=k}} \phi \mu_{ij}^\theta + \mu_t^T Var[\eta_t] \mu_t \quad t = k + 1, \dots, 2k.$$

El error de predicción se obtiene como la raíz cuadrada del MSE en cada caso.

2.4.- *Bootstrap*

La metodología *bootstrap* fue propuesta por Efron (1979) y aplicada en la estimación de las provisiones técnicas por Lowe (1994).

Esta técnica ofrece estimaciones del error estadístico sin imponer restricciones significativas sobre las variables aleatorias analizadas, convirtiéndose en un procedimiento de amplia aplicabilidad, independientemente del estadístico considerado (Solanas y Olivera, 1992).

El principio básico consiste en generar B muestras *bootstrap* mediante selección con reemplazo del conjunto de datos observado. Para cada una de ellas se calcula la magnitud de interés y se pueden evaluar los percentiles a partir de los B valores resultantes (ver por ejemplo Davidson y Hinkley, 1997). En el caso de modelos de regresión, existen dos posibles enfoques:

- el remuestreo de residuos (*bootstrap residuals* en inglés), donde cada muestra *bootstrap* del vector de respuestas de dimensión n se obtiene a partir de n residuos remuestreados,
- y el remuestreo de pares o casos (*bootstrap pairs* en inglés), donde cada muestra *bootstrap* consiste en n pares de respuesta-predicción tomados de los datos originales.

La diferencia entre ambos métodos radica en que en el remuestreo de residuos se consideran fijas las variables predictoras del modelo de regresión. Se asume que el modelo de regresión básico es correcto y que los residuos pueden considerarse iguales. Si esto no es así, por ejemplo, cuando los residuos tienen diferentes varianzas o cuando hay errores en los predictores, el remuestreo de residuos llevará a resultados incorrectos. En cambio, el remuestreo de pares es menos sensible a supuestos erróneos del modelo. Además, si se cumplen los supuestos subyacentes del remuestreo de residuos, el remuestreo de pares proporcionará también resultados aproximadamente similares.

2.5 Aplicación de *bootstrap* en el cálculo de provisiones no vida

England y Verrall (1999) proponen un método de estimación de provisiones técnicas utilizando *bootstrap* como una forma sencilla de evaluar el riesgo de reserva. Este método hace uso de los factores de desarrollo del método Chain-Ladder.

England (2002) propone modificaciones a la propuesta inicial de England y Verrall (1999) en relación con los dos primeros momentos. Introduce un modelo dividido en dos etapas: *bootstrap* para estimar el error de estimación y simulación para calcular el error de proceso. Este enfoque nos permite obtener la distribución completa de predicción en cada remuestreo, en lugar de solo los dos primeros momentos.

En el cálculo de reservas, en este trabajo utilizamos el GLM, a partir del cual calcularemos magnitudes como los pagos futuros totales o los pagos futuros por años de calendario.

En el caso del *bootstrap pairs*, hemos generado las multiplicidades de los individuos en cada muestra. Estas multiplicidades, que suman n , actúan a modo de pesos en el modelo. Para cada muestra se estima el modelo GLM con el que se calculan las magnitudes.

Para el caso particular de *bootstrap residuals* la metodología es la siguiente:

- Se inicia con el modelo GLM para el triángulo de pagos por siniestros original, para calcular los residuos de Pearson:

$$r_{ij}^P = \frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\hat{c}_{ij}^\theta}}.$$

- Se calcula el parámetro de escala:

$$\hat{\phi}^P = \frac{1}{n - 2k - 1} \sum_{ij} \frac{(c_{ij} - \hat{c}_{ij})^2}{\hat{c}_{ij}^\theta}.$$

- Los residuos escalados se remuestran B veces, r_{ij}^{P*} , y de la fórmula original de los residuos se despeja:

$$c_{ij}^* = r_{ij}^{P*} \hat{c}_{ij}^{\theta^{1/2}} + \hat{c}_{ij}$$

y se construyen B muestras de triángulos de cuantías no acumuladas c_{ij}^* .

- Se procede a realizar la estimación del GLM para cada uno de los B triángulos de cuantías.

Finalmente, el resultado es que tenemos B valores respecto a los pagos futuros por años de calendario, las provisiones por años de origen y totales, lo que nos proporciona la distribución predictiva de estas reservas.

2.5.1 Errores de predicción

A partir de los B valores obtenidos para las distintas magnitudes con el procedimiento de *bootstrap*, $(\hat{c}_{ij}^{boot}, \hat{P}_i^{boot}, \hat{P}^{boot}, \hat{P}_t^{boot})$ podremos utilizar la desviación estándar de estos valores en las fórmulas para el cálculo del error cuadrático medio (MSE) en cada caso y, a su vez, los errores de predicción asociados calculados como la raíz del MSE.

En concreto, para el GLM, la estimación del MSE, siguiendo las expresiones de Boj y Costa (2018), se obtiene a partir de las expresiones que se detallan a continuación.

Error cuadrático medio para cada estimación de las cuantías:

$$MSE^{boot}(c_{ij}) \approx \hat{\Phi}^P \hat{c}_{ij}^\theta + SE(\hat{c}_{ij}^{boot})^2, i, j = 1, \dots, k.$$

Error cuadrático medio para las reservas por años de origen:

$$MSE^{boot}(P_i) \approx \sum_{\substack{j=1, \dots, k \\ i+j > k}} \hat{\Phi}^P \hat{c}_{ij}^\theta + SE(\hat{P}_i^{boot})^2, i = 1, \dots, k.$$

Error cuadrático medio para la estimación de la reserva total:

$$MSE^{boot}(P) \approx \sum_{\substack{i, j=1, \dots, k \\ i+j > k}} \hat{\Phi}^P \hat{c}_{ij}^\theta + SE(\hat{P}^{boot})^2.$$

Error cuadrático medio para los pagos futuros por años de calendario:

$$MSE^{boot}(P_t) \approx \sum_{\substack{i, j=1, \dots, k \\ i+j=t}} \hat{\Phi}^P \hat{c}_{ij}^\theta + SE(\hat{P}_t^{boot})^2, t = k+1, \dots, 2k.$$

2.5.2 Distribuciones predictivas y VaR

Con la aplicación del método *bootstrap* podremos calcular diversos estadísticos que resultan fundamentales en el análisis de riesgos, a partir de la distribución predictiva de las distintas magnitudes con los B valores generados.

Entre estos estadísticos se encuentran:

- la media de los datos, que proporciona un valor promedio representativo de la distribución

- la desviación estándar, que nos muestra la dispersión de los valores respecto a la media
- el coeficiente de variación, que nos está indicando la variabilidad relativa de los valores en relación a la media
- el coeficiente de asimetría, que nos mide el grado de asimetría de los datos respecto a la media
- el cociente de curtosis, que nos indica la pesadez de la cola de la distribución
- el percentil 50% (mediana), que significa que el 50% de los valores serán menores que su valor y que el otro 50% serán mayores
- los percentiles 75%, 90%, 95%, 99,5%, que nos permiten obtener aquellos valores que acumulan una probabilidad del 75%, 90%, 95% y 99,5%, respectivamente.

Por último, la distribución predictiva también permite definir el VaR, que es una medida estadística que cuantifica el riesgo de pérdida potencial para un determinado nivel de confianza. Se define como el mayor valor de una variable tal que la probabilidad de obtener un valor superior sea al menos igual al nivel de confianza que se fije.

El VaR a un nivel de confianza α para una variable, VaR_α , se calcula como el percentil $1 - \alpha$ de los valores de dicha variable. En nuestro caso, si lo calculamos para los pagos futuros del año de calendario t , se definiría como:

$$Pr[P_t > VaR_\alpha] = \alpha .$$

2.6.- Provisiones: valor actual de los pagos futuros

En apartados anteriores hemos definido la provisión total como la suma de las cuantías futuras estimadas:

$$P = \sum_{i=1}^k \sum_{j=k-i+1}^k \hat{c}_{ij} ,$$

que a su vez coincide con la suma de los pagos futuros por año de calendario:

$$P = \sum_{t=k+1, \dots, 2k} \hat{P}_t .$$

En este trabajo se calcularán los valores actuales de los pagos futuros, por año de calendario, teniendo en cuenta que estas cuantías constituyen una renta anual vencida de pagos.

Además, se calculará también el valor actual del VaR de los pagos futuros.

Los valores actuales se determinarán utilizando los tipos de interés proporcionados por la Autoridad Europea de Seguros y Pensiones de Jubilación (EIOPA).

Las fórmulas usadas para calcular el valor actual de los pagos futuros y del VaR son las siguientes:

$$VA = \sum_{i=1}^k \hat{P}_t \cdot (1 + EIOPA_i)^{-i},$$
$$VA = \sum_{i=1}^k VaR(\hat{P}_t) \cdot (1 + EIOPA_i)^{-i},$$

siendo $t = k + i$.

Representa la suma del vector de pagos futuros multiplicado por los factores de actualización con el tipo de interés correspondiente $EIOPA_i$ según el número de años i a actualizar, lo mismo sería con el valor actual del VaR.

3.- Caso práctico

3.1.- Análisis código de R

El código en R se basará en los siguientes pasos:

Introducción de datos:

- Datos acumulados por año de origen: Se definen vectores que contienen los montos acumulados de pagos por siniestros para cada año de origen.
- Creación de triángulos de datos acumulados: Se crean matrices que representan los triángulos de datos acumulados. Cada fila representa un año de origen y cada columna representa un año de desarrollo. Los valores no disponibles (NA) se agregan para completar el triángulo de datos.
- Cálculo de datos no acumulados para cada año de origen: Se calculan los datos no acumulados restando de los valores acumulados de cada año de desarrollo el valor acumulado correspondiente al año anterior.

- Creación de triángulos de datos no acumulados: Se crean matrices que representan los triángulos de datos no acumulados. Al igual que con los datos acumulados, se agregan valores NA para completar el triángulo de datos.

Elección de la distribución:

Se elige la distribución estadística a utilizar para modelar los datos. Se proporcionan dos opciones (ϕ_1 =poisson y ϕ_2 =gamma):

- Poisson: Se utiliza cuando se modelan eventos que ocurren de manera independiente a una tasa constante en un intervalo de tiempo fijo.
- Gamma: Se emplea para modelar datos positivos y sesgados, como los costos de reclamaciones, que pueden no ser adecuadamente modelados por una distribución Poisson.

Cálculo de las provisiones:

Para calcular las provisiones, se siguen los siguientes pasos:

- Se generan vectores de índices i y j , que corresponden a los años de origen y años de desarrollo, respectivamente.
- Se ajusta un GLM a los datos de siniestralidad c_{ij} utilizando los factores i y j como variables explicativas.
- Se extraen los coeficientes del modelo ajustado y se calculan las cuantías futuras estimadas, que permiten obtener tanto la reserva total como por año de origen y por año de calendario, sumando las predicciones ajustadas que corresponden a posiciones futuras en la matriz de predicciones.
- Se calculan los residuos de Pearson, se determinan los valores de n , p y ϕ . P para calcular los residuos escalados.
- Se calcula la matriz de varianzas y covarianzas.
- Se calcula el error de predicción utilizando la fórmula tanto para la reserva total, como para los años de origen y de calendario.

Error de predicción con *bootstrap residuals* para las provisiones:

- Se establece una semilla aleatoria.
- Se inicializa el número de muestras *bootstrap*.
- Se realizan iteraciones de *bootstrap*:
 - Se toma una muestra aleatoria con reemplazo de los residuos escalados ajustados.
 - Se ajustan los valores de las muestras aleatorias.
 - Se ajusta un nuevo modelo generalizado lineal utilizando las muestras aleatorias ajustadas.
 - Se calculan los coeficientes del modelo ajustado y las predicciones ajustadas.
 - Se calculan las reservas totales, por año de origen y por año de calendario.
 - Se guardan las reservas y los pagos totales en vectores.
- Se calcula el error de predicción con *bootstrap* para la reserva total, por año de origen y por año de calendario.

Error de predicción con *bootstrap pairs* para las provisiones:

- Se establece una semilla aleatoria para asegurar la reproducibilidad de los resultados.
- Se inicializa el número de iteraciones de *bootstrap*, son 1000 iteraciones.
- Se realiza un bucle sobre cada iteración de *bootstrap*:
 - Se toma una muestra aleatoria con reemplazo de los valores de la matriz *multi* para la iteración actual. La matriz *multi* tiene 55 filas y 1000 columnas, representando diferentes combinaciones de años de origen y años de desarrollo.
 - Se filtran los valores mayores que 0 y se utilizan como pesos en el ajuste del modelo.

- Se ajusta un GLM ponderado utilizando como predictores los años de origen y los años de desarrollo.
- Se generan predicciones ajustadas para todas las combinaciones posibles de años de origen y años de desarrollo.
- Se calculan las reservas totales, por año de origen y por año de calendario.
- Se almacenan los resultados de las reservas y pagos totales en vectores.
- Se calcula el error de predicción con *bootstrap* para la reserva total, por año de origen y por año de calendario.

Distribución predictiva de los pagos:

- Se definen los percentiles de la distribución de los pagos.
- Se calcula la media, la desviación estándar y el coeficiente de variación de los pagos.
- Se calcula la distribución de los pagos y se analiza la simetría y la curtosis de la distribución.
- Se realiza un análisis gráfico de los datos utilizando histogramas y *boxplots*.

Valor actual de los pagos futuros por año de calendario:

- Se define el nivel de confianza.
- Se calcula el VaR de los pagos por año de calendario.
- Se calcula el valor actual del VaR de los pagos por año de calendario.
- Se calcula el valor actual de los pagos por año de calendario.

3.2.- Datos

Los datos que se van a utilizar en este trabajo, han sido extraídos de Taylor y Ashe (1983) y se han usado en diversos trabajos y aplicaciones del ámbito actuarial, tanto a nivel teórico como práctico.

El triángulo de cuantías acumuladas es:

Figura 3: Triángulo de cuantías acumuladas.

Años de origen	Años de desarrollo									
	1	2	3	4	5	6	7	8	9	10
1	357.848	1.124.788	1.735.330	2.218.270	2.745.596	3.319.994	3.466.336	3.606.286	3.833.515	3.901.463
2	352.118	1.236.139	2.170.033	3.353.322	3.799.067	4.120.063	4.647.867	4.914.039	5.339.085	
3	290.507	1.292.306	2.218.525	3.235.179	3.985.995	4.132.918	4.628.910	4.909.315		
4	310.608	1.418.858	2.195.047	3.757.447	4.029.929	4.381.982	4.588.268			
5	443.160	1.136.350	2.128.333	2.897.821	3.402.672	3.873.311				
6	396.132	1.333.217	2.180.715	2.985.752	3.691.712					
7	440.832	1.288.463	2.419.861	3.483.130						
8	359.480	1.421.128	2.864.498							
9	376.686	1.363.294								
10	344.014									

Fuente: Elaboración propia.

Se obtiene el triángulo de cuantías no acumuladas:

Figura 4: Triángulo de cuantías no acumuladas.

Años de origen	Años de desarrollo									
	1	2	3	4	5	6	7	8	9	10
1	357.848	766.940	610.542	482.940	527.326	574.398	146.342	139.950	227.229	67.948
2	352.118	884.021	933.894	1.183.289	445.745	320.996	527.804	266.172	425.046	
3	290.507	1.001.799	926.219	1.016.654	750.816	146.923	495.992	280.405		
4	310.608	1.108.250	776.189	1.562.400	272.482	352.053	206.286			
5	443.160	693.190	991.983	769.488	504.851	470.639				
6	396.132	937.085	847.498	805.037	705.960					
7	440.832	847.631	1.131.398	1.063.269						
8	359.480	1.061.648	1.443.370							
9	376.686	986.608								
10	344.014									

Fuente: Elaboración propia.

4.- Análisis resultados

En primer lugar, aplicando GLM con la distribución Poisson sobredispersa y la función de enlace logarítmica, se obtiene los siguientes coeficientes en el modelo:

Tabla 1: Valores de los coeficientes del modelo GLM para la distribución Poisson.

	1	2	3	4	5	6	7	8	9	10
α_i	270.061,4	376.125	372.325,3	366.724	336.287,3	353.798,1	391.841,7	469.647,5	390.560,8	344.014
β_j	1	2,490	2,608	2,789	1,545	1,083	0,993	0,674	1,009	0,251

Fuente: Elaboración propia.

Obtenemos la siguiente matriz con la predicción de las cuantías futuras:

Figura 6: Matriz completa de pagos no acumulados para la distribución Poisson.

Años de origen	Años de desarrollo									
	1	2	3	4	5	6	7	8	9	10
1	357.848	766.940	610.542	482.940	527.326	574.398	146.342	139.950	227.229	67.948
2	352.118	884.021	933.894	1.183.289	445.745	320.996	527.804	266.172	425.046	94.633,81
3	290.507	1.001.799	926.219	1.016.654	750.816	146.923	495.992	280.405	375.833,5	93.677,80
4	310.608	1.108.250	776.189	1.562.400	272.482	352.053	206.286	247190	370.179,3	92.268,49
5	443.160	693.190	991.983	769.488	504.851	470.639	334.148,1	226.674,1	339.455,9	84.610,55
6	396.132	937.085	847.498	805.037	705.960	383.286,6	351.547,5	238.477,3	357.131,7	89.016,32
7	440.832	847.631	1.131.398	1.063.269	605548.1	424.501	389.349,1	264.120,5	395.533,7	98.588,16
8	359.480	1.061.648	1.443.370	1.310.258,2	725788.5	508.791,9	466.660	316.565,5	474.072,7	118.162,7
9	376.686	986.608	1.018.834,1	1.089.616	603568.6	423.113,4	388.076,4	263.257,2	394.240,8	98.265,88
10	344.014	856.803,5	897.410,1	959.756,3	531635.7	372.687	341.825,7	231.882,4	347.255,4	86.554,62

Fuente: Elaboración propia.

Cabe destacar que las cuantías estimadas que se obtienen en este caso son exactamente las mismas que se obtendrían aplicando el método Chain-Ladder.

Se obtiene una cuantía para la reserva total de 18.680.856.

El error de predicción obtenido con fórmula para el modelo es de 2.945.659, y expresado como porcentaje del valor de la reserva total, el error porcentualmente hablando es del 15,77% aproximadamente.

El error de predicción aplicando *bootstrap residuals* para la reserva total, tiene un valor de 2.941.689 y al ponerlo en porcentaje en comparación a la reserva total, da un error de un 15,75%.

El error de predicción aplicando *bootstrap pairs* para la reserva total, tiene un valor de 2.389.777 y al ponerlo en porcentaje en comparación a la reserva total, da un error de un 12,79%.

Calculamos la distribución predictiva para estimar la incertidumbre asociada a los pagos futuros, obtenemos los estadísticos básicos y los percentiles de la distribución y hacemos la representación gráfica del histograma de los valores obtenidos.

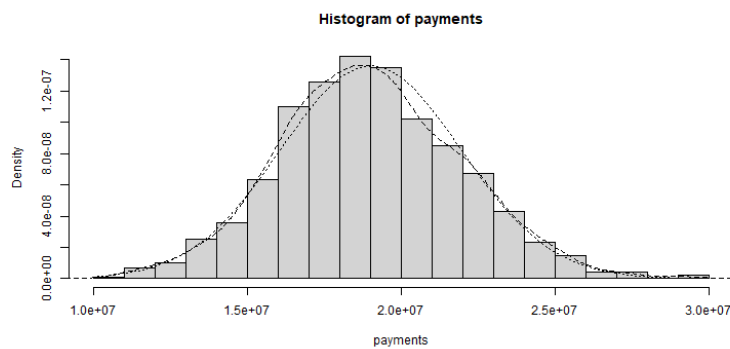
Los resultados que se han obtenido, para el caso de *bootstrap residuals*, son los siguientes:

Tabla 2: Percentiles y estadísticos básicos de la reserva total para la distribución Poisson con *bootstrap residuals*.

Percentiles	50%	75%	90%	95%	99,5%
	18.883.889	20.935.342	22.828.991	23.936.250	27.037.626
Estadísticos básicos	Media	SD	CV	Asimetría	Curtosis
	19.006.660	2.933.797	15,43563 %	0,1750084	0,06236347

Fuente: Elaboración propia.

Gráfico 1: Histograma de la reserva total para la distribución Poisson con *bootstrap residuals*.



Fuente: Elaboración propia, output de R.

La media de la reserva total se sitúa en poco más de 19 millones y la variabilidad entorno a la media es de un 15,44%.

El coeficiente de asimetría tiene un valor de 0,175, lo que nos está indicando que hay una asimetría positiva, por lo cual en los datos existe una cola hacia la derecha en la distribución, como se puede ver en el gráfico, y la mediana es menor que la media.

El valor del coeficiente de curtosis es de 0,062, por lo cual la distribución tiene una mayor concentración de pagos alrededor de la media en comparación con la distribución Normal.

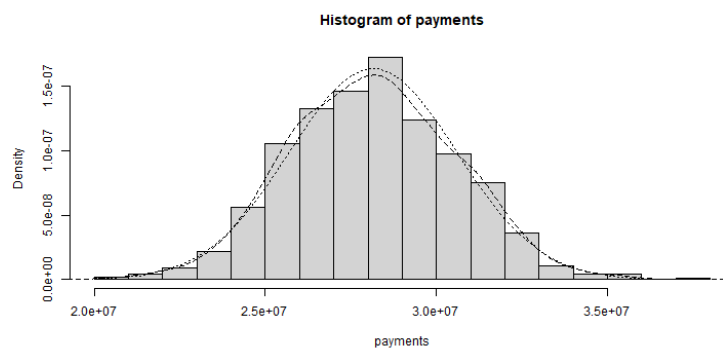
Para *bootstrap pairs* se han obtenido los siguientes resultados:

Tabla 3: Percentiles y estadísticos básicos de la reserva total para la distribución Poisson con *bootstrap pairs*.

Percentiles	50%	75%	90%	95%	99,5%
	28.141.728	29.824.972	31.350.411	32.139.432	34.718.477
Estadísticos básicos	Media	SD	CV	Asimetría	Curtosis
	28.145.779	2.442.209	8,676998%	0,0897	-0,00069

Fuente: Elaboración propia.

Gráfico 2: Histograma de la reserva total para la distribución Poisson con *bootstrap pairs*.



Fuente: Elaboración propia, output de R.

La media de la reserva total se sitúa en poco más de 28 millones y hay una variabilidad relativa de los datos del 8,67% alrededor de esta media.

Se observa una asimetría positiva en el gráfico, el coeficiente de asimetría tiene un valor de 0,0897, y la mediana es inferior a la media.

El valor del coeficiente de curtosis es de -0,00069, por lo cual la distribución tiene una ligera menor concentración de pagos alrededor de la media en comparación con la distribución Normal.

En la siguiente tabla mostramos los pagos por año de calendario junto a los errores de predicción y los coeficientes de variación pertinentes aplicando tanto fórmula como *bootstrap residuals* y *bootstrap pairs*:

Tabla 4: Pagos por año de calendario, errores de predicción y coeficientes de variación para la distribución Poisson.

		Fórmula		<i>Bootstrap residuals</i>		<i>Bootstrap pairs</i>	
Año	Pagos por año de calendario	Error de predicción	CV	Error de predicción	CV	Error de predicción	CV
1	5.226.535,83	747.369,6	12,995%	742.708,8	14,210%	1.058.740,3	20,570%
2	4.179.394,44	710.144,6	16,991%	706.410,6	16,902%	950.884,4	22,751%
3	3.131.667,52	644.139,5	20,568%	641.730,5	20,491%	885.321,4	28,269%
4	2.127.271,92	479.125,6	22,523%	476.527,5	22,400%	795.642,8	37,402%
5	1.561.878,91	404.967,7	25,282%	403.423,5	25,829%	706.768,2	45,251%
6	1.177.743,69	364.294,9	30,931%	364.265,7	30,929%	633.661	53,802%
7	744.287,39	294.424,6	39,557%	293.417,6	39,422%	536.967,2	72,145%
8	445.521,29	250.986,8	56,335%	253.033,7	56,794%	442.639,2	99,353%
9	86.554,62	108.268,8	125,087%	107.583,8	124,295%	333.676,4	385,509%

Fuente: Elaboración propia.

Los tres métodos muestran resultados muy similares en términos de la evolución del error de predicción y coeficiente de variación a lo largo de los años, el coeficiente de variación más bajo en los años iniciales se consigue con la estimación mediante fórmula.

Bootstrap residuals tiende a tener un error de predicción ligeramente menor y un coeficiente de variación ligeramente más bajo en comparación con *bootstrap pairs*.

El último año muestra un aumento significativo en el error de predicción y en el coeficiente de variación, lo que indica una mayor variabilidad en las estimaciones. Para concluir los cálculos con la distribución Poisson, se adjuntan los valores de los estadísticos básicos y de los percentiles de la distribución predictiva de los pagos por año de calendario, tanto para *bootstrap residuals* como para *bootstrap pairs*.

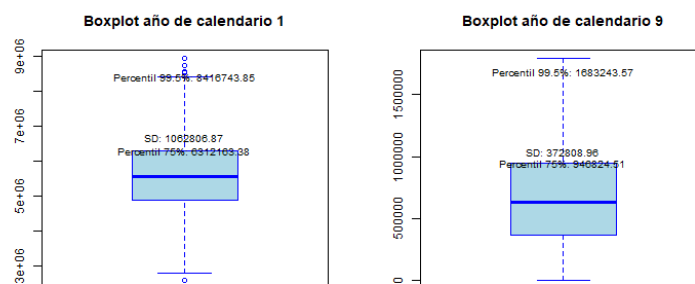
También se representa el *boxplot* de los pagos del primer año de calendario y del último año de calendario, y para mejorar la interpretación se han añadido los valores de la media, la mediana y los percentiles del 75%, y del 99,5%. Para el resto de años de calendario los gráficos están disponibles en el anexo del trabajo.

Tabla 5: Percentiles y estadísticos básicos de los pagos por año de calendario para la distribución Poisson con *bootstrap residuals*.

Año calendario	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	5.207.535	5.733.548	7.154.048	5.257.769	13,852%	0,2075037	-0,1106865
2	4.208.109	4.681.521	6.049.683	4.228.886	16,229%	0,2488105	0,04717294
3	3.156.082	3.576.893	5.049.994	3.171.020	20,079%	0,4259835	0,4410754
4	2.104.054	2.472.264	3.629.494	2.155.709	22,418%	0,4225176	0,4151118
5	1.578.041	1.841.048	2.787.872	1.587.562	26,007%	0,4564245	0,2904018
6	1.157.230	1.420.237	2.314.460	1.206.044	30,708%	0,6103596	0,8597823
7	736.419,1	946.824,5	1.788.446,3	760.457,9	40,018%	0,7080904	0,8488666
8	420.810,9	631.216,3	1.262.695,7	462.050,4	56,003%	0,8554886	1,005433
9	52.601,36	105.202,72	526.013,62	88.580,69	120,214%	1,862255	4,404386

Fuente: Elaboración propia.

Gráfico 3: *Boxplot* de los pagos por año de calendario para la distribución Poisson con *bootstrap residuals*.



Fuente: Elaboración propia, output de R.

Ambos años muestran diferencias tanto en la distribución como en la variabilidad de los pagos por año de calendario.

En el año 1, la distribución está centrada alrededor de 5.207.535, con una media ligeramente superior de 5.257.769. Tiene un coeficiente de variación del 13,85%, indicando una variabilidad moderada. La asimetría es leve y positiva 0,208, con una curtosis negativa -0,111.

En el año 9, la distribución está sesgada hacia valores bajos, con una mediana y percentiles bajos. La media está influenciada por valores extremadamente altos (*outliers*), alcanzando 88.580,69. Tiene un coeficiente de variación extremadamente alto, del 120,21%, indicando una gran variabilidad. La asimetría es pronunciadamente positiva 1,862, indicando una distribución muy sesgada hacia la derecha. La curtosis es muy elevada, 4,404, lo cual significa que hay cola pesada y concentración alrededor de la media.

Se puede ver perfectamente el análisis con los *boxplots*, y se ha de remarcar que, en ambos años, la diferencia entre la mediana y la media sugiere la presencia de asimetría en la distribución de los pagos por año de calendario.

El alto coeficiente de variación en el año 9 refleja una mayor incertidumbre y variabilidad en los pagos, posiblemente debido a fluctuaciones significativas o eventos atípicos.

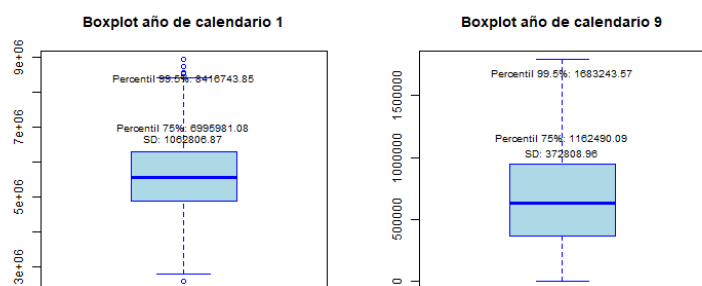
Tabla 6: Percentiles y estadísticos básicos de los pagos por año de calendario para la distribución Poisson con *bootstrap pairs*.

Año calendario	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	5.575.744	6.312.163	8.416.744	5.628.503	18,882%	0,1662191	-0,1288433
2	4.997.129	5.680.947	7.891.256	5.027.533	20,205%	0,2976072	0,01685517
3	4.208.109	4.839.325	6.680.636	4.213.790	21,557%	0,2032706	0,4168376
4	3.734.697	4.313.312	6.259.825	3.745.796	22,529%	0,1315296	0,07437741
5	3.156.082	3.682.095	5.049.731	3.173.177	23,843%	0,2102473	-0,2878135
6	2.472.264	2.998.278	4.209.161	2.521.288	27,098%	0,1807008	-0,2777828

7	1.841.048	2.261.859	3.524.291	1.864.771	30,883%	0,3606988	-0,0711293
8	1.315.034	1.643.793	2.788.135	1.341.861	36,801%	0,352077	0,163337
9	631.216,3	946.824,5	1.683.243,6	668.352,9	55,780%	0,3821065	-0,6658682

Fuente: Elaboración propia.

Gráfico 4: *Boxplot* de los pagos por año de calendario para la distribución Poisson con *bootstrap pairs*.



Fuente: Elaboración propia, output de R.

Ambos años muestran diferencias significativas en la distribución y variabilidad de los pagos por año de calendario.

En el año 1, la distribución está centrada alrededor de 5.575.744, con una media de 5.628.503 y un coeficiente de variación del 18,88%. La asimetría es leve y positiva 0,166, con una curtosis negativa -0,129.

En el año 9, la distribución está sesgada hacia valores bajos, con una mediana y percentiles bajos. La media es de 668.352,9, tiene un coeficiente de variación alto, del 55,78%. La asimetría es pronunciadamente positiva 0,382, con una curtosis negativa -0,666, indicando una distribución muy sesgada hacia la derecha y leptocúrtica.

En el año 1 los *outliers* no son tan evidentes y la distribución es más equilibrada, en el año 9 los *outliers* tienen un impacto significativo en la distribución, haciendo que sea sesgada hacia valores bajos con una variabilidad mucho más alta.

Por último, se ha calculado el valor actual de los pagos por año de calendario mediante la curva de tipo de interés de EIOPA, los datos fueron extraídos el mes de abril de 2024:

Tabla 7: Tipos de interés de EIOPA publicados el mes de abril de 2024.

<i>I1</i>	<i>I2</i>	<i>I3</i>	<i>I4</i>	<i>I5</i>	<i>I6</i>	<i>I7</i>	<i>I8</i>
3,514%	3,035%	2,783%	2,637%	2,549%	2,502%	2,477%	2,466%

Fuente: Elaboración propia.

El valor actual de los pagos por año de calendario es de 18.291.167.

También se ha calculado el valor actual del VaR de los pagos por año de calendario al nivel de confianza del 99,5%, el resultado obtenido para *bootstrap residuals* es de 29.777.911 y para *bootstrap pairs* es de 45.054.435.

Resultados con la distribución Gamma:

A continuación, se muestran los resultados con la distribución Gamma. Los valores de los coeficientes del modelo GLM son:

Tabla 8: Valores de los coeficientes del modelo GLM para la distribución Gamma.

Año	1	2	3	4	5	6	7	8	9	10
α_i	284.798,3	391.127,1	378.115,2	336.032,5	358.656,9	374.203,2	405.084,3	452.010,5	387.194,8	344.014
β_j	1	2,4807621	2,5384701	2,7115715	1,5136597	1,1171972	0,9472358	0,6378393	0,9422926	0,2385829

Fuente: Elaboración propia.

La matriz de cuantías no acumuladas es la siguiente:

Figura 7: Matriz completa de pagos no acumulados para la distribución Gamma.

Años de origen	Años de desarrollo									
	1	2	3	4	5	6	7	8	9	10
1	357.848	766.940	610.542	482.940	527.326	574.398	146.342	139.950	227.229	67.948
2	352.118	884.021	933.894	1.183.289	445.745	320.996	527.804	266.172	425.046	93.316,25
3	290.507	1.001.799	926.219	1.016.654	750.816	146.923	495.992	280.405	356.295,1	90.211,82
4	310.608	1.108.250	776.189	1.562.400	272.482	352.053	206.286	214.334,7	316.640,9	80.171,60
5	443.160	693.190	991.983	769.488	504.851	470.639	339.732,6	228.765,4	337.959,7	85.569,40
6	396.132	937.085	847.498	805.037	705.960	418.058,8	354.458,7	238.681,5	352.608,9	89.278,49
7	440.832	847.631	1.131.398	1.063.269	613.159,8	452.559	383.710,3	258.378,7	381.707,9	61.315,8
8	359.480	1.061.648	1.443.370	684.190	504.984,8	508.791,9	428.160,5	288.310	425.926,1	107.841,98
9	376.686	986.608	982.882,5	1.049.906,4	586.081,2	432.572,9	366.764,8	246.968,1	364.850,8	92.378,07
10	344.014	853.416,9	873.269,3	932.818,6	520.720,1	384.331,5	325.862,4	219.425,6	324.161,8	82.075,86

Fuente: Elaboración propia.

Finalmente, el valor obtenido para la reserva total para la distribución Gamma es de 18.085.805.

El valor del error de predicción de la reserva total con fórmula para esta distribución es de 2.702.710. Por lo cual, el error de predicción porcentualmente hablando respecto a la reserva total es del 14,94% aproximadamente.

El error de predicción con *bootstrap residuals* para la reserva total, tiene un valor de 1.061.701 y al ponerlo en porcentaje en comparación a la reserva total, da un error de un 5,87%, por tanto se observa que hay una diferencia notable entre el error de predicción con fórmula y el que se obtiene aplicando *bootstrap residuals*, dicha diferencia nos dice que el método *bootstrap residuals* para la distribución Gamma es mucho más eficiente y fiable en la precisión en la predicción de los resultados.

El error de predicción aplicando *bootstrap pairs* para la reserva total, tiene un valor de 2.436.438 y al ponerlo en porcentaje en comparación a la reserva total, da un error de un 13,47%, de ello se deriva que hay una diferencia notable entre el error de predicción aplicando *bootstrap residuals* y aplicando *bootstrap pairs*. El *bootstrap residuals* es mucho más eficiente para la distribución Gamma.

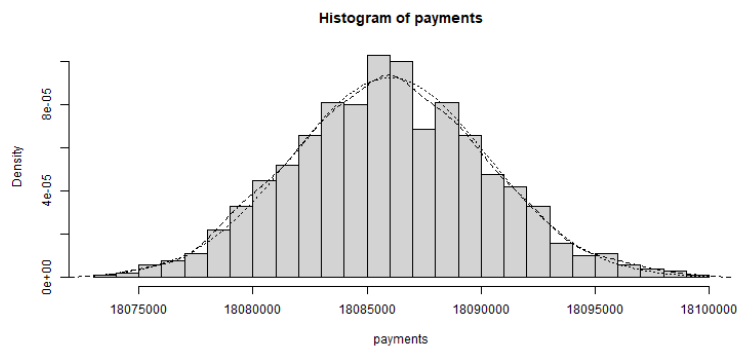
Calculamos la distribución predictiva tanto en el caso de *bootstrap residuals* como en el caso de *bootstrap pairs*, para estimar la incertidumbre asociada a la reserva total. Se obtienen los percentiles de la distribución y los estadísticos básicos y se hace la representación gráfica del histograma de la reserva total.

Tabla 9: Percentiles y estadísticos básicos de la reserva total para la distribución Gamma con *bootstrap residuals*.

Percentiles	50%	75%	90%	95%	99,5%
	18.085.929	18.088.806	18.091.517	18.093.040	18.097.547
Estadísticos básicos	Media	SD	CV	Asimetría	Curtosis
	18.085.989	4.305,748	0,02380709%	0,09963367	-0,05499622

Fuente: Elaboración propia.

Gráfico 5: Histograma de la reserva total para la distribución Gamma con *bootstrap residuals*.



Fuente: Elaboración propia, output de R.

La reserva total en términos medios tiene un valor de alrededor de 18 millones y hay muy poca variabilidad de los datos en relación con la media, solo del 0,024%. La mediana también tiene un valor muy cercano a la media. Hay una ligera asimetría positiva y el coeficiente de curtosis es negativo, con un valor de -0,055.

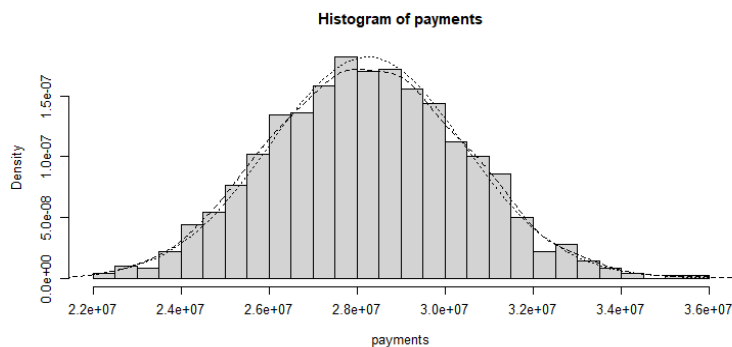
En el caso de aplicar *bootstrap pairs*, para la distribución predictiva de la reserva total se obtienen los siguientes resultados:

Tabla 10: Percentiles y estadísticos básicos de la reserva total para la distribución Gamma con *bootstrap pairs*.

Percentiles	50%	75%	90%	95%	99,5%
	28.169.495	29.688.231	31.110.802	31.767.132	33.781.180
Estadísticos básicos	Media	SD	CV	Asimetría	Curtosis
	28.229.271	2.192.980	7,768462%	0,08093949	-0,1683142

Fuente: Elaboración propia.

Gráfico 6: Histograma de la reserva total para la distribución Gamma con *bootstrap pairs*.



Fuente: Elaboración propia, output de R.

En este caso se obtiene una reserva total media que se sitúa por encima de los 28 millones, con una variabilidad relativa del 7,77%. La mediana tiene un valor inferior a la media y la distribución presenta una asimetría en la cola derecha y un coeficiente de curtosis negativo de -0,168.

En cuanto a los pagos futuros por año de calendario, aplicando tanto fórmula como *bootstrap residuals* y *bootstrap pairs*, se han obtenido los siguientes errores de predicción, expresados en valores absolutos y en porcentaje:

Tabla 11: Pagos por año de calendario, errores de predicción y coeficientes de variación para la distribución Gamma.

		Fórmula		Bootstrap residuals		Bootstrap pairs	
Año	Pagos por año de calendario	Error de predicción	CV	Error de predicción	CV	Error de predicción	CV
1	5.096.855,28	847.281,60	16,623%	652.964,77	12,811%	1.124.529,3	22,063%
2	4.050.001,53	749.549,79	18,507%	545.647,89	13,472%	990.456,9	24,455%
3	3.064.407,69	628.140,96	20,497%	434.825,65	14,189%	907.992,2	29,630%
4	2.078.010,53	431.885,83	20,783%	297.581,31	14,320 %	792.159	38,121%

5	1.510.392,68	345.880,74	22,900%	233.914,65	15,487%	698.571,1	46,250%
6	1.095.402,72	292.255,69	26,680%	194.252,15	17,733%	617.043,6	56,330%
7	692.118,2	220.057,76	31,794%	142.601,47	20,603%	523.345,3	75,6149%
8	416.539,91	181.226,50	43,507%	109.441,63	26,273%	425.398,6	102,126%
9	82.075,86	47.918,09	58,382%	26.649,16	32,468%	325.694,6	396,821%

Fuente: Elaboración propia.

Los tres métodos muestran resultados similares en términos de la evolución del error de predicción y coeficiente de variación a lo largo de los años, con el coeficiente de variación más bajo logrado en los primeros años con la estimación mediante fórmula. *Bootstrap residuals* tiende a tener un error de predicción menor y un coeficiente de variación ligeramente más bajo en comparación con los de fórmula y *bootstrap pairs*. La diferencia en este caso es notable, *bootstrap residuals* realiza estimaciones más fiables, con menores errores de predicción, y por lo cual los coeficientes de variación tienen valores muy menores respecto a los otros dos métodos.

El año 9 muestra un aumento significativo en el error de predicción y en el coeficiente de variación, lo que indica una mayor variabilidad en las estimaciones para ese año.

Para completar el análisis con la distribución Gamma, se adjuntan los valores de los estadísticos básicos y de los percentiles de la distribución predictiva de los pagos por año de calendario, tanto para *bootstrap residuals* como para *bootstrap pairs*. También se representa el *boxplot* de los pagos del primer año de calendario y del último año de calendario. Para mejorar la interpretación, se han añadido los valores de la media, la mediana y los percentiles del 75% y del 99,5%. Los gráficos de los años intermedios están disponibles en el anexo del trabajo.

En la siguiente tabla se muestran los estadísticos básicos y percentiles para la distribución obtenidos en el caso de aplicar *bootstrap residuals*.

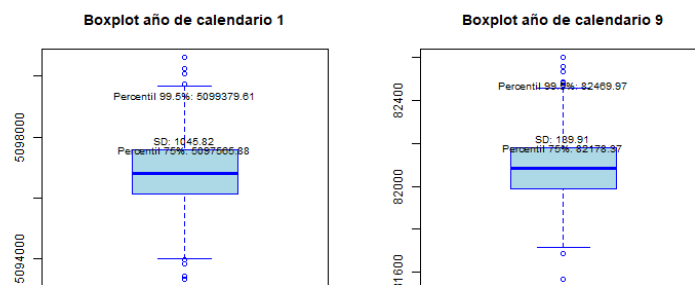
Tabla 12: Percentiles y estadísticos básicos de los pagos por año de calendario para la distribución Gamma con *bootstrap residuals*.

Año calendario	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	5.096.798	5.097.566	5.099.380	5.096.842	0,0205%	0,04420634	0,03617091
2	4.050.015	4.050.674	4.052.515	4.050.019	0,0242%	0,01101901	0,06052413
3	3.064.348	3.064.918	3.066.672	3.064.372	0,0286%	0,07154984	-0,05156381
4	2.077.977	2.078.441	2.079.656	2.077.995	0,0310%	0,01466922	-0,1007629
5	1.510.381	1.510.731	1.511.828	1.510.387	0,0363%	0,06324552	0,2469075
6	1.095.429	1.095.755	1.096.628	1.095.403	0,046%	-0,08457788	-0,250257
7	692.126,4	692.384,1	693.116,1	692.124,4	0,0556%	0,02419637	-0,01927999
8	416.545,6	416.766,6	417.573,3	416.549,1	0,0824%	0,2229111	0,2498785
9	82.081,86	82.178,37	82.469,97	82.084,37	0,170%	0,1011123	0,224467

Fuente: Elaboración propia.

El *boxplot* de los pagos por año de calendario del primer y del último año quedan representados de la siguiente manera:

Gráfico 7: *Boxplot* de los pagos por año de calendario para la distribución Gamma con *bootstrap residuals*.



Fuente: Elaboración propia, output de R.

En el año 1, la distribución está centrada alrededor de 5.096.842, con una mediana de 5.097.566 y percentiles 75% y 99,5% muy cercanos. La media es de 5.096.842, y el coeficiente de variación es extremadamente bajo, del 0,0205%. La asimetría es leve y positiva, 0,044, con una curtosis ligeramente positiva, 0,036.

En el año 9, la distribución muestra una mayor variabilidad y sesgo hacia valores bajos. La mediana y los percentiles son significativamente más bajos que la media, que es de 82.084,37. El coeficiente de variación es del 0,170%, indicando una dispersión de los datos mayor que en el primer año. La asimetría es positiva, de 0,101 igual que la curtosis, de 0,224, lo que sugiere una distribución sesgada hacia la derecha y leptocúrtica.

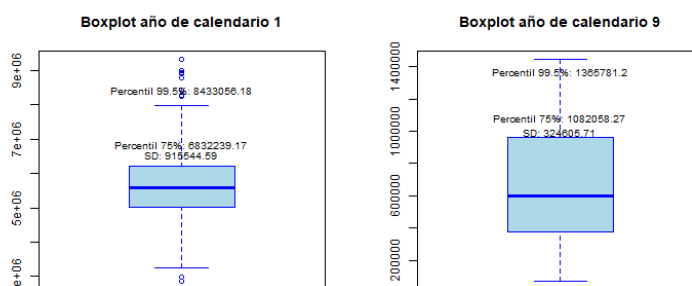
Tabla 13: Percentiles y estadísticos básicos de los pagos por año de calendario para la distribución Gamma con *bootstrap pairs*.

Año calendario	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	5.574.581	6.224.448	8.433.056	5.629.137	16,264%	0,3209107	0,5121843
2	5.020.392	5.590.693	7.151.461	5.036.231	16,413%	0,1170409	-0,09553141
3	4.231.665	4.774.673	6.508.803	4.233.019	18,830%	0,1197773	0,1469395
4	3.757.500	4.272.704	5.643.562	3.770.101	19,472%	0,07549443	-0,3337787
5	3.183.811	3.654.381	4.873.115	3.199.344	20,573%	0,1359776	-0,2548117
6	2.495.838	2.932.942	3.864.882	2.506.753	23,363%	0,02603992	-0,3725532
7	1.831.259	2.216.083	3.230.505	1.864.660	27,003%	0,3110886	-0,134892
8	1.316.482	1.598.474	2.311.682	1.323.562	31,057%	0,1732945	-0,6011536
9	595.947,7	959.699,1	1.365.781,2	666.471,3	48,705%	0,1683729	-1,297668

Fuente: Elaboración propia.

El *boxplot* de los pagos por año de calendario del primer y del último año quedan representados de la siguiente manera:

Gráfico 8: *Boxplot* de los pagos por año de calendario para la distribución Gamma con *bootstrap pairs*.



Fuente: Elaboración propia, output de R.

En el año 1, la distribución está centrada alrededor de 5.629.137, con una mediana de 5.574.581 y percentiles del 75% y del 99,5% muy cercanos. El coeficiente de variación es del 16,26%, la asimetría es positiva, 0,321, con una curtosis positiva, 0,512.

En el año 9, la distribución muestra una mayor variabilidad y sesgo hacia valores bajos. La mediana y los percentiles son significativamente más bajos que la media, que es de 666.471,3. El coeficiente de variación es alto, del 48,70%, indicando una gran dispersión de los datos. La asimetría es positiva, 0,168, con una curtosis negativa, -1,297, lo que sugiere una distribución muy sesgada hacia la izquierda y leptocúrtica.

Para el cálculo del valor actual de los pagos por año de calendario se ha utilizado la misma la curva de tipo de interés de EIOPA citada en el caso de la distribución Poisson.

Ahora, para la distribución Gamma, el valor actual de los pagos por año de calendario es de 17.714.684. El valor actual del VaR al nivel de confianza del 99,5% obtenido para *bootstrap residuals* es de 17.728.276 y para *bootstrap pairs* es de 42.070.523.

Análisis comparativo de resultados:

Una vez vistos todos los resultados, se van a realizar comparaciones entre las dos distribuciones.

Inicialmente, comparando los resultados obtenidos, la reserva total para la distribución Poisson es de 18.680.856, mientras que para la distribución Gamma es de 18.085.805. Poisson es mayor respecto a Gamma en 595.051 o, visto porcentualmente, en un 3,18%.

Seguidamente, los errores predicción obtenidos para la distribución Poisson son los siguientes:

- El error de predicción mediante fórmula para la reserva total es del 15,76% aproximadamente.
- El error de predicción con *bootstrap residuals* para la reserva total es del 15,75%.
- El error de predicción con *bootstrap pairs* para la reserva total es del 12,79%.

Los errores de predicción obtenidos para la distribución Gamma son los siguientes:

- El error de predicción mediante fórmula para la reserva total es del 14,94%.
- El error de predicción con *bootstrap residuals* para la reserva total es del 5,87%.
- El error de predicción con *bootstrap pairs* para la reserva total es del 13,47%.

En primer lugar, se puede observar que en las dos distribuciones el error de predicción es menor al aplicar *bootstrap* en comparación con aplicar fórmula, por lo tanto, con dicha metodología se mejoran las predicciones, independientemente de la distribución que se aplique.

Para la distribución Poisson se consigue un menor error de predicción con *bootstrap pairs* mientras que para la distribución Gamma el menor porcentaje de error se obtiene con *bootstrap residuals*.

El menor error de predicción se ha conseguido con *bootstrap residuals* para la distribución Gamma, es un 5,87%, en todos los casos restantes el porcentaje es más del doble de este valor.

La diferencia a destacar al comparar las dos distribuciones sería que, aunque tanto para la distribución Poisson como para la distribución Gamma los resultados obtenidos son bastante similares en el caso de fórmula y de *bootstrap pairs*, la distribución Gamma proporciona unas mejores estimaciones que la distribución de Poisson debido al resultado obtenido en *bootstrap residuals*, que con diferencia es el caso con un menor error de predicción respecto al resto.

Si analizamos los errores de predicción para los pagos por año de calendario, cabe destacar que, obviando el caso de *bootstrap residuals* para la distribución Gamma, las tendencias en los coeficientes de variación son muy similares, los incrementos son

exponenciales, llegando a valores altos en los últimos tres años, destacando los últimos dos particularmente.

Los coeficientes de variación obtenidos tanto para Poisson como para Gamma en *bootstrap pairs*, para estos años finales, tienen unos valores muy elevados. Concretamente destacan para Poisson el último año para el coeficiente de variación con fórmula con 125%, en *bootstrap residuals*, el último año también con un coeficiente de variación del 125% y para *bootstrap residuals* los últimos dos años con coeficientes de variación del 100% y 385% aproximadamente. Para la distribución Gamma los coeficientes de variación obtenidos para *bootstrap pairs* en los últimos dos años tienen valores del 102% y de 396% aproximadamente.

En el caso de la distribución Gamma con *bootstrap residuals*, los coeficientes de variación en los pagos por año de calendario empiezan teniendo un valor de un 13-14% en los primeros cuatro años y para los dos últimos años tienen un valor del 26% y 32%, respectivamente.

En el apartado de los estadísticos básicos para *bootstrap residuals*, a partir de los resultados obtenidos para ambas distribuciones, se observan diferencias significativas en las características estadísticas de la reserva total. Mientras que para la distribución Poisson la media de la reserva total se sitúa en más de 19 millones, para la distribución Gamma la media se sitúa ligeramente por encima de los 18 millones, una diferencia razonable. Además, el coeficiente de variación en el conjunto de datos Poisson es del 15,43%, indicando una mayor variabilidad relativa en comparación con el valor obtenido para la distribución Gamma que es de 0,024%, por lo cual la variabilidad es menor para el tipo de datos utilizado usando la distribución Gamma. En cuanto a la asimetría, la distribución Poisson muestra una asimetría positiva más pronunciada 0,175 en comparación con la distribución Gamma 0,0997, lo que indica una cola positiva más evidente en la distribución de la reserva total. Por último, el cociente de curtosis muestra que las distribuciones han obtenido unos resultados muy parejos, Poisson tiene un valor de 0,062 y Gamma de -0,055, las distribuciones no tienen colas pesadas, el coeficiente negativo en Gamma nos indica que las colas serán menos pesadas, aunque los valores obtenidos son pequeños en ambos casos.

En el apartado de los estadísticos básicos de la reserva total para *bootstrap pairs*, no hay diferencias significativas entre los resultados de la distribución Poisson y de la

distribución Gamma, los estadísticos obtenidos son muy similares. Sin embargo, al comparar los resultados de la reserva total media, en este caso tanto para la distribución de Poisson como para la distribución Gamma se ha obtenido un valor superior a 28 millones, mucho más alto que el que se obtiene con estas distribuciones al aplicar *bootstrap residuals*.

Si analizamos los resultados de los estadísticos básicos para los pagos por año de calendario, se mantiene que aplicando *bootstrap residuals* para la distribución Gamma se obtienen unos coeficientes de variación mínimos, inferiores al 0,5% para todos los años, por lo tanto, las estimaciones son muy similares al valor medio. En cambio, para la distribución Poisson y aplicando *bootstrap pairs* para la distribución Gamma va aumentando el valor de la dispersión alrededor de la media a medida que avanzamos en los años de calendario.

Es destacable resaltar que los valores medios de los pagos futuros por año de calendario obtenidos con *bootstrap pairs* son importes bastante más elevados que los obtenidos con *bootstrap residuals*, tanto para la distribución Poisson como para la distribución Gamma. Esta misma tendencia se observa al comparar los percentiles de la distribución predictiva.

Finalmente, si comparamos los valores actuales de los pagos futuros por año de calendario para la distribución Gamma el valor es de unos 17,71 millones, inferior al obtenido para la distribución Poisson, que es de 18,29 millones, tal como ya sucedía con el importe de la reserva total, cuando no teníamos en cuenta la rentabilidad.

En el caso del valor actual del VaR al 99,5% de nivel de confianza de los pagos futuros por año de calendario, aplicando *bootstrap residuals* se obtiene un valor actual muy similar para la distribución Gamma, de 17,73 millones, pero aumenta hasta 29,78 millones para la distribución Poisson.

Aplicar *bootstrap pairs* hace disparar el valor actual del VaR de los pagos futuros por año de calendario, llegando a más de 45 millones para la distribución de Poisson y a más de 42 millones para la distribución Gamma.

5.- Conclusiones

El trabajo ha tratado un aspecto respecto al cual aún no se han realizado grandes investigaciones, el cálculo de la estimación *bootstrap* del VaR para el cálculo de reservas no vida, centrándose en los cálculos por año de calendario.

Para realizar los cálculos, se ha utilizado un código propio de R. Este código ha sido diseñado específicamente para obtener todos los resultados numéricos y gráficos del trabajo.

En el trabajo se han utilizado dos distribuciones, la Poisson y la Gamma, junto con dos tipos de *bootstrap*, *residuals* y *pairs*. Como se ha visto durante el trabajo, el menor error de predicción en el cálculo de la reserva total para los datos analizados se ha obtenido con la distribución Gamma, aplicando *bootstrap residuals*. El coeficiente de variación es de un 5% aproximadamente, comparándolo con las otras distribuciones y metodologías las diferencias son muy grandes.

Si se analiza solamente el método *bootstrap residuals* para la distribución Gamma y sus coeficientes de variación año a año para los pagos por año de calendario, se ve perfectamente como tanto esta distribución como esta metodología *bootstrap* se adaptan perfectamente a los datos, obteniendo coeficientes de variación ínfimos.

Seguidamente, en el trabajo se calcula el valor actual de los pagos por año de calendario y del VaR, los tipos de interés utilizados para actualizar las cuantías son los de EIOPA. Los resultados obtenidos corroboran las diferencias vistas en los coeficientes de variación, debido a que cuanto menor sea la diferencia entre el valor actual de los pagos futuros y del VaR, mejor se ha ajustado tanto la distribución como la metodología *bootstrap* a los datos.

En este caso, *bootstrap residuals* en el cálculo del valor actual del VaR para la distribución de Poisson y Gamma obtiene valores más cercanos al valor actual de los pagos futuros, mientras que las predicciones que realiza *bootstrap pairs* obtiene resultados elevados, llegando a duplicar el valor actual de los pagos futuros.

Cabe destacar que, para la distribución Gamma, el valor actual de los pagos por año de calendario es de 17.714.684, mientras que el valor actual del VaR al nivel de confianza del 99,5% obtenido para *bootstrap residuals* es de 17.728.276. Este resultado corrobora los resultados obtenidos previamente, por lo cual, para esta serie de datos, se llega a la conclusión que para los datos analizados la distribución y la metodología *bootstrap* que realizan las predicciones más ajustadas y con menores coeficientes de variación son la distribución Gamma junto con *bootstrap residuals*.

6.- Bibliografía

Boj, E., & Costa, T. (2018). Cálculo de provisiones para prestaciones pendientes de declaración con modelos lineales generalizados: formulación del error de predicción por año de calendario. *Cuadernos De Gestión*, 17(2), 157–174.

Boj, E., Claramunt, M.M. & Costa, T. (2021). Tarificación y provisiones (tercera edición). *Depósito Digital de la Universidad de Barcelona, colección OMADO*.

Boj, E., Costa, T., & Espejo, J. (2014). Provisiones técnicas por años de calendario. *Anales del Instituto de Actuarios Españoles, Tercera Época*, 20, 83-116.

Clarke, T.G. & Harland, N. (1974). A practical statistical method of estimating claims liability and claims cash flow. *ASTIN Bulletin*, 8, 26-37.

Davidson, R., Hinkley, D. (1997). Bootstrap methods and their application. Cambridge, UK: Cambridge University Press.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7, 1-26.

Gesmann, M., Murphy, D., Zhang, Y., Carrato, A., Crupi, G., Dutang, C., La-Coume, A., Charpentier, A., Wüthrich, M., Concina, F., & Dale. (2023). Statistical methods and models for claims reserving in general insurance: Package ChainLadder. *Versión 0.2.18*.

Hinkley, D. (2017). GIRA - General Insurance Reserving Application.

Kramreiter, H. & Straub, E. (1973). On the calculacation of IBNR reserves II. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 73, 177-190.

Lowe, J. (1994). A practical guide to measuring reserve variability using: Bootstrapping, Operational Time and a distribution free approach. *Proceeding of the 1994 General Insurance Convention, Institute of Actuaries and Faculty of Actuaries*.

Paule-Vinaez, J., Coca-Pérez, J. L., & Granado-Sánchez, M. (2020). Análisis comparativo de metodologías para la cuantificación de provisiones técnicas en entidades aseguradoras. Adaptación a Solvencia II. *Revista mexicana de economía y finanzas*, 15(3), 313-329.

Paule-Vinaez, J., Granado-Sánchez, M. & Coca-Pérez, J. L. (2018). Método Bootstrap para el cálculo de provisiones técnicas por siniestros. *Cuadernos De Gestión*, 18(2), 163–184.

Pinheiro, P. J. R., Andrade e Silva, J. M. & Centeno, M. D. L. (2003). Bootstrap methodology in claim reserving. *The Journal of Risk and Insurance*, 4, 701-714.

R Development Core Team (2024). R: A language and environment for statistical computing. Vienna (Austria).

Skurnick, D. (1973). A survey of loss reserving methods. *Proceedings of the Casualty Actuarial Society*, 60(113), 16-58.

Solanas, A., & Olivera, V. (1992). Bootstrap: fundamentos e introducción a sus aplicaciones. *Anuario de psicología/The UB Journal of psychology*, número 55.

Tarbell, T.F. (1934). Incurred but not reported claims reserves. *Proceedings of the Casualty Actuarial Society*, 20, 275-280.

Taylor, G. & Ashe, F.R. (1983). Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics*, 23, 37-61

7.- Anexos

Cálculo de las reservas por año de origen Poisson:

El modelo y los resultados son los mismos que para el cálculo de las reservas totales, la única diferencia está en el valor del cálculo de las reservas y en el a distribución Poisson. error de predicción de las reservas.

Tabla 14: Pagos por año de origen, errores de predicción y coeficientes de variación para la distribución Poisson.

Año	Pagos por año de origen	Fórmula		Bootstrap residuals		Bootstrap pairs	
		Error de predicción	CV	Error de predicción	CV	Error de predicción	CV
1	94.633,81	110.099,6	116,34280%	106.707,2	112,75797%	333.673,6	352,59446%
2	469.511,29	216.042,8	46,01439%	215.841,1	45,97144%	427.108,0	90,96863%
3	709.637,82	260.871,3	36,76119%	261.259,8	36,81593%	576.599,8	81,25269%
4	984.888,64	303.549,1	30,82065%	302.739,4	30,73844%	652.846,3	66,28630%
5	1.419.459,46	375.012,8	26,41941%	374.492,6	26,38276%	661.460,7	46,59948%
6	2.177.640,62	495.376,8	22,74833%	493.433,5	22,65909%	1.031.234,4	47,35558%
7	3.920.301,01	789.959,7	20,15048%	799.064,2	20,38273%	990.903,0	25,27620%
8	4.278.972,26	1.046.512,6	24,45710%	1.094.562,7	25,58004%	941.835,1	22,0107%
9	4.625.810,69	1.980.100,7	42,80549%	2.014.502,6	43,54918%	1.079.577,7	23,33813%

Fuente: Elaboración propia.

Percentiles y estadísticos básicos con bootstrap residuals.

Tabla 15: Percentiles y estadísticos básicos de los pagos por año de origen para la distribución Poisson con bootstrap residuals.

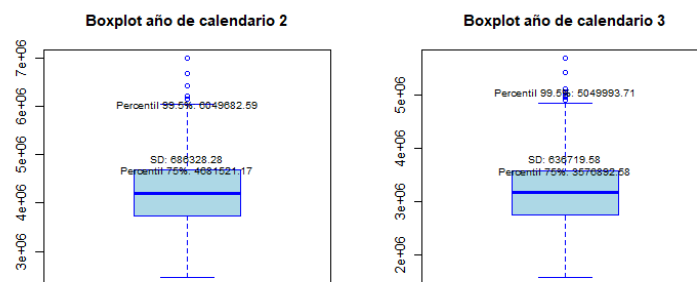
Año origen	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	526.013.6	841.621.8	1.736.107.9	605.020,9	62,50699%	0,8822599	0,3043512
2	1.578.041	1.893.649	2.998.541	1.594.189	30,24189%	0,349332	0,1411282

3	1.420.237	1.841.048	3.156.345	1.441.277	42,434%	0,3536267	-0,2635548
4	2.419.663	2.840.474	4.365.913	2.408.038	28,86236%	0,2082749	-0,1393587
5	3.208.683	3.747.847	5.154.933	3.257.445	22,16522%	0,1654263	0,08707212
6	4.313.312	4.997.129	6.733.237	4.221.943	25,47519%	-0,1405811	-0,2682342
7	3.682.095	4.471.116	6.575.433	3.780.775	26,31775%	0,4932029	0,00394
8	4.839.325	5.628.346	7.469.656	4.907.286	20,27588%	0,1188833	-0,023339
9	5.943.954	6.785.576	8.784.953	6.014.282	18,53988%	0,0522	-0,1502

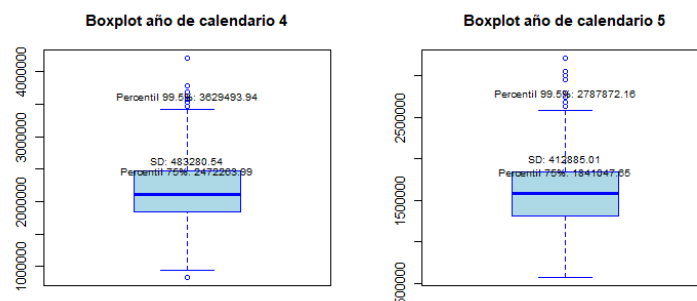
Fuente: Elaboración propia.

Boxplots con bootstrap residuals para años de calendario.

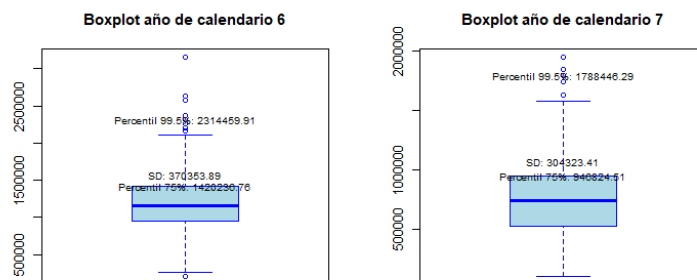
Gráfico 9: Boxplots de los pagos por año de calendario para la distribución Poisson con *bootstrap residuals*.



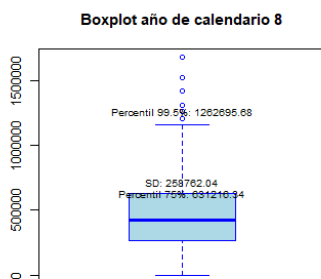
Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.

Percentiles y estadísticos básicos con *bootstrap pairs*.

Tabla 16: Percentiles y estadísticos básicos de los pagos por año de calendario para la distribución Poisson con *bootstrap pairs*.

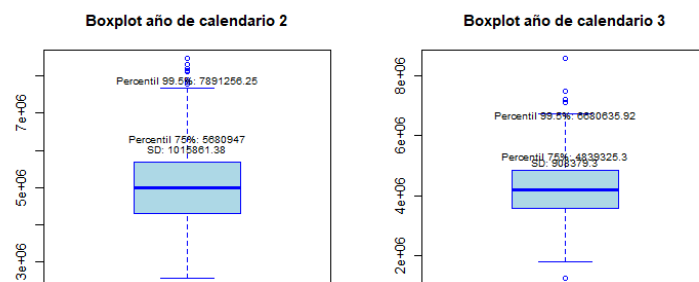
Año origen	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	52.601,36	157.804,08	5.262.76,62	104.939,7	106,9389%	1,473014	2,656645
2	473.412,3	631.216,3	1.157.230	490.981,1	45,50218%	0,5891382	0,4720241
3	683.817,7	894.223,1	1.578.040,8	727.108,6	35,93524%	0,6718	0,8001
4	999.425,9	1.209.831,3	1.893.649	1.005.370	30,41433%	0,4227357	0,2153855
5	1.420.237	1.683.244	2.419.926	1.434.597	25,9323%	0,2036089	0,000358
6	2.156.656	2.472.264	3.630.020	2.168.702	22,87799%	0,3449427	0,1374727
7	3.892.501	4.418.514	6.627.772	3.944.050	19,87356%	0,4914701	0,7166621

8	4.313.312	5.049.731	7.838.129	4.371.699	25,85594%	0,5151221	0,4114172
9	4.628.920	5.891.352	10.628.894	4.760.160	42,38164%	0,6156434	0,9018681

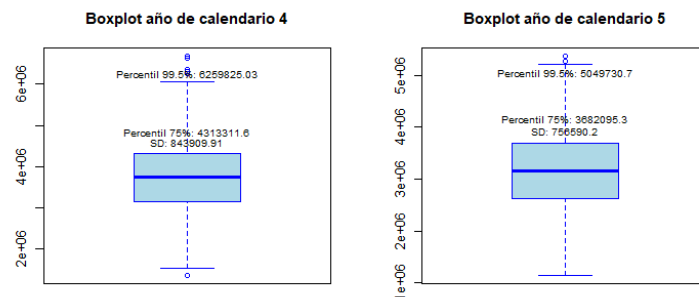
Fuente: Elaboración propia.

Boxplots con *bootstrap pairs* para años de calendario.

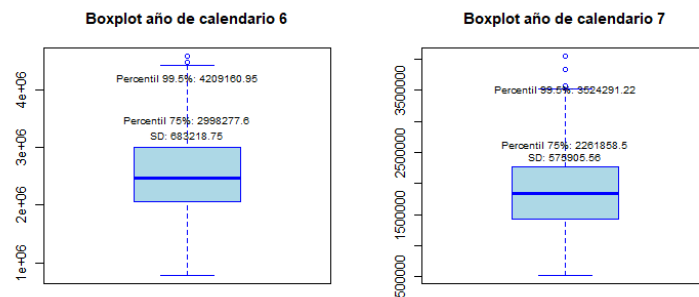
Gráfico 10: *Boxplot* de los pagos por año de calendario para la distribución Poisson con *bootstrap pairs*.



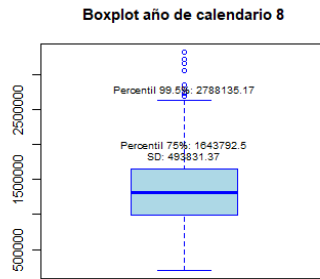
Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R..



Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.

Cálculo de las reservas por año de origen Gamma:

El modelo y los resultados son los mismos que para el cálculo de las reservas totales, la única diferencia está en el valor del cálculo de las reservas y en el error de predicción de las reservas.

Tabla 17: Pagos por año de origen, errores de predicción y coeficientes de variación para la distribución Gamma.

		Fórmula		<i>Bootstrap residuals</i>		<i>Bootstrap pairs</i>	
Año	Pagos por año de origen	Error de predicción	CV	Error de predicción	CV	Error de predicción	CV
1	93.316,25	45.166,37	48,40%	30.298,75	32,46%	305.100,4	326,95%
2	446.506,96	160.557,17	35,95%	119.334,89	26,72%	411.260,0	92,10%
3	611.147,19	177.624,61	29%	126.847,64	20,75%	580.352,7	94,96%
4	992.027,16	254.470,93	25,65%	174.635,32	17,60%	643.141,5	64,83%
5	1.453.086,32	351.334,26	24,17%	227.208,15	15,63%	668.236,9	46%
6	2.186.161,87	526.287,87	24,07%	316.433,90	14,47%	1.017.218,6	46,52%
7	3.665.072,13	941.322,25	25,68%	532.016,41	14,51%	1.016.800,2	27,74%
8	4.122.404,71	1.175.945,87	28,52%	556.356,58	13,50%	990.074,5	24%
9	4.516.082,02	1.667.392,40	36,92%	566.682,32	12,55%	1.150.801,2	25,50%

Fuente: Elaboración propia.

Percentiles y estadísticos básicos con *bootstrap residuals*.

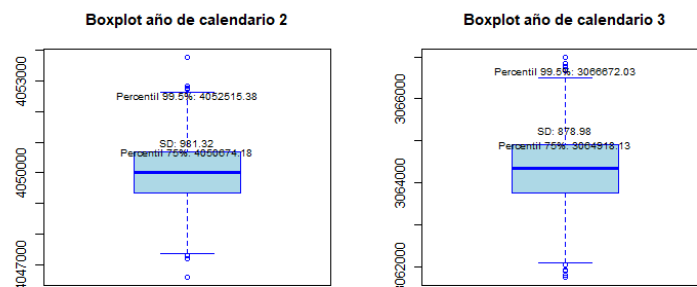
Tabla 18: Percentiles y estadísticos básicos de los pagos por año de origen para la distribución Gamma con *bootstrap residuals*.

Año origen	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	93.324,20	93.419,02	93.773,84	93.325,84	0,1651346%	0,2057103	0,3382918
2	446.504,6	446.720,6	447.309,2	446.517,1	0,06978305%	0,07729288	0,05203582
3	611.154,9	611.374,5	612.080,1	611.153,3	0,05636338%	0,0574301	0,2872668
4	992.041,2	992.332,8	993.291,8	992.043,4	0,04412157%	0,07011	0,00768
5	1.453.105	1.453.464	145.4498	1.453.096	0,03809189%	-0,008054	-0,049321
6	2.186.125	2.186.643	2.187.996	2.186.139	0,0341582%	-0,078658	0,070148
7	3.665.001	3.665.788	3.668.295	3.665.067	0,0306504%	0,3171637	0,185765
8	4.122.432	4.123.454	4.126.116	4.122.437	0,03666165%	-0,00525816	-0,07017722
9	4.516.081	4.517.708	4.524.523	4.516.170	0,06017742%	0,326867	0,8282004

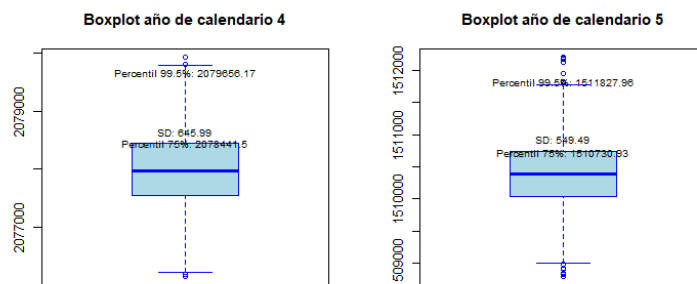
Fuente: Elaboración propia.

Boxplots con *bootstrap residuals* para año de calendario.

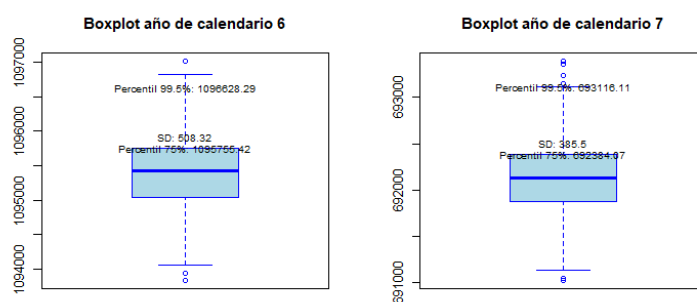
Gráfico 11: Boxplots de los pagos por año de calendario para la distribución Gamma con *bootstrap residuals*.



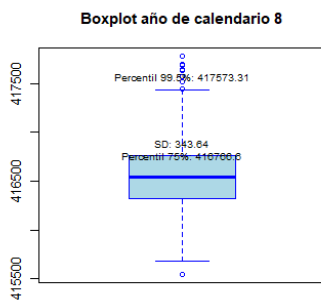
Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.

Percentiles y estadísticos básicos con *bootstrap pairs*.

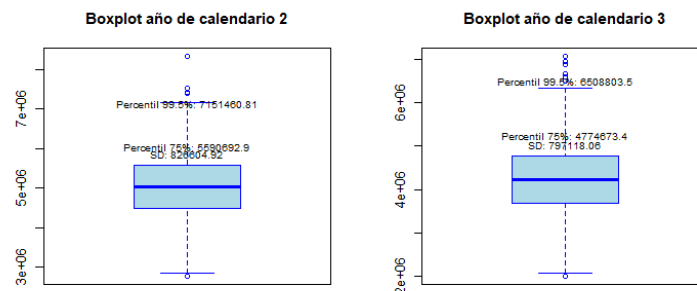
Tabla 19: Percentiles y estadísticos básicos de los pagos por año de origen para la distribución Gamma con *bootstrap pairs*.

Año origen	50%	75%	99,5%	Media	CV	Asimetría	Curtosis
1	4.386.86,8	839.015,5	1.338.733,1	576.303,1	52,68011%	0,6178804	-0,8571652
2	1.578.011	1.906.304	2.522.485	1.610.079	24,4446%	-0,02462315	-0,2457538
3	1.478.878	1.819.902	2.967.305	1.457.574	38,85443%	0,2709954	-0,2804718
4	2.364.536	2.795.750	3.988.346	2.376.669	26,04423%	0,1576944	-0,3859193
5	3.278.822	3.675.909	5.047.584	3.257.803	19,29046%	0,1582739	-0,02050543
6	4.358.910	4.958.985	6.122.298	4.241.101	22,79518%	-0,4438729	-0,2796579
7	3.724.216	4.399.649	6.139.600	3.815.004	22,71384%	0,3371543	-0,3147145
8	4.824.610	5.391.762	6.971.610	4.856.531	16,86333%	0,1220506	-0,04837382
9	6.030.510	6.717.054	8.457.570	6.038.215	16,58804%	0,01747538	-0,235783

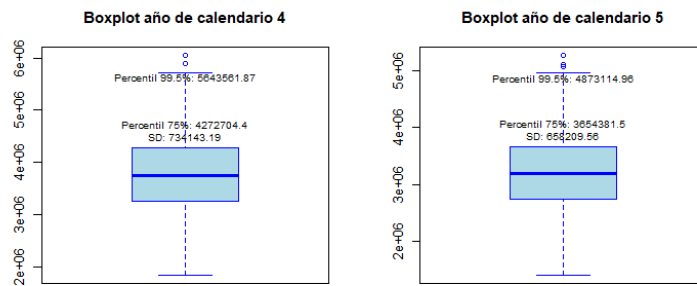
Fuente: Elaboración propia.

Boxplots con *bootstrap pairs* para años de calendario.

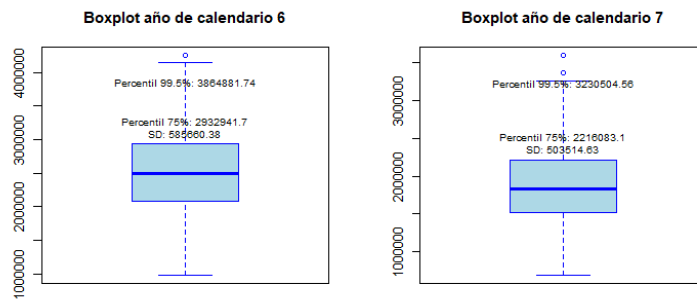
Gráfico 12: Boxplots de los pagos por año de calendario para la distribución Gamma con *bootstrap pairs*.



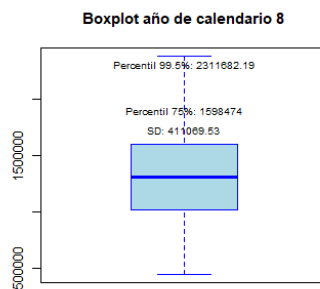
Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.



Fuente: Elaboración propia, output de R.

Código R

#datos acumulados año de origen

```
C1<-c(357848,1124788,1735330,2218270,2745596,3319994,3466336,3606286,3833515,3901463)
```

```
C2<-c(352118,1236139,2170033,3353322,3799067,4120063,4647867,4914039,5339085)
```

```
C3<-c(290507,1292306,2218525,3235179,3985995,4132918,4628910,4909315)
```

```
C4<-c(310608,1418858,2195047,3757447,4029929,4381982,4588268)
```

```
C5<-c(443160,1136350,2128333,2897821,3402672,3873311)
```

```
C6<-c(396132,1333217,2180715,2985752,3691712)
```

```
C7<-c(440832,1288463,2419861,3483130)
```

```
C8<-c(359480,1421128,2864498)
```

```
C9<-c(376686,1363294)
```

```
C10<-344014
```

#triángulo datos acumulados

```
T1<-C1
```

```
T2<-c(C2,NA)
```

```
T3<-c(C3,rep(NA,2))
```

```
T4<-c(C4,rep(NA,3))
```

```
T5<-c(C5,rep(NA,4))
```

```
T6<-c(C6,rep(NA,5))
```

```
T7<-c(C7,rep(NA,6))
```

```
T8<-c(C8,rep(NA,7))
```

```
T9<-c(C9,rep(NA,8))
```

```
T10<-c(C10,rep(NA,9))
```

```
triangleacum<-rbind(T1,T2,T3,T4,T5,T6,T7,T8,T9,T10)
```

```
triangleacum
```

#datos no acumulados años de origen

```
c1<-c(C1[1],diff(C1))
```

```
c2<-c(C2[1],diff(C2))
```

```

c3<-c(C3[1],diff(C3))
c4<-c(C4[1],diff(C4))
c5<-c(C5[1],diff(C5))
c6<-c(C6[1],diff(C6))
c7<-c(C7[1],diff(C7))
c8<-c(C8[1],diff(C8))
c9<-c(C9[1],diff(C9))
c10<-C10

#triángulo datos no acumulados

t1<-c1
t2<-c(c2,NA)
t3<-c(c3,rep(NA,2))
t4<-c(c4,rep(NA,3))
t5<-c(c5,rep(NA,4))
t6<-c(c6,rep(NA,5))
t7<-c(c7,rep(NA,6))
t8<-c(c8,rep(NA,7))
t9<-c(c9,rep(NA,8))
t10<-c(c10,rep(NA,9))

trianglenoacum<-rbind(t1,t2,t3,t4,t5,t6,t7,t8,t9,t10)

trianglenoacum

#Vector de cuantías no acumuladas

cij<-c(c1,c2,c3,c4,c5,c6,c7,c8,c9,c10)

## Selección distribución:

# Caso Poisson

ro=1

fam = quasipoisson

# Caso Gamma

ro=2

```

```

fam = Gamma(link = "log")

## Reserva total #####

# Cálculo de la reserva

n<-length(cij)

k<-trunc(sqrt(2*n))

i<-rep(1:k,k:1);i<-as.factor(i)

j<-sequence(k:1);j<-as.factor(j)

orig.CL<-glm(cij~i+j,fam)

coefs<-exp(as.numeric(coef(orig.CL)))

alpha<-c(1,coefs[2:k])*coefs[1]

beta<-c(1,coefs[(k+1):(2*k-1)])

orig.fits<-alpha%*%t(beta)

future<-row(orig.fits)+col(orig.fits)-1>k

orig.reserve<-sum(orig.fits[future])

orig.reserve

# cálculo residuos de Pearson

Prs.resid<-(cij-fitted(orig.CL))/sqrt(fitted(orig.CL)^ro)

# cálculo n, p i phi.p

n<-(k*(k+1))/2

p<-2*k-1

phi.P<-sum(Prs.resid^2)/(n-p)

# cálculo residuos escalados

Adj.Pr.resid<-Prs.resid*sqrt(n/(n-p))

# Matriz de varianza y covarianza

cij.l<-xtabs(cij~i+j)

cij.v<-as.vector(cij.l)

ii<-row(cij.l);jj<-col(cij.l)

futurebis<-as.numeric(ii+jj-1>k)

ii<-as.factor(ii)

```



```

jj<-as.factor(jj)

Cov.eta <- model.matrix(cij.v~ii+jj)%*%vcov(orig.CL)%*%t(model.matrix(cij.v~ii+jj))

mu.hat<-as.vector(orig.fits*future)

# Matriz completa pagos no acumulados

matrixnoacum<-orig.fits

matrixnoacum

#Matriz completa pagos acumulados

matrixacum<-t(apply(matrixnoacum,1,cumsum))

matrixacum

## Error de predicción con fórmula para la reserva total

PEfor<-sqrt(phi.P*sum(mu.hat^ro)+t(mu.hat)%*%Cov.eta%*%mu.hat)

PEfor

PEfor/orig.reserve*100

## Error de predicción con bootstrap para la reserva total

set.seed(6345789)

nBoot<-1000;payments<-reserves<-numeric(nBoot)

for(boots in 1:nBoot){

  Ps.cij<-sample(Adj.Prs.resid,n,replace=TRUE)

  Ps.cij<-Ps.cij*sqrt(fitted(orig.CL))+fitted(orig.CL)

  Ps.cij<-pmax(Ps.cij,0)

  Ps.CL<-glm(Ps.cij~i+j,fam)

  coefs<-exp(as.numeric(coef(Ps.CL)))

  Ps.alpha<-c(1,coefs[2:k])*coefs[1]

  Ps.beta<-c(1,coefs[(k+1):(2*k-1)])

  Ps.fits<-Ps.alpha%*%t(Ps.beta)

  Ps.reserve<-sum(Ps.fits[future])

  Ps.totpayments<-phi.P*rpois(1,Ps.reserve/phi.P)

  reserves[boots]<-Ps.reserve

  payments[boots]<-Ps.totpayments

```

```

}

PEbs<-sqrt(phi.P*sum(orig.fits[future]^ro)+sd(reserves)^2)

PEbs

PEbs/orig.reserve*100

# Distribución predictiva
quantile(payments, c(0.5,0.75,0.90,0.95,0.995))

mean(payments)

sd(payments)

# coeficiente de variación en %
cv<-(sd(payments)/mean(payments))*100

pp<-(payments-mean(payments))/sd(payments)

#assimetria
sum(pp^3)/(nBoot-1)

#Kurtosis
sum(pp^4)/(nBoot-1) -3

hist(payments,breaks=21,prob=TRUE)

lines(density(payments), lty="dashed")

curve(dnorm(x, mean = mean(payments), sd = sd(payments)),
lty="dotted", add=TRUE)

## Reservas años de origen#####

# Cálculo de las reservas
orig.CL<-glm(cij~i+j,fam)
coefs<-exp(as.numeric(coef(orig.CL)))
alpha<-c(1,coefs[2:k])*coefs[1]
beta<-c(1,coefs[(k+1):(2*k-1)])
orig.fits<-alpha%*%t(beta)
orig.prov<-numeric(k-1)
for (orig in 1:(k-1)){
  orig.prov[orig]<-sum(orig.fits[orig+1,(k-(orig-1)):k])

```

```

}

orig.prov

## Error de predicción con fórmula para las reservas por año de origen

mu.hat.m<-matrix(mu.hat,k,k)

mu.hat.orig<-matrix(0,k,k)

mu.hat.vec<-numeric(k*k)

PEfororig<-numeric(k-1)

for(orig in 1:(k-1)){

  mu.hat.orig<-rbind(matrix(0,orig,k),mu.hat.m[orig+1,],matrix(0,k-(orig+1),k))

  mu.hat.vec<-as.vector(mu.hat.orig)

  PEfororig[orig]<-sqrt(phi.P*sum(mu.hat.vec^ro)+t(mu.hat.vec)%*%Cov.eta%*%mu.hat.vec)

}

PEfororig

PEfororig/orig.prov*100

## Error de predicción con bootstrap para las reservas por año de origen

set.seed(6345789)

nBoot<-1000;payments<-reservesorig<-matrix(0,nBoot,k-1)

for(boots in 1:nBoot){

  Ps.cij<-sample(Adj.Prs.resid,n,replace=TRUE)

  Ps.cij<-Ps.cij*sqrt(fitted(orig.CL))+fitted(orig.CL)

  Ps.cij<-pmax(Ps.cij,0)

  Ps.CL<-glm(Ps.cij~i+j,fam)

  coefs<-exp(as.numeric(coef(Ps.CL)))

  Ps.alpha<-c(1,coefs[2:k])*coefs[1]

  Ps.beta<-c(1,coefs[(k+1):(2*k-1)])

  Ps.fits<-Ps.alpha%*%t(Ps.beta)

  provor<-numeric(k-1)

  payori<-numeric(k-1)

  for(orig in 1:(k-1)){

```

```

    provor[orig]<-sum(Ps.fits[orig+1,(k-(orig-1)):k])
    payori[orig]<- phi.P*rpois(1,provor[orig]/phi.P)
  }
  reservesorig[boots,]<-provor
  payments[boots,]<-payori
}
PEbsorig<-numeric(k-1)
for(orig in 1:(k-1)){
  PEbsorig[orig]<-sqrt(phi.P*sum((orig.fits[orig+1,(k-(orig-1)):k])^ro)+sd(reservesorig[,orig])^2)
}
PEbsorig
PEbsorig/orig.prov*100
# Distribución predictiva años de origen
orig=9
quantile(payments[,orig],c(0.5,0.75,0.90,0.95,0.995))
mean(payments[,orig])
sd(payments[,orig])
# coeficiente de variación en %
cv<-(sd(payments[,orig])/mean(payments[,orig]))*100;cv
pp<-(payments[,orig]-mean(payments[,orig]))/sd(payments[,orig])
#assimetria
sum(pp^3)/(nBoot-1)
#Kurtosis
sum(pp^4)/(nBoot-1) -3
hist(payments[,orig],breaks=21,prob=TRUE)
lines(density(payments[,orig]), lty="dashed")
curve(dnorm(x, mean = mean(payments[,orig]), sd = sd(payments[,orig])),
lty="dotted", add=TRUE)

```

```

##pagos por año de calendario#####

# Cálculo de los pagos por años de calendario

orig.CL<-glm(cij~i+j,fam)

coefs<-exp(as.numeric(coef(orig.CL)))

alpha<-c(1,coefs[2:k])*coefs[1]

beta<-c(1,coefs[(k+1):(2*k-1)])

orig.fits<-alpha%*%t(beta)

future<-row(orig.fits)+col(orig.fits)-1>k

cal.reserve<-matrix(orig.fits,k,k)

cal.d<-matrix(0,k,k)

cal.prov<-numeric(k-1)

for(cal in 2:k){

  cal.d[row(cal.reserve)+col(cal.reserve)==k+cal]<-
cal.reserve[row(cal.reserve)+col(cal.reserve)==k+cal]

  cal.prov[cal-1]<-sum(cal.d)

  cal.d<-matrix(0,k,k)

}

cal.prov

## Error de predicción con fórmula para los pagos por año de calendario

mu.hat.m<-matrix(mu.hat,k,k)

mu.hat.d<-matrix(0,k,k)

mu.hat.vec<-numeric(k*k)

PEforcal<-numeric(k-1)

for(cal in 2:k){

mu.hat.d[row(mu.hat.m)+col(mu.hat.m)==k+cal]<-
mu.hat.m[row(mu.hat.m)+col(mu.hat.m)==k+cal]

mu.hat.vec<-as.vector(mu.hat.d)

```

```

PEforcal[cal-1]<-sqrt(phi.P*sum(mu.hat.vec^ro)+t(mu.hat.vec)%*%Cov.eta%*%mu.hat.vec)
mu.hat.d<-matrix(0,k,k)
}
PEforcal
PEforcal/cal.prov*100
## Error de predicción con bootstrap para los pagos por año de calendario
set.seed(6345789)
nBoot<-1000;payments<-reservescal<-matrix(0,nBoot,k-1)
for(boots in 1:nBoot){
  Ps.cij<-sample(Adj.Prs.resid,n,replace=TRUE)
  Ps.cij<-Ps.cij*sqrt(fitted(orig.CL))+fitted(orig.CL)
  Ps.cij<-pmax(Ps.cij,0)
  Ps.CL<-glm(Ps.cij~i+j,fam)
  coefs<-exp(as.numeric(coef(Ps.CL)))
  Ps.alpha<-c(1,coefs[2:k])*coefs[1]
  Ps.beta<-c(1,coefs[(k+1):(2*k-1)])
  Ps.fits<-Ps.alpha%*%t(Ps.beta)
  matres.m<-matrix(Ps.fits,k,k); matres.d<-matrix(0,k,k);provc<-numeric(k-1);paycal<-
  numeric(k-1)
  for(cal in 2:k){
    matres.d[row(matres.m)+col(matres.m)==k+cal]<-
    matres.m[row(matres.m)+col(matres.m)==k+cal]
    provc[cal-1]<-sum(matres.d)
    paycal[cal-1]<- phi.P*rpois(1,provc[cal-1]/phi.P)
    matres.d<-matrix(0,k,k)
  }
  reservescal[boots,]<-provc
  payments[boots,]<-paycal
}

```

```

PEbscal<-numeric(k-1)

for(cal in 1:k-1){

  cal.d[row(cal.reserve)+col(cal.reserve)==k+cal+1]<-
cal.reserve[row(cal.reserve)+col(cal.reserve)==k+cal+1]

  PEbscal[cal]<-sqrt(phi.P*sum(cal.d^ro)+sd(reservescal[,cal])^2)

  cal.d<-matrix(0,k,k)

}

PEbscal

PEbscal/cal.prov*100

# Distribución predictiva de los pagos por año de calendario

cal=8

quantile(payments[,cal], c(0.5,0.75,0.90,0.95,0.995))

mean(payments[,cal])

sd(payments[,cal])

#coeficient de variació en %

cv<-(sd(payments[,cal])/mean(payments[,cal]))*100 ;cv

pp<-(payments[,cal]-mean(payments[,cal]))/sd(payments[,cal])

#assimetria

sum(pp^3)/(nBoot-1)

#Kurtosis

sum(pp^4)/(nBoot-1) -3

# hist(payments[,cal],breaks=21,prob=TRUE)

hist(payments[,cal],nclass=20,prob=TRUE)

lines(density(payments[,cal]), lty="dashed")

curve(dnorm(x, mean = mean(payments[,cal]), sd = sd(payments[,cal])),

lty="dotted", add=TRUE)

BOOTSTRAP PAIRS:

## Reserva total #####

```

```

# Cálculo de la reserva

n<-length(cij)

k<-trunc(sqrt(2*n))

i<-rep(1:k,k:1);i<-as.factor(i)

j<-sequence(k:1);j<-as.factor(j)

orig.CL<-glm(cij~i+j,fam)

coefs<-exp(as.numeric(coef(orig.CL)))

alpha<-c(1,coefs[2:k])*coefs[1]

beta<-c(1,coefs[(k+1):(2*k-1)])

orig.fits<-alpha%*%t(beta)

future<-row(orig.fits)+col(orig.fits)-1>k

orig.reserve<-sum(orig.fits[future])

orig.reserve

# Cálculo residuos pearson

Prs.resid<-(cij-fitted(orig.CL))/sqrt(fitted(orig.CL)^ro)

n<-(k*(k+1))/2

p<-2*k-1

phi.P<-sum(Prs.resid^2)/(n-p)

Adj.Pr.resid<-Prs.resid*sqrt(n/(n-p))

cij.l<-xtabs(cij~i+j)

cij.v<-as.vector(cij.l)

ii<-row(cij.l);jj<-col(cij.l)

futurebis<-as.numeric(ii+jj-1>k)

ii<-as.factor(ii)

jj<-as.factor(jj)

Cov.eta <- model.matrix(cij.v~ii+jj)%*%vcov(orig.CL)%*%t( model.matrix(cij.v~ii+jj))

mu.hat<-as.vector(orig.fits*future)

## Error de predicción con bootstrap para la reserva total

multi1 <- read.table(file="multi1.txt")

```



```

multi1 <- as.matrix(multi1)

multi2 <- read.table(file="multi2.txt")

multi2 <- as.matrix(multi2)

multi<-matrix(c(multi1,multi2),nrow=55,ncol=1000)

nBoot=1000

payments<-reserves<-numeric(nBoot)

set.seed(6345789)

for(boots in 1:nBoot){

  Ps.w<-multi[,boots]

  index<-Ps.w>0; index

  Ps.wara<-Ps.w[index]; Ps.wara

  cijara<- cij[index];cijara

  iara<-i[index]; iara<-as.factor(iara)

  jara<-j[index];jara<-as.factor(jara)

  Ps.CL<-glm(cijara~iara+jara,fam,weights=Ps.wara)

  jpred<-rep(1:k,k) ; jpred

  ipred<-rep(1:k, each = k)

  jpred<-as.factor(jpred)

  ipred<-as.factor(ipred)

  newdata <- data.frame(i = ipred, j= jpred)

  Ps.fits<-predict(Ps.CL,newdata,type="response")

  Ps.fits<-matrix(Ps.fits,nrow=k,ncol=k)

  Ps.fits<-t(Ps.fits)

  Ps.reserve<-sum(Ps.fits[future])

  Ps.totpayments<-phi.P*rpois(1,Ps.reserve/phi.P)

  reserves[boots]<-Ps.reserve

  payments[boots]<-Ps.totpayments

}

PEbs<-sqrt(phi.P*sum(orig.fits[future]^ro)+sd(reserves)^2)

```

```

PEbs

PEbs/orig.reserve*100

# Distribución predictiva

quantile(payments, c(0.5,0.75,0.90,0.95,0.995))

mean(payments)

sd(payments)

cv<-(sd(payments)/mean(payments))*100; cv  ## c.v. in %

pp<-(payments-mean(payments))/sd(payments)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments,breaks=21,prob=TRUE,main="Histogram of payments")

lines(density(payments), lty="dashed")

curve(dnorm(x, mean = mean(payments), sd = sd(payments)),

lty="dotted", add=TRUE)

## Reservas por año de origen #####

# Cálculo de las reservas

orig.CL<-glm(cij~i+j,fam)

jpred<-rep(1:k,k)

ipred<-rep(1:k, each = k)

jpred<-as.factor(jpred)

ipred<-as.factor(ipred)

newdata <- data.frame(i = ipred, j= jpred)

orig.fits<-predict(orig.CL,newdata,type="response")

orig.fits<-matrix(orig.fits,nrow=k,ncol=k)

orig.fits<-t(orig.fits)

orig.prov<-numeric(k-1)

for (orig in 1:(k-1)){

orig.prov[orig]<-sum(orig.fits[orig+1,(k-(orig-1)):k])

}

```

```

orig.prov

## Error de predicción con bootstrap para las reservas por año de origen

nBoot<-1000;payments<-reservesorig<-matrix(0,nBoot,k-1)

set.seed(6345789)

for(boots in 1:nBoot){

  Ps.w<-multi[,boots]

  index<-Ps.w>0; index

  Ps.wara<-Ps.w[index]

  cijara<- cij[index];cijara

  iara<-i[index]; iara<-as.factor(iara)

  jara<-j[index];jara<-as.factor(jara)

  Ps.CL<-glm(cijara~iara+jara,fam,weights=Ps.wara)

  jpred<-rep(1:k,k)

  ipred<-rep(1:k, each = k)

  jpred<-as.factor(jpred)

  ipred<-as.factor(ipred)

  newdata <- data.frame(i = ipred, j= jpred)

  Ps.fits<-predict(Ps.CL,newdata,type="response")

  Ps.fits<-matrix(Ps.fits,nrow=k,ncol=k)

  Ps.fits<-t(Ps.fits)

  provor<-numeric(k-1)

  payori<-numeric(k-1)

  for(orig in 1:(k-1)){

    provor[orig]<-sum(Ps.fits[orig+1,(k-(orig-1)):k])

    payori[orig]<- phi.P*rpois(1,provor[orig]/phi.P)

  }

  reservesorig[boots,]<-provor

  payments[boots,]<-payori

}

```

```

PEbsorig<-numeric(k-1)

for(orig in 1:(k-1)){

  PEbsorig[orig]<-sqrt(phi.P*sum((orig.fits[orig+1,(k-(orig-
1)):k])^ro)+sd(reservesorig[,orig])^2)

}

PEbsorig

PEbsorig/orig.prov*100

# Distribución predictiva por año de orig en

orig=1

payments.orig.1<-payments[,orig]

quantile(payments.orig.1, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.1)

sd(payments.orig.1)

cv<-(sd(payments.orig.1)/mean(payments.orig.1))*100; cv ## c.v. in %

pp<-(payments.orig.1-mean(payments.orig.1))/sd(payments.orig.1)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.1,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 1)")

lines(density(payments.orig.1), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.1), sd = sd(payments.orig.1)),

lty="dotted", add=TRUE)

orig=2

payments.orig.2<-payments[,orig]

quantile(payments.orig.2, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.2)

sd(payments.orig.2)

cv<-(sd(payments.orig.2)/mean(payments.orig.2))*100; cv ## c.v. in %

pp<-(payments.orig.2-mean(payments.orig.2))/sd(payments.orig.2)

```

```

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.2,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 2)")

lines(density(payments.orig.2), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.2), sd = sd(payments.orig.2)),
lty="dotted", add=TRUE)

orig=3

payments.orig.3<-payments[,orig]

quantile(payments.orig.3, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.3)

sd(payments.orig.3)

cv<-(sd(payments.orig.3)/mean(payments.orig.3))*100; cv ## c.v. in %

pp<-(payments.orig.3-mean(payments.orig.3))/sd(payments.orig.3)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.3,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 3)")

lines(density(payments.orig.3), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.3), sd = sd(payments.orig.3)),
lty="dotted", add=TRUE)

orig=4

payments.orig.4<-payments[,orig]

quantile(payments.orig.4, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.4)

sd(payments.orig.4)

cv<-(sd(payments.orig.4)/mean(payments.orig.4))*100; cv ## c.v. in %

pp<-(payments.orig.4-mean(payments.orig.4))/sd(payments.orig.4)

sum(pp^3)/(nBoot-1) # estimates the skewness

```

```

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.4,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 4)")

lines(density(payments.orig.4), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.4), sd = sd(payments.orig.4)),
lty="dotted", add=TRUE)

orig=5

payments.orig.5<-payments[,orig]

quantile(payments.orig.5, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.5)

sd(payments.orig.5)

cv<-(sd(payments.orig.5)/mean(payments.orig.5))*100; cv ## c.v. in %

pp<-(payments.orig.5-mean(payments.orig.5))/sd(payments.orig.5)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.5,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 5)")

lines(density(payments.orig.5), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.5), sd = sd(payments.orig.5)),
lty="dotted", add=TRUE)

orig=6

payments.orig.6<-payments[,orig]

quantile(payments.orig.6, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.6)

sd(payments.orig.6)

cv<-(sd(payments.orig.6)/mean(payments.orig.6))*100; cv ## c.v. in %

pp<-(payments.orig.6-mean(payments.orig.6))/sd(payments.orig.6)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

```

```

hist(payments.orig.6,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 6)")

lines(density(payments.orig.6), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.6), sd = sd(payments.orig.6)),
lty="dotted", add=TRUE)

orig=7

payments.orig.7<-payments[,orig]

quantile(payments.orig.7, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.7)

sd(payments.orig.7)

cv<-(sd(payments.orig.7)/mean(payments.orig.7))*100; cv  ## c.v. in %

pp<-(payments.orig.7-mean(payments.orig.7))/sd(payments.orig.7)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.7,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 7)")

lines(density(payments.orig.7), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.7), sd = sd(payments.orig.7)),
lty="dotted", add=TRUE)

orig=8

payments.orig.8<-payments[,orig]

quantile(payments.orig.8, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.8)

sd(payments.orig.8)

cv<-(sd(payments.orig.8)/mean(payments.orig.8))*100; cv  ## c.v. in %

pp<-(payments.orig.8-mean(payments.orig.8))/sd(payments.orig.8)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.8,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 8)")

```

```

lines(density(payments.orig.8), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.8), sd = sd(payments.orig.8)),

lty="dotted", add=TRUE)

orig=9

payments.orig.9<-payments[,orig]

quantile(payments.orig.9, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.orig.9)

sd(payments.orig.9)

cv<-(sd(payments.orig.9)/mean(payments.orig.9))*100; cv  ## c.v. in %

pp<-(payments.orig.9-mean(payments.orig.9))/sd(payments.orig.9)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.orig.9,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(accident year 9)")

lines(density(payments.orig.9), lty="dashed")

curve(dnorm(x, mean = mean(payments.orig.9), sd = sd(payments.orig.9)),

lty="dotted", add=TRUE)

# Reservas por año de calendario #####

# Cálculo de las reservas

orig.CL<-glm(cij~i+j,fam)

jpred<-rep(1:k,k)

ipred<-rep(1:k, each = k)

jpred<-as.factor(jpred)

ipred<-as.factor(ipred)

newdata <- data.frame(i = ipred, j= jpred)

orig.fits<-predict(orig.CL,newdata,type="response")

orig.fits<-matrix(orig.fits,nrow=k,ncol=k)

orig.fits<-t(orig.fits)

future<-row(orig.fits)+col(orig.fits)-1>k

```



```

cal.reserve<-matrix(orig.fits,k,k)

cal.d<-matrix(0,k,k)

cal.prov<-numeric(k-1)

for(cal in 2:k){

  cal.d[row(cal.reserve)+col(cal.reserve)==k+cal]<-
cal.reserve[row(cal.reserve)+col(cal.reserve)==k+cal]

  cal.prov[cal-1]<-sum(cal.d)

  cal.d<-matrix(0,k,k)

}

cal.prov

## Error de predicción con bootstrap para las reservas por año de calendario

nBoot<-1000;payments<-reservescal<-matrix(0,nBoot,k-1)

set.seed(6345789)

for(boots in 1:nBoot){

  Ps.w<-multi[,boots]

  index<-Ps.w>0; index

  Ps.wara<-Ps.w[index]

  cijara<- cij[index]

  iara<-i[index]; iara<-as.factor(iara)

  jara<-j[index];jara<-as.factor(jara)

  Ps.CL<-glm(cijara~iara+jara,fam,weights=Ps.wara)

  jpred<-rep(1:k,k)

  ipred<-rep(1:k, each = k)

  jpred<-as.factor(jpred)

  ipred<-as.factor(ipred)

  newdata <- data.frame(i = ipred, j= jpred)

  Ps.fits<-predict(Ps.CL,newdata,type="response")

  Ps.fits<-matrix(Ps.fits,nrow=k,ncol=k)

  Ps.fits<-t(Ps.fits)

```

```

matres.m<-matrix(Ps.fits,k,k); matres.d<-matrix(0,k,k);provc<-numeric(k-1);paycal<-
numeric(k-1)

for(cal in 2:k){

  matres.d[row(matres.m)+col(matres.m)==k+cal]<-
matres.m[row(matres.m)+col(matres.m)==k+cal]

  provc[cal-1]<-sum(matres.d)

  paycal[cal-1]<- phi.P*rpois(1,provc[cal-1]/phi.P)

  matres.d<-matrix(0,k,k)

}

reservescal[boots,]<-provc

payments[boots,]<-paycal

}

PEbscal<-numeric(k-1)

for(cal in 1:k-1){

  cal.d[row(cal.reserve)+col(cal.reserve)==k+cal+1]<-
cal.reserve[row(cal.reserve)+col(cal.reserve)==k+cal+1]

  PEbscal[cal]<-sqrt(phi.P*sum(cal.d^ro)+sd(reservescal[,cal])^2)

  cal.d<-matrix(0,k,k)

}

PEbscal

PEbscal/cal.prov*100

# Distribución predictiva para años de calendario

cal=1

payments.cal.10<-payments[,cal]

quantile(payments.cal.10, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.10)

sd(payments.cal.10)

cv<-(sd(payments.cal.10)/mean(payments.cal.10))*100; cv ## c.v. in %

pp<-(payments.cal.10-mean(payments.cal.10))/sd(payments.cal.10)

sum(pp^3)/(nBoot-1) # estimates the skewness

```

```

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.10,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 10)")

lines(density(payments.cal.10), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.10), sd = sd(payments.cal.10)),
lty="dotted", add=TRUE)

cal=2

payments.cal.11<-payments[,cal]

quantile(payments.cal.11, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.11)

sd(payments.cal.11)

cv<-(sd(payments.cal.11)/mean(payments.cal.11))*100; cv ## c.v. in %

pp<-(payments.cal.11-mean(payments.cal.11))/sd(payments.cal.11)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.11,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 11)")

lines(density(payments.cal.11), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.11), sd = sd(payments.cal.11)),
lty="dotted", add=TRUE)

cal=3

payments.cal.12<-payments[,cal]

quantile(payments.cal.12, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.12)

sd(payments.cal.12)

cv<-(sd(payments.cal.12)/mean(payments.cal.12))*100; cv ## c.v. in %

pp<-(payments.cal.12-mean(payments.cal.12))/sd(payments.cal.12)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

```

```

hist(payments.cal.12,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 12)")

lines(density(payments.cal.12), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.12), sd = sd(payments.cal.12)),
lty="dotted", add=TRUE)

cal=4

payments.cal.13<-payments[,cal]

quantile(payments.cal.13, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.13)

sd(payments.cal.13)

cv<-(sd(payments.cal.13)/mean(payments.cal.13))*100; cv ## c.v. in %

pp<-(payments.cal.13-mean(payments.cal.13))/sd(payments.cal.13)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.13,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 13)")

lines(density(payments.cal.13), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.13), sd = sd(payments.cal.13)),
lty="dotted", add=TRUE)

cal=5

payments.cal.14<-payments[,cal]

quantile(payments.cal.14, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.14)

sd(payments.cal.14)

cv<-(sd(payments.cal.14)/mean(payments.cal.14))*100; cv ## c.v. in %

pp<-(payments.cal.14-mean(payments.cal.14))/sd(payments.cal.14)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.14,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 14)")

```

```

lines(density(payments.cal.14), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.14), sd = sd(payments.cal.14)),
lty="dotted", add=TRUE)

cal=6

payments.cal.15<-payments[,cal]

quantile(payments.cal.15, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.15)

sd(payments.cal.15)

cv<-(sd(payments.cal.15)/mean(payments.cal.15))*100; cv ## c.v. in %

pp<-(payments.cal.15-mean(payments.cal.15))/sd(payments.cal.15)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.15,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 15)")

lines(density(payments.cal.15), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.15), sd = sd(payments.cal.15)),
lty="dotted", add=TRUE)

cal=7

payments.cal.16<-payments[,cal]

quantile(payments.cal.16, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.16)

sd(payments.cal.16)

cv<-(sd(payments.cal.16)/mean(payments.cal.16))*100; cv ## c.v. in %

pp<-(payments.cal.16-mean(payments.cal.16))/sd(payments.cal.16)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.16,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 16)")

lines(density(payments.cal.16), lty="dashed")

```

```

curve(dnorm(x, mean = mean(payments.cal.16), sd = sd(payments.cal.16)),
lty="dotted", add=TRUE)

cal=8

payments.cal.17<-payments[,cal]

quantile(payments.cal.17, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.17)

sd(payments.cal.17)

cv<-(sd(payments.cal.17)/mean(payments.cal.17))*100; cv ## c.v. in %

pp<-(payments.cal.17-mean(payments.cal.17))/sd(payments.cal.17)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.17,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 17)")

lines(density(payments.cal.17), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.17), sd = sd(payments.cal.17)),
lty="dotted", add=TRUE)

cal=9

payments.cal.18<-payments[,cal]

quantile(payments.cal.18, c(0.5,0.75,0.90,0.95,0.995))

mean(payments.cal.18)

sd(payments.cal.18)

cv<-(sd(payments.cal.18)/mean(payments.cal.18))*100; cv ## c.v. in %

pp<-(payments.cal.18-mean(payments.cal.18))/sd(payments.cal.18)

sum(pp^3)/(nBoot-1) # estimates the skewness

sum(pp^4)/(nBoot-1) -3 # estimates the kurtosis

hist(payments.cal.18,breaks=21,prob=TRUE,main="Predictive distribution of the reserve
(calendar year 18)")

lines(density(payments.cal.18), lty="dashed")

curve(dnorm(x, mean = mean(payments.cal.18), sd = sd(payments.cal.18)),

```

```

lty="dotted", add=TRUE)

## cálculo valor actual de los pagos por año de calendario #####

# nivel de confianza

alfa = 0.995

VaR <- numeric(k-1)

meanIBNR <- numeric(k-1)

sdIBNR <- numeric(k-1)

for(cal in 1:(k-1)){

  VaR[cal]<-quantile(payments[,cal],alfa)

  meanIBNR[cal] <- mean(payments[,cal])

  sdIBNR[cal] <- sd(payments[,cal])

}

VaR

meanIBNR

sdIBNR

# tipo de interés

I<-c(0.139,0.273,0.459,0.672,0.894,1.114,1.323,1.516)

i.renta<-(1+I/100)^-(1:(k-1));i.renta

# Valor actual de los pagos por años de calendario

VA <- sum(cal.prov*i.renta); VA

# Valor actual del VaR

VAalfa <- sum(VaR*i.renta); VAalfa

#gráfico boxplot

par(mfrow = c(1, 2))

# Calcular los valores estadísticos por año

percentil_99_5 <- quantile(payments[, cal], 0.995)

percentil_75 <- quantile(payments[, cal], 0.75)

media <- mean(payments[, cal])

desviacion_estandar <- sd(payments[, cal])

```

```
# Crear el boxplot con los estadísticos
```

```
boxplot(payments[, cal],  
        main = paste("Boxplot año de calendario", cal),  
        col = "lightblue",  
        border = "blue",  
        ylim = c(min(payments[, cal]), max(payments[, cal])))
```

```
# Agregar texto para representar la media y la desviación estándar
```

```
text(1, mean(payments[, cal]) + sd(payments[, cal]), paste("SD:", round(desviacion_estandar,  
2)), col = "black", cex = 0.8) # Desviación estándar
```

```
text(1, percentil_75, paste("Percentil 75%:", round(percentil_75, 2)), col = "black", cex = 0.8) #  
Percentil 90%
```

```
text(1, percentil_99_5, paste("Percentil 99.5%:", round(percentil_99_5, 2)), col = "black", cex =  
0.8) # Percentil 99.5%
```