

TRABAJO FIN DE MÁSTER

Título: Provisiones técnicas de prestaciones pendientes: el método Chain-Ladder estocástico desde un punto de vista práctico en Solvencia II.

Autor: Juan Espejo Fernández.

Tutora: Eva Boj del Val.

Curso: 2013 – 2014.

Facultad de Economía y Empresa

Universidad de Barcelona

Trabajo Fin de Máster

Máster en Ciencias Actuariales y Financieras

**PROVISIONES TÉCNICAS DE PRESTACIONES
PENDIENTES: EL MÉTODO CHAIN-LADDER
ESTOCÁSTICO DESDE UN PUNTO DE VISTA
PRÁCTICO EN SOLVENCIA II.**

Autor:

JUAN ESPEJO FERNÁNDEZ

Tutora:

EVA BOJ DEL VAL

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.



Agradecimientos:

Doy mis más sinceras gracias a todas aquellas personas que me han dedicado parte de su tiempo y que han aportado su granito de arena en la realización de este trabajo:

*A Eva Boj del Val y a M. Teresa Costa Cor,
por la orientación y el seguimiento continuado del trabajo, así como sus consejos.*

*A mis padres, Gerardo y Teodora,
y a mi hermano Ángel,
por haber estado ahí en todo momento.*

A todos ellos, MUCHAS GRACIAS.

RESUMEN

Título: Provisiones técnicas de prestaciones pendientes: el método Chain-Ladder estocástico desde un punto de vista práctico en Solvencia II.

Autor: Juan Espejo Fernández

Tutora: Eva Boj del Val

La inminente entrada en vigor del nuevo marco de solvencia para el sector asegurador denominado Solvencia II justifica, sin duda, la elaboración del presente documento. En él se hace una breve introducción sobre los modelos lineales generalizados (MLG) para acabar estudiando un caso particular: el método Chain-Ladder desde el punto de vista estocástico. Además, con el fin de crear una herramienta útil para el sector profesional, se lleva a cabo una aplicación con RExcel (basada en código de R y VBA), analizando casos prácticos, con la que se obtienen resultados de cálculos exigidos por la nueva Directiva de Solvencia II.

Palabras clave: *Chain-Ladder, Modelo Lineal Generalizado, provisiones técnicas, RExcel, Solvencia II.*

ABSTRACT

Title: Claims reserving: the Chain-Ladder stochastic method from a practical point of view in Solvency II.

Author: Juan Espejo Fernández

Tutor: Eva Boj del Val

The imminent entry into force of the new framework of solvency for the insurance industry called Solvency II justifies, undoubtedly, the elaboration of this document. In it is done a brief introduction on the generalized linear models (GLM) to end up by studying a particular case: the method Chain-Ladder from the point of view stochastic. Furthermore in order to create a useful tool for the professional sector, it is performed an RExcel application (based on R and VBA codes) by analyzing practical cases, with which are obtained results of calculations required by the new Solvency II Directive.

Keywords: *Chain-Ladder, Generalized Linear Model, claims reserving, RExcel, Solvency II.*

ÍNDICE GENERAL**CAPÍTULO 1**

INTRODUCCIÓN. OBJETIVOS Y METODOLOGÍA.....	1
1.1. Introducción.....	1
1.2. Objetivos.....	1
1.3. Metodología.....	2

CAPÍTULO 2

CONTEXTUALIZACIÓN DEL CÁLCULO DE LA PROVISIÓN TÉCNICA EN EL MARCO DE SOLVENCIA II	3
2.1. Las provisiones técnicas de prestaciones.....	3
2.1.1. Provisión de prestaciones pendientes de liquidación o pago	4
2.1.2. Provisión de siniestros pendientes de declaración.....	4
2.1.3. Provisión de gastos internos de liquidación de siniestros.....	5
2.2. Antecedentes a Solvencia II.....	5
2.3. La Directiva de Solvencia II.....	6
2.3.1. Provisiones técnicas.....	7
2.3.2. Capital requerido.....	8

CAPÍTULO 3

REVISIÓN DE MODELOS ESTADÍSTICOS. EL MÉTODO CHAIN-LADDER Y SUS GENERALIZACIONES	10
3.1. Ámbito normativo para el uso de métodos estadísticos	10
3.2. Clasificación de los métodos estadísticos.....	11
3.3. Aspectos generales	12
3.4. Métodos deterministas	13
3.4.1. Grossing-up	13
3.4.2. Link ratio.....	14
3.4.3. Chain-Ladder clásico.....	15
3.4.4. Variantes de Chain-Ladder.....	16
3.4.5. Mínimos cuadrados de De Vylder	16
3.4.6. Separación de Taylor.....	17
3.4.6.1. Método de separación aritmética	18

3.4.6.2. Método de separación geométrica.....	18
3.5. Métodos estocásticos	19
3.5.1. Bornhuetter-Ferguson.....	20
CAPÍTULO 4	
EL MODELO LINEAL GENERALIZADO. FORMULACIÓN Y EL CASO PARTICULAR CHAIN-LADDER ESTOCÁSTICO.....	22
4.1. El modelo lineal generalizado	22
4.1.1. Descripción del modelo	22
4.1.2. Estimación de parámetros	24
4.1.3. Desvianza y residuos.....	25
4.1.4. Función de enlace	26
4.1.5. Familia paramétrica de distribuciones	27
4.1.6. Caso particular: Modelo Poisson sobredisperso	27
4.2. Chain-Ladder estocástico: Modelo de Mack	28
CAPÍTULO 5	
ERROR CUADRÁTICO MEDIO DE LAS RESERVAS TOTALES, POR AÑO DE ORIGEN Y POR AÑO DE CALENDARIO	32
5.1. Introducción al error de predicción.....	32
5.2. Error de predicción en el modelo Chain-Ladder con Poisson sobredisperso	32
CAPÍTULO 6	
DISTRIBUCIÓN PREDICTIVA DE LAS RESERVAS. ESTIMACIÓN <i>BOOTSTRAP</i>	34
6.1. Introducción al remuestreo o <i>bootstrap</i>	34
6.2. Error de predicción estimado mediante <i>bootstrap</i>	35
CAPÍTULO 7	
CASOS PRÁCTICOS CON REXCEL	37
7.1. Introducción a RExcel	37
7.2. Desarrollo de una aplicación informática.....	38
7.3. Presentación de resultados.....	41

CAPÍTULO 8

CONCLUSIONES. SINOPSIS Y APORTACIONES 46

GLOSARIO

PALABRAS CLAVES EN EL ÁMBITO DE SOLVENCIA II 48

REFERENCIAS 50

ANEXO 1

CÓDIGO DE R UTILIZADO EN LA FORMULACIÓN DE LA APLICACIÓN 53

ANEXO 2

HISTOGRAMAS. DISTRIBUCIÓN PREDICTIVA BOOTSTRAP 60

ÍNDICE DE FIGURAS**CAPÍTULO 2**

Figura 2.1. Definición de las provisiones de prestaciones según el art. 39 del ROSSP ..	3
Figura 2.2. Los tres pilares en los que se sustenta Solvencia II	6

CAPÍTULO 3

Figura 3.1. Estructura general del triángulo run-off	12
--	----

CAPÍTULO 7

Figura 7.1. Vista inicial al abrir el archivo de la aplicación "ChainLadderEstocástico.xlsm"	39
Figura 7.2. Ventana sobre requisitos previos para ejecutar la aplicación	40
Figura 7.3. Información sobre la autoría de la aplicación	40
Figura 7.4. Elección de la metodología y el cálculo de las reservas	40
Figura 7.5. Selección y estadísticos a visualizar para el caso 4 (GLM – Bootstrap, Reservas por años de calendario) de la aplicación.....	45

ÍNDICE DE TABLAS**CAPÍTULO 4**

Tabla 4.1. Funciones de enlace más frecuentes	24
---	----

CAPÍTULO 7

Tabla 7.1. Triángulo de desarrollo formado por cuantías acumuladas (C_{ij})	41
Tabla 7.2. Triángulo de desarrollo formado por cuantías desacumuladas (c_{ij}).....	41
Tabla 7.3. Rectángulo de desarrollo formado por cuantías acumuladas (C_{ij}) para el modelo de Mack obtenido con la función MackChainLadder del paquete ChainLadder de R	42
Tabla 7.4. Resumen de parámetros obtenidos para el modelo de Mack hallado con la función MackChainLadder del paquete ChainLadder de R	42
Tabla 7.5. Factores de desarrollo (f_k) obtenidos para el modelo de Mack con la función MackChainLadder del paquete ChainLadder de R	42
Tabla 7.6. Resumen de parámetros obtenidos para el Modelo Lineal Generalizado (Chain-Ladder) utilizando la opción mse.method = "formula" mediante la función glmReserve del paquete ChainLadder de R.....	42

Tabla 7.7. Valor de los factores del predictor lineal para el GLM Poisson sobredisperso de Chain-Ladder obtenidos, utilizando la opción <i>mse.method = "formula"</i> mediante la función <i>glmReserve</i> del paquete <i>ChainLadder</i> de R	43
Tabla 7.8. Rectángulo de desarrollo formado por cuantías desacomuladas (c_{ij}) obtenido por la aplicación desarrollada	43
Tabla 7.9. Resultados de la aplicación sobre reservas totales y por año de origen (por fórmula).....	44
Tabla 7.10. Resultados de la aplicación sobre reservas por año de calendario (por fórmula).....	44
Tabla 7.11. Resultados de la aplicación sobre reservas totales y por año de origen (por bootstrap).....	44
Tabla 7.12. Resultados de la aplicación sobre reservas por año de calendario (por bootstrap).....	45

ÍNDICE DE GRÁFICOS

ANEXO 2

Gráfico A.1. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva total	60
Gráfico A.2. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 6	60
Gráfico A.3. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 7	61
Gráfico A.4. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 8	61
Gráfico A.5. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 9	62
Gráfico A.6. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 10	62

CAPÍTULO 1

INTRODUCCIÓN. OBJETIVOS Y METODOLOGÍA

1.1. Introducción

Durante los últimos años, muchos han sido los autores que han tratado los métodos de cálculo de las provisiones técnicas (reservas en la literatura científica actuarial) sobre prestaciones pendientes. La abundante bibliografía generada, junto a la importancia suscitada en la actualidad dentro del marco normativo de la Directiva de Solvencia II, ha provocado que sea uno de los temas de investigación en seguros más importantes.

Si bien es cierto que hasta ahora la estimación de las provisiones técnicas se llevaba a cabo de forma muy mecánica, utilizando uno de los diversos métodos disponibles con el objetivo de obtener el nivel esperado de reservas que fuera capaz de hacer frente a los pagos futuros, asociados a siniestros que se hubieran producido durante el ejercicio de cálculo y en los anteriores, con la implantación inminente de Solvencia II se incorporan nuevos criterios. Por lo tanto, en este nuevo contexto, adquiere importancia tratar de predecir los siniestros que puedan acontecer en el futuro y saber relacionarlo con el nivel de reservas estimado previamente.

Así pues, la necesidad de una armonización europea a nivel internacional a la hora de la valoración de las provisiones técnicas, fundamentales para conocer la posición de solvencia de una compañía aseguradora, ha provocado que se plantee la incorporación de elementos estocásticos que, por un lado, justifiquen la metodología de cálculo del nivel de reservas y, por el otro, permitan valorar cuantitativamente la incertidumbre asociada a estos mecanismos.

En este sentido, el nuevo contexto normativo sobre el cálculo del nivel de capital, basado en los riesgos soportados, invita a la aplicación de métodos estocásticos para su estimación. Actualmente, se dispone de múltiples posibilidades y todas ellas igualmente válidas siempre que se lleven a cabo de forma correcta. No obstante, la metodología más popular en el sector asegurador seguramente sea la de Chain-Ladder; razón por la cual el presente trabajo se centra en ella.

Todo ello motiva la elaboración de este documento, el cual se basa en estudiar la metodología Chain-Ladder, bajo una perspectiva estocástica, desde el punto de vista del nuevo marco normativo europeo y centrado en el cálculo de los márgenes de solvencia, relativo a las provisiones técnicas para el sector de seguros no-vida.

1.2. Objetivos

Dentro de la contextualización del ámbito de Solvencia II y del cálculo de márgenes de solvencia, en el trabajo se pretende revisar con detalle las fórmulas relativas al modelo lineal generalizado con distribución de Poisson sobredispersa y función de enlace logarítmica en referencia al cálculo de errores cuadráticos medios para la reserva total, por año de origen y por año de calendario. Realizar el estudio por año de calendario resulta un aspecto de gran relevancia ya que posibilita el cálculo de prestaciones pendientes en un entorno financiero-actuarial.

Por otro lado, se trata la distribución predictiva de las reservas estimada mediante *bootstrap*, desde una visión meramente práctica.

A fin de ilustrar los cálculos y hacer las correspondientes y oportunas comprobaciones, se utiliza la aplicación RExcel.

1.3. Metodología

Para lograr los objetivos expuestos, se ha seguido la siguiente metodología:

- Búsqueda bibliográfica en el marco teórico y normativo.
- Determinación de criterios matemáticos y formulación a seguir.
- Uso de RExcel para la ilustración práctica de resultados obtenidos.

Con el conjunto de información obtenida de las tareas mencionadas, se han desarrollado los contenidos que conforman la tesina:

1. *Introducción. Objetivos y metodología.* Breve exposición del marco en que se contextualiza el trabajo, motivación, objetivos y detalle de la estructura seguida.
2. *Contextualización del cálculo de la provisión técnica en el marco de Solvencia II.* Situación y adaptación al nuevo marco normativo por lo que al cálculo del nivel de reservas se refiere.
3. *Revisión de modelos estadísticos. El método Chain-Ladder y sus generalizaciones.* Estado del arte de las distintas metodologías existentes para el cálculo de provisiones técnicas, haciendo especial hincapié en el modelo de Chain-Ladder.
4. *El modelo lineal generalizado. Formulación y caso particular Chain-Ladder estocástico.* Introducción a los modelos lineales generalizados y teorización sobre el método Chain-Ladder estocástico.
5. *Error cuadrático medio de las reservas totales, por año de origen y por año de calendario.* Análisis y determinación de errores en el cálculo de los niveles de reserva, según las tres posibilidades citadas en el título del capítulo.
6. *Distribución predictiva de las reservas. Estimación bootstrap.* Determinación y análisis del error de predicción en el cálculo de las provisiones técnicas mediante la técnica de remuestreo *bootstrap*.
7. *Casos prácticos con RExcel.* Ilustración de los cálculos y resultados obtenidos mediante la aplicación informática desarrollada basada en RExcel.
8. *Conclusiones. Sinopsis y aportaciones.* Recapitulación de las distintas valoraciones extraídas de cada apartado, así como las aportaciones realizadas.

Finalmente, se incluye un glosario de palabras clave en el ámbito de Solvencia II, un apartado de referencias y el código de R utilizado (y algunos resultados) en forma de anexos que complementa el contenido del estudio. Además, junto al trabajo se entrega el archivo correspondiente a la aplicación. Toda la información del documento está debidamente referenciada.

CAPÍTULO 2

CONTEXTUALIZACIÓN DEL CÁLCULO DE LA PROVISIÓN TÉCNICA EN EL MARCO DE SOLVENCIA II

2.1. Las provisiones técnicas de prestaciones

El concepto de “provisión técnica” aparece recogido en el artículo 29 del vigente Reglamento de Ordenación y Supervisión de los Seguros Privados (en adelante, ROSSP) de 20 de noviembre de 1998. En él, se cita textualmente que *“las provisiones técnicas deberán reflejar en el balance de las entidades aseguradoras el importe de las obligaciones asumidas que se derivan de los contratos de seguros y reaseguros. Se deberán constituir y mantener por un importe suficiente para garantizar, atendiendo a criterios prudentes y razonables, todas las obligaciones derivadas de los referidos contratos, así como para mantener la necesaria estabilidad de la entidad aseguradora frente a oscilaciones aleatorias o cíclicas de la siniestralidad o frente a posibles riesgos especiales [...] En el caso de que se elaboren balances con periodicidad diferente a la anual, el cálculo y la constitución de las provisiones técnicas se efectuarán aplicando los criterios establecidos en este reglamento con la adaptación temporal necesaria”*.

Nótese, que la definición anterior incluye dos aspectos muy relevantes a la hora de la determinación del nivel de reservas de una compañía aseguradora: el principio de suficiencia y la periodicidad.

Por otro lado, y dentro del ámbito de las provisiones técnicas de los seguros no-vida, el artículo 39 del ROSSP define la provisión de prestaciones como el importe total de las obligaciones pendientes del asegurador derivadas de los siniestros ocurridos con anterioridad a la fecha del cierre del ejercicio, cumpliéndose lo indicado en la Figura 2.1:

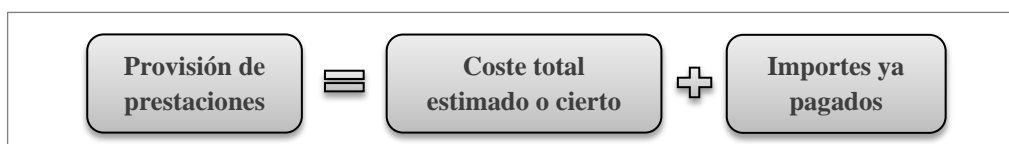


Figura 2.1. Definición de las provisiones de prestaciones según el art. 39 del ROSSP. [Fuente: Elaboración propia]

Además, el cálculo de la provisión de prestaciones debe tener en cuenta los aspectos siguientes:

- Factores y circunstancias que influyan en su coste final.
- Los siniestros se clasificarán por años de ocurrencia.
- El cálculo de la provisión se realizará, al menos, por ramos de seguro.
- Valoración individual de los siniestros o utilización de métodos estadísticos de cálculo global.
- Si la información es insuficiente y no puede estimarse de forma adecuada la provisión, ésta deberá dotarse, como mínimo, por la diferencia entre las primas de riesgo devengadas en el ejercicio, en la parte imputable al mismo, y los pagos por siniestros ocurridos en el ejercicio.

- La provisión de prestaciones estará integrada por tres componentes, cada una de ellas definida en los artículos 40, 41 y 42 del ROSSP, respectivamente:
 1. Provisión de prestaciones pendientes de liquidación o pago.
 2. Provisión de siniestros pendientes de declaración.
 3. Provisión de gastos internos de liquidación de siniestros.

2.1.1. Provisión de prestaciones pendientes de liquidación o pago

Se trata del importe de todos aquellos siniestros ocurridos y declarados antes del cierre del ejercicio. Incluye los gastos de carácter externo inherentes a la liquidación de siniestros y, en su caso, los intereses de demora y las penalizaciones legalmente establecidas en las que haya incurrido la entidad.

2.1.2. Provisión de siniestros pendientes de declaración

Recoge el importe estimado de los siniestros ocurridos antes del cierre del ejercicio y no declarados en esa fecha. Su cálculo debe realizarse mediante métodos estadísticos.

Ahora bien, si la entidad no dispone de métodos estadísticos o éstos no son los adecuados, deberá utilizarse una metodología subsidiaria contemplada en la normativa; así pues, la provisión se deberá determinar multiplicando el número de siniestros pendientes de declaración por el coste medio de los mismos, estimados ambos de la siguiente manera:

- El número de siniestros pendientes de declaración N se calcula como:

$$N_t = \frac{N_{t-1} + N_{t-2} + N_{t-3}}{P_{t-1} + P_{t-2} + P_{t-3}} \cdot P_t \quad (2.1)$$

siendo:

t : ejercicio que se cierra.

N_t : número de siniestros pendientes de declaración a final del ejercicio t .

P_t : primas devengadas en el ejercicio t .

- El coste medio C de los siniestros pendientes de declaración se determina como:

$$C_t = \frac{C_{t-1} + C_{t-2} + C_{t-3}}{Q_{t-1} + Q_{t-2} + Q_{t-3}} \cdot Q_t \quad (2.2)$$

siendo:

C_t : coste medio de los siniestros pendientes de declaración al final del ejercicio t .

Q_t : coste medio de los siniestros ocurridos en el ejercicio t , ya declarados dentro de dicho ejercicio.

Con todo ello, resulta:

$$\{\text{Provisión de siniestros pendientes de declaración}\}_t = N_t \cdot C_t \quad (2.3)$$

No obstante, si no se dispone de los datos de los tres últimos ejercicios, se dota esta provisión aplicando un 5% a la provisión de prestaciones pendientes de liquidación o pago del seguro directo. El porcentaje se eleva al 10% para el coaseguro y el reaseguro aceptado.

2.1.3. Provisión de gastos internos de liquidación de siniestros

Esta reserva está constituida por el importe suficiente para afrontar los gastos internos de la entidad, necesarios para la total finalización de los siniestros que han de incluirse en la provisión de prestaciones tanto del seguro directo como del reaseguro aceptado.

Sólo en el caso de que la entidad no disponga de un método propio para el cálculo de esta provisión, deberá determinarse en función de la relación existente entre los gastos internos imputables a las prestaciones, resultantes de la reclasificación de gastos por destino establecida en el Plan de Contabilidad de las Entidades Aseguradoras y el importe de las prestaciones pagadas en el ejercicio que se cierra corregido por la variación de la provisión para prestaciones pendientes de liquidación o pago y de siniestros pendientes de declaración. El porcentaje resultante deberá multiplicarse, al menos, por el 50% del importe de la provisión para prestaciones pendientes de liquidación o pago más el 100% del importe de la provisión de siniestros pendientes de declaración.

2.2. Antecedentes a Solvencia II

En Europa, las normas en materia de solvencia se compilaron en dos directivas en 1973 y 1979¹, en las que se impuso la creación de un fondo de capital suficiente para absorber los resultados debidos a cambios imprevistos.

La importancia de esta regulación aumentó con la apertura de los mercados en el marco de la tercera generación de directivas comunitarias de seguros, decretadas a mediados de 1994, con las que se abolió el control de precios y productos en toda la Unión Europea (UE). El control de solvencia tenía como objeto permitir a las autoridades de supervisión detectar a tiempo los casos problemáticos entre los aseguradores y poder así proteger mejor a los asegurados. Las normas se concibieron como un estándar mínimo común, y los estados miembros eran libres de imponer regulaciones más estrictas. La regulación de solvencia sólo experimentó cambios mínimos hasta que se adoptaron las Directivas de Solvencia I en febrero de 2002².

En esencia, se trataba de un sistema de cálculo del capital exigido a partir de una serie de ratios que relacionan dichas exigencias con algún aspecto del volumen de negocio, tales como primas o siniestros, eligiéndose el caso más desfavorable.

¹ Para los seguros de no vida, la Primera Directiva del Consejo 73/239/CEE y para los seguros de vida, la Primera Directiva del Consejo 79/267/CEE.

² Para los seguros no vida, Directiva 2002/13/CE del Parlamento Europeo y del Consejo, de 5 de marzo de 2002 y para los seguros de vida la Directiva 2002/12/CE del Parlamento Europeo y del Consejo, de 5 de marzo de 2002.

Aún tratándose de un sistema sencillo de poner en práctica que permite la comparación de resultados entre entidades, comporta algunas limitaciones. No contempla riesgos distintos de los técnicos (como los de inversión), aspectos referidos a la diversificación del riesgo o su transferencia no suponen reducciones en la cantidad exigida de capital y tampoco se valora por completo la calidad del reaseguro ni los porcentajes cedidos.

2.3. La Directiva de Solvencia II

El nuevo marco normativo para el sector asegurador, denominado Solvencia II, apareció publicado el 17 de diciembre de 2009 en el Diario Oficial de la Unión Europea como Directiva 2009/138/CE del Parlamento Europeo y del Consejo, de 25 de noviembre de 2009, sobre el seguro de vida, el acceso a la actividad del seguro y de reaseguro y su ejercicio.

Con un total de 312 artículos y siete anexos, la Directiva representa un ambicioso proyecto de revisión legislativa que culminará, según lo previsto, el 1 de enero de 2016 con su entrada en vigor definitiva. Por su parte, en España desde el pasado 1 de enero de 2014 son efectivas unas medidas temporales de adaptación a Solvencia II en forma de orden ministerial, que tendrán validez hasta el 31 de diciembre de 2015. En estas pautas se indica la manera de proceder en la fase preparatoria previa a la aplicación de la Directiva de Solvencia II.

Enmarcado dentro de una estrategia de mejora de la regulación, el objetivo perseguido por este proyecto legislativo es el de incrementar la protección de los tomadores y beneficiarios de seguros, modernizar el modelo de supervisión y contribuir a intensificar la integración del mercado asegurador, potenciando la competitividad de las entidades que lo configuran.

Como bien es sabido en el sector de los seguros, este nuevo sistema de solvencia de la UE para las compañías de seguros se basa en el enfoque de tres pilares (Figura 2.2).

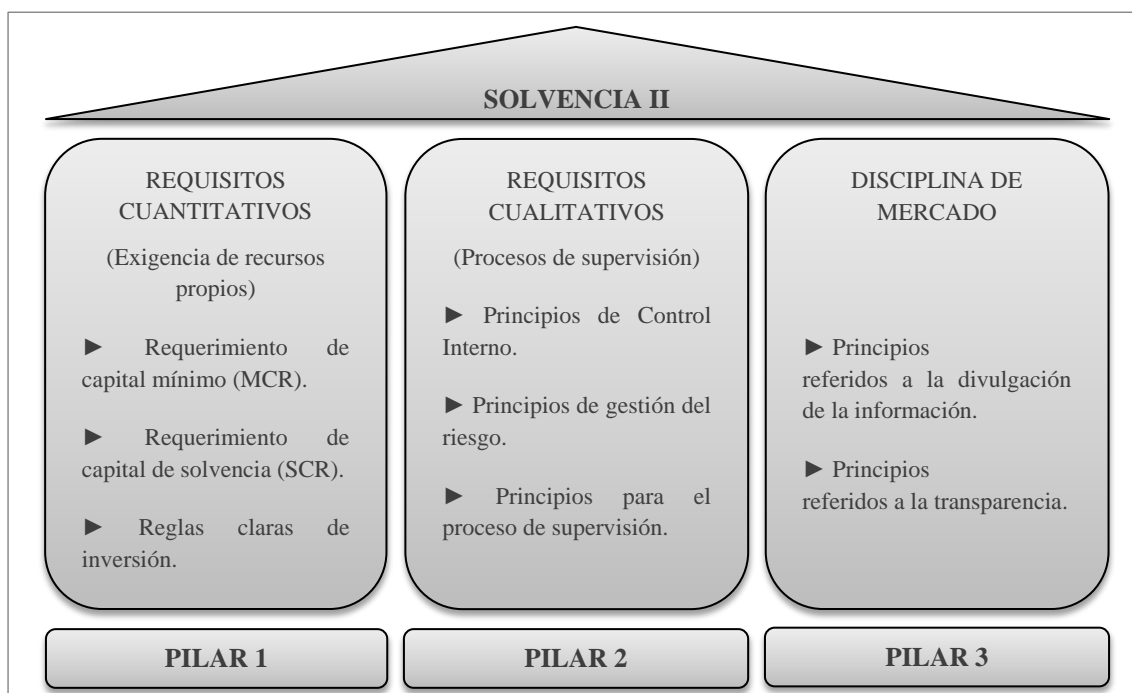


Figura 2.2. Los tres pilares en los que se sustenta Solvencia II. [Fuente: Elaboración propia]

Dado que el trabajo que se presenta trata sobre el cálculo de las provisiones técnicas, se centrará la atención en el primer pilar ya que es donde se establecen los principios cuantitativos del sistema, abordando aspectos ligados con los niveles de reserva y de capital.

2.3.1. Provisiones técnicas

La parte de la Directiva dedicada al cálculo de las provisiones técnicas se halla en la sección 2 del capítulo VI, concretamente en los artículos 77 a 86, ambos incluidos. No obstante, el artículo 76 recoge las disposiciones generales; en particular, se establece que el valor de las provisiones será el valor actual que las empresas de seguros o reaseguros tendrían que pagar si transfirieran sus obligaciones de pago a otra empresa de seguros o reaseguros (artículo 76.2). Su cálculo se realizará utilizando la información facilitada por los mercados financieros y con los datos disponibles sobre riesgos de suscripción (artículo 76.3).

Luego, de manera más precisa, el artículo 77 establece lo siguiente:

1. *El valor de las provisiones técnicas será igual a la suma de la mejor estimación y de un margen de riesgo, con arreglo a lo previsto en los apartados 2 y 3.*
2. *La mejor estimación se corresponderá con la media de los flujos de caja futuros ponderada por su probabilidad, teniendo en cuenta el valor temporal del dinero (valor actual esperado de los flujos de caja futuros) mediante la aplicación de la pertinente estructura temporal de tipos de interés sin riesgo.*
3. *El margen de riesgo será tal que se garantice que el valor de las provisiones técnicas sea equivalente al importe que las empresas de seguros y de reaseguros previsiblemente necesitarían para poder asumir y cumplir las obligaciones de seguro y reaseguro.*
4. *Las provisiones técnicas se calcularán como la suma de la mejor estimación de los pagos futuros más un margen de riesgo, calculándose cada uno de ellos por separado.*
5. *El margen de riesgo será igual al coste de financiación de un volumen de fondos propios admisibles igual al capital de solvencia obligatorio necesario para asumir las obligaciones de seguro y reaseguro durante su periodo de vigencia. La tasa utilizada para determinar el coste financiero del citado importe de fondos propios admisibles será la misma para todas las entidades y será igual al tipo adicional, por encima del tipo libre de riesgo que tendría que satisfacer una aseguradora o reaseguradora por mantener un importe de fondos propios admisibles igual al capital de solvencia obligatorio necesario para asumir las obligaciones de seguro y reaseguro durante el periodo de vigencia de las mismas.*

Además, para la realización de los cálculos deberán considerarse todos los gastos en que se incurrirán para hacer frente a las obligaciones, la inflación y los pagos que estén previstos realizar a los tomadores y beneficiarios de seguros, incluyendo las futuras participaciones discrecionales en beneficios (artículo 78), así como el valor de las garantías financieras y de las posibles opciones contractuales incluidas en los contratos

de seguros y reaseguro (artículo 79). Todo el cálculo tendrá que segmentarse en grupos de riesgo homogéneos, siendo el ramo la unidad máxima admisible (artículo 80).

Por lo que respecta a los artículos 81 a 85, éstos se refieren a cuestiones no directamente vinculadas a la forma de cálculo de las provisiones técnicas, sino que hacen especial mención a los resultados que puedan obtenerse.

Por último, el artículo 86 expone las medidas de ejecución que adoptará la Comisión en temas tales como:

- a) Metodologías actuariales y estadísticas para calcular la mejor estimación.
- b) Estructura temporal de tipos de interés sin riesgo aplicable al cálculo anterior.
- c) Circunstancias en las que las provisiones técnicas deberán calcularse como un todo o como la suma de la mejor estimación y el margen de riesgo.
- d) Métodos e hipótesis que deberán utilizarse en el cálculo del margen de riesgo.
- e) Ramos o agrupación de ellos en que puede dividirse la cartera para calcular el nivel de provisiones técnicas.
- f) Normas y requisitos que deben cumplir los datos que se utilicen en los cálculos.
- g) Casos en los que podría recurrirse al uso de aproximaciones.
- h) Metodologías que deberán utilizarse al calcular el ajuste por incumplimiento de la contraparte a que se refiere el artículo 81.
- i) Métodos y técnicas simplificados para calcular las provisiones técnicas, a fin de garantizar que las metodologías actuariales y estadísticas sean proporcionados a la naturaleza, el volumen y la complejidad de los riesgos soportados por las empresas de seguros y de reaseguros.

2.3.2. Capital requerido

Aunque el estudio presentado se centra en el cálculo de provisiones técnicas, dada la relación existente entre el nivel de reservas para hacer frente a los compromisos adquiridos y la exigencia de recursos propios en forma de capital, se procede a introducir los aspectos generales referidos a los requerimientos de capital.

Así pues, para afrontar los riesgos asumidos por una compañía, ésta debe saber estimar dos cantidades: el capital de solvencia obligatorio (en inglés, *Solvency Capital Requirement* o *SCR*) y el capital mínimo obligatorio (en inglés, *Minimum Capital Requirement* o *MCR*).

El *SCR* se corresponde con el capital económico necesario que ha de disponer la aseguradora para cumplir con todas las obligaciones frente a los tomadores y beneficiarios de seguros en los doce meses siguientes, con un nivel de confianza del 99,5%. Dicha cuantía debe calcularse sobre la base del verdadero perfil de riesgo de la empresa, teniendo en cuenta la incidencia de las posibles técnicas de reducción del riesgo, así como los efectos de diversificación.

Además, la determinación del *SCR* puede hacerse mediante el uso del modelo estándar, común para todas las compañías, o bien, mediante modelos internos siempre que se autorice por el correspondiente organismo supervisor.

Por tanto, para evaluar dicha cifra de capital, se precisan varios elementos:

1. Una medida del riesgo.
2. Una cierta probabilidad o nivel de confianza.
3. Un horizonte temporal (se utiliza el año).

Conviene destacar el uso del *Value at Risk (VaR)* al 99,5% (pérdida esperada más probable, la cual se espera sobrepasar uno de cada 200 años) como medida de riesgo en el modelo estándar; no obstante, debe citarse también la posibilidad de utilizar en un modelo interno el *Tail Value at Risk (TVaR)* como medida de riesgo, por presentar las siguientes ventajas frente al *VaR*:

- No sólo considera la probabilidad de insolvencia, sino también la pérdida esperada por la misma.
- *TVaR* es subaditivo, es decir, las exigencias de capital por dos o más riesgos combinados es menor que las exigencias asociadas a cada uno por separado. Sin embargo, esta propiedad no se da, por lo general, en el *VaR*, sobre todo en casos de colas muy pesadas o en distribuciones muy asimétricas.

En cuanto al *MCR*, denominado fondo de garantía en Solvencia I, representa un nivel de capital debajo del cual los intereses de los tomadores y beneficiarios se verían seriamente dañados si le fuese permitido a la entidad aseguradora continuar operando. En el supuesto de que el nivel de *MCR* fuese vulnerado, se han previsto acciones para que el supervisor actúe retirando la autorización para operar a la entidad.

Entre otros aspectos, la Directiva de Solvencia II establece que el cálculo del *MCR* se efectuará como una función lineal de un conjunto o subconjuntos de las siguientes variables: las provisiones técnicas, las primas suscritas, los capitales en riesgo, los impuestos diferidos y los gastos de administración de la empresa. Además, estipula que el *MCR* no será inferior al 25% ni excederá el 45% del *SCR* de la empresa.

CAPÍTULO 3

REVISIÓN DE MODELOS ESTADÍSTICOS. EL MÉTODO CHAIN-LADDER Y SUS GENERALIZACIONES

3.1. Ámbito normativo para el uso de métodos estadísticos

El artículo 43 del ROSSP deja bien claros aquellos aspectos legales referentes a la utilización de técnicas estadísticas, los cuales se detallan a continuación.

Así, las entidades aseguradoras pueden emplear métodos estadísticos para el cálculo de la provisión por prestaciones que incluyan tanto los siniestros pendientes de liquidación o pago como los siniestros pendientes de declaración, en cuyo caso no es necesario efectuar el desglose de la provisión entre ambos componentes.

Por su parte, las entidades aseguradoras deben recibir autorización de la Dirección General de Seguros y Fondos de Pensiones (en adelante, DGSFP) sobre:

- Los métodos estadísticos a aplicar, así como las hipótesis contempladas para los mismos.
- Las modificaciones de los métodos o hipótesis utilizados.
- Justificación detallada de los contrastes de bondad.
- Periodo de obtención de información.

La estimación del importe final de la provisión debe hacerse tomando en consideración los resultados de, al menos, dos métodos pertenecientes a grupos de métodos estadísticos diferentes (considerándose pertenecientes al mismo grupo aquellos métodos que se basen en las mismas hipótesis o que obtengan sus resultados a partir de las mismas magnitudes o variables).

Además, la determinación de la provisión de prestaciones mediante técnicas estadísticas requiere que:

- La entidad tenga un volumen de siniestros suficiente para permitir la inferencia estadística y que disponga de información histórica relativa a los mismos.
- Datos homogéneos procedentes de estadísticas fiables.
- Posibilidad de excluirse siniestros que vayan a provisionarse de forma individual.
- La entidad debe realizar, al menos anualmente, un contraste de bondad de los cálculos realizados.

Por otro lado, El Ministro de Economía y Hacienda podrá establecer que, en determinados ramos o riesgos, la provisión de prestaciones se calcule por métodos estadísticos en su conjunto o en alguna de sus partes. En tal caso, la DGSFP dará publicidad a los métodos estadísticos que serán obligatorios en ausencia de otros propuestos por la entidad.

También, la entidad podrá solicitar de la DGSFP la no aplicación de métodos estadísticos cuando pueda acreditar que el método utilizado de estimación siniestro a siniestro ha conducido a resultados suficientes durante los últimos siete ejercicios.

Por último, la DGSFP podrá obligar, mediante resolución motivada, a que el importe de la provisión se determine por otros métodos estadísticos si considera que el importe estimado por la entidad, utilizando un método de valoración individual o un método estadístico, resulta insuficiente y puede comprometer su solvencia.

3.2. Clasificación de los métodos estadísticos

Los métodos estadísticos utilizados para el cálculo de las provisiones técnicas pueden clasificarse, según De Vylder (1996) y Taylor (1986), de la siguiente manera:

- Primera clasificación:
 - Micro-modelos: los importes de los siniestros se tratan por separado.
 - Macro-modelos: los importes de los siniestros se tratan de forma agregada.
- Segunda clasificación:
 - Modelos deterministas: no intervienen variables aleatorias.
 - Modelos estocásticos: los siniestros son considerados relaciones de variables aleatorias. Dentro de este grupo, pueden distinguirse:
 - Modelos paramétricos: las distribuciones implicadas se suponen conocidas (Normal, Poisson, etc.), pero deben estimarse algunos parámetros.
 - Modelos de distribución libre: no se realizan hipótesis particulares sobre las distribuciones implicadas.

Además, existen otros criterios de clasificación que consideran:

- Si el modelo tiene en cuenta (o no) el número de siniestros como variable dependiente.
- Las variables explicativas incluidas en el modelo: año de desarrollo, volumen de exposición, inflación, velocidad de finalización de los siniestros, etc.

Cabe mencionar que en la práctica actuarial, los métodos deterministas han sido los utilizados tradicionalmente para el cálculo de las provisiones técnicas. No obstante, los recientes cambios normativos del sector asegurador producidos por la inminente entrada en vigor de la Directiva de Solvencia II, en los que se eleva el nivel de exigencia en las reservas a las empresas aseguradoras, ha puesto de manifiesto que los métodos estocásticos pueden resultar, sin duda, un instrumento adecuado para alcanzar los niveles de precisión exigibles en el ámbito asegurador europeo.

3.3. Aspectos generales

Si bien es cierto que cada metodología de cálculo de las provisiones técnicas para siniestros pendientes utiliza unos datos de partida distintos, pueden destacarse algunos datos básicos necesarios en todas ellas. Por lo general, éstos son:

- c_{ij} : cuantía pagada en el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i .
- C_{ij} : cuantía acumulada pagada hasta (e incluido) el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i .³
- n_i : número de siniestros ocurridos en el año de origen i (estimación al final del año i).
- s_{ij} : cuantía media pagada en el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i .
- S_{ij} : cuantía media acumulada pagada hasta (e incluido) el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i .

Además, para el análisis de los siniestros ocurridos y no reportados (en inglés, *incurred but not reported* o *IBNR*), la información disponible se presenta en el denominado triángulo de siniestros (en inglés, *run-off triangle*) que puede expresarse tal y como se muestra en la Figura 3.1.

		Año de desarrollo					
		j = 0, 1, 2, ..., k					
		0	1	2	...	k-1	k
Año de origen u ocurrencia	0	$c_{0,0}$	$c_{0,1}$	$c_{0,2}$...	$c_{0,k-1}$	$c_{0,k}$
	1	$c_{1,0}$	$c_{1,1}$	$c_{1,2}$...	$c_{1,k-1}$	
	2	$c_{2,0}$	$c_{2,1}$	$c_{2,2}$...		
		
	k-1	$c_{k-1,0}$	$c_{k-1,1}$				
	k	$c_{k,0}$					

Figura 3.1. Estructura general del triángulo run-off. [Fuente: Elaboración propia]

En donde las filas hacen referencia al año de ocurrencia, origen o notificación del siniestro y las columnas, al año de pago o desarrollo. A efectos del presente estudio, el número máximo de años de los cuales se tiene información es k y, dado que tanto filas como columnas se inician en cero⁴ debe verificarse para todos los datos conocidos, $i + j \leq k$. Además, se emplea la notación I y J para referirse al número máximo de años de origen y desarrollo, respectivamente.

Finalmente, comentar que la literatura que hay al respecto suele tomar periodos anuales, pero no existe ninguna diferencia esencial si éstos son meses, trimestres, semestres, etc.

³ Nótese que: $C_{in} = \sum_{j=0}^n c_{ij}$

⁴ Existen autores que sitúan el origen temporal en 1. En tal caso, debe cumplirse entonces que $i + j \leq k + 1$.

3.4. Métodos deterministas

Como ya se ha comentado con anterioridad, los métodos deterministas no utilizan explícitamente ningún supuesto probabilístico para el cálculo de la reserva; sin embargo, por ser de los primeros planteados y de sencilla aplicación, han resultado hasta ahora ser los más utilizados por las compañías de seguros y reaseguros.

Estos métodos se caracterizan por ser modelos en los que, mediante un cálculo más o menos complejo basado en la información histórica recogida en forma de triángulos, tratan de estimar el valor que deberían tener las provisiones técnicas para hacer frente al valor de los siniestros futuros.

De las metodologías pertenecientes a este grupo, en este apartado se introducen de forma muy general las siguientes:

- Grossing-up.
- Link ratio.
- Chain-Ladder clásico.
- Variantes de Chain-Ladder.
- Mínimos cuadrados de De Vylder.
- Separación de Taylor (aritmética y geométrica).

3.4.1. Grossing-up

También conocido como método de crecimiento, se basa en el cálculo de la proporción acumulada de siniestros para cada periodo de desarrollo con respecto al total reportado, para cada año de origen. Así, un dato fundamental para aplicar esta metodología es la estimación de la siniestralidad total del primer año considerado (el más antiguo); es decir, la cuantía total que la compañía deberá pagar, correspondiente al primer año del que se dispongan datos. Dicho importe se divide en dos partes:

- Cuantías conocidas por el paso del tiempo: $C_{0,j}$ o bien $c_{0,j}$, $\forall j = 0, 1, \dots, J$.
- Cuantía estimada del total de gastos correspondientes a los siniestros de ese año, que además cumplirá: $C_{0,\infty} \leq C_{0,j}$.

Por tanto, para este primer año tomado como referencia se trata de obtener los cocientes p_j , definidos según la expresión matemática:

$$p_j = \frac{C_{0,j}}{C_{0,\infty}} \quad (3.1)$$

A continuación, se incluye el supuesto implícito de que estas proporciones se mantienen constantes, sea cual sea el año de desarrollo, aplicándose así al resto de los años de ocurrencia a fin de obtener las correspondientes $\hat{c}_{0,j}$. Estas cuantías se obtienen como se indica en (3.2).

$$\hat{C}_{i,\infty} = \frac{\text{último } C_{ij} \text{ disponible}}{p_i \text{ correspondiente}} \quad (3.2)$$

Con todo, las reservas estimadas se obtendrán como la suma de las diferencias entre ese nivel $\hat{C}_{0,j}$ y la última cifra conocida de provisión, $C_{i,J-i}$, tal y como sigue:

$$PT = \sum_{i=0}^k (\hat{C}_{i,\infty} - C_{i,k-i}) \quad (3.3)$$

3.4.2. Link ratio

Este método parte del triángulo de siniestros acumulados y obtiene las tasas de variación de la siniestralidad en un ejercicio de ocurrencia (porcentajes de crecimiento, denominados también *link ratios*) entre un periodo de desarrollo y el siguiente. El cálculo de dichas ratios se calcula según la expresión matemática:

$$f_{ij} = \frac{C_{i,j+1}}{C_{ij}} \quad i = 0, 1, \dots, k-1 \quad j = 0, 1, \dots, k-1 \quad (3.4)$$

Tras obtener cada una de las tasas, deben de determinarse los factores de proyección que constituyen, a su vez, el elemento básico de cálculo para esta metodología. Dichos factores no son más que los productos de las ratios previamente calculadas, dando lugar a una estimación de cuál será la siniestralidad de un ejercicio suponiendo que en el paso de un año a otro, dicha variable crecerá al ritmo fijado por la *link ratio* o el producto de éstas, si se tratase de un periodo superior a un ejercicio. La manera analítica de expresarlo sería como figura a continuación:

$$F_k = \prod_{j=k}^h \hat{f}_j \quad h \text{ variando de } 0 \text{ a } k-1 \quad (3.5)$$

siendo k el ejercicio considerado y \hat{f}_j la *link ratio* aplicada en el cálculo.

Finalmente, debe comentarse que en función del criterio elegido para la determinación de la *link ratio* (forma de enlazar unas columnas del triángulo de desarrollo con otras) hará que se emplee una u otra de las distintas variantes existentes a esta metodología. Entre ellas, algunas de las cuales se explican posteriormente con mayor detalle, las más conocidas resultan ser:

- El método del máximo o prudente: se elige la mayor f_{ij} de cada columna.
- El método del mínimo: se escoge la mínima f_{ij} de cada columna.
- El método media aritmética: se realiza una media de las f_{ij} de cada columna.
- El método Chain-Ladder: si para la obtención de las ratios, en lugar de coger cada C_{ij} y dividirla por la anterior, se coge la suma total de la columna j y se divide por la suma de la columna $j-1$ sin tener en cuenta el último valor de ésta, para que en ambas columnas haya el mismo número de términos.

3.4.3. Chain-Ladder clásico

En general el método Chain-Ladder utiliza un factor para “suavizar” los datos y, en base a éstos, realizar interpolaciones con el objetivo de estimar los siniestros agregados para cada año de ocurrencia y, posteriormente, la reserva correspondiente.

El supuesto básico de esta metodología es que las columnas del triángulo de desarrollo son proporcionales. La sustentación del supuesto depende en buena medida, tanto del tipo de negocio que se trate, como de la homogeneidad y tamaño de la cartera. Además, se supone que el esquema de evolución de los pagos es estable en el tiempo, a lo largo de los años de desarrollo.

Para estimar la proporción de cambio de un ejercicio a otro (f_{ij}) se calcula una media ponderada de los factores de conexión (*link ratio*), en donde cada valor se pondera con la siniestralidad que le precede. Nótese que f_{ij} viene definido según la expresión matemática (3.4) vista con anterioridad.

Luego, para el ejercicio j de desarrollo, la tasa de modificación de la liquidación de siniestros (f_j) viene dada por la siguiente formulación:

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{ij} \cdot f_{ij}}{\sum_{i=0}^{I-j-1} C_{ij}} = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{ij}} \quad j = 0, 1, \dots, k-1 \quad (3.6)$$

A partir de las proporciones obtenidas, se calculan los factores de proyección (F_k) tal y como se ha indicado en la expresión (3.5).

Con estos factores se estimarán los valores que faltan en el triángulo de desarrollo hasta completarlo. Dichos valores se hallan según la siguiente expresión:

$$\hat{C}_{ij} = C_{i,k-i} \cdot \prod_{h=k-i}^{j-1} \hat{f}_h = C_{i,k-i} \cdot F_{k-i} \quad (3.7)$$

La provisión técnica total vendrá dada por la diferencia entre la suma de estas estimaciones y los importes ya dotados.

Por último, deben mencionarse algunos aspectos importantes a considerar en esta metodología, tal y como destacan las autoras M. M. Claramunt y T. Costa (2003):

- Los datos del triángulo de desarrollo deben ser positivos.
- Puede aplicarse bajo el supuesto que en el año de origen y los k siguientes no se paga el 100% de los siniestros de un año de origen.
- Válido en situaciones con tasas de inflación constantes.
- Aplicable sobre las cuantías medias acumuladas pagadas hasta (e incluido) el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i (S_{ij}).
- Aunque no se trata de un método estocástico, surge como un caso particular de diferentes modelos estocásticos.

3.4.4. Variantes de Chain-Ladder

Varios han sido los intentos de depurar la metodología clásica de Chain-Ladder, dando lugar a variantes del modelo. No obstante, la técnica permanece esencialmente igual, manteniendo el supuesto principal idéntico. La diferencia radica en la ponderación de los factores de desarrollo.

En este sentido, Van Eeghen (1981) describe diferentes variantes del método Chain-Ladder. En todas ellas, parte de un triángulo *run-off* formado por los factores de desarrollo d_{ij} , obtenidos de la siguiente manera:

$$d_{ij} = \frac{C_{i,j+1}}{C_{ij}} \quad i = 0, 1, \dots, k-1 \quad j = 0, 1, \dots, k-1 \quad (3.8)$$

Además, cada variante utiliza una hipótesis alternativa. De entre todas ellas, se destacan dos por su simplicidad matemática y por ser las más habituales en la literatura relacionada con ello:

- I. Ajuste de tendencias lineales (por mínimos cuadrados) a cada columna de d_{ij} .
- II. Supuesto de que el factor de desarrollo es constante en cada columna, estimándose mediante una media aritmética ponderada de los factores de desarrollo empíricos; esto es:

$$\hat{d}_j = \frac{\sum_{i=0}^{k-j-1} w_{ij} \cdot d_{ij}}{\sum_{i=0}^{k-j-1} w_{ij}} \quad (3.9)$$

con pesos w_{ij} :

- $w_{ij} = 1$.
- $w_{ij} = C_{ij}$ (Chain-Ladder clásico).
- $w_{ij} = i+j+1$.
- $w_{ij} = (i+j+1)^2$.

Por último, citar que otras variantes se basan en ajustar una curva exponencial a cada columna del triángulo de desarrollo, teniendo en cuenta de algún modo la inflación de cada periodo. No obstante, resultan algo complejas y poco prácticas.

3.4.5. Mínimos cuadrados de De Vylder

Esta metodología supone que la fracción de siniestros reportados hasta el periodo j resulta independiente del año de ocurrencia i . De este modo, puede usarse un modelo multiplicativo para la representación de las reclamaciones.

Además, por lo expuesto anteriormente, los datos de partida del triángulo de desarrollo deben ser cuantías no acumuladas (c_{ij}).

Así pues, la hipótesis principal del modelo es:

$$c_{ij} = x_i \cdot p_j \quad \text{cumpliéndose} \quad \sum_{j=0}^k p_j = 1 \quad (3.10)$$

siendo:

x_i : cuantía total a pagar de los siniestros ocurridos en el año de origen i .

p_j : proporción de x_i que se paga en el año de desarrollo j .

Luego, los estimadores anteriores se obtienen por mínimos cuadrados; es decir, se trata de determinar los valores de x_i y p_j tales que $\sum (c_{ij} - x_i \cdot p_j)^2$ sea mínima, donde la suma se realiza sobre el conjunto de subíndices del que se tenga información.

El sistema de ecuaciones resultante es:

$$x_i = \frac{\sum_{\forall j \in I_i} c_{ij} \cdot p_j}{\sum_{\forall j \in I_i} p_j^2} \quad \text{y} \quad p_j = \frac{\sum_{\forall i \in I_j} c_{ij} \cdot x_i}{\sum_{\forall i \in I_j} x_i^2} \quad (3.11)$$

Finalmente, deben hacerse algunos comentarios respecto a esta metodología:

- Las columnas del triángulo *run-off* con las cuantías, ya sean acumuladas o no, resultan proporcionales.
- El método es válido también en el supuesto de una tasa de inflación constante.

3.4.6. Separación de Taylor

El modelo propuesto por Taylor (1986) se basa en un triángulo de desarrollo compuesto por las cuantías medias pagadas en el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i . Esto es:

$$s_{ij} = \frac{c_{ij}}{n_i} \quad i = 0, 1, \dots, k \quad j = 0, 1, \dots, k \quad (3.12)$$

Luego, la hipótesis considerada es que cada elemento del triángulo considerado está formado por dos parámetros, r y λ , que juntos definen el desarrollo de los siniestros. Por tanto:

$$s_{ij} = r_j \cdot \lambda_{i+j} \quad (3.13)$$

siendo:

r_j : proporción de la cuantía de los siniestros que se paga en el año de desarrollo j .

λ_{i+j} : índice de las influencias externas exógenas (e.g. inflación), correspondiendo el mismo valor a aquellas cuantías que se sitúan temporalmente en el mismo momento.

En función de qué supuestos adicionales o restricciones se empleen para la determinación de los parámetros anteriormente definidos, se dan dos posibilidades:

- Método de separación aritmética.
- Método de separación geométrica.

Antes de proceder a la explicación de cada una de las variantes del método, cabe destacar dos aspectos de gran relevancia:

- Para completar el rectángulo de desarrollo (rellenando el triángulo opuesto al de *run-off*) hay que estimar, fuera del modelo, los valores para $h > k$ de $\hat{\lambda}_h$.
- Se estiman por separado los efectos del año de pago (e.g. inflación).

3.4.6.1. Método de separación aritmética

Esta metodología utiliza la siguiente condición adicional para estimar los parámetros del modelo:

$$\sum_{j=0}^k r_j = 1 \quad (3.14)$$

Con ello, las expresiones analíticas que sirven para determinar los estimadores de los parámetros para $h = 0, 1, \dots, k$ son, por un lado:

$$\hat{r}_h = \frac{v_h}{\sum_{j=h}^k \lambda_j} \quad (3.15)$$

siendo:

$$v_j \text{ la suma de } s_{ij} \text{ por la columna } j: v_j = \sum_{i=0}^{k-j} s_{ij}$$

Y por el otro,

$$\hat{\lambda}_h = \frac{d_h}{1 - \sum_{j=h+1}^k \hat{r}_j} \quad (3.16)$$

siendo:

$$d_j \text{ la suma de } s_{ij} \text{ por la diagonal } j: d_j = \sum_{i=0}^j s_{i,j-i}$$

3.4.6.2. Método de separación geométrica

En este caso, la restricción usada para hallar los parámetros del modelo es la siguiente:

$$\prod_{j=0}^k r_j = 1 \quad (3.17)$$

Así, siguiendo el mismo esquema anterior, la formulación matemática que sirve para determinar los estimadores de los parámetros para $h = 0, 1, \dots, k$ resulta ser, por un lado:

$$\hat{r}_h = \left(\frac{w_h}{\prod_{j=h}^k \hat{\lambda}_j} \right)^{\frac{1}{k-h+1}} \quad (3.18)$$

siendo:

$$w_j \text{ los productos de } s_{ij} \text{ por la columna } j: w_j = \prod_{i=0}^{k-j} s_{ij}$$

Y por otro lado se tiene,

$$\hat{\lambda}_h = \left(e_h \cdot \prod_{j=h+1}^k \hat{r}_j \right)^{\frac{1}{h+1}} \quad (3.19)$$

siendo:

$$e_j \text{ los productos de } s_{ij} \text{ por la diagonal } j: e_j = \prod_{i=0}^j s_{i,j-i}$$

3.5. Métodos estocásticos

El principal inconveniente que presentan los métodos anteriores es la de no ofrecer límites de confianza para la estimación de la reserva correspondiente. Por ello, se introducen técnicas estadísticas en el desarrollo de las metodologías con el propósito de obtener este tipo de estimación por intervalos y, eventualmente, la distribución completa del estimador.

En este sentido, se presenta una breve recopilación de estos métodos, tratándolos de la manera más general posible:

- Modelo Lineal Generalizado (en adelante, *GLM*).
- Modelo de Mack (generalización de Chain-Ladder).
- Bornhuetter-Ferguson (estocástico en sentido bayesiano).

A continuación se explicará únicamente el modelo de Bornhuetter-Ferguson ya que en el siguiente capítulo se desarrollarán con detalle los otros dos mencionados.

3.5.1. Bornhuetter-Ferguson

Este modelo (Bornhuetter y Ferguson, 1972), estocástico en sentido bayesiano tal y como ya se ha mencionado anteriormente, combina dos metodologías: la *loss ratio* con otra basada en la experiencia de la compañía. Así, mientras el primer método da una siniestralidad a priori, el segundo se basa en la liquidación de siniestros (recogida en la diagonal del triángulo de desarrollo, los valores $C_{i,I-i}$).

La expresión matemática general que resulta para la estimación de las cuantías futuras acumuladas puede formularse como:

$$\hat{C}_{ij} = \hat{C}_{i,k-i} + (\hat{\gamma}_j - \hat{\gamma}_{k-i}) \cdot \hat{\alpha}_i \quad \text{con } i + j > k \quad (3.20)$$

siendo:

$\hat{\gamma}$: vector de los estimadores a priori de las cuotas acumuladas.

$\hat{\alpha}$: vector de los estimadores a priori de las pérdidas finales esperadas.

Con el objetivo de entender mejor la expresión (3.20), a continuación se enuncian las dos grandes hipótesis que asume el método de Bornhuetter-Ferguson extendido, tal y como se explica en Schmidt y Zocher (2008). Dichos supuestos son:

- I. Existen los vectores $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k)$ y $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_k)$ con $\gamma_k = 1$, tales que la identidad $E[C_{ij}] = \gamma_j \cdot \alpha_i$ sirve para $\forall i, j \in [0, 1, \dots, k]$. Entonces, $E[C_{ij}] = \alpha_i$ y $\gamma_j = \frac{E[C_{ij}]}{E[C_{ik}]}$, lo que significa que es el desarrollo de las cuotas acumuladas.
- II. Los vectores $\hat{\gamma} = (\hat{\gamma}_0, \hat{\gamma}_1, \dots, \hat{\gamma}_k)$ de estimadores a priori de las cuotas acumuladas (con $\hat{\gamma}_k = 1$) y el vector $\hat{\alpha} = (\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_k)$ vienen dados.

En este sentido, cabe mencionar que por lo general, los mencionados estimadores a priori pueden obtenerse a partir de información proporcionada por una o ambas de las siguientes fuentes:

- **Información interna:** es aquella contenida en el triángulo de desarrollo de la cartera analizada.
- **Información externa:** referida a aquella información que no está contenida en el triángulo de desarrollo. Puede obtenerse a partir de estadísticas de los mercados, de otras carteras consideradas con características similares a la analizada, o de las primas u otras medidas de volumen de la cartera analizada.

Por otro lado, puede hallarse una relación entre el patrón de desarrollo de las cuotas acumuladas (γ_{ih}) y el patrón de desarrollo de los factores de desarrollo (d_{ih}). Así:

$$\text{▪ Cuotas acumuladas individuales: } \gamma_{ih} = \frac{E[C_{ih}]}{E[C_{ik}]} \quad 0 \leq h \leq k \quad (3.21)$$

• Factores de desarrollo:
$$d_{ih} = \frac{E[C_{i,h+1}]}{E[C_{ih}]} \quad 0 \leq h \leq k-1 \quad (3.22)$$

A partir de las expresiones (3.21) y (3.22):

$$\gamma_{ih} = \frac{E[C_{ih}]}{E[C_{ik}]} = \frac{E[C_{i,k-1}]}{E[C_{ik}]} \cdot \frac{E[C_{i,k-2}]}{E[C_{i,k-1}]} \cdots \frac{E[C_{i,h+1}]}{E[C_{i,h+2}]} \cdot \frac{E[C_{ih}]}{E[C_{i,h+1}]} = \prod_{j=h}^{k-1} \frac{1}{d_{ij}} \quad (3.23)$$

Por lo tanto, si se utilizan los patrones γ_h y d_h (y en concreto, el estimador de Chain-Ladder para d_h), se tiene que:

$$\hat{\gamma}_h = \prod_{s=h}^{k-1} \frac{1}{\hat{f}_s} \quad (3.24)$$

Finalmente, comentar que el modelo de Bornhuetter-Ferguson original propone utilizar información externa para el estimador a priori de las $\hat{\alpha}$ e información interna para el estimador a priori de las $\hat{\gamma}$. En concreto:

$$\hat{\alpha}_i = \text{Primas}_i \times \text{Ratio de perdidas esperado} \quad (3.25)$$

$\hat{\gamma}_h$ equivalente con el estimador Chain-Ladder de los factores de desarrollo (3.24).

CAPÍTULO 4

EL MODELO LINEAL GENERALIZADO. FORMULACIÓN Y EL CASO PARTICULAR CHAIN-LADDER ESTOCÁSTICO

4.1. El modelo lineal generalizado

4.1.1. Descripción del modelo

Los modelos lineales generalizados (en inglés, *Generalized Linear Models* o *GLM*) se utilizan para evaluar y cuantificar la relación existente entre una variable, denominada respuesta o dependiente (endógena), y un conjunto de variables explicativas o independientes (exógenas). A diferencia del modelo clásico de regresión lineal por mínimos cuadrados ordinarios (MCO), dicha relación se realiza mediante una función de enlace y, además, se permite que la variable endógena tenga una distribución no Normal.

Así, para tratar de hacer más comprensible la metodología al lector, a continuación se describe el modelo aportando algunas expresiones analíticas del mismo.

Supóngase la variable aleatoria $\mathbf{Y}_{(n \times 1)}$, con (y_i) para $i = 1, 2, \dots, n$ observaciones independientes, que recoge la siniestralidad a explicar y que representa el papel de variable respuesta en el modelo. Por otro lado, supóngase P predictores o factores potenciales de la estructura de riesgo $\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_p$, vectores $(n \times 1)$: (f_{ij}) para $i = 1, 2, \dots, n$ y $j = 1, 2, \dots, P$. También, se asumen unos errores centrados $E[\varepsilon_i] = 0$ y un predictor lineal determinista simbolizado por η_i tal que:

$$\eta_i = \beta_0 + \sum_{j=1}^P \beta_j \cdot f_{ij} \quad (4.1)$$

Aunque ya han sido citadas con anterioridad las dos grandes diferencias entre el *GLM* y el modelo clásico, a continuación se remarcan dada su importancia. Éstas son:

1. La variable respuesta (\mathbf{Y}) no tiene porqué seguir una distribución Normal. Habitualmente, se considera que pueda ser una distribución derivada de la familia exponencial de McCullagh y Nelder [para mayor detalle, véase McCullagh y Nelder (1989) pág. 28]. Dichas distribuciones suelen expresar su función de densidad en un punto como:

$$f(y_i; \theta_i, \phi_i) = \exp \left\{ \frac{y_i \cdot \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i) \right\} \quad (4.2)$$

para funciones especificadas $a(\cdot)$, $b(\cdot)$, $c(\cdot)$, donde θ_i se denomina el parámetro canónico y ϕ_i , parámetro de dispersión.

Además, para el caso de la familia exponencial resulta interesante ver la relación que se da entre la media y la varianza. Así, sea $I(y_i; \theta_i, \phi_i)$ el logaritmo neperiano de la función de verosimilitud de la observación i -ésima. Por tanto:

$$I(y_i; \theta_i, \phi_i) = \ln(f(y_i; \theta_i, \phi_i)) = \frac{y_i \cdot \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i) \quad (4.3)$$

Sea U la función de primeras derivadas de I respecto a θ_i , definida como:

$$U = \frac{\partial \ln(f(y_i; \theta_i, \phi_i))}{\partial \theta_i} = \frac{y_i - b'(\theta_i)}{a(\phi_i)} \quad (4.4)$$

A partir de aquí, resulta fácilmente deducible que:

$$E[U] = 0 \Rightarrow E[y_i] = \mu_i = b'(\theta_i) \quad (4.5)$$

$$Var[U] = -E[U^2] \Rightarrow Var[y_i] = a(\phi_i) \cdot b''(\theta_i) \quad (4.6)$$

Por tanto, la varianza se obtiene como producto de dos componentes:

- La primera, $a(\phi_i)$, que depende sólo del parámetro de dispersión ϕ_i y a menudo se escribe como $a(\phi_i) = \frac{\phi}{w_i}$, con parámetro de dispersión constante para todas las observaciones, ϕ , y unos pesos especificados a priori, w_i , que varían de una observación a otra.
- La segunda, $b''(\theta_i)$, únicamente depende de θ_i (y por tanto, de la esperanza μ_i). Esta componente es denominada como función de varianza y se explicita su dependencia respecto de la esperanza: $b''(\theta_i) = V(\mu_i)$.

Considerando lo anterior, puede escribirse entonces:

$$Var[y_i] = \frac{\phi V(\mu_i)}{w_i} \quad (4.7)$$

donde ϕ es el parámetro de dispersión, $V(\mu_i)$ es la función de varianza y w_i es el posible peso especificado a priori de la observación i . Debe notarse que bajo el supuesto de $\phi = 1$, los recíprocos de los pesos pueden interpretarse como parámetros de escala no constantes: $\frac{1}{w_i} = \phi_i$.

2. Existe una relación lineal entre la variable explicativa y una transformación de la media de la variable dependiente o respuesta. Esto es:

$$E[y_i] = \mu_i = F(\eta_i) = F\left(\beta_0 + \sum_{j=1}^p \beta_j \cdot f_{ij}\right) \quad (4.8)$$

Para despejar de (4.8) la variable respuesta, debe hallarse la función inversa de F , $g = F^{-1}$, a la que se denomina función de enlace (en inglés, *link function*), que es la que relaciona la respuesta con el predictor lineal. A la función de enlace (g) se le exige únicamente que sea monótona y diferenciable. Así, se tiene que:

$$g(E[y_i]) = g(\mu_i) = \eta_i = \beta_0 + \sum_{j=1}^P \beta_j \cdot f_{ij} \quad (4.9)$$

Para algunas distribuciones de la familia exponencial, existen funciones de enlace denominadas naturales o canónicas. Para dichos enlaces, el parámetro canónico coincide con el predictor lineal; es decir: $\theta(\mu_i) = \eta_i$.

Las funciones más habituales son las que se recogen en la Tabla 4.1:

Función de enlace	$g(\mu)$	Canónico para:
Identidad	μ	Normal o Gaussiana
Logarítmica	$\ln(\mu)$	Poisson
Recíproco	μ^{-1}	Gamma
-	μ^{-2}	Inversa Gaussiana
Logit	$\ln\left(\frac{\mu}{1-\mu}\right)$	Binomial

Tabla 4.1. Funciones de enlace más frecuentes. [Fuente: Elaboración propia]

No obstante, puede utilizarse otro enlace que no sea el canónico para la modelización a realizar. En general, suele utilizarse uno derivado de la familia de enlaces paramétricos:

$$\eta_i = g(\mu_i) = \begin{cases} \mu_i^\lambda & \text{para } \lambda \neq 0 \\ \log(\mu_i) & \text{para } \lambda = 0 \end{cases} \quad (4.10)$$

Luego, un modelo lineal generalizado tiene tres componentes básicos:

- **Componente aleatoria:** identifica la variable respuesta y su distribución de probabilidad.
- **Componente sistemática:** especifica las variables explicativas (independientes o predictoras) utilizadas en la función predictora lineal.
- **Función de enlace o link:** se trata de una función del valor esperado de Y , $E[Y]$, como una combinación lineal de las variables predictoras.

4.1.2. Estimación de parámetros

Para estimar los parámetros β_j del predictor lineal se busca maximizar el logaritmo de la función de verosimilitud:

$$l = \sum_{i=1}^n l_i = \sum_{i=1}^n \left\{ \frac{y_i \cdot \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i) \right\} \quad (4.11)$$

Debe destacarse que bajo el supuesto de una distribución del error Normal y el uso de la función Identidad como función de enlace, se obtiene como caso particular la solución por MCO del modelo clásico.

Por otro lado, la versión más general del *GLM* no exige que la distribución del error de \mathbf{Y} pertenezca a la familia de distribuciones exponenciales caracterizadas por (4.2). En este contexto, la estimación de los parámetros β_j se lleva a cabo maximizando los logaritmos de las funciones de cuasi-verosimilitud; es decir:

$$q(\mathbf{y}; \boldsymbol{\mu}) = \sum_{i=1}^n q_i = \sum_{i=1}^n w_i \int_{y_i}^{\mu_i} \frac{y_i - s}{\phi V(s)} ds \quad (4.12)$$

Tal maximización conlleva resolver el sistema de ecuaciones lineales mostrado en (4.13) mediante algún método numérico. Para el caso general de la familia exponencial, las cuasi-verosimilitudes juegan el papel de verosimilitudes.

$$\sum_{i=1}^n w_i \frac{y_i - \mu_i}{\phi V(\mu_i)} \frac{\partial \mu_i}{\partial \beta_j} = 0 \quad \forall j \in \{1, 2, \dots, P\} \quad (4.13)$$

4.1.3. Desvianza⁵ y residuos

Una manera de evaluar si un modelo resulta o no adecuado es mediante la comparación con otro en el que se estima el número máximo posible de parámetros (tantos como observaciones haya). A este último se le denomina modelo saturado, máximo o completo y cumple que $\hat{\mu}_i = y_i$ para $i = 1, 2, \dots, n$.

Así, la variabilidad no explicada por un modelo (fijada una función de enlace, una distribución del error y unos predictores) se plasma en la desvianza escalada como:

$$D^*(\mathbf{y}; \hat{\boldsymbol{\mu}}) = -2 \cdot \log \left(\frac{L(\mathbf{y}; \hat{\boldsymbol{\mu}})}{L(\mathbf{y}; \mathbf{y})} \right) \quad (4.14)$$

donde $L(\mathbf{y}; \hat{\boldsymbol{\mu}})$ denota la función de verosimilitud del modelo en estudio y $L(\mathbf{y}; \mathbf{y})$ la función de verosimilitud del modelo saturado.

La desvianza escalada disminuye a mayor número de predictores incluidos en el modelo, hasta llegar a explicar por completo la variabilidad total de los datos. Luego, en términos de desvianzas no escaladas se tiene que:

$$D(\mathbf{y}; \hat{\boldsymbol{\mu}}) = \phi \cdot D^*(\mathbf{y}; \hat{\boldsymbol{\mu}}) \quad (4.15)$$

Y escribiendo las desvianzas en función de las cuasi-verosimilitudes descritas en (4.12), se obtiene:

⁵ Se trata de la medida denominada *deviance*, traducida también por algunos autores como divergencia para no ser confundida con la tradicional desviación.

$$D(\mathbf{y}; \hat{\boldsymbol{\mu}}) = \sum_{i=1}^n d_i = \sum_{i=1}^n 2 \cdot w_i \cdot \frac{y_i - \hat{\mu}_i}{V(\hat{\mu}_i)} d(s) = -2\phi \cdot q(\mathbf{y}; \hat{\boldsymbol{\mu}}) \quad (4.16)$$

Para el cálculo de las desviaciones asociadas a un modelo, tan sólo se precisa disponer de los dos primeros momentos.

Con todo, tras la estimación del modelo parece oportuno analizar si es adecuado para recoger el fenómeno estudiado. Para ello, debe discernirse sobre las elecciones hechas por la función de varianza, la función de enlace y por los regresores incluidos. La manera de hacerlo más común es mediante el estudio de los residuos.

La definición de residuo en los GLM resulta muy amplia fruto de la multiplicidad de distribuciones que pueden utilizarse en este contexto. Es por ello que a continuación sólo se muestran definidos dos tipos de residuos, tratándose a su vez de los más usuales:

- Residuos de Pearson:
$$r_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\frac{V(\hat{\mu}_i)}{w_i}}} \quad (4.17)$$

- Residuos de desviación:
$$r_i^D = \text{signo}(y_i - \hat{\mu}_i) \cdot \sqrt{d_i} \quad (4.18)$$

(donde d_i es la i -ésima componente de (4.16)).

Por otro lado, y en relación a lo anterior, la estimación del parámetro de dispersión ϕ de un modelo con $P+1$ coeficientes puede obtenerse mediante la fórmula:

$$\hat{\phi}^D = \frac{1}{n-P-1} \sum_{i=1}^n (r_i^D)^2 = \frac{1}{n-P-1} \sum_{i=1}^n d_i \quad (4.19)$$

Y, alternativamente, el estimador de momentos basado en los residuos generalizados de Pearson:

$$\hat{\phi}^P = \frac{1}{n-P-1} \sum_{i=1}^n (r_i^P)^2 = \frac{1}{n-P-1} \sum_{i=1}^n w_i \cdot \frac{(y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)} \quad (4.20)$$

Por último, comentar que el modelo más apropiado para unos datos determinados es el que resulta con una desviación menor.

4.1.4. Función de enlace

Las funciones de enlace que tienen un interés especial para el presente estudio son: el enlace identidad y el logarítmico. Así, de la combinación de cualquier distribución del error con el primer enlace citado resulta un modelo aditivo; en cambio, si se utiliza el segundo, se obtiene un modelo multiplicativo.

Luego, atendiendo a (4.9) que relaciona la respuesta con el predictor lineal, se tiene que:

▪ Para el enlace identidad:
$$\mu_i = \eta_i = \beta_0 + \sum_{j=1}^P \beta_j \cdot f_{ij} \quad (4.21)$$

▪ Para el enlace logarítmico:
$$\ln(\mu_i) = \eta_i = \beta_0 + \sum_{j=1}^P \beta_j \cdot f_{ij}$$

$$\mu_i = \exp\left(\beta_0 + \sum_{j=1}^P \beta_j \cdot f_{ij}\right) = e^{\beta_0} \cdot e^{\beta_1 \cdot f_{i1}} \cdot e^{\beta_2 \cdot f_{i2}} \cdot \dots \cdot e^{\beta_P \cdot f_{iP}} \quad (4.22)$$

Obsérvese que si se parte de la base β_0 , en el modelo aditivo debe sumarse las P cantidades $\sum_{j=1}^P \beta_j \cdot f_{ij}$; por su parte, si se parte de e^{β_0} en el multiplicativo, debe multiplicarse por las P cantidades de incremento $\prod_{j=1}^P e^{\beta_j \cdot f_{ij}}$. En el caso particular de predictores binarios, partiendo del efecto global, se procederá a sumar β_j o a multiplicar por e^{β_j} , sólo cuando para el individuo i se dé la característica j , para $j = 1, 2, \dots, P$, ya que en tal caso se cumple $f_{ij} = 1$.

4.1.5. Familia paramétrica de distribuciones

Finalmente, no podía hacerse una introducción a los *GLM* sin contar con este apartado, de gran relevancia por permitir el uso de la familia paramétrica en la distribución del error, tal y como también era posible en la función de enlace. Así pues:

$$V(\mu_i) = \mu_i^\zeta \quad (4.23)$$

Nótese que con esta familia de distribuciones de error se pueden obtener algunos casos particulares de la familia exponencial, como por ejemplo:

- Si $\zeta = 1$, la estructura de error es Poisson.
- Si $\zeta = 2$, la estructura de error es Gamma.
- Si $\zeta = 3$, la estructura de error es Inversa Gaussiana.

4.1.6. Caso particular: Modelo Poisson sobredisperso

De la utilización del *GLM* para el cálculo de la provisión técnica, cabe mencionar que pueden resultar varios casos particulares. Aunque el más conocido sea el Chain-Ladder estocástico, en función de qué supuestos se asuman, también aparecen como caso particular del *GLM* el método de mínimos cuadrados de De Vylder y la metodología de separación de Taylor, tanto la aritmética como la geométrica; ambos, comentados en el capítulo anterior como modelos deterministas. Para mayor detalle, se propone consultar Kaas, R., Goovaerts, M., Dhaene, J. y M. Denuit (2008). *Modern actuarial risk theory: using R. Second edition*. Springer-Verlag. Heidelberg (Alemania). Páginas 271–277.

Así, centrándose en el supuesto en el que se emplea el GLM para la modelización del cálculo de pagos futuros, aplicando como distribución del error una Poisson sobredispersa y junto con la función de enlace canónica logarítmica, se obtiene como caso particular el Chain-Ladder clásico determinista; es decir, la estimación de las reservas coincide en ambos modelos.

Este modelo, el de Poisson sobredisperso, aplicado a las cuantías del triángulo de desarrollo ya conocido, supone los valores siguientes para la distribución:

- Media: $\mu_{ij} = E[c_{ij}]$.
- Varianza: $Var[c_{ij}] = \frac{\phi V(\mu_{ij})}{w_{ij}} = \frac{\phi \mu_{ij}}{w_{ij}}$.
- Parámetro de dispersión: $\phi > 1$.
- Pesos: $w_{ij} = 1$.

Y lo combina con la función de enlace logarítmica; esto es: $\ln \mu_{ij} = \eta_{ij}$.

Luego, el predictor lineal es de la forma:

$$\eta_{ij} = c_0 + \alpha_i + \beta_j \quad (4.24)$$

Siendo α_i el factor correspondiente a los años de ocurrencia con $i=1,2,\dots,k$, β_j el factor correspondiente a los años de desarrollo con $j=1,2,\dots,k$, y c_0 el importe de los siniestros ocurridos en el año 0 y pagados en dicho año.

Con todo lo expuesto, ya pueden realizarse las estimaciones de las cuantías del resto del rectángulo de desarrollo mediante la expresión:

$$\hat{c}_{ij} = \exp(c_0 + \alpha_i + \beta_j) \quad (4.25)$$

A partir de aquí, se calculan las reservas totales (\hat{R}), como la suma de individuales; las reservas por año de origen (\hat{R}_i), sumando las cuantías estimadas por filas para $i=1,2,\dots,k$; y las reservas por año de calendario (\hat{R}_t), sumando las cuantías estimadas de cada contradiagonal del rectángulo de desarrollo completado para $t=1,2,\dots,k$.

4.2. Chain-Ladder estocástico: Modelo de Mack

En este apartado se presenta quizá la metodología estadística más utilizada para calcular las provisiones técnicas de prestaciones pendientes. El modelo aparece por primera vez en Mack (1993) y su principal objetivo es, además de obtener el valor esperado de las reservas, hallar la estimación del error cuadrático medio de las reservas estimadas, total y por año de origen, mediante el método de Chain-Ladder.

Que el modelo de Mack se base en la metodología de Chain-Ladder es algo que merece la pena explicar. Si bien es cierto que Chain-Ladder se caracteriza por ser un método de distribución libre y, por tanto, susceptible de ser usado con cualquier conjunto de datos, también es verdad que las estimaciones realizadas mediante dicho modelo para los años de ocurrencia más recientes, dan como resultado valores más sensibles a los cambios para los datos observados.

Así pues, en este modelo, el factor de cambio de columna (f_j) de Chain-Ladder se corresponde con el valor esperado del factor de desarrollo, supuesto conocido el importe total pagado hasta (e incluido) el periodo $j-1$ con $j=0,1,\dots,n$. Los estimadores de dicho factor de cambio de columna y de la cantidad total pagada hasta un año determinado propuestos por Chain-Ladder se obtienen fácilmente de este modelo de Mack.

A modo de recordatorio, a continuación se muestran las expresiones ya vistas en el Capítulo 3, cuando se hablaba de la metodología clásica de Chain-Ladder, del cálculo de los factores de cambio y del coste de los valores futuros, respectivamente:

▪ Factores de cambio:
$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{ij}} \quad 0 \leq j \leq k-1 \quad (4.26)$$

▪ Valores futuros:
$$\hat{C}_{ij} = C_{i,k-i} \cdot \prod_{h=k-i}^{j-1} \hat{f}_h \quad i, j \in F = \{i, j | 0 \leq i \leq k, 0 \leq j \leq k, i+j > k\} \quad (4.27)$$

En cuanto a los datos de origen para aplicar el modelo, se parte como en la mayoría de metodologías ya explicadas, del triángulo de desarrollo formado por las cuantías acumuladas pagadas hasta (e incluido) el año de desarrollo j , respecto de los siniestros ocurridos en el año de origen i (C_{ij}).

Por lo que a los supuestos utilizados por el modelo de Mack se refiere, éste plantea que las tres hipótesis básicas del método de Chain-Ladder son las siguientes:

▪ Hipótesis 1:
$$E \left[\frac{C_{i,j+1}}{C_{ij}} \middle| C_{ij} \right] = f_j \quad i = 0,1,\dots,k-1 \quad j = 0,1,\dots,k-1 \quad (4.28)$$

O, de forma equivalente,

$$E [C_{i,j+1} | C_{ij}] = C_{ij} \cdot f_j \quad i = 0,1,\dots,k-1 \quad j = 0,1,\dots,k-1 \quad (4.29)$$

▪ Hipótesis 2:

$$\{C_{i,1}, C_{i,2}, \dots, C_{i,l}\}, \{C_{j,1}, C_{j,2}, \dots, C_{j,l}\} \text{ con } i \neq j, \text{ son independientes} \Rightarrow \hat{f}_j \text{ es insesgado}$$

▪ Hipótesis 3:
$$\text{Var} \left[\frac{C_{i,j+1}}{C_{ij}} \middle| C_{ij} \right] = \sigma_j^2 \quad i = 0, 1, \dots, k-1 \quad j = 0, 1, \dots, k-1 \quad (4.30)$$

O, de forma equivalente,

$$\text{Var} [C_{i,j+1} | C_{ij}] = C_{ij} \cdot \sigma_j^2 \quad i = 0, 1, \dots, k-1 \quad j = 0, 1, \dots, k-1 \quad (4.31)$$

De las hipótesis expuestas, puede apreciarse la necesidad de determinar dos estimadores en el modelo, \hat{f}_j y $\hat{\sigma}_j^2$.

El primero de los estimadores mencionados del modelo de Mack responde a la expresión (4.26). Éste, que coincide con el del método de Chain-Ladder, es una media aritmética ponderada de los factores de desarrollo siendo las ponderaciones C_{ij} .

El segundo estimador se define según la expresión:

$$\hat{\sigma}_j^2 = \frac{1}{k-j-1} \sum_{i=0}^{k-j-1} C_{i,j} \cdot \left(\frac{C_{i,j+1}}{C_{ij}} - \hat{f}_j \right)^2 \quad j = 0, 1, \dots, k-2 \quad (4.32)$$

Dado que la formulación anterior no es válida para $j = k-1$ ya que entonces el denominador se igualaría a cero y el valor de la varianza sería infinito, Mack propone para este caso las posibilidades siguientes:

- a) Si $\hat{f}_{k-1} = 1$ y se espera que el desarrollo de los siniestros finalice tras $k-1$ años, entonces $\sigma_{k-1}^2 = 0$.
- b) Si lo anterior no ocurre, es necesario extrapolar el siguiente valor de la secuencia $\{\hat{\sigma}_0^2, \hat{\sigma}_1^2, \dots, \hat{\sigma}_{k-2}^2\}$, que suele ser una sucesión de valores exponencialmente decrecientes. Para ello se sugieren, a su vez, dos posibilidades:

- Mediante el uso de una regresión log-lineal del tipo $\ln \hat{\sigma}_j^2 = \hat{\alpha} + \hat{\beta} \cdot j$, lo que implica que $\hat{\sigma}_{k-1}^2 = \exp \{ \hat{\alpha} + \hat{\beta} \cdot (k-1) \}$.
- Imponiendo que el ritmo de caída de la varianza se mantenga en las últimas observaciones.

Con todo lo anterior y suponiendo que las tres hipótesis planteadas son igualmente válidas para los valores futuros, éstos se determinan según la expresión (4.27). A partir de estos valores estimados y los datos originales se calculan las reservas estimadas por año de origen; esto es:

$$\hat{R}_i = \hat{C}_{i,I} - C_{i,I-i} \quad i = 0, 1, \dots, k \quad (4.33)$$

Luego, como consecuencia de las hipótesis 1 y 2 se obtiene que:

$$E[\hat{C}_{il}] = E[C_{il}] \quad (4.34)$$

Por su parte, gracias a los supuestos planteados en cada una de las tres hipótesis, puede procederse al cálculo del error cuadrático medio (sus siglas en inglés, *MSE*) de las reservas por año de origen y el error típico o estándar (en inglés, *standard error*, *s.e.*).

Así, se propone el siguiente estimador para el error cuadrático medio de las reservas por año de origen:

$$MSE(\hat{R}_i) = \hat{C}_{ik}^2 \cdot \sum_{s=k-i}^{k-1} \frac{\hat{\sigma}_s^2}{\hat{f}_s^2} \left(\frac{1}{\hat{C}_{is}} + \frac{1}{\sum_{q=0}^{k-s} \hat{C}_{qs}} \right) \quad i = 1, 2, \dots, k \quad (4.35)$$

Por su parte, también se propone un estimador para el error cuadrático medio de las reservas totales, que de nuevo sólo tiene sentido cuando las reservas se calculan sin considerar rentabilidad sobre dichas reservas. Dicho estimador viene dado por:

$$MSE(\hat{R}) = \sum_{i=1}^k \left\{ \left(s.e.(\hat{R}_i) \right)^2 + \hat{C}_{ik} \cdot \left(\sum_{j=i+1}^k \hat{C}_{jk} \right) \cdot \sum_{h=k-i}^{k-1} \frac{\frac{2 \cdot \hat{\sigma}_h^2}{\hat{f}_h^2}}{\sum_{n=0}^{k-h} \hat{C}_{nh}} \right\} \quad (4.36)$$

donde $s.e.(\hat{R}_i)$ corresponde al error estándar de la reserva con año de origen i y se refiere a la raíz cuadrada de la estimación del error cuadrático medio anterior.

CAPÍTULO 5

ERROR CUADRÁTICO MEDIO DE LAS RESERVAS TOTALES, POR AÑO DE ORIGEN Y POR AÑO DE CALENDARIO

5.1. Introducción al error de predicción

En los capítulos 3 y 4 se han presentado, con mayor o menor detalle, diferentes metodologías que buscan estimar el valor de los pagos futuros a partir de unos datos iniciales conocidos; lo que los convierte, pues, en modelos predictivos.

En dichos métodos se utiliza el valor esperado como la predicción y su variabilidad viene dada por la raíz cuadrada de la predicción, también conocida como error de predicción.

En este sentido, considérese una variable aleatoria y , y su correspondiente estimación \hat{y} . Luego, el error cuadrático medio de predicción resulta ser:

$$MSE(\hat{y}) = E[(y - \hat{y})^2] \approx E[(y - E[y])^2] + E[(\hat{y} - E[\hat{y}])^2] \quad (5.1)$$

Lo que expresado en forma verbal es equivalente a:

$$\text{Varianza de la predicción} = \text{Varianza del proceso} + \text{Varianza de la estimación}$$

Por otro lado, merece destacar la importancia que tiene diferenciar entre el error de predicción y el error estándar. Estrictamente, el error estándar es entendido como la raíz cuadrada de la varianza de la estimación; en cambio, el error de predicción se refiere a la variabilidad de una previsión, teniendo en cuenta la incertidumbre en el parámetro estimación y fijándose, también, en la variabilidad inherente de los datos estimados.

5.2. Error de predicción en el modelo Chain-Ladder con Poisson sobredisperso

En este apartado se presentan las expresiones analíticas que permiten calcular, de forma muy aproximada, el error de predicción cometido al utilizar el modelo estocástico de Chain-Ladder con distribución del error según un modelo de Poisson sobredisperso, el cual ya se ha explicado en el capítulo anterior y sobre el que se centra la aplicación práctica objeto del presente trabajo.

En este contexto, resulta indispensable referirse a England y Verrall (1999, 2002 y 2006) para adentrarse con más detalle en el cálculo del error de predicción sobre las reservas totales y por año de origen.

Ahora bien, en este documento se aporta también la correspondiente formulación para el cálculo del error de predicción por año de calendario. Si bien es cierto que no es algo novedoso por sí mismo, lo cierto es que hasta el momento la literatura científica existente a penas lo ha estudiado. En este sentido, y dado el marco normativo actual de Solvencia II, se ha considerado que tratar el cálculo de las reservas por año de origen puede ser algo especialmente útil de cara al sector asegurador en los próximos años.

Así pues, tal y como se ha introducido con anterioridad, para determinar el error de predicción se precisa hallar el error cuadrático medio (*MSE*) de predicción, ya que es quien recoge dicha variabilidad. Éste se define como:

$$MSE(\hat{c}_{ij}) = E\left[\left(c_{ij} - \hat{c}_{ij}\right)^2\right] \cong Var\left[c_{ij}\right] + Var\left[\hat{c}_{ij}\right] \quad (5.2)$$

Claramente, se observa que el error de predicción es la suma de dos componentes, algo que ya se ha denotado en la introducción del capítulo. Así, mientras la primera componente se refiere a la variabilidad de los datos (varianza del proceso), la segunda lo hace a la variabilidad de la estimación (varianza de la estimación).

En cuanto a la estimación de cada una de las dos componentes, para la primera se utilizará la fórmula de la distribución y para la segunda, la siguiente aproximación:

$$Var(\hat{c}_{ij}) = \left|\frac{\partial \mu_{ij}}{\partial \eta_{ij}}\right|^2 Var\left[\eta_{ij}\right] \quad (5.3)$$

Con lo que se obtiene:

- Error cuadrático medio para cada estimación de las cuantías:

$$MSE(\hat{c}_{ij}) \approx Var\left[c_{ij}\right] + Var\left[\hat{c}_{ij}\right] = \phi \cdot \hat{\mu}_{ij} + \mu_{ij}^2 \cdot Var\left[\hat{\eta}_{ij}\right] \quad (5.4)$$

- Error cuadrático medio para la estimación de la reserva total:

$$MSE(\hat{R}) \approx \sum_{i,j \in \Delta} \phi \cdot \hat{\mu}_{ij} + \sum_{i,j \in \Delta} \mu_{ij}^2 \cdot Var\left[\hat{\eta}_{ij}\right] + 2 \cdot \sum_{\substack{i_1, j_1 \in \Delta \\ i_2, j_2 \in \Delta \\ i_1 \neq i_2, j_1 \neq j_2}} \hat{\mu}_{i_1, j_1} \cdot \hat{\mu}_{i_2, j_2} \cdot Cov\left[\hat{\eta}_{i_1, j_1}, \hat{\eta}_{i_2, j_2}\right] \quad (5.5)$$

donde Δ se refiere a los elementos estimados del triángulo.

- Error cuadrático medio para la estimación de las reservas por año de origen:

$$MSE(\hat{R}_i) \approx \sum_{j \in \Delta_i} \phi \cdot \hat{\mu}_{ij} + \sum_{j \in \Delta_i} \mu_{ij}^2 \cdot Var\left[\hat{\eta}_{ij}\right] + 2 \cdot \sum_{\substack{j_1, j_2 \in \Delta_i \\ j_2 > j_1}} \hat{\mu}_{i, j_1} \cdot \hat{\mu}_{i, j_2} \cdot Cov\left[\hat{\eta}_{i, j_1}, \hat{\eta}_{i, j_2}\right] \quad (5.6)$$

donde Δ_i se refiere a los elementos estimados del triángulo de la fila i .

- Error cuadrático medio para la estimación de la reserva por año de calendario:

$$MSE(\hat{R}_t) \approx \sum_{i, j \in \Delta_t} \phi \cdot \hat{\mu}_{ij} + \sum_{i, j \in \Delta_t} \mu_{ij}^2 \cdot Var\left[\hat{\eta}_{ij}\right] + 2 \cdot \sum_{\substack{i_1, j_1 \in \Delta_t \\ i_2, j_2 \in \Delta_t \\ i_1 \neq i_2, j_1 \neq j_2}} \hat{\mu}_{i_1, j_1} \cdot \hat{\mu}_{i_2, j_2} \cdot Cov\left[\hat{\eta}_{i_1, j_1}, \hat{\eta}_{i_2, j_2}\right] \quad (5.7)$$

donde Δ_t se refiere a los elementos estimados de una misma contradiagonal del triángulo, indicando así los años de calendario. Con $t = 1, 2, \dots, k$ años de calendario futuros, haciendo variar de manera simultánea $i = 1, 2, \dots, k$ y $j = k - i + 1$; es decir, para $t = 1$, $\Delta_{t=1} = \{c_{1,k}, c_{2,k-1}, \dots, c_{k,1}\}$.

CAPÍTULO 6

DISTRIBUCIÓN PREDICTIVA DE LAS RESERVAS. ESTIMACIÓN BOOTSTRAP

6.1. Introducción al remuestreo o *bootstrap*

El *bootstrap* o remuestreo es una técnica desarrollada para llevar a cabo ciertas inferencias. A grandes rasgos, se trata de reemplazar el sistema clásico de cálculo de ciertas estadísticas, mediante la aplicación de expresiones analíticas complejas, por un mecanismo basado en la aplicación de un algoritmo numérico de simulación. Éste consiste en generar un elevado número de muestras (habitualmente un número de iteraciones superior a 1000) para poder realizar esas inferencias.

Además, aparte de poder estimar los errores estándares de predicción como alternativa a la formulación comentada en el capítulo 5, con esta metodología también puede estimarse la distribución predictiva de las reservas totales, por año de origen y por año de calendario (algo que como se ha comentado ya, es de gran relevancia en el presente trabajo).

En el contexto del cálculo de reservas, este método de remuestreo surge como solución para hallar un modelo que reproduzca los resultados de Chain-Ladder estocástico, permitiendo conocer el error de predicción cometido, y basándose en la distribución de Poisson sobredispersa.

Así, a continuación se presenta el procedimiento general de la metodología *bootstrap*, refiriéndose al caso particular denominado *bootstrapping residuals* y siendo el modelo seguido en el desarrollo de la aplicación práctica del presente trabajo, tal y como se comentará en el capítulo 7:

1. Se aplica el modelo Chain-Ladder clásico al triángulo *run-off* de los datos originales de cuantías acumuladas, se desacumulan y se calculan los residuos de Pearson, tal y como sigue:

$$r_{ij}^P = \frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\hat{c}_{ij}}} \quad (6.1)$$

2. Los citados residuos se remuestran N veces y con ellos se obtienen N nuevas muestras de triángulos de desarrollo de cuantías no acumuladas, teniendo presente la expresión (6.1) de la que se despeja:

$$c_{ij}^* = r_{ij}^{P*} \cdot \sqrt{\hat{c}_{ij}} + \hat{c}_{ij} \quad (6.2)$$

3. Se procede a la estimación del modelo lineal generalizado Poisson sobredisperso (válido también para el Chain-Ladder clásico) a cada una de las N muestras y con él las reservas individuales, por año de origen y total. Además, como novedad, en este estudio, se ha añadido la estimación de las reservas por año de calendario tal y como se viene diciendo.
4. El resultado es que se dispone de N valores de las reservas totales, por año de origen y por año de calendario, lo que proporciona la distribución predictiva de

las citadas reservas. Al conocerse dicha distribución, pueden calcularse estadístico de información tales como la media, la desviación estándar o cuantiles. En este sentido, merece destacar la posibilidad de poderse calcular, también, algunas medidas de riesgo basadas en la determinación de un cuantil concreto de la distribución, como puede ser el VaR_α ⁶, ya mencionado en el capítulo 2. La aplicación práctica de lo expuesto se verá con mayor detalle en el capítulo 7 donde se desarrollan casos prácticos.

Alternativamente, en el proceso *bootstrap* puede trabajarse con los residuos de Pearson escalados, los cuales incluyen en el denominador también el parámetro de dispersión que multiplica a la varianza de la variable aleatoria. Esto es:

$$\frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\hat{\phi} \cdot \hat{c}_{ij}}} \quad (6.3)$$

Por otro lado, puede incluirse además la corrección por sesgo del mismo modo que cuando se estimaba el parámetro de escala:

$$\sqrt{\frac{n}{n-p}} \cdot \frac{c_{ij} - \hat{c}_{ij}}{\sqrt{\hat{c}_{ij}}} \quad (6.4)$$

En tal caso, no sería necesario incluir la corrección por sesgo en el cálculo del error estándar estimado con *bootstrap*.

6.2. Error de predicción estimado mediante *bootstrap*

El error de predicción estimado mediante esta metodología de remuestreo corresponde con la raíz cuadrada de la suma de la varianza de las reservas según la distribución supuesta (en el presente caso, Poisson) más la varianza estimada del error mediante *bootstrap*, tal y como sigue:

- Error cuadrático medio para cada estimación de las cuantías:

$$MSE_{bootstrap}(c_{ij}) = \hat{\phi}^p \cdot \hat{c}_{ij} + \frac{n}{n-p} (SE_{bootstrap}(c_{ij}))^2 \quad (6.5)$$

- Error cuadrático medio para la estimación de la reserva total:

$$MSE_{bootstrap}(R) = \hat{\phi}^p \cdot \hat{R}_i + \frac{n}{n-p} (SE_{bootstrap}(R))^2 \quad (6.6)$$

- Error cuadrático medio para la estimación de las reservas por año de origen:

$$MSE_{bootstrap}(R_i) = \hat{\phi}^p \cdot \hat{R}_i + \frac{n}{n-p} (SE_{bootstrap}(R_i))^2 \quad i = 1, 2, \dots, k \quad (6.7)$$

⁶ Véase la definición formal de *VaR* para un determinado nivel de confianza α : $VaR_\alpha(X) = \inf \{x : P[X \leq x] \geq \alpha\}$

- Error cuadrático medio para la estimación de la reserva por año de calendario:

$$MSE_{bootstrap}(R_t) = \hat{\phi}^p \hat{R}_t + \frac{n}{n-p} (SE_{bootstrap}(R_t))^2 \quad t = 1, 2, \dots, k \quad (6.8)$$

En donde el parámetro de dispersión se estima con la expresión (4.20) ya vista en el capítulo 4.

Por su parte, $SE_{bootstrap}$ hace referencia al error estándar de la distribución *bootstrap* considerada (reservas estimadas, totales, por año de origen y por año de calendario).

Nótese que para la estimación del error estándar se emplea una estimación corregida por sesgo, siendo $n-p$ el número de grados de libertad, con $p = 2k + 1$ el número de parámetros del modelo.

Finalmente, hay que mencionar que en el proceso de simulación de residuos y de estimación de los N modelos involucrados en el mecanismo de remuestreo, debe tenerse en cuenta la distribución del error supuesta (si es Poisson como en el caso presentado o, por ejemplo, si es Gamma); de esta forma, se aplicará la correspondiente formulación en la función de varianza.

CAPÍTULO 7

CASOS PRÁCTICOS CON REXCEL

7.1. Introducción a RExcel

En este capítulo se pretende justificar el desarrollo de una aplicación informática, útil para las compañías aseguradoras del ramo no-vida, mediante el uso de RExcel; un complemento para Microsoft Excel que permite el acceso al paquete estadístico de R desde el propio Excel.

Tal y como indican sus autores (T. Baier y E. Neuwirth, 2003), las principales características de dicho complemento son:

- La transferencia de datos entre R y Excel en ambas direcciones.
- Ejecución de código R directamente desde rangos de Excel.
- Programación de macros en código VBA que llaman a R para realizar cálculos sin exponer R al usuario.
- Llamar a funciones de R directamente desde las celdas de Excel, utilizando el modo de actualización automática de Excel para activar el nuevo cálculo de R.
- Uso de Excel como una interfaz gráfica de usuario de R, por lo que la funcionalidad de R resulta accesible a través de los menús y cuadros de diálogo, en lugar de un estilo de programación orientada a la línea de comandos.

Todo ello, conlleva a que RExcel pueda resultar atractivo para su utilización a nivel profesional dentro del sector asegurador; especialmente, por permitir trabajar con las múltiples posibilidades que ofrece R a través de sus paquetes, utilizando la interfaz gráfica de usuario del propio Excel, la cual facilita mucho la introducción de datos.

Así pues, dado que el presente trabajo se engloba en la especialidad de Modelos Actuariales y Financieros Aplicados del Máster en Ciencias Actuariales y Financieras, se ha creído oportuno desarrollar un aplicativo para el cálculo de provisiones mediante la metodología de Chain-Ladder estocástico. Todo lo referido a dicha aplicación se explica con mayor detalle en el siguiente apartado.

Por otro lado, debe mencionarse que RExcel funciona en Microsoft Windows (XP, Vista y 7) y únicamente con las versiones de 32 bits de Excel 2003, 2007, 2010 y 2013; no obstante, es posible instalarlo en las versiones de Windows de 64 bits.

Comentar también que la versión utilizada en este trabajo, RExcel 3.2.14, requiere una versión de 32 bits, igual o posterior, a R 2.15.0. Para mayor detalle de los requisitos informáticos y/o descarga de versiones, consultar los enlaces siguientes:

- RExcel y sus requisitos de instalación: <http://rcom.univie.ac.at>.
- Paquete estadístico R: <http://www.r-project.org>.

7.2. Desarrollo de una aplicación informática

Como ya se ha indicado, el desarrollo de la herramienta informática para el cálculo de provisiones técnicas se ha llevado a cabo gracias a RExcel. Su implementación en un libro de Excel se ha realizado mediante código de programación de *Visual Basic for Applications* (VBA); no obstante, la formulación utilizada, referida a la metodología de cálculo empleada, se basa en código de R. El archivo que permite la ejecución de dicho aplicativo se entrega adjunto a este documento con el nombre y la extensión siguientes: *ChainLadderEstocástico.xlsm*.

Sin querer entrar en detalles de la programación, en este apartado del trabajo se explica el objetivo perseguido por la aplicación, su estructura y la interacción que se plantea para con el usuario. De todos modos, al final del trabajo se adjunta en forma de anexo el código de R utilizado para los cálculos correspondientes del aplicativo desarrollado, por si el lector quisiera reproducirlos y comprobar que la aplicación funciona correctamente (véase Anexo 1).

Así, el objetivo principal que se busca con la creación de dicha aplicación es el de permitir al usuario calcular la estimación de reservas totales, por año de origen y por año de calendario, partiendo de un triángulo de desarrollo completo, formado por cuantías no acumuladas (c_{ij}) y con igual número de años de origen (i) que de desarrollo (j). La metodología que se utiliza es la de Chain Ladder estocástico, utilizando un *GLM* con Poisson sobredisperso y enlace logarítmico, bajo dos posibilidades:

- Mediante fórmula analítica.
- Mediante *bootstrapping*.

Además, para el caso de las reservas por año de calendario, aparte de estimar los pagos futuros, se incluyen como resultados adicionales:

- Valor actual de las reservas estimadas (IBNR).
- Valor actual de las reservas añadiéndoles un margen de solvencia elegido por el usuario, correspondiente a un porcentaje del error de predicción estimado; esto es, se calcula el valor actual de IBNR + % P.E.
- Valor actual del VaR_{α} de las reservas, donde el nivel de confianza (α) lo determina el usuario. Este resultado sólo se incluye en la metodología *bootstrap* ya que es donde se hace un remuestreo de la distribución predictiva y, por lo tanto, es donde tiene sentido calcularlo.

Lógicamente, es preciso introducir la rentabilidad futura que se espera obtener para el cálculo de los valores actuales. Para ello, el usuario libremente puede introducir cualquier tipo de interés efectivo anual para que la aplicación haga los correspondientes cálculos. En este sentido, la aplicación queda limitada a poder introducir únicamente un interés efectivo anual, sin poder jugar con rentabilidades de duraciones mayores a la anual, de forma directa.

Por lo que a la estructura del libro de Excel de la aplicación se refiere, éste contiene las siguientes pestañas:

1. *INPUTS*. Pestaña donde el usuario debe introducir el triángulo de desarrollo con los datos, expresados en cuantías no acumuladas tal y como ya se ha dicho. Además, en esta hoja del aplicativo se halla el botón que permite la ejecución de la aplicación.
2. *GLM 1*. Pestaña donde se muestran los resultados obtenidos de las reservas totales y por año de origen mediante fórmula analítica.
3. *GLM 2*. Pestaña donde se muestran los resultados obtenidos de las reservas por año de calendario mediante fórmula analítica.
4. *Bootstrap 1*. Pestaña donde se muestran los resultados obtenidos de las reservas totales y por año de origen mediante *bootstrap*.
5. *Bootstrap 2*. Pestaña donde se muestran los resultados obtenidos de las reservas por año de calendario mediante *bootstrap*.

Cabe destacar que el contenido que haya en cualquier pestaña de resultados es eliminado por la aplicación antes de ofrecer otro nuevo en la misma pestaña.

A continuación, se ilustran los principales mensajes que el usuario se va a encontrar cuando ejecute la aplicación. En primer lugar, al abrirla aparecerá un libro de Excel, con las pestañas ya descritas y, concretamente, en la pestaña *INPUTS* aparecerá el ejemplo que se utiliza más adelante para la presentación de resultados, así como el botón de ejecución de la macro (véase Figura 7.1).

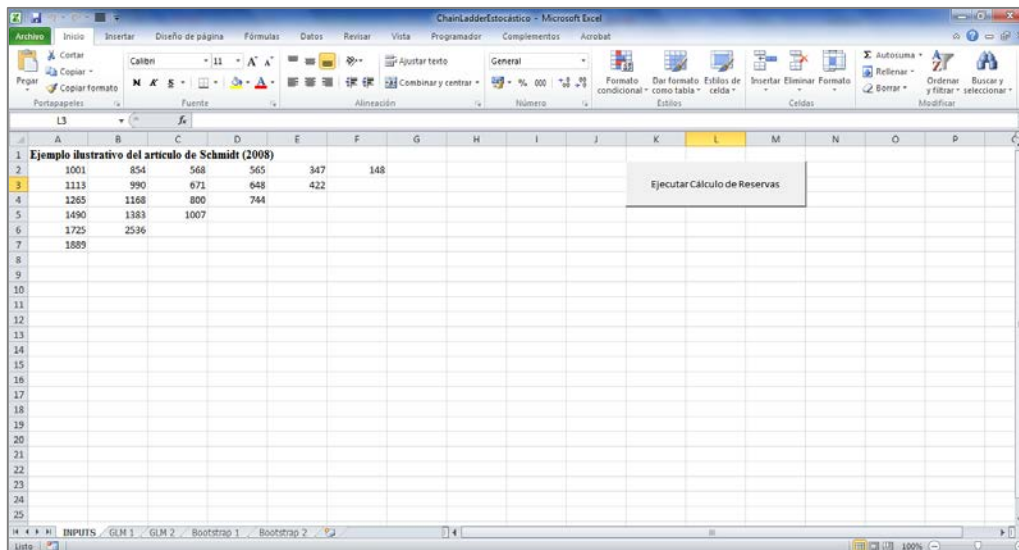


Figura 7.1. Vista inicial al abrir el archivo de la aplicación "ChainLadderEstocástico.xlsm". [Fuente: Elaboración propia]

Apretando en el botón que aparece en el libro de Excel, empieza la ejecución de la aplicación. Inmediatamente, aparece el mensaje informativo ilustrado en la Figura 7.2, advirtiendo al usuario que para ejecutar correctamente la aplicación es preciso tener instalado el complemento RExcel y el paquete estadístico de R. El mensaje posterior simplemente pide al usuario si desea o no proceder con la ejecución de la macro.

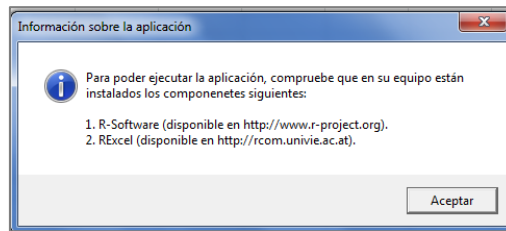


Figura 7.2. Ventana sobre requisitos previos para ejecutar la aplicación. [Fuente: Elaboración propia]

Si el usuario decide ejecutar, aparece una ventana informativa sobre la autoría de la aplicación y su contextualización académica, tal y como se aprecia en la Figura 7.3).

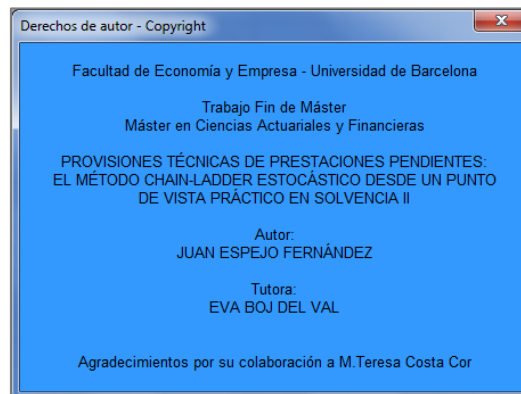


Figura 7.3. Información sobre la autoría de la aplicación. [Fuente: Elaboración propia]

A continuación, la siguiente ventana emergente que aparece es la ilustrada en la Figura 7.4. En ella, se pide al usuario que mediante la introducción del número correspondiente haga la selección de aquello que desea calcular y de qué manera.

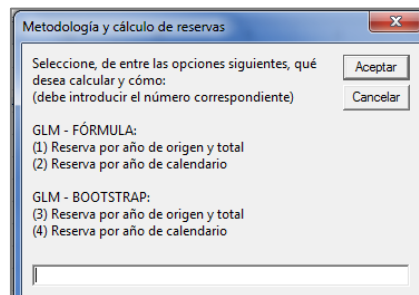


Figura 7.4. Elección de la metodología y el cálculo de las reservas. [Fuente: Elaboración propia]

Tras introducir la correspondiente elección, la aplicación indicará que se seleccione un rango de datos de la pestaña *INPUTS*. Tras ello, se pedirá al usuario que introduzca un valor, cuando proceda, para asignar a las siguientes variables:

- Una rentabilidad para los pagos futuros (tipo de interés efectivo anual).
- Un porcentaje que se asignará al error de predicción para sumar al IBNR a modo de margen de solvencia en uno de los cálculos de valor actual ya comentados.
- Número de iteraciones a realizar en *bootstrap*.
- Nivel de confianza α para el cálculo del $VarR_\alpha$.

Finalmente, se activará la pestaña correspondiente del libro de Excel y se mostrarán los resultados obtenidos.

7.3. Presentación de resultados

Para llevar a cabo la ilustración de los resultados obtenidos por la aplicación desarrollada a partir de unos datos iniciales, se ha utilizado un ejemplo muy recurrente en la literatura actuarial. Se trata del ejemplo utilizado en el artículo de Klaus D. Schmidt y M. Zocher (2008) citado en el apartado de referencias del trabajo.

Así pues, el triángulo de desarrollo planteado es el siguiente:

Año de origen i	Año de desarrollo j					
	0	1	2	3	4	5
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	
2	1265	2433	3233	3977		
3	1490	2873	3880			
4	1725	4261				
5	1889					

Tabla 7.1. Triángulo de desarrollo formado por cuantías acumuladas (C_{ij}). [Fuente: Klaus D. Schmidt y M. Zocher (2008)]

Nótese que se trata de un triángulo de desarrollo formado por cuantías acumuladas (C_{ij}). Por tanto, para poder utilizar dichos datos en la aplicación presentada resulta necesario desacumular las cuantías; es decir, convertir el triángulo de datos anterior en otro formado, ahora, por cuantías no acumuladas (c_{ij}). Se obtiene, pues:

Año de origen i	Año de desarrollo j					
	0	1	2	3	4	5
0	1001	854	568	565	347	148
1	1113	990	671	648	422	
2	1265	1168	800	744		
3	1490	1383	1007			
4	1725	2536				
5	1889					

Tabla 7.2. Triángulo de desarrollo formado por cuantías desacumuladas (c_{ij}). [Fuente: Klaus D. Schmidt y M. Zocher (2008)]

A continuación, y antes de mostrar cualquier resultado obtenido con la macro de Excel desarrollada, se presentan los resultados que se obtienen de aplicar a los datos de la Tabla 7.1 el método de Chain-Ladder y el modelo de Mack, utilizando la función MackChainLadder del *package* “ChainLadder” de R.

Así, las siguientes tablas (Tabla 7.3, Tabla 7.4 y Tabla 7.6) recogen los aspectos más relevantes obtenidos de la aplicación de ambas metodologías citadas en el párrafo anterior.

Año de origen i	Año de desarrollo j					
	0	1	2	3	4	5
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	4014,588
2	1265	2433	3233	3977	4454,116	4651,780
3	1490	2873	3880	4780,731	5354,269	5591,880
4	1725	4261	5662,015	6976,436	7813,390	8160,132
5	1889	3874,541	5148,489	6343,693	7104,741	7420,034

Tabla 7.3. Rectángulo de desarrollo formado por cuantías acumuladas (C_{ij}) para el modelo de Mack obtenido con la función *MackChainLadder* del paquete *ChainLadder* de R. [Fuente: Elaboración propia]

Año de origen i	Año de desarrollo j			
	Ultimate	IBNR	Mack S.E.	CV (IBNR)
0	3483	0	0,00	-
1	4015	171	9,46	0,0555
2	4652	675	26,30	0,0390
3	5592	1712	31,39	0,0183
4	8160	3899	111,81	0,0287
5	7420	5531	140,14	0,1748
Totales	33321,41	11987,41	980,86	0,08

Tabla 7.4. Resumen de parámetros obtenidos para el modelo de Mack hallado con la función *MackChainLadder* del paquete *ChainLadder* de R. [Fuente: Elaboración propia]

Además, para el caso de Mack se obtienen los factores de desarrollo siguientes (siendo coincidentes con los de la metodología clásica de Chain-Ladder):

k	Factor desarrollo (f_k)
0	2,051107
1	1,328800
2	1,232147
3	1,119969
4	1,044378
5	1,000000

Tabla 7.5. Factores de desarrollo (f_k) obtenidos para el modelo de Mack con la función *MackChainLadder* del paquete *ChainLadder* de R [Fuente: Elaboración propia]

Año de origen i	Año de desarrollo j			
	Ultimate	IBNR	S.E.	CV (IBNR)
0	3483	0	0.00	-
1	4015	171	82,960	0,4851
2	4652	675	160,004	0,2370
3	5592	1712	270,821	0,1582
4	8160	3899	477,307	0,1224
5	7420	5531	737,732	0.1334
Totales	29838	11987	1167,056	0,0974

Tabla 7.6. Resumen de parámetros obtenidos para el Modelo Lineal Generalizado (Chain-Ladder) utilizando la opción *mse.method = "formula"* mediante la función *glmReserve* del paquete *ChainLadder* de R. [Fuente: Elaboración propia]

Por otro lado, los factores que se obtienen del predictor lineal al aplicar GLM con Poisson sobredispersión, son:

c_0	6,78751
α_1	0,14204
α_2	0,28936
α_3	0,47342
α_4	0,85137
α_5	0,75629
β_1	0,04984
β_2	-0,39393
β_3	-0,45773
β_4	-0,90911
β_5	-1,79030

Tabla 7.7. Valor de los factores del predictor lineal para el GLM Poisson sobredisperso de Chain-Ladder obtenidos, utilizando la opción `mse.method = "formula"` mediante la función `glmReserve` del paquete `ChainLadder` de R.

[Fuente: Elaboración propia]

Ahora sí, se procede a mostrar los resultados obtenidos mediante la aplicación propia desarrollada, para cada una de las cuatro posibilidades de cálculo implementadas. Cabe mencionar que los datos introducidos supuestos toman los valores siguientes, siendo utilizados cuando proceda:

1. Rentabilidad para los pagos futuros (tipo de interés efectivo anual): 0,3%.
2. Porcentaje del error de predicción: 25%.
3. Número de iteraciones a realizar en *bootstrap*: 1000.
4. Nivel de confianza α para el cálculo del VaR_α : 75%.

Además, el rectángulo de desarrollo, formado por cuantías des acumuladas (c_{ij}), obtenido en todos los casos es el que sigue:

Año de origen i	Año de desarrollo j					
	0	1	2	3	4	5
0	1001	854	568	565	347	148
1	1113	990	671	648	422	170,588
2	1265	1168	800	744	477,116	197,664
3	1490	1383	1007	900,731	573,539	237,611
4	1725	2536	1401,015	1314,420	836,955	346,741
5	1889	1985,541	1273,948	1195,207	761,046	315,293

Tabla 7.8. Rectángulo de desarrollo formado por cuantías des acumuladas (c_{ij}) obtenido por la aplicación desarrollada.

[Fuente: Elaboración propia]

Luego, los resultados concretos para cada uno de las cuatro posibilidades que ofrece la macro desarrollada, se presentan a continuación nombradas por su pestaña de resultados:

1. GLM 1. Reservas totales y por año de origen:

Años de origen	IBNR	P.E.	C.V.
0	0	0	-
1	170,588306	82,959836	0,48631608
2	674,779827	160,003724	0,23711199
3	1711,87999	270,820512	0,15820064
4	3899,13151	477,307109	0,1224137
5	5531,03429	737,731548	0,1333804
PT Total	11987,4139	1167,05581	0,09735676

Tabla 7.9. Resultados de la aplicación sobre reservas totales y por año de origen (por fórmula). [Fuente: Elaboración propia]

2. GLM 2. Reservas por año de calendario:

Años de calendario	IBNR	P.E.	C.V.
6	4934,99152	440,797315	0,08932078
7	3359,57066	379,501103	0,11296119
8	2269,77214	331,884075	0,14621912
9	1107,78673	244,241108	0,22047665
10	315,292873	139,453771	0,44229916
Tipo de interés (tanto por uno): 0,003 % P.E. (tanto por uno): 0,25 Valor Actual IBNR: 11914,39871 Valor Actual IBNR + 25% P.E.: 12295,48392			

Tabla 7.10. Resultados de la aplicación sobre reservas por año de calendario (por fórmula). [Fuente: Elaboración propia]

3. Bootstrap 1. Reservas totales y por año de origen:

Años de origen	IBNR	P.E.	C.V.
0	0	0	-
1	170,588306	81,3294014	0,47675836
2	674,779827	163,706997	0,24260802
3	1711,87999	263,794507	0,15409638
4	3899,13151	473,690494	0,12148615
5	5531,03429	732,82986	0,13249418
PT Total	11987,4139	1207,77949	0,10075397

Tabla 7.11. Resultados de la aplicación sobre reservas totales y por año de origen (por bootstrap). [Fuente: Elaboración propia]

4. *Bootstrap* 2. Reservas por año de calendario:

Años de calendario	IBNR	P.E.	C.V.	VaR
6	4934,99152	448,599947	0,09090187	5096,583
7	3359,57066	386,579795	0,11506821	4737,66871
8	2269,77214	337,81103	0,14883037	5222,20301
9	1107,78673	247,589377	0,22349914	4289,02584
10	315,292873	144,500638	0,45830607	4486,4287

Tipo de interés (tanto por uno):	0,003
% P.E. (tanto por uno):	0,25
Valor Actual	
IBNR:	
11914,39871	
Valor Actual IBNR + 25% P.E.:	
12302,72639	
Valor Actual VaR al 75%:	
23623,86782	

Tabla 7.12. Resultados de la aplicación sobre reservas por año de calendario (por bootstrap). [Fuente: Elaboración propia]

A la vista de los resultados ofrecidos en las cuatro tablas anteriores, debe explicarse que en función del nivel de confianza al que se quiera calcular el VaR_α y del porcentaje del error de predicción que se asigne, los resultados de los valores actuales oscilarán con mayor o menor grado hacia una postura que conlleve más o menos riesgo (en función de lo alejado que se encuentre de la media). El llamado “Valor Actual IBNR” representa un umbral mínimo exigible, pues se corresponde con la media de la distribución.

Por otro lado, cuando proceda la posibilidad de visualizar histogramas y estadísticos referidos a la distribución predictiva, aparecerá la ventana que se ilustra en la Figura 7.5. En ella, el usuario puede seleccionar entre las reservas por años de calendario y total.

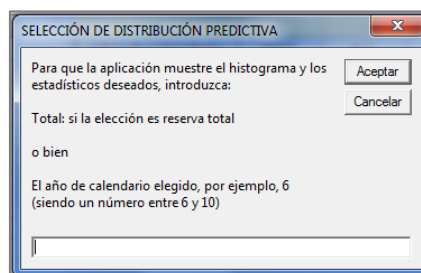


Figura 7.5. Selección de histogramas y estadísticos a visualizar para el caso 4 (GLM - Bootstrap, Reservas por años de calendario) del aplicativo. [Fuente: Elaboración propia]

Finalmente, en el Anexo 2 se muestran los histogramas relativos a la distribución predictiva de las reservas por años de calendario y total junto con algunos estadísticos de dichas distribuciones. Con esta descripción estadística de las distribuciones predictivas de las reservas el usuario puede tomar mejores decisiones sobre el nivel de confianza a seleccionar en el cálculo del VaR .

CAPÍTULO 8

CONCLUSIONES. SINOPSIS Y APORTACIONES

La elaboración del presente estudio ha representado una excelente oportunidad para adentrarse en los distintos métodos estadísticos existentes para el cálculo de provisiones técnicas de prestaciones pendientes. Lógicamente, no se ha podido abordar la totalidad de modelos que se dan ya que de algún modo resultaba necesario acotar el alcance del estudio, pero sí ha servido para introducir los más habituales, haciendo especial mención al modelo Chain-Ladder estocástico, derivado como caso particular del *GLM*.

Además, la inminente entrada en vigor de la Directiva europea de Solvencia II ha justificado, sin duda alguna, que el estudio aportara un enfoque más práctico y pensando en la adaptación de las compañías dentro de la citada contextualización legislativa del sector asegurador.

Con todo ello, se ha considerado que la realización de esta tesina ofrecía la posibilidad de plantearse la creación de una aplicación informática que combinara, por un lado, la teoría y metodología estadística más popular para el cálculo de reservas y, por el otro, que ofreciera resultados realmente útiles en el nuevo marco normativo.

En este sentido, durante el transcurso de la elaboración del documento se ha puesto de manifiesto el hecho que la mayoría de autores que han escrito sobre esta temática actuarial se centrasen, básicamente, en la estimación de los pagos futuros para las reservas totales y/o por año de origen; no obstante, apenas se ha escrito algo sobre el cálculo de reservas por año de calendario.

Así, la principal aportación de este trabajo es la formulación, así como su posterior implementación en una aplicación informática, del cálculo de las reservas por año de calendario en el *GLM*; lo cual resulta especialmente útil por posibilitar el cálculo de prestaciones pendientes en un entorno financiero, aspecto que se contempla en Solvencia II.

La otra gran aportación realizada es la creación de una herramienta informática basada en los códigos de programación de R y *Visual Basic for Applications* de Microsoft Excel, los cuales han sido posibles complementar entre ellos gracias al programa RExcel, ejecutable como complemento en Excel. En dicha aplicación se permite al usuario calcular las reservas totales, por año de origen y por año de calendario partiendo de un triángulo de desarrollo formado por las cuantías no acumuladas de los siniestros. A través de un botón el usuario puede ejecutar la aplicación y calcular sus reservas mediante dos metodologías: fórmula analítica y *bootstrap*. Además, el usuario es quien decide con qué tipo de interés anual se calcula el Valor Actual de los pagos futuros, qué porcentaje del error de predicción se quiere añadir al valor del IBNR como margen de solvencia para hallar su actualización financiera y, además, en el caso de *bootstrap* se puede introducir el número de iteraciones a realizar para llevar a cabo el remuestreo y el nivel de confianza al que se desea calcular el *VaR* de las reservas por año de calendario.

Con todo ello, y a fin de poder corroborar que los resultados obtenidos por la aplicación creada son los correctos, se ha creído conveniente ilustrarlos a partir de un ejemplo utilizado en la literatura actuarial por otros autores y, por lo tanto, ofreciendo la posibilidad de ser contrastados los resultados ofrecidos en este estudio.

Por último, dado que la elaboración de este trabajo se enmarca dentro de la especialidad de modelos actuariales aplicados del Máster en Ciencias Actariales y Financieras, se ha considerado suficientemente justificado el desarrollo práctico mediante una aplicación informática que pudiera ser utilizada por cualquier compañía de seguros para el cálculo de sus reservas. No obstante, también se ha considerado oportuno hacer una contextualización teórica sobre qué se iba a utilizar y por qué; pues el fundamento teórico no se entendería sin una visión teórica de lo aplicado y, posiblemente, al revés tampoco.

GLOSARIO

PALABRAS CLAVES EN EL ÁMBITO DE SOLVENCIA II

Dado que en el presente estudio se ha utilizado terminología específica del contexto normativo de Solvencia II, y con el objetivo de facilitar la comprensión del documento a lectores que no estén inmersos plenamente en el sector asegurador, a continuación se presenta un breve glosario de conceptos fundamentales citados a lo largo del presente estudio, extraídos todos ellos de la propia Directiva.

Capital de solvencia obligatorio (en inglés, *Solvency Capital Requirement* o *SCR*). Se corresponde con el capital económico que han de poseer las empresas de seguros y de reaseguros para limitar la probabilidad de ruina a un caso por cada 200 o, de forma alternativa, que las empresas todavía estén en situación, con una probabilidad del 99,5% como mínimo, de cumplir sus obligaciones frente a los tomadores y beneficiarios de seguros en los doce meses siguientes. Ese capital económico debe calcularse sobre la base del verdadero perfil de riesgo de dichas empresas, teniendo en cuenta la incidencia de las posibles técnicas de reducción del riesgo, así como los efectos de diversificación.

Capital mínimo obligatorio (en inglés, *Minimum Capital Requirement* o *MCR*). Hace referencia al nivel de fondos propios por debajo del cual los tomadores y los beneficiarios, en caso de continuar las empresas de seguros y de reaseguros su actividad, estarían expuestos a un nivel de riesgo inaceptable. La insuficiencia del capital mínimo obligatorio activa la intervención en última instancia de las autoridades de supervisión. Anteriormente, fondo de garantía.

Efectos de diversificación (en inglés, *diversification effects*). Reducción de la exposición al riesgo de las empresas y grupos de seguros y de reaseguros relacionada con la diversificación de sus actividades, y resultante de la posibilidad de compensar el resultado negativo de un riesgo con el resultado más favorable de otro riesgo, cuando no exista una total correlación entre dichos riesgos

Horizonte temporal (en inglés, *time horizon*). Período durante el cual cualquier cantidad de capital requerido se lleva a cabo con el fin de cubrir pérdidas, dentro de un nivel de tolerancia al riesgo dado.

Margen de solvencia (en inglés, *solvency margin*). Conjunto de recursos constituidos por patrimonio propio no comprometido, coincidente en cierta medida con el patrimonio neto contable que, como mínimo, deben tener las entidades aseguradoras, en todo momento, para garantizar económicamente al máximo los compromisos con sus asegurados.

Medida de riesgo (en inglés, *risk measure*). Función matemática que asigna un valor monetario a una determinada previsión de distribución de probabilidad y que crece monótonicamente con el nivel de exposición al riesgo subyacente a esa previsión de distribución de probabilidad.

Mejor estimación (en inglés, *best estimate*). Valor que corresponde con la media de los flujos de caja futuros ponderada por su probabilidad, teniendo en cuenta el valor temporal del dinero (valor actual esperado de los flujos de caja futuros) mediante la aplicación de la pertinente estructura temporal de tipos de interés sin riesgo.

Modelo interno (en inglés, *internal model*). Aquel sistema de gestión de riesgos de una aseguradora para el análisis de la situación general de riesgos de la empresa de seguros, para cuantificar los riesgos y/o para determinar la exigencia de capital sobre la base del perfil específico del riesgo de la compañía.

Nivel de confianza (en inglés, *confidence level*). Medida estadística del nivel de certeza con respecto a un resultado. Típicamente se expresa como el valor de probabilidad ($1-\alpha$) asociado con un intervalo de confianza.

Provisiones técnicas (en inglés, *technical provisions*). Aprovisionamientos que reflejan el valor cierto o estimado de las obligaciones contraídas por razón de los contratos de seguros y de reaseguros suscritos, así como el de los gastos relacionados con el cumplimiento de dichas obligaciones. Forman parte del pasivo de la entidad aseguradora.

Solvencia (en inglés, *solvency*). Proceso por el cual una entidad aseguradora no sólo demuestra su capacidad presente de respuesta a factores de riesgo, sino también la futura; y no sólo frente a los riesgos puramente derivados de su actividad (siniestros) sino frente a todos aquéllos a los que está sometida, desde la desviación de la siniestralidad hasta quebrantos derivados de una gestión deficiente.

Tail-Value-at-Risk o **TVaR** (también denominado en castellano, *valor en riesgo en la cola*). Medida de riesgo coherente que cuantifica, para un nivel de confianza $1-\alpha$, las pérdidas promedio por encima del umbral definido (normalmente establecido como el *VaR* para un determinado cuantil).

Value-at-Risk o **VaR** (también denominado en castellano, *valor en riesgo*). Medida de riesgo definida como el cuantil que recoge la pérdida máxima esperada que se puede lograr bajo un determinado nivel de confianza dentro de un horizonte temporal (normalmente un año). En la práctica, el nivel de probabilidad asumido se suele interpretar también como un indicador de frecuencia, de forma que un *VaR* al 99% se define como la pérdida que sólo se excederá una vez cada cien años.

REFERENCIAS

- Albarrán, I. y P. Alonso (2010). *Métodos estocásticos de estimación de las provisiones técnicas en el marco de Solvencia II*. Fundación MAPFRE. Madrid (España). http://www.mapfre.com/documentacion/publico/i18n/catalogo_imagenes/grupo.cmd?path=1062009.
- Alonso, P. (2007). Solvencia II: ejes del proyecto y diferencias con Basilea II. *Anales Instituto de Actuarios Españoles Tercera Época* 13, 37-56. [http://www.actuarios.org/espa/anales/2007/art 37-56.pdf](http://www.actuarios.org/espa/anales/2007/art%2037-56.pdf).
- Baier, T. y E. Neuwirth (2007). Excel :: COM :: R. *Computational Statistics* 22:1, 91-108. <http://rcom.univie.ac.at>.
- Boj, E. y T. Costa (2014). Modelo lineal generalizado y cálculo de la provisión técnica. Textos docentes. Departamento de Matemática Económica, Financiera y Actuarial. Facultad de Economía y Empresa, Universidad de Barcelona. <http://diposit.ub.edu/dspace/bitstream/2445/49068/6/glmbootsprovest.pdf>.
- Boj, E., Claramunt, M. M. y J. Fortiana (2004). *Análisis multivariante aplicado a la selección de factores de riesgo en la tarificación*. Cuadernos de la Fundación MAPFRE, núm. 88. Fundación MAPFRE Estudios, Madrid (España). http://www.mapfre.com/documentacion/publico/i18n/catalogo_imagenes/grupo.cmd?path=1050569.
- Claramunt, M. M. y T. Costa (2003). *Matemática Actuarial No Vida. Un enfoque práctico*. Colección de Publicaciones del Departamento de Matemática Económica, Financiera y Actuarial de la Universidad de Barcelona, nº 63.
- De Vylder, F.E. (1996). *Advanced Risk Theory*. Editions de l'Université de Bruxelles. Swiss Association of Actuaries.
- El Parlamento Europeo y el Consejo de la Unión Europea (2009). Directiva 2009/138/CE del Parlamento Europeo y del Consejo, de 25 de noviembre de 2009. Diario Oficial de la Unión Europea, L 335:1–155. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:335:0001:0155:es:PDF>.
- England, P. D. y R. J. Verrall (1999). Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics* 25, 281-293.
- England, P. D. y R. J. Verrall (2002). Stochastic claims reserving in general insurance (with discussion). *British Actuarial Journal* 8, 443-544.
- England, P. D. (2002). Addendum to “Analytic and bootstrap estimates of prediction errors in claim reserving. *Insurance: Mathematics and Economics* 31, 461-466.
- England, P. D. y R. J. Verrall (2006). Predictive Distributions of Outstanding Liabilities in General Insurance. *Annals of Actuarial Science* 1 (II), 221-270.

- Gesmann, M., Murphy, D. y W. Zhang (2013). Claims reserving with R: ChainLadder-0.1.7 Package Vignette DRAFT. <http://cran.r-project.org/web/packages/ChainLadder/vignettes/ChainLadder.pdf>.
- Gesmann, M., Murphy, D. y W. Zhang (2013). Package 'ChainLadder'. Reference Manual. <http://cran.r-project.org/web/packages/ChainLadder/ChainLadder.pdf>.
- Gesmann, M., Murphy, D. y W. Zhang (2013). *ChainLadder: Statistical methods for the calculation of outstanding claims reserves in general insurance*. Package on CRAN. Version 0.1.7. Publicado 28-09-2013. <http://cran.r-project.org/web/packages/ChainLadder/index.html>.
- González-Quevedo, F. (2002). Método global de cálculo de la provisión de siniestros pendientes, a partir de la utilización de la información histórica de que se dispone, completa e incompleta. *Anales Instituto de Actuarios Españoles Tercera Época* 8, 87-109. <http://www.actuarios.org/espa/anales/2002/quevedo2002.pdf>.
- Haberman, S. y A. E. Renshaw (1996). Generalized Linear Models and Actuarial Science. *Journal of the Royal Statistical Society. Series D (The Statistician)* 45:4, 407-436. <http://www.jstor.org/stable/2988543>.
- Hartl, T. (2010). Fitting a GLM to Incomplete Development Triangles. CAS E-Forum, Fall 2010. https://www.casact.org/pubs/forum/10fforum/Hartl_Fitting.pdf.
- Jessen, A. H. (2009). *Claims reserving and other topics in non-life insurance mathematics*. A Thesis in the Department of Mathematical of Sciences. University of Copenhagen. <https://www.math.ku.dk/noter/filer/phd09ahj.pdf>.
- Kaas, R., Goovaerts, M., Dhaene, J. y M. Denuit (2008). *Modern actuarial risk theory: using R. Second edition*. Springer-Verlag. Heidelberg (Alemania).
- Mack, T. (1993). Distribution-free calculation of the standard error of chain-ladder reserve estimates. *Astin Bulletin* 23:2, 213–225.
- McCullagh, P. y J. Nelder (1989). *Generalized linear models* (2nd edition). Chapman and Hall. Londres (Reino Unido).
- Ministerio de Economía y Competitividad. Reglamento de Ordenación y Supervisión de los Seguros Privados (1998). Real Decreto 2486/1998, de 20 de noviembre de 1998. Boletín Oficial del Estado, de 25 de noviembre de 1998. [http://www.dgsfp.mineco.es/sector/documentos/legislacion/2013/R.D.2486.1998 REGLAMENTO DE ORDENACION Y SUPERVISION DE LOS SEGUROS PRIVADOS.16.10.2013.pdf](http://www.dgsfp.mineco.es/sector/documentos/legislacion/2013/R.D.2486.1998%20REGLAMENTO%20DE%20ORDENACION%20Y%20SUPERVISION%20DE%20LOS%20SEGUROS%20PRIVADOS.16.10.2013.pdf).
- Moreno Gamazo, F. P. (2013). Jornada sobre las Directrices de EIOPA de preparación a Solvencia II. Dirección General de Seguros y Fondos de Pensiones. [http://www.dgsfp.mineco.es/sector/documentos/Jornada Directrices EIOPA de preparacion Solvencia II_10-12-2013/Fernando Moreno_Jornada Directrices EIOPA. DGSFP-UNESPA.pdf](http://www.dgsfp.mineco.es/sector/documentos/Jornada%20Directrices%20EIOPA%20de%20preparacion%20Solvencia%20II_10-12-2013/Fernando%20Moreno_Jornada%20Directrices%20EIOPA.DGSFP-UNESPA.pdf).

- Price Waterhouse Coopers (2007). Solvencia II. Principios incluidos en la propuesta de Directiva. Chile. <http://www.pwc.com/cl/es/publicaciones/assets/solvencia2.pdf>.
- Schmidt, K. D. y M. Zocher (2008). The Bornhuetter-Ferguson Principle. *Variance. Advancing the Science of Risk* 2:1, 85-110. <http://www.variancejournal.org/issues/02-01/85.pdf>.
- UNESPA (2007). El Modelo Español de Solvencia paso a paso. Madrid (España). http://www.unespa.es/adjuntos/fichero_2590_20080227.pdf.
- Van Eeghen, J. (1981). *Loss Reserving Methods*. Surveys of Actuarial Studies, nº 1. National Nederlanden.
- Taylor, G.C. (1986). *Claim reserving in non-life insurance*. North Holland. Insurance Series 1.
- Zhou, J. y J. Garrido (Thesis Supervisor) (2011). *Theory and applications of Generalized Linear Models in insurance*. A Thesis in the Department of Mathematics and Statistics. Concordia University. Montréal (Canadá). http://spectrum.library.concordia.ca/35732/1/Zhou_PhD_F2011.pdf.
- Zhou, J. y J. Garrido (2009). A loss reserving method based on Generalized Linear Models. Technical Report No. 2/09. Concordia University. Montréal (Canadá). http://www.mathstat.concordia.ca/resources-and-links/technical-reports/REPORTS/2_09.pdf.

ANEXO 1

CÓDIGO DE R UTILIZADO EN LA FORMULACIÓN DE LA APLICACIÓN

A continuación se muestra el código de programación utilizado en la formulación que utiliza la aplicación informática desarrollada. Cabe mencionar que para salvaguardar la autoría del proyecto, sólo se adjunta la parte correspondiente a la implementación con R, dejando sin especificar el código utilizado en VBA. Así pues, el lector puede encontrar en falta los cálculos de los valores actuales y del coeficiente de variación, los cuales se han implementado en lenguaje de programación VBA directamente, así como lo relativo a los histogramas de la función predictiva.

```
#### GLM - Fórmula: Reserva por año de origen y total
## Instrucciones para hallar mediante R el rectángulo de desarrollo
## Sea Xij.l la matriz con los elementos Xij del triángulo y 0 en el resto de posiciones

n<-nrow(Xij.l) # número de años de observación, que a su vez es el número de filas y columnas del
               # rectángulo de desarrollo

ii<-row(Xij.l) # indicador del número de cada fila de la matriz Xij.l

jj<-col(Xij.l) # indicador del número de cada columna de la matriz Xij.l

Xij.v<-as.vector(Xij.l) # vector con los elementos de la matriz Xij.l

future<-as.numeric(ii+jj-1>n) # vector en el que se ha asignado 1 a las posiciones de los pagos futuros del
                              # triángulo y 0 a las posiciones de los pagos del triángulo original

ii<-as.factor(ii) # conversión de los elementos de la matriz ii en factor

jj<-as.factor(jj) # conversión de los elementos de la matriz jj en factor

Full.CL<-glm(Xij.v~ii+jj,family=quasipoisson,wei=1-future) #aplicación del GLM

Sig<-vcov(Full.CL) # cálculo de la matriz de varianzas y covarianzas de los coeficientes del modelo

X<-model.matrix(Full.CL) # matriz de diseño

Cov.eta<-X%*%Sig%*%t(X) # cálculo de la matriz de varianzas y covarianzas del predictor lineal

mu.hat<-fitted(Full.CL)*future # obtención de los valores estimados de los pagos futuros

pagos.futuros<-(matrix(mu.hat,nrow=n,ncol=n))

rect.desarrollo<-Xij.l+pagos.futuros

## Obtención IBNR total

reserva.total<-sum(mu.hat)

## Obtención IBNR por año de origen

reserva.origen<-numeric(n-1)

for(i in 1:(n-1)){
```

```
reserva.origen[i]<-sum(pagos.futuros[i,])
```

```
}
```

```
## Error cuadrático medio de la predicción de las reservas totales
```

```
phi.P<-summary(Full.CL)$dispersion
```

```
pe2<-phi.P*sum(mu.hat)+t(mu.hat)%*%Cov.eta%*%mu.hat # Cálculo del error cuadrático medio para la  
# provisión total.
```

```
pe<-sqrt(pe2) # Cálculo del error de predicción
```

```
## Error de predicción de las reservas por año de origen
```

```
mu.hat.m<-matrix(mu.hat,n,n)
```

```
mu.hat<-matrix(0,n,n)
```

```
mu.hat.vec<-numeric(n*n)
```

```
pe2orig<-numeric(n-1)
```

```
reserva.orig<-numeric(n-1)
```

```
for(i in 1:(n-1)){
```

```
mu.hat<-rbind(matrix(0,i,n),mu.hat.m[i+1,],matrix(0,n-(i+1),n))
```

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

```
reserva.orig[i]<-sum(mu.hat.vec)
```

```
pe2orig[i]<-phi.P*sum(mu.hat.vec)+t(mu.hat.vec)%*%Cov.eta%*%mu.hat.vec
```

```
}
```

```
reserva.origen<-matrix(reserva.orig,n-1,1)
```

```
pe2.origen<-matrix(sqrt(pe2orig),n-1,1)
```

```
#### GLM - Fórmula: Reserva por año de calendario
```

```
## Obtención IBNR por año de calendario
```

```
mu.hat.m<-matrix(mu.hat,n,n)
```

```
mu.hat.d<-matrix(0,n,n)
```

```
provdia<-numeric(n-1)
```

```
for (i in 2:n){
```

```
mu.hat.d[row(mu.hat.m)+col(mu.hat.m)==n+i]<-mu.hat.m[row(mu.hat.m)+col(mu.hat.m)==n+i]
```

```
provdia[i-1]<-sum(mu.hat.d);mu.hat.d<-matrix(0,n,n)
```



```
}
prov.año.calendario<-matrix(provdiag,n-1,1)

## Error de predicción de las reservas por año de calendario

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

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

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

pe2diag<-numeric(n-1)

for (i in 2:n){

mu.hat.d[row(mu.hat.m)+col(mu.hat.m)==n+i]<-mu.hat.m[row(mu.hat.m)+col(mu.hat.m)==n+i]

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

pe2diag[i-1]<-phi.P*sum(mu.hat.vec)+t(mu.hat.vec)%*%Cov.eta%*%mu.hat.vec

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

}

pe2.calendario<-matrix(sqrt(pe2diag),n-1,1)

#### GLM - Bootstrap: Reserva por año de origen y total
## Instrucciones para hallar mediante R el rectángulo de desarrollo
## Sea Xij.l la matriz con los elementos Xij del triángulo y NA en el resto de posiciones

Xij.v<-as.vector(t(Xij.l))

cij<-na.omit(Xij.v)

## Cálculo con GLM para la Reserva Total

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,quasipoisson)

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])

pagos.futuros<-future*orig.fits[future]

past<-row(orig.fits)+col(orig.fits)-2<k

Xij.v[is.na(Xij.v)]<-0

triang.sup<-Xij.v*past

rect.desarrollo<-triang.sup+pagos.futuros

## Cálculo residuos de Pearson

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

## Cálculo n, p y 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))

## Bootstrap para Reserva Total

nBoot<-as.numeric(mBoot) # mBoot = núm iteraciones introducido por el usuario

payments<-reserves<-numeric(nBoot)

for(boots in 1:nBoot){

  Ps.cij<-sample(Adj.Pr.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,quasipoisson)

  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
```

```
}
```

```
## Error predicción bootstrap para reserva total
```

```
PEbs<-sqrt(phi.P*orig.reserve+sd(reserves)^2)
```

```
## Cálculos con GLM para la Reserva por Año de Origen
```

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

```
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])
```

```
}
```

```
## Bootstrap para la obtención de las Reservas por Año de Origen
```

```
nBoot<-as.numeric(mBoot) # mBoot = núm iteraciones introducido por el usuario
```

```
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,quasipoisson)
```

```
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  
}  
  
## Error de predicción bootstrap por año de origen  
PEbsorig <- numeric(k-1)  
for(orig in 1:k-1){  
  PEbsorig[orig] <- sqrt(phi.P*provor[orig]+sd(reservesorig[,orig])^2)  
}  
PE.orig <- matrix(PEbsorig,k-1,1)  
  
#### GLM - Bootstrap: Reserva por año de calendario  
## Cálculos con GLM para la Reserva por Año de Calendario  
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)  
}  
prov.calendario <- matrix(cal.prov,k-1,1)  
  
## Bootstrap para Reservas Año de Calendario  
nBoot <- as.numeric(mBoot) # mBoot = núm iteraciones introducido por el usuario  
nBoot <- as.numeric(mBoot)  
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,quasipoisson)
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
}
```

Error predicción bootstrap para Reserva por Año de Calendario

```
PEbscal<-numeric(k-1)
for(cal in 1:k-1){
  PEbscal[cal]<-sqrt(phi.P*cal.provc[cal]+sd(reservescal[,cal])^2)
}
PE.calendario<-matrix(PEbscal,k-1,1)
```

Distribución predictiva

```
confianza2<-as.numeric(confianza) # confianza = valor introducido por el usuario de 0-100
alfa<-confianza2/100
for(cal in 1:k-1){
  quantile(payments[,cal], alfa)
}
VaR<-matrix(payments,k-1,1)
```

ANEXO 2

HISTOGRAMAS. DISTRIBUCIÓN PREDICTIVA BOOTSTRAP

A continuación, se muestran los histogramas relativos a la distribución predictiva de las reservas por años de calendario y total, junto con algunos estadísticos de dichas distribuciones que la aplicación desarrollada genera por *bootstrap*, para el caso ilustrado de 1000 iteraciones y $\alpha = 75\%$.

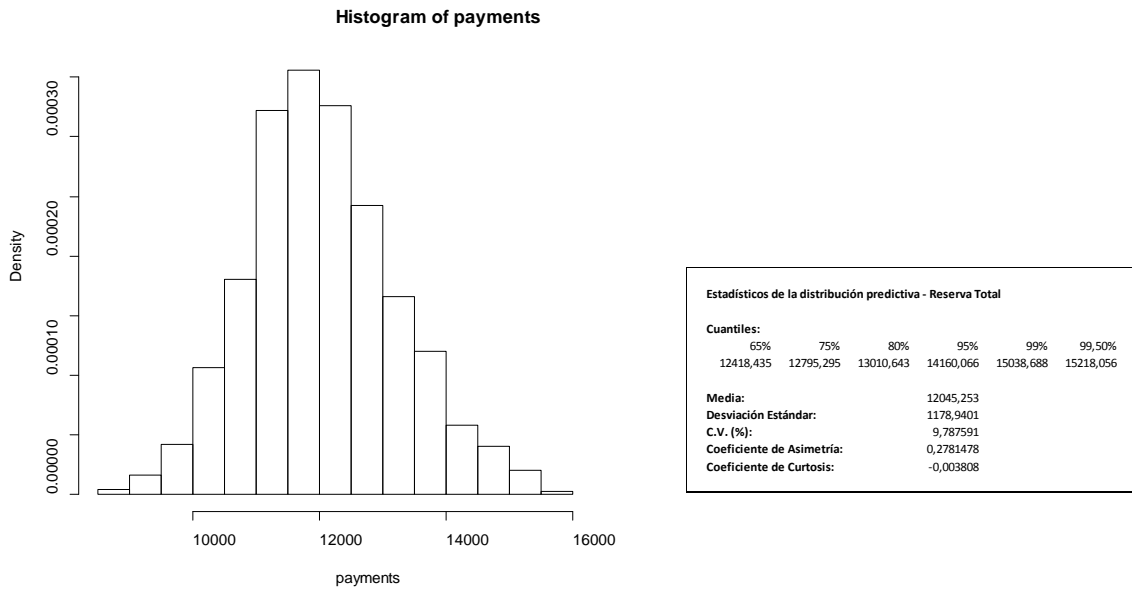


Gráfico A.1. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva total. [Fuente: Elaboración propia]

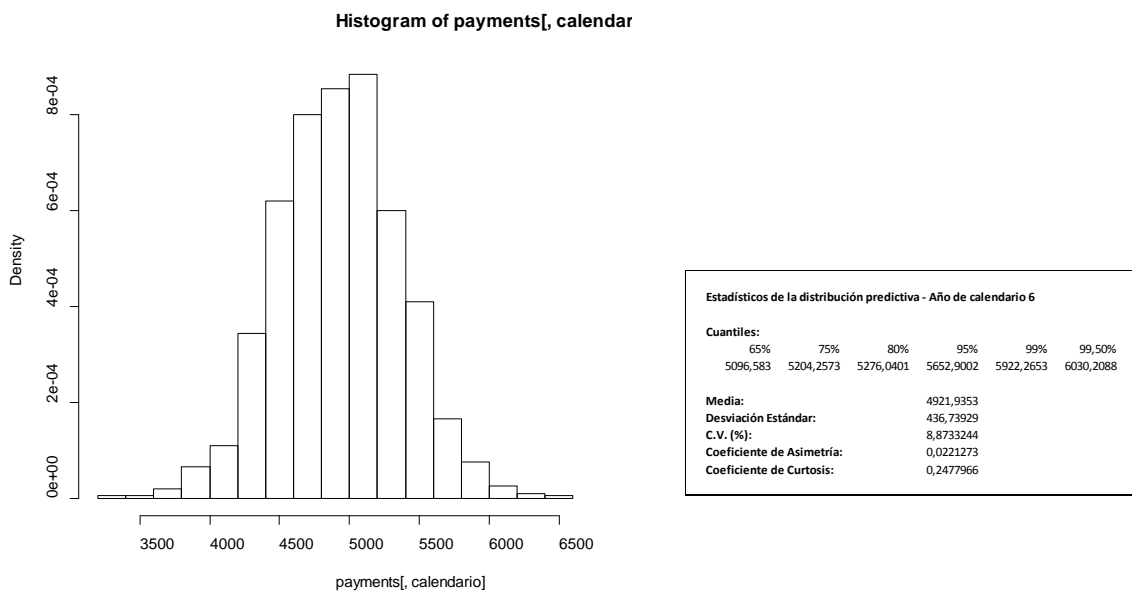


Gráfico A.2. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 6. [Fuente: Elaboración propia]

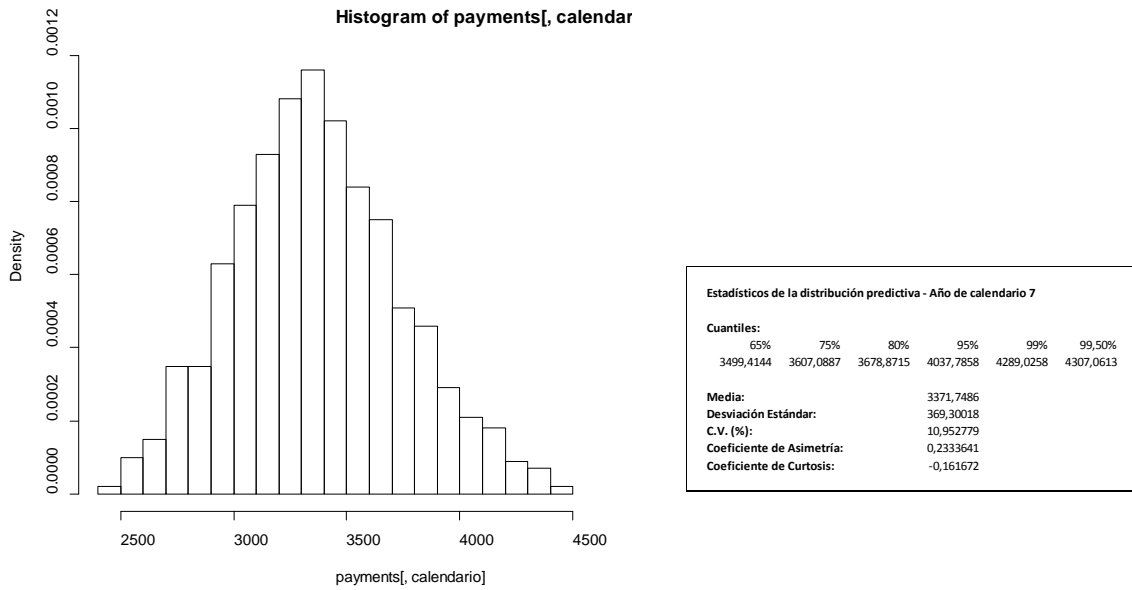


Gráfico A.3. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 7. [Fuente: Elaboración propia]

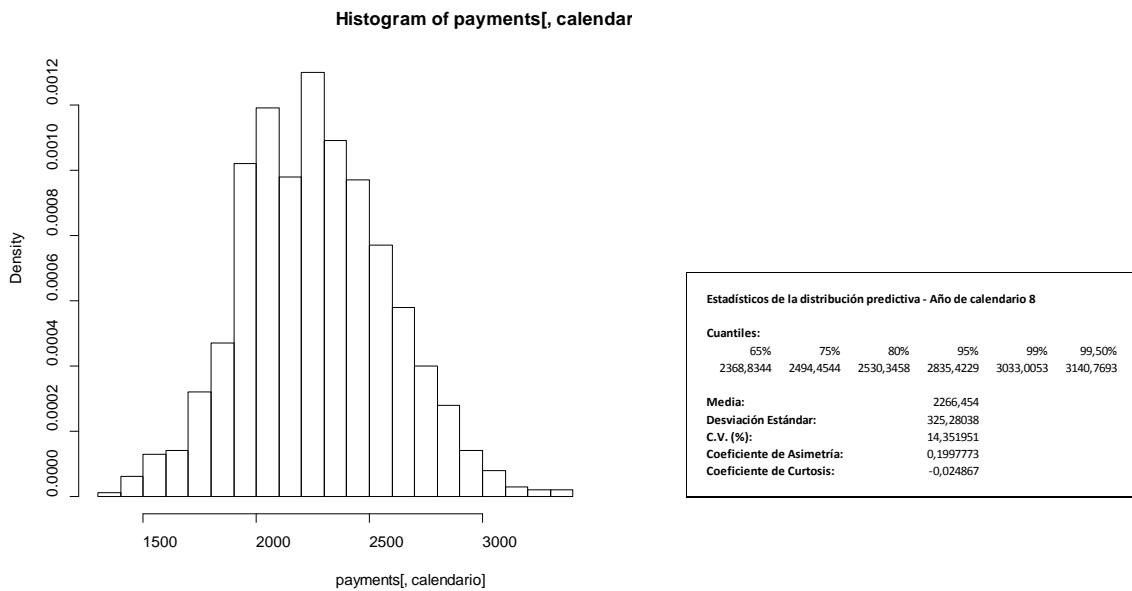


Gráfico A.4. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 8. [Fuente: Elaboración propia]

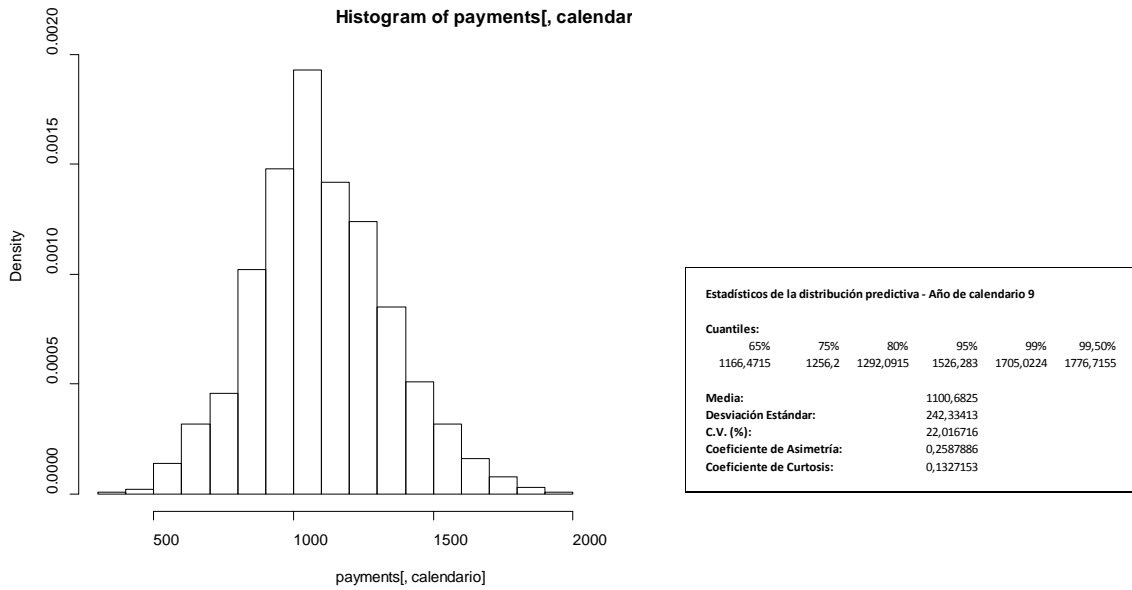


Gráfico A.5. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 9. [Fuente: Elaboración propia]

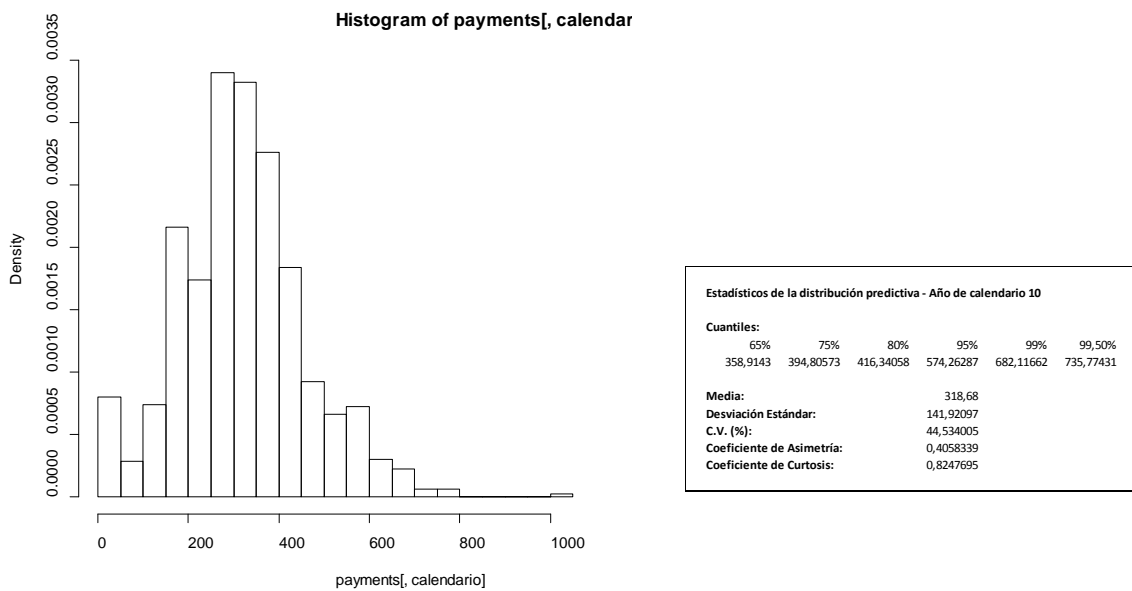


Gráfico A.6. Resultados de la aplicación sobre la descripción estadística de la distribución predictiva para la reserva del año de calendario 10. [Fuente: Elaboración propia]