



Cláudia Sofia Ribeiro da Conceição

Licenciada em Matemática Aplicada à Economia e à Gestão

Modelos Determinísticos e Estocásticos Aplicados ao Cálculo de Provisões para Sinistros

Dissertação para obtenção do Grau de
Mestre em Matemática e Aplicações

Orientador: Rui M. R. Cardoso,
Prof. Auxiliar, Universidade Nova de Lisboa

Júri:

Presidente: Manuel Leote Tavares Inglês Esquível

Arguente: Gracinda Rita Diogo Guerreiro

Vogal: Rui Manuel Rodrigues Cardoso



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

SETEMBRO, 2014

Modelos Determinísticos e Estocásticos Aplicados ao Cálculo de Provisões para Sinistros

Copyright © Cláudia Sofia Ribeiro da Conceição, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

*"O único sítio onde o sucesso vem antes do trabalho é no
dicionário, por força da ordem alfabética!"
Albert Einstein (1879-1955)*

Agradecimentos

Ao meu orientador, Prof. Dr. Rui Manuel Rodrigues Cardoso, por todo o apoio, paciência, conhecimento partilhado e por toda a disponibilidade ao longo destes 2 anos.

À minha supervisora interna da empresa Actuariado, Dra. Carmen Pereira Oliveira, pela oportunidade, pelo apoio, por todos os conselhos e pelo tempo que disponibilizou para que finalizasse esta etapa.

Ao Prof. Dr. Manuel Esquível pelo apoio e pela orientação no início deste processo.

A todos os meus colegas de mestrado que partilharam todos os altos e baixos ao longo destes anos, bem como aos meus amigos que me apoiaram e que me acompanharam.

Aos meus pais, às minhas irmãs, às minhas sobrinhas e ao meu namorado por toda a paciência, carinho, apoio, dedicação e acima de tudo por acreditarem sempre que seria capaz, mesmo quando eu mesma não o sabia.

Resumo

A estimação das provisões para sinistros é de extrema importância na atividade de uma seguradora do ramo Não Vida, porque, por um lado, são necessários recursos suficientes para garantir o pagamento de qualquer sinistro que ocorra e, por outro, o excesso de provisões pode condicionar a rentabilidade da companhia. Assim, faz parte do papel do atuário garantir a adequação destas provisões técnicas.

Com esse propósito, é habitualmente utilizado o método determinístico *Chain Ladder* para a estimação das provisões para sinistros. Contudo, de forma a contornar as limitações desta metodologia, recorre-se a modelos estocásticos, como o modelo proposto por *Thomas Mack*, uma aplicação dos Modelos Lineares Generalizados, a técnica *Bootstrap*, entre outros. Dos modelos estocásticos é de salientar um dos mais recentemente estudados, *Double Chain Ladder*, que está diretamente associado ao método *Chain Ladder*, trazendo algumas vantagens sobre esse último e possibilitando também a aplicação da técnica *Bootstrap* e uma simples associação ao método *Bornhuetter-Ferguson*.

O objetivo desta dissertação consiste na aplicação e análise destes modelos na estimação de provisões para sinistros no ramo Não Vida.

Palavras-chave: Provisão de Sinistros, Modelos Determinísticos, Modelos Estocásticos, *Chain Ladder*, *Link ratios*, *Grossing Up*, *Bornhuetter-Ferguson*, Teoria da Credibilidade, *Thomas Mack*, Modelos Lineares Generalizados, Técnica *Bootstrap*, *Double Chain Ladder*

Abstract

The estimation of claims reserving is very important in the activity of a non-life insurance company because sufficient resources are required to guarantee payment of any claim occurring and also, the excess reserves can affect the company's profitability. One function the actuary is to ensure the adequacy of the technical provisions of insurance.

For this purpose, it is commonly used deterministic *Chain Ladder* method for the estimation of claims provisions. However, in order to bypass the limitations of this methodology, one resorts to stochastic models, such as proposed by *Thomas Mack*, an application of Generalized Linear Models, the *Bootstrap* technique, among others. One of the most recent stochastic methods is *Double Chain Ladder*, which is directly associated with *Chain Ladder* method, bringing some advantages over the last and also enabling the application of the *Bootstrap* technique and a simple association with the *Bornhuetter-Ferguson* method.

The objective of this dissertation is the application and analysis of these models in the estimation of reserves for outstanding claims in non-life business.

Keywords: Claims reserving, Deterministic Models, Stochastic Models, Chain Ladder, Link Ratios, Grossing Up, Bornhuetter-Ferguson, Credibility Theory, Thomas Mack, Generalized Linear Models, Bootstrap Method, Double Chain Ladder

Conteúdo

1	Introdução	1
2	Conceitos Preliminares	3
2.1	Processo de Sinistro	3
2.2	Nomenclatura	4
3	Modelos Determinísticos	9
3.1	Método <i>Chain Ladder</i>	9
3.1.1	Metodologia	9
3.1.2	Vantagens e Limitações	10
3.1.3	Outros métodos baseados em <i>Link Ratios</i>	11
3.2	Método <i>Grossing Up</i>	11
3.3	Método <i>Bornhuetter-Ferguson</i>	12
3.3.1	Metodologia	13
3.4	Fator Cauda	14
4	Teoria da Credibilidade e Provisionamento	15
4.1	Generalidades	15
4.2	Metodologia	16
5	Modelo <i>Thomas Mack</i>	19
5.1	Pressupostos	19
5.1.1	1º Pressuposto - Proporcionalidade	19
5.1.2	2º Pressuposto - Independência	21
5.1.3	3º Pressuposto - Variabilidade mínima	22
5.2	Medidas de Variabilidade	22
6	Modelos Lineares Generalizados	25
6.1	Generalidades	25
6.1.1	Componentes	25

6.1.2	Família Exponencial	26
6.2	MLG na estimação da Provisão para Sinistros	27
6.2.1	Pressupostos	27
6.2.2	Modelo de <i>Sobre-dispersão de Poisson</i>	28
7	Simulação <i>Bootstrap</i>	31
7.1	Metodologia	31
7.2	Medidas de Variabilidade	32
8	Método <i>Double Chain Ladder</i>	35
8.1	Pressupostos	35
8.2	Estimação dos Parâmetros	37
8.3	Estimação das Reservas	38
8.4	Fator Cauda	40
8.5	<i>Bootstrap</i>	40
8.6	<i>Bornhuetter-Ferguson and Double Chain Ladder</i>	42
8.7	Comentários	43
9	Aplicação Prática	45
9.1	Dados Utilizados	45
9.2	<i>Chain Ladder</i>	46
9.3	Outros <i>Link Ratios</i>	47
9.4	Grossing Up	48
9.5	<i>Bornhuetter-Ferguson</i>	49
9.6	Fator Cauda	50
9.7	Teoria da Credibilidade	53
9.8	<i>Thomas Mack</i>	54
9.8.1	Validação dos Pressupostos	54
9.8.2	Medidas de Variabilidade	57
9.9	Aplicação dos Modelos Lineares Generalizados	58
9.10	Simulação <i>Bootstrap</i>	60
9.11	<i>Double Chain Ladder</i>	62
9.11.1	Fator Cauda	64
9.11.2	<i>Bootstrap</i>	64
9.11.3	<i>Bornhuetter-Ferguson and Double Chain Ladder</i>	66
9.12	Comentários sobre os Resultados	67
10	Conclusão	71
A	Resultados do R	75
A.1	Dados Utilizados	75
A.2	<i>Chain Ladder</i> e <i>Thomas Mack</i>	76

A.3	<i>Bootstrap</i>	77
A.4	<i>Double Chain Ladder</i>	77
A.4.1	<i>Fator Cauda</i>	80
A.4.2	<i>Bootstrap</i>	80
A.4.3	<i>Bornhuetter-Ferguson and Double Chain Ladder</i>	83



Lista de Figuras

2.1	Evolução de um processo de sinistro. Fonte: Borginho (2001)	3
8.1	Conjunto de índices dos sinistros agregados, assumindo um atraso máximo de $d = n - 1$	39
9.1	Projeção dos Coeficientes de desenvolvimento no método <i>Chain Ladder</i> . .	51
9.2	Projeção das taxas de sinistralidade no método <i>Bornhuetter-Ferguson</i>	52
9.3	Ajustamento dos dados aos fatores de desenvolvimento	54
9.4	Resíduos ponderados	57
9.5	Gráfico dos Resíduos de Pearson corrigidos	59
9.6	Histograma das reservas totais aplicando a técnica <i>Bootstrap</i>	62

Lista de Tabelas

2.1	Matriz de informação conhecida	5
2.2	Matriz de informação conhecida - Maior histórico	5
2.3	Matriz de informação conhecida - Histórico reduzido	5
4.1	Proposta para Z_i	17
6.1	Funções de ligação	26
9.1	Montantes Incrementais Pagos (em u.m.)	45
9.2	Montantes Acumulados Pagos (em u.m.)	46
9.3	Número de Sinistros	46
9.4	Fatores de desenvolvimento individuais	46
9.5	Aplicação do método <i>Chain Ladder</i>	47
9.6	Coefficientes de desenvolvimento de outros <i>Link Ratios</i>	47
9.7	Aplicação de outros <i>Link Ratios</i>	48
9.8	Percentagens para aplicação de <i>Grossing Up</i>	48
9.9	Aplicação de <i>Grossing Up</i>	49
9.10	Prémios Emitidos	49
9.11	Taxas de sinistralidade no método <i>Bornhuetter-Ferguson</i>	49
9.12	Aplicação do método <i>Bornhuetter-Ferguson</i>	50
9.13	Coefficientes de desenvolvimento projetados no método <i>Chain Ladder</i>	51
9.14	Resultados com os coeficientes projetados no método <i>Chain Ladder</i>	51
9.15	Taxas de sinistralidade projetadas no método <i>Bornhuetter-Ferguson</i>	52
9.16	Resultados com as taxas projetados no método <i>Bornhuetter-Ferguson</i>	52
9.17	Taxas de sinistralidade	53
9.18	Fator de Credibilidade	53
9.19	Resultados da aplicação de Teoria da Credibilidade	53
9.20	Matriz R	55
9.21	Matriz S	55

9.22	Coeficientes de correlação de <i>Spearman</i>	55
9.23	Conjustos <i>S</i> e <i>L</i>	56
9.24	Momentos das variáveis Z_j e Z	56
9.25	Estimativas para a variância e erro padrão	57
9.26	Estimativas para o intervalo de confiança	58
9.27	Montantes Incrementais Pagos Ajustados (em u.m.)	58
9.28	Resíduos de <i>Pearson</i>	58
9.29	Resíduos de <i>Pearson</i> corrigidos	59
9.30	Estimativas dos parâmetros para o MLG - <i>Sobre-dispersão de Poisson</i>	60
9.31	Estimativas das reservas pelo MLG - <i>Sobre-dispersão de Poisson</i>	60
9.32	Resíduos de <i>Pearson</i> corrigidos simulados (exemplo)	61
9.33	Triângulo de pseudo-dados (exemplo)	61
9.34	Estimativas da provisão (exemplo)	61
9.35	Resultados da aplicação da técnica <i>bootstrap</i>	62
9.36	Triângulo incremental do número de sinistros ajustado	63
9.37	Estimativas dos parâmetros do método <i>Double Chain Ladder</i>	63
9.38	Estimativas dos parâmetros do método <i>Double Chain Ladder</i>	63
9.39	Resultados pelo método <i>Double Chain Ladder</i> sem fator cauda	64
9.40	Resultados pelo método <i>Double Chain Ladder</i> com fator cauda	64
9.41	Resultados pelo método <i>Bootstrap - Double Chain Ladder</i> sem fator cauda	65
9.42	Resultados pelo método <i>Bootstrap - Double Chain Ladder</i> com fator cauda	65
9.43	Custos totais incrementais(em u.m.)	66
9.44	Parâmetro γ' do método <i>Bornhuetter-Ferguson and Double Chain Ladder</i>	66
9.45	Resultados pelo método <i>Bornhuetter-Ferguson and Double Chain Ladder</i> sem fator cauda	66
9.46	Resultados pelo método <i>Bornhuetter-Ferguson and Double Chain Ladder</i> com fator cauda	67
9.47	Resultados sem fator cauda	69
9.48	Resultados com fator cauda	69

Listagens

A.1	Estatísticas dos montantes acumulados	75
A.2	Resultados da aplicação do modelo proposto por Mack	76
A.3	Resultados da aplicação da técnica <i>Bootstrap</i>	77
A.4	Parâmetros dos montantes pagos aplicando DCL	77
A.5	Parâmetros do número de sinistros aplicando DCL	78
A.6	Parâmetros do modelo DCL	79
A.7	Resultados do modelo DCL sem fator cauda	79
A.8	Resultados do modelo DCL com fator cauda	80
A.9	Resultados do modelo <i>Bootsatrap-DCL</i> sem fator cauda	80
A.10	Resultados do modelo <i>Bootsatrap-DCL</i> com fator cauda	82
A.11	Parâmetros do modelo <i>BF-DCL</i>	83
A.12	Resultados do modelo <i>BF-DCL</i> sem fator cauda	83
A.13	Resultados do modelo <i>BF-DCL</i> com fator cauda	84

Siglas e Acrónimos

BDCL	<i>Bornhuetter-Ferguson and Double Chain Ladder</i>
DCL	<i>Double Chain Ladder</i>
EQM	Erro Quadrático Médio
IBNER	<i>Incurred But Not Enough Reported</i>
IBNR	<i>Incurred But Not Reported</i>
IC	Intervalo de Confiança
ISP	Instituto de Seguros de Portugal
MLG	Modelo Linear Generalizado
RBNS	<i>Reported But Not Settled</i>
u.m.	unidade monetária

Símbolos

A	Número de sinistros declarados acumulados
α	Nível de significância
C	Montantes pagos incrementais
D	Montantes pagos acumulados
f	Coefficiente de desenvolvimento
N	Número de sinistros declarados incremental
pe	Prémios emitidos
r	Resíduos
θ	Taxa de sinistralidade
X	Histórico de dados fornecido pela empresa
Y	Pagamentos individuais
Z	Fator de credibilidade





Introdução

Segundo o Instituto de Seguros de Portugal (ISP), em 2012, as provisões técnicas do conjunto das empresas de seguros chegaram aos 21.529,02 milhões de euros. Este montante deve, em qualquer momento, ser suficiente para permitir às seguradoras cumprirem as suas responsabilidades com riscos estabelecidos nos contratos de seguro.

As provisões técnicas, que devem ser asseguradas pela empresa de seguros, incluem a provisão para prémios não adquiridos, a provisão para riscos em curso, a provisão matemática do ramo Vida, a provisão para envelhecimento, a provisão para sinistros, a provisão para participação nos resultados e a provisão para desvios de sinistralidade.

Sabe-se também que ascenderam a 12.200,16 milhões de euros os custos brutos com sinistros em 2012, representando 114,85% dos prémios brutos emitidos. Assim sendo, uma das provisões mais relevantes para uma seguradora é a provisão para sinistros.

Segundo o artigo 73º do Decreto-Lei n.º 94-B/98, de 17 de Abril:

"A provisão para sinistros corresponde ao custo total estimado que a empresa de seguros suportará para regularizar todos os sinistros que tenham ocorrido até ao final do exercício, quer tenham sido comunicados ou não, após dedução dos montantes já pagos respeitantes a esses sinistros."

A estimação desta provisão é efetuada através de dois tipos de métodos, os métodos determinísticos e os estocásticos. Os principais métodos determinísticos utilizados na prática atuarial são os métodos que se baseiam em *Link Ratio*, categoria na qual se insere o método *Chain Ladder*, os *Grossing Up* e o método *Bornheutter-Ferguson*. As estimativas obtidas através destes métodos são pontuais e muito sensíveis a pequenas alterações nos dados. Com o intuito de obter medidas de erro e intervalos de confiança para essas estimativas alguns autores propõem que se utilizem modelos estocásticos. Deste tipo de modelos são de salientar o modelo proposto por Mack (1999), uma aplicação dos modelos lineares generalizados a técnicas de provisão, a técnica *Bootstrap* e o mais recente método

Double Chain Ladder.

A presente dissertação é constituída por oito capítulos com o objetivo de aplicar alguns modelos determinísticos e estocásticos à estimação das provisões para sinistros.

O primeiro capítulo faz referência a conceitos preliminares acerca de provisões para sinistros, bem como ao desenvolvimento do processo de um sinistro.

No segundo capítulo introduz-se alguns modelos determinísticos, tais como o método *Chain Ladder* e outros métodos baseados em *link ratios*, o método *Grossing Up* e o método *Bornheutter-Ferguson*, onde também se explora a possibilidade de integração do fator cauda nestes modelos.

O terceiro capítulo refere-se a uma aplicação de Teoria da Credibilidade, iniciando-se com uma breve introdução desta teoria e partindo para a aplicação na estimação da provisão para sinistros.

O quarto capítulo estuda o modelo proposto por Mack (1999), que possibilita a obtenção de estatísticas de interesse a partir das estimativas obtidas no método *Chain Ladder*, começando por referir os pressupostos necessários para a sua aplicação.

No quinto capítulo faz-se uma breve introdução aos modelos lineares generalizados e, posteriormente, estuda-se a sua aplicação na estimação da provisão para sinistros.

No sexto capítulo aplica-se a técnica de *Bootstrap*, obtendo estimativas de erros de previsão e intervalos de confiança para as provisões, referindo também as suas limitações.

No sétimo capítulo introduz-se o método *Double Chain Ladder* que, tal como o nome indica, tem por base o modelo *Chain Ladder*. Refere-se também as suas vantagens e o que inova em relação aos outros métodos. De seguida considera-se a possibilidade de integrar o fator cauda, bem como a aplicação da técnica de *Bootstrap* a este modelo e a mais recente versão de *Bornhuetter-Ferguson and Double Chain Ladder*.

No último capítulo é feita uma aplicação prática da teoria exposta anteriormente, permitindo uma comparação entre os diferentes métodos utilizados.

Em anexo é apresentado o *Software R*, onde se utilizam algumas das suas funções como suporte aos valores obtidos na aplicação prática.



Conceitos Preliminares

O contrato de seguro é um compromisso entre a seguradora e o tomador de seguro (pessoa ou entidade que celebra o seguro), no qual este último paga um prêmio para que a seguradora assuma a cobertura de determinados riscos explicitados nesse contrato.

De forma a garantir o pagamento do capital seguro sempre que ocorra um sinistro nos termos do contrato, a seguradora constitui uma provisão técnica, da qual a que tem mais peso é a provisão para sinistros. É uma das funções do Atuário proceder à estimação destas provisões.

Ao longo deste capítulo serão explicitados alguns conceitos e nomenclaturas inerentes a este estudo, bem como o desenvolvimento de um processo de sinistro.

2.1 Processo de Sinistro

O desenvolvimento do processo de um sinistro é ilustrado através da Figura 2.1,

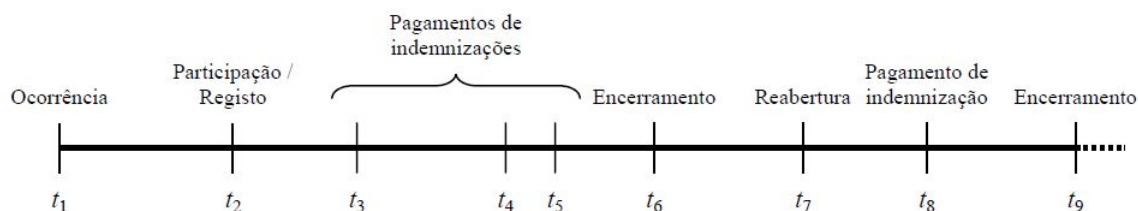


Figura 2.1: Evolução de um processo de sinistro. Fonte: Borginho (2001)

onde,

- t_1 : Momento no qual ocorre o sinistro que trará responsabilidade para a seguradora;
- t_2 : O tomador de seguro informa a seguradora do sucedido;
- t_3, t_4, t_5 : Instantes de pagamentos de indenizações por parte da seguradora;
- t_6 : Encerramento do sinistro;
- t_7 : Reabertura do processo, que poderá surgir devido a novas informações relevantes relacionadas com o sinistro;
- t_8 : Pagamentos que regularizarão a situação;
- t_9 : Novo encerramento do processo.

De forma a assegurar todas estas etapas, podemos classificar as provisões necessárias em dois tipos, as provisões para sinistros ocorridos mas não declarados, usualmente designados por *IBNR* (Incurred But Not Reported), e as provisões para sinistros ocorridos e já declarados mas com indenizações ainda por regularizar, ou seja, para sinistros declarados e que ainda não foram encerrados, conhecidos por *RBNS* (Reported But Not Settled) ou por *IBNER* (Incurred But Not Enough Reported).

Note-se que, normalmente, estas duas provisões não se distinguem, sendo o cálculo efetuado de forma agregada. No entanto, no Capítulo 8 é apresentada uma metodologia que estima estas duas provisões individualmente.

2.2 Nomenclatura

O momento em que ocorre um sinistro, designado de período de ocorrência, poderá não coincidir com o momento em que se desenvolve o processo usual de regularização de um sinistro, tal como ilustrado na Figura 2.1. O tempo decorrido entre estes dois momentos é denominado de período de desenvolvimento. Assim, por norma, a informação relativa aos sinistros é disposta numa matriz cujas linhas correspondem aos períodos de ocorrência e as colunas aos períodos de desenvolvimento dos sinistros. Esta matriz é dividida em dois triângulos, o triângulo superior esquerdo, que contém a informação já conhecida e o triângulo inferior direito, que corresponde à provisão que se pretende estimar. O primeiro triângulo pode estar totalmente preenchido, ou seja, tem-se informação sobre o encerramento de sinistros para todos os anos de ocorrência considerados, como ilustrado na Tabela 2.1, onde $X_{i,j}$, com $0 \leq i \leq n$ e $0 \leq j \leq n - i$, representa o histórico fornecido pela seguradora, relativo ao período de ocorrência i e período de desenvolvimento j . Os dados podem estar organizados por anos, semestres, trimestres, etc.

Tabela 2.1: Matriz de informação conhecida

Período de ocorrência	Período de desenvolvimento							
	0	1	2	...	k	...	$n-1$	n
0	$X_{0,0}$	$X_{0,1}$	$X_{0,2}$...	$X_{0,k}$...	$X_{0,n-1}$	$X_{0,n}$
1	$X_{1,0}$	$X_{1,1}$	$X_{1,2}$...	$X_{1,k}$...	$X_{1,n-1}$	
2	$X_{2,0}$	$X_{2,1}$	$X_{2,2}$...	$X_{2,k}$...		
...			
i	$X_{i,0}$	$X_{i,1}$	$X_{i,2}$...	$X_{i,k}$			
...				
$n-1$	$X_{n-1,0}$	$X_{n-1,1}$						
n	$X_{n,0}$							

Note-se que, no entanto, nem sempre o número de anos de desenvolvimento é igual ao número de anos de ocorrência, e a informação conhecida pode tomar outros formatos que não um triângulo superior esquerdo. Existem casos em que há mais informação conhecida, como se pode observar na Tabela 2.2, e casos em que o histórico é mais reduzido, como na Tabela 2.3.

Tabela 2.2: Matriz de informação conhecida - Maior histórico

Período de ocorrência	Período de desenvolvimento							
	0	1	2	...	k	...	$n-1$	n
0	$X_{0,0}$	$X_{0,1}$	$X_{0,2}$...	$X_{0,k}$...	$X_{0,n-1}$	$X_{0,n}$
1	$X_{1,0}$	$X_{1,1}$	$X_{1,2}$...	$X_{1,k}$...	$X_{1,n-1}$	$X_{1,n}$
2	$X_{2,0}$	$X_{2,1}$	$X_{2,2}$...	$X_{2,k}$...	$X_{2,n-1}$	$X_{2,n}$
...			
i	$X_{i,0}$	$X_{i,1}$	$X_{i,2}$...	$X_{i,k}$			
...				
$m-1$	$X_{m-1,0}$	$X_{m-1,1}$						
m	$X_{m,0}$							

Tabela 2.3: Matriz de informação conhecida - Histórico reduzido

Período de ocorrência	Período de desenvolvimento						
	0	1	2	...	$k-2$	$k-1$	k
0							$X_{0,k}$
1						$X_{1,k-1}$	$X_{1,k}$
2					$X_{2,k-2}$	$X_{2,k-1}$	$X_{2,k}$
...		
i	$X_{i,0}$	$X_{i,1}$	$X_{i,2}$...	$X_{i,k-2}$		
...			
$n-1$	$X_{n-1,0}$	$X_{n-1,1}$					
n	$X_{n,0}$						

É necessário ter-se em conta que a utilização de informação histórica em pouca ou em demasiada quantidade poderá comprometer a viabilidade dos resultados obtidos, visto

que, pouca informação poderá comprometer a significância dos métodos utilizados, enquanto que o excesso de informação poderá introduzir eventuais tendências passadas que não explicariam a realidade atual. Por exemplo, se considerarmos um histórico com um elevado número de anos de ocorrência em que, em anos mais remotos, a regularização dos sinistros era muito demorada comparada com a atual situação, esta tendência agravará a estimativa de anos de ocorrência mais recentes.

Normalmente, obtêm-se resultados mais fiáveis quando se trata de um conjunto de dados com características semelhantes, permitindo obter estimativas futuras mais consistentes. Assim, torna-se essencial uma análise *a priori* dos dados, contornando estes aspetos.

Considere-se a seguinte nomenclatura que se pode adotar para a variável $X_{i,j}$:

- $N_{i,j}$, quando os dados se referem ao número de sinistros declarados,
- $C_{i,j}$, quando os dados se referem a montantes pagos.

Esta designação diz respeito a informação disponível que se apresenta sob a forma de dados incrementais. Tem-se sob a forma de dados cumulativos:

$$\begin{aligned} A_{i,j} &= \sum_{k=0}^j N_{i,k} \\ D_{i,j} &= \sum_{k=0}^j C_{i,k} \end{aligned} \quad (2.1)$$

Tal como já referido anteriormente, o objetivo fundamental é inferir sobre os montantes a pagar/número de sinistros que ocorrerão no futuro, ou seja, estimar o preenchimento do triângulo inferior direito.

Depois de completa a matriz, obtêm-se um valor estimado para as responsabilidades atuais, associadas ao ano de ocorrência i , denotando-se por \hat{R}_i , tal que,

$$\hat{R}_i = \hat{D}_{i,n} - D_{i,n-i}, \quad 0 \leq i \leq n \quad (2.2)$$

onde $\hat{D}_{i,n}$ é a estimativa de $D_{i,n}$. Portanto, tem-se como valor estimado da provisão total para sinistros \hat{R} , onde,

$$\hat{R} = \sum_{i=0}^n \hat{R}_i$$

Note-se que, em ramos de grande dimensão e muito desenvolvidos, não se deve considerar que o triângulo está encerrado, isto é, que os sinistros serão todos declarados e encerrados ao fim de n anos de desenvolvimento, portanto, surge a necessidade de integrar na estimação um fator cauda que representa os montantes que se estimam pagar

depois do último ano de desenvolvimento conhecido. Nestes casos, o valor da estimativa da reserva no ano de ocorrência i toma a seguinte forma:

$$\hat{R}_i = \hat{D}_{i,\infty} - D_{i,n-i}, \quad i = 0, \dots, n \quad (2.3)$$

onde $\hat{D}_{i,\infty}$ é a estimativa total dos montantes acumulados que vão ser pagos, incluindo os pagamentos que serão feitos depois do último ano de desenvolvimento do triângulo, referente ao ano de ocorrência i .

Ao longo do presente relatório assume-se que os períodos de ocorrência e de desenvolvimento são anuais, sem perda de generalidade.

Apesar da inflação influenciar os custos com sinistros, usualmente aplica-se o método de estimação escolhido e posteriormente atualizam-se os valores estimados ao ano corrente, recorrendo a uma estimativa da taxa de inflação futura. Neste relatório não se considera o fator inflação.

3

Modelos Determinísticos

São considerados modelos determinísticos aqueles em que as estimativas obtidas baseiam-se apenas nos dados observados no passado e não nalguma distribuição de probabilidade. Assim, obtêm-se estimativas pontuais, ou seja, estimativas que não nos informam sobre a variabilidade da estimativa nem sobre os seus erros.

Neste capítulo serão abordadas algumas metodologias determinísticas para a estimação da provisão para sinistros, tais como, o método *Chain Ladder*, outros métodos que se baseiam em *link ratios*, modelos *Grossing Up* e o modelo *Bornhuetter-Ferguson*, bem como a inclusão do fator cauda em alguns destes modelos.

3.1 Método *Chain Ladder*

O método *Chain Ladder* insere-se na categoria de modelos com *link ratios*, ou seja, é um dos métodos que se baseia num conjunto de rácios que relacionam os montantes de um determinado ano com os montantes do ano seguinte. Esta é uma das técnicas atuariais mais antiga e amplamente aplicada pelas seguradoras na estimação da provisão para sinistros.

3.1.1 Metodologia

Tal como já referido, o objetivo é inferir sobre os valores dos montantes a pagar no futuro, ou seja, estimar o preenchimento do triângulo inferior.

Assumindo que existe proporcionalidade entre os anos de desenvolvimento, designa-se por coeficiente de desenvolvimento ou fator de proporcionalidade, f_j , o fator que determina o aumento médio dos montantes pagos entre anos de desenvolvimento consecutivos, tal que o seu estimador é definido por,

$$\hat{f}_j = \frac{\sum_{i=0}^{n-j-1} D_{i,j+1}}{\sum_{i=0}^{n-j-1} D_{i,j}}, \quad j = 0, \dots, n-1$$

e, caso existam pagamentos para além do ano de desenvolvimento n ,

$$\hat{f}_n = \frac{D_{0,\infty}}{D_{0,n}}$$

caso contrário, $\hat{f}_n = 1$.

Com base nos montantes acumulados conhecidos e nos coeficientes de desenvolvimento estimados tem-se a seguinte estimativa para preencher o triângulo inferior:

$$\hat{D}_{i,j} = \hat{D}_{i,j-1} \times \hat{f}_{j-1}, \quad i + j > n$$

sendo que os montantes estimados para a primeira diagonal desconhecida têm por base os valores da última diagonal conhecida.

Assim, estima-se o valor das responsabilidades atuais associadas ao ano de ocorrência i através da equação (2.3).

Este método pressupõe que existe proporcionalidade entre os anos de desenvolvimento, independência entre os diferentes anos de ocorrência e que os estimadores dos fatores de desenvolvimento obtidos correspondem aos que têm menor variância. Estes pressupostos serão descritos detalhadamente no Capítulo 5.

3.1.2 Vantagens e Limitações

Esta metodologia é de fácil aplicação, geralmente, dá-nos bons resultados e adapta-se bem aos dados em estudo. No entanto é de se apontar algumas críticas, tais como, a falta de robustez, ou seja, a elevada sensibilidade a pequenas alterações nos dados, em que pequenas flutuações nos dados observados poderá provocar grandes variações nas provisões estimadas. É ainda de se apontar que, quando, num determinado ano de ocorrência, ainda não existem pagamentos, este método indica que a provisão a constituir deve ser nula, o que não é realista, visto que pode ser um ano ainda pouco desenvolvido (com poucos sinistros).

3.1.3 Outros métodos baseados em *Link Ratios*

Considera-se como coeficiente de desenvolvimento individual, $f_{i,j}$, tal que o seu estimador é definido por,

$$\hat{f}_{i,j} = \frac{D_{i,j+1}}{D_{i,j}}, \quad i = 0, \dots, n-1, \quad j = 0, \dots, n-i-1 \quad (3.1)$$

$$\hat{f}_{0,n} = \hat{f}_n$$

Outros métodos baseados em *link ratios* diferem do método *Chain Ladder* apenas na estimação dos coeficientes de desenvolvimento, \hat{f}_j . Apresentam-se de seguida dois exemplos,

$$\hat{f}_j = \frac{\sum_{i=0}^{n-j-1} f_{i,j}}{n-j}, \quad j = 0, \dots, n \quad (3.2)$$

e,

$$\hat{f}_j = \min_i \{f_{i,j}\}, \quad j = 0, \dots, n. \quad (3.3)$$

Estes métodos são normalmente designados por, respetivamente, média dos fatores (*Average factor*) e pelo pior fator (*Worst factor*).

3.2 Método *Grossing Up*

Esta metodologia baseia-se em percentagens para estimar a parte inferior da matriz e é conhecida por técnica *Iceberg*, apresentada por Salzmann (1984). Esta analogia deve-se ao facto do montante total pago ser estimado por proporção da parte visível (sinistros pagos até ao momento). Por isso, a proporção invisível do *iceberg* (montantes que faltam pagar) é calculada por subtração das anteriores.

Considerando o primeiro ano de ocorrência e sabendo o valor de $D_{0,\infty}$, ou seja, quanto é que foi pago na totalidade referente a sinistros do primeiro ano, calcula-se a percentagem de montantes pagos em cada ano de desenvolvimento referentes a esse mesmo período, tal que,

$$\hat{g}_{0,j} = \frac{D_{0,j}}{D_{0,\infty}}, \quad j = 0, \dots, n.$$

Para os restantes anos de ocorrência considera-se que o último valor acumulado conhecido corresponde a uma média das percentagens estimadas anteriormente para o ano a que se refere esse valor, estimando as restantes percentagens em função desta, ou seja, para o segundo ano conhecido faz-se,

1. $\hat{g}_{1,n-1} = \hat{g}_{0,n-1}$
2. $\hat{D}_{1,\infty} = \frac{D_{1,n-1}}{\hat{g}_{1,n-1}}$

$$3. \hat{g}_{1,j} = \frac{D_{1,j}}{\hat{D}_{1,\infty}}, \quad j = 0, \dots, n-2$$

e, identicamente, para o ano $i = 2, \dots, n$, seguem-se os seguintes passos,

$$1. \hat{g}_{i,n-i} = \text{média}(\hat{g}_{0,n-i}; \hat{g}_{1,n-i}; \dots; \hat{g}_{i-1,n-i})$$

$$2. \hat{D}_{i,\infty} = \frac{D_{i,n-i}}{\hat{g}_{i,n-i}}$$

$$3. \hat{g}_{i,j} = \frac{D_{i,j}}{\hat{D}_{i,\infty}}, \quad j = 0, \dots, n-i-1$$

Outras perspetivas podem ser adotadas para a obtenção de $\hat{g}_{i,n-i}$, como por exemplo, estimando o pior caso possível, conhecido por *Grossing Up - Worst Case*, onde se substitui a média pelo mínimo no primeiro passo, bem como também se pode utilizar qualquer função $h(\hat{g}_{0,n-i}; \hat{g}_{1,n-i}; \dots; \hat{g}_{i-1,n-i})$.

Estimam-se assim as percentagens de montantes que são pagos em cada ano de desenvolvimento,

$$\hat{g}_j = \hat{g}_{j,n-j}, \quad j = 0, \dots, n$$

com,

$$\hat{g}_\infty = 100\%$$

Naturalmente, uma vez que os coeficientes de desenvolvimento não são mais do que o aumento médio de montantes pagos de um determinado ano de desenvolvimento para o seguinte, obtém-se a seguinte associação,

$$\hat{f}_j = \frac{\hat{g}_{j+1}}{\hat{g}_j}, \quad j = 0, \dots, n-1$$

e,

$$\hat{f}_n = \frac{\hat{g}_\infty}{\hat{g}_n}$$

Outras metodologias *Grossing Up* podem ser aplicadas, consultar, por exemplo, Institute and Faculty of Actuaries (1997).

3.3 Método *Bornhuetter-Ferguson*

Este método, inicialmente apresentado por Bornhuetter e Ferguson (1972) e posteriormente explorado e desenvolvido por Neuhaus (1992), baseia-se numa combinação da informação relativa aos montantes de indemnizações com fatores externos, tais como, a taxa de sinistralidade e o volume de prémios emitidos.

3.3.1 Metodologia

Este método apresenta uma formulação semelhante à do método de *Chain Ladder*, no entanto, a estimação dos coeficientes de desenvolvimento requer uma medida de exposição ao risco, tal como o volume de prémios emitidos.

Sejam pe_i , $i = 0, \dots, n$, o montante de prémios emitidos no ano de ocorrência i e θ_j , $j = 0, \dots, n$, a taxa de sinistralidade associada ao ano de desenvolvimento j , definida como o rácio entre montantes pagos e prémios emitidos. Considera-se que

$$S_j = \sum_{i=0}^{n-j} C_{i,j}, \quad j = 0, \dots, n$$

e que

$$p_j = \sum_{i=0}^{n-j} pe_i, \quad j = 0, \dots, n.$$

Então, assumindo que as taxas de sinistralidade são constantes para todos os anos de desenvolvimento e identicamente iguais a θ , tem-se o seguinte estimador,

$$\hat{\theta} = \sum_{j=0}^n \left(\frac{S_j}{p_j} \right) \quad (3.4)$$

onde se pode concluir que, se o nível dos prémios for adequado face às responsabilidades assumidas, então este estimador deve tender para 1.

As estimativas dos montantes incrementais futuros são calculadas recorrendo apenas aos prémios emitidos e às taxas de sinistralidade,

$$\hat{C}_{i,j} = pe_i \frac{S_j}{p_j}, \quad i + j > n$$

Portanto, as estimativas das provisões são obtidas por,

$$\hat{R}_i = \sum_{j=n+1-i}^n \hat{C}_{i,j}, \quad 1 \leq i \leq n \quad (3.5)$$

considerando que $R_0 = 0$.

Note-se que a introdução de mais informação para a obtenção de estimativas para a reserva total transmite mais fiabilidade do que os métodos anteriormente apresentados. Também é de referir que, ao contrário do método *Chain Ladder*, este método não falha quando, para um determinado ano de ocorrência, ainda não existem pagamentos, mas existem prémios, permitindo estimar uma reserva para esse ano. Assim sendo, este método pode ser mais útil em carteiras mais pequenas e com pouco histórico.

3.4 Fator Cauda

Quando não se considera que o triângulo está encerrado, ou seja, quando a evolução dos pagamentos justifica a utilização de um número de anos de desenvolvimento maior do que o considerado, deverá ser integrado um fator cauda ao triângulo que representará os montantes que serão pagos depois do último ano de desenvolvimento conhecido, n . Este fator pode ser simplesmente um valor que a Seguradora provisionou, caso o Atuário o considere justificável. Caso contrário, a metodologia mais utilizada para projetar os coeficientes de desenvolvimento é a descrita por Neuhaus (2004).

Segundo este procedimento, a projeção dos coeficientes de desenvolvimento é definida por,

$$\hat{f}_j = 1 + \delta(\hat{f}_{j-1} - 1), j > n - 1$$

onde δ varia entre 0 e 1, escolhendo-se o valor que melhor se ajusta ao padrão dos coeficientes de desenvolvimento, \hat{f}_j , já estimados, para $j < n$.

Note-se que quando se opta pela aplicação do método *Bornhuetter-Ferguson*, a projeção da cauda não pode ser estimada da mesma forma, visto que as taxas de sinistralidade, $\frac{S_j}{p_j}$, não têm o mesmo comportamento de \hat{f}_j . Assim, pode-se aplicar um modelo regressivo do tipo $y = a + b \ln(x)$ aos quocientes $\frac{S_j}{p_j}$, $j = 0, \dots, n$.

4

Teoria da Credibilidade e Provisionamento

A Teoria da Credibilidade é, principalmente, aplicada na área de tarifação, no entanto, pode também ser aplicada na estimação da provisão para sinistros.

Para mais detalhes acerca desta teoria pode, por exemplo, consultar-se Bühlmann (1967), Bühlmann e Straub (1970), Waters (1987), Norberg (2004), entre outros.

Neste capítulo é apresentado o conceito desta teoria e a sua aplicação na estimação da provisão para sinistros.

4.1 Generalidades

Essencialmente, a teoria da credibilidade assenta na aplicação de um fator de credibilidade $Z \in [0, 1]$ na fórmula de cálculo de prémios,

$$Z\bar{X} + (1 - Z)\mu$$

onde,

- \bar{X} é a média dos montantes dos sinistros observados com base nos dados do risco;
- μ é o valor esperado dos montantes dos sinistros com base em dados externos ao risco.

Note-se que, quanto mais credível for a informação da própria seguradora, maior deverá ser o fator de credibilidade, e quanto maior for a credibilidade da informação dos dados externos, menor deverá ser o fator.

4.2 Metodologia

A metodologia aqui explorada tem por base o método apresentado em Benktander (1976), posteriormente adaptado por Neuhaus (1992) e por Mack (2000).

Será feita uma breve descrição do método e, para mais detalhes, recomenda-se a leitura de, por exemplo, Hürlimann (2009).

Considere-se a taxa de sinistralidade esperada global da equação (3.4). Tem-se para cada ano de desenvolvimento j a respetiva taxa de sinistralidade,

$$\hat{\theta}_j = \frac{S_j}{p_j}, \quad j = 0, \dots, n$$

Segundo Hürlimann (2009) considere-se,

$$z_i = \frac{\sum_{j=0}^{n-i} \hat{\theta}_j}{\hat{\theta}}, \quad i = 0, \dots, n$$

representa a proporção de taxa de sinistralidade verificada até ao momento $n - i$ sobre a taxa de sinistralidade esperada total, por ano de ocorrência.

As reservas que se baseiam apenas nos montantes pagos conhecidos são designadas por reservas *individuais* e são dadas por,

$$\hat{R}_i^{ind} = \frac{D_{i,n-i}}{z_i} - D_{i,n-i} = (1 - z_i) \frac{D_{i,n-i}}{z_i}, \quad i = 0, \dots, n. \quad (4.1)$$

Por outro lado, têm-se as reservas globais já estimadas pelo método *Bornhuetter-Ferguson*, referido na Secção 3.3, que são agora dadas por,

$$\hat{R}_i^{col} = (1 - z_i) p e_i \hat{\theta}, \quad i = 0, \dots, n \quad (4.2)$$

também conhecidas por reservas *coletivas*, por depender da experiência da empresa em estudo.

Dependendo do valor que se toma para o fator de credibilidade associado a \hat{R}_i^{ind} , Z_i , estimam-se as reservas para cada ano de ocorrência i , que são uma combinação das reservas estimadas em (4.1) e em (4.2), dadas por

$$\hat{R}_i = Z_i \hat{R}_i^{ind} + (1 - Z_i) \hat{R}_i^{col} \quad (4.3)$$

Para o fator de credibilidade foram apresentadas diferentes propostas por diferentes autores, tais como as que são apresentadas na Tabela 4.1,

Tabela 4.1: Proposta para Z_i

Fator de Credibilidade - Z_i	Método
z_i	<i>Benktander</i>
$z_i \times \hat{\theta} = \sum_{j=0}^{n-i} \hat{\theta}_j$	<i>Neuhaus</i>
$\frac{z_i}{z_i + \sqrt{z_i}}$	<i>Optimal Credibility</i>

para cada ano de ocorrência, com $i = 0, \dots, n$.

Sendo este método uma combinação de duas reservas distintas, uma que se baseia apenas na informação dos montantes pagos e outra que recorre a informação extra, permite-nos escolher o fator de credibilidade que melhor se adequa à carteira em estudo, ou seja, se se tratar de uma carteira grande e com muito histórico aconselha-se um fator de credibilidade mais próximo de 0, caso contrário opta-se por um fator de credibilidade mais próximo de 1.

5

Modelo *Thomas Mack*

O modelo introduzido em Mack (1993) e posteriormente desenvolvido em Mack (1999) possibilita a obtenção de medidas de erro e de intervalos de confiança, partindo das estimativas obtidas pelo método *Chain Ladder* na Secção 3.1.

Neste capítulo apresentam-se os pressupostos inerentes a esta metodologia e posteriormente obtêm-se as estatísticas de interesse para as estimativas já obtidas.

5.1 Pressupostos

Este modelo assenta na verificação de três pressupostos:

1. $E(D_{i,j+1}|D_{i,0}, \dots, D_{i,j}) = D_{i,j}f_j, 0 \leq i \leq n, 0 \leq j \leq n - 1;$
2. As variáveis $\{D_{i,0}, \dots, D_{i,n}\}$ e $\{D_{k,0}, \dots, D_{k,n}\}$, para diferentes anos de ocorrência, $i \neq k$, são independentes;
3. $Var(D_{i,j+1}|D_{i,0}, \dots, D_{i,j}) = D_{i,j}\sigma_j^2, 0 \leq i \leq n, 0 \leq j \leq n - 1.$

5.1.1 1º Pressuposto - Proporcionalidade

O pressuposto de que $E(D_{i,j+1}|D_{i,0}, \dots, D_{i,j}) = D_{i,j}f_j, 0 \leq i \leq n, 0 \leq j \leq n - 1$ implica a verificação de dois testes, existência de proporcionalidade entre os anos de desenvolvimento e inexistência de correlação entre os fatores de desenvolvimento.

Uma forma intuitiva de testar a existência de proporcionalidade entre os anos de desenvolvimento, ou seja, entre as colunas da matriz de pagamentos, é construir um gráfico com os pares ordenados $(D_{i,j}, D_{i,j+1})$ e traçar uma reta que passe na origem, com declive

\hat{f}_j . Caso se verificarem desvios significativos rejeita-se a hipótese e deve-se procurar estimadores para \hat{f}_j que melhor se ajustem aos dados.

Para verificar a hipótese de inexistência de correlação entre os fatores de desenvolvimento recorre-se ao teste de *Spearman*.

Este teste inicia-se com a construção de duas matrizes, $[r_{i,j}]$ e $[s_{i,j}]$, com os seguintes elementos:

- $r_{i,j}$ é o número de ordem atribuído ao fator de desenvolvimento individual $\hat{f}_{i,j}$;
- $s_{i,j}$ é o número de ordem atribuído ao fator de desenvolvimento individual precedente, $\hat{f}_{i,j-1}$.

onde, $\hat{f}_{i,j}$ é estimado segundo a equação (3.1).

De seguida, estima-se o coeficiente de correlação de *Spearman*, T_j :

$$T_j = 1 - 6 \sum_{i=0}^{n-j-1} \frac{(r_{i,j} - s_{i,j})^2}{(n-j)^3 - n + j}, \quad 1 \leq j \leq n-2 \quad (5.1)$$

sendo T_j um valor situado entre -1 e 1 . Quanto mais próximo de zero, menor é a correlação entre os fatores de desenvolvimento dos anos $j-1$ e j e entre os anos j e $j+1$. Caso contrário, verifica-se a existência de correlação positiva ou negativa.

Na ausência de correlações, tem-se que,

$$E(T_j) = 0, \quad 1 \leq j \leq n-2$$

e que,

$$V(T_j) = \frac{1}{n-j-1}, \quad 1 \leq j \leq n-2.$$

De forma a obter um estimador de variância mínima, Mack (1993) propõe:

$$T = \frac{\sum_{j=1}^{n-2} (n-j-1)T_j}{\sum_{j=1}^{n-2} (n-j-1)} \quad (5.2)$$

com,

$$E(T) = 0$$

e,

$$V(T) = \frac{1}{\frac{(n-1)(n-2)}{2}}$$

Tendo em conta que T é estimado pela soma ponderada das variáveis aleatórias T_j 's não correlacionadas e com distribuição aproximadamente Normal, pode-se assumir que

a distribuição de T também se aproxima da distribuição Normal. Assim, se o valor observado de T se situar dentro do seguinte intervalo de confiança,

$$\left] E(T) - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sqrt{V(T)}; E(T) + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sqrt{V(T)} \right[\quad (5.3)$$

aceita-se a hipótese de ausência de correlação entre os fatores de desenvolvimento, onde $\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)$ corresponde ao quantil de probabilidade $\left(1 - \frac{\alpha}{2} \right)$ da distribuição Normal Padrão e α é o nível de significância do teste.

5.1.2 2º Pressuposto - Independência

Este pressuposto consiste na verificação da hipótese de independência entre os anos de ocorrência, ou seja, entre as linhas da matriz de pagamentos.

Para se confirmar que as variáveis $\{D_{i,0}, \dots, D_{i,n}\}$ e $\{D_{k,0}, \dots, D_{k,n}\}$, $i \neq k$, são independentes há que distinguir os fatores de desenvolvimento individuais em dois conjuntos, os mais elevados e os mais baixos, em cada ano de desenvolvimento j , e verificar se existem diagonais na matriz de pagamentos com supremacia de elementos de um dos dois conjuntos. O conjunto L é constituído pelos fatores de desenvolvimento mais elevados de cada ano de desenvolvimento e o conjunto S pelos mais baixos. Ambos os conjuntos têm de ter o mesmo número de elementos em cada ano j , portanto, nos anos em que se verificar um número ímpar de fatores de desenvolvimento ignora-se o fator de valor mediano.

Considere-se a diagonal D_j de fatores de desenvolvimento individuais tal que,

- $D_0 = \{ \hat{f}_{0,0} \};$
- $D_1 = \{ \hat{f}_{0,1}; \hat{f}_{1,0} \};$
- ...
- $D_n = \{ \hat{f}_{0,n}; \hat{f}_{1,n-1}; \hat{f}_{2,n-2}; \dots; \hat{f}_{k,n-k}; \dots; \hat{f}_{n,0} \}.$

De seguida, contabiliza-se, para cada diagonal D_j de fatores de desenvolvimento, o número de elementos de S e de L . Se o número de elementos nos dois conjuntos for, aproximadamente, o mesmo, significa que não existe influência de um determinado ano nos dados em causa.

Considerando que L_j é o número de elementos de L que pertencem a D_j e que S_j é o número de elementos de S em D_j , se $Z_j = \min(L_j; S_j)$ for significativamente menor que $\frac{L_j + S_j}{2}$, então existe uma supremacia de fatores de desenvolvimento elevados ou baixos na diagonal j de fatores de desenvolvimento, levando à rejeição da hipótese nula.

Note-se que, atendendo à hipótese que se quer testar, S_j e L_j têm distribuição Binomial de parâmetros $z = L_j + S_j$ e $p = 0,5$.

Assim, assumindo uma aproximação à distribuição Normal, aceita-se a hipótese de que existe independência entre os anos de ocorrência com um nível de significância α , se a variável $Z = \sum_{j=0}^n Z_j$ se situar dentro do intervalo

$$\left] E(Z) - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sqrt{V(Z)}; E(Z) + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sqrt{V(Z)} \right[$$

sendo os dois primeiros momentos da variável Z_j dados por,

$$E(Z_j) = \frac{z}{2} - C_m^{z-1} \frac{z}{2^z} \quad (5.4)$$

e por,

$$V(Z_j) = \frac{z(z-1)}{4} - C_m^{z-1} \frac{z(z-1)}{2^z} + E(Z_j) - [E(Z_j)]^2 \quad (5.5)$$

onde m é o inteiro, aproximado por defeito, de $\frac{z-1}{2}$.

Assim, os dois primeiros momentos da variável Z são estimados da seguinte forma,

$$E(Z) = \sum_{j=0}^{n-1} E(Z_k) \quad (5.6)$$

$$V(Z) = \sum_{j=0}^{n-1} V(Z_k) \quad (5.7)$$

5.1.3 3º Pressuposto - Variabilidade mínima

O pressuposto $Var(D_{i,j+1}|D_{i,0}, \dots, D_{i,j}) = D_{i,j}\sigma_j^2$, $0 \leq i \leq n$, $0 \leq j \leq n-1$, com σ_j^2 constante de proporcionalidade desconhecida, exige que os estimadores dos fatores de desenvolvimento estimados correspondam aos que têm menor variância.

Segundo Mack (1993), analisando o gráfico com os pares ordenados $\left(D_{i,j}, \frac{D_{i,j+1} - D_{i,j}\hat{f}_j}{\sqrt{D_{i,j}}} \right)$, para cada ano de desenvolvimento j , $j = 0, \dots, n$, caso não se verifique nenhuma tendência nesses resíduos, ou seja, caso sejam aparentemente aleatórios, admite-se como válido este pressuposto.

5.2 Medidas de Variabilidade

Após a análise dos pressupostos estudados, prossegue-se então para a estimação do erro padrão e dos intervalos de confiança das reservas estimadas pelo método *Chain Ladder* na Secção 3.1.

Com o intuito de se saber qual a distância entre a estimativa $\hat{D}_{i,n}$ e o valor que de facto se vai observar $D_{i,n}$, recorre-se ao *Erro Quadrático Médio* (EQM),

$$EQM(\hat{D}_{i,n}) = E \left[(D_{i,n} - \hat{D}_{i,n})^2 | D \right]$$

onde $D = \{D_{i,j} | i+j \leq n+1\}$ é o conjunto de todos os dados conhecidos até ao momento. É importante ter-se em conta que o objetivo é saber o erro que advém da estimação, ou

seja, o erro quadrático médio é estimado na condição de se conhecer todos os dados observados.

O erro quadrático médio de $\hat{D}_{i,n}$, $EQM(\hat{D}_{i,n})$, coincide com o erro quadrático médio de \hat{R}_i , $EQM(\hat{R}_i)$, visto que,

$$EQM(\hat{R}_i) = E((R_i - \hat{R}_i)^2 | D) = E((D_{i,n} - \hat{D}_{i,n})^2 | D) = EQM(\hat{D}_{i,n}). \quad (5.8)$$

Segundo Mack (1999) tem-se,

$$EQM(\hat{D}_{i,j+1}) = \hat{D}_{i,j}^2 \left[EQM\left(\frac{\hat{D}_{i,j+1}}{\hat{D}_{i,j}}\right) + EQM(\hat{f}_j) \right] + EQM(\hat{D}_{i,j}) \hat{f}_j^2 \quad (5.9)$$

em que,

$$EQM\left(\frac{\hat{D}_{i,j+1}}{\hat{D}_{i,j}}\right) = \frac{\hat{\sigma}_j^2}{\hat{D}_{i,j}}$$

onde a constante de proporcionalidade $\hat{\sigma}_j^2$ é estimada através de:

$$\hat{\sigma}_j^2 = \frac{1}{n-j-1} \sum_{i=0}^{n-j-1} D_{i,j} \left(\frac{D_{i,j+1}}{D_{i,j}} - \hat{f}_j \right)^2, \quad 0 \leq j \leq n-2. \quad (5.10)$$

Para $j = n-1$ tem-se a seguinte hipótese de estimativa,

$$\hat{\sigma}_{n-1}^2 = \text{Min} \left\{ \frac{\hat{\sigma}_{n-2}^4}{\hat{\sigma}_{n-3}^2}, \text{Min} \{ \hat{\sigma}_{n-3}^2, \hat{\sigma}_{n-2}^2 \} \right\} \quad (5.11)$$

Para $j < n$ tal que $\hat{f}_{j-1} < \hat{f}_n < \hat{f}_j$, foi proposto o seguinte estimador de σ_n^2 , estimado por interpolação numérica,

$$\hat{\sigma}_n^2 = \left[1 - \left(\frac{\hat{f}_{j-1} - \hat{f}_n}{\hat{f}_{j-1} - \hat{f}_j} \right) \right] \hat{\sigma}_{j-1}^2 + \left(\frac{\hat{f}_{j-1} - \hat{f}_n}{\hat{f}_{j-1} - \hat{f}_j} \right) \hat{\sigma}_j^2. \quad (5.12)$$

Tenha-se em atenção que, quando $\hat{f}_{n-1} = 1$ ou $\hat{f}_n = 1$, considere-se $\hat{\sigma}_{n-1}^2 = 0$ e $\hat{\sigma}_n^2 = 0$, respetivamente, visto que, nestes casos, não se espera que ocorram mais sinistros relativos a esses anos, assumindo-se assim que não há variabilidade possível.

O erro quadrático médio de cada fator de desenvolvimento é dado por,

$$EQM(\hat{f}_j) = \frac{\hat{\sigma}_j^2}{\sum_{i=0}^{n-j-1} D_{i,j}}, \quad 0 \leq i \leq n, \quad 0 \leq j \leq n-1 \quad (5.13)$$

e, em particular,

$$E\hat{Q}M(\hat{f}_n) = E\hat{Q}M(\hat{f}_j) + \left[E\hat{Q}M(\hat{f}_{j-1}) - E\hat{Q}M(\hat{f}_j) \right] \left(\frac{\hat{f}_n - \hat{f}_j}{\hat{f}_{j-1} - \hat{f}_j} \right) \quad (5.14)$$

onde, $\hat{f}_{j-1} < \hat{f}_n < \hat{f}_j$.

Através da equação (5.8), facilmente se conclui que,

$$E\hat{Q}M(\hat{R}) = E\hat{Q}M\left(\sum_{i=0}^n \hat{R}_i\right) = E\hat{Q}M\left(\sum_{i=0}^n \hat{D}_{i,n}\right) \quad (5.15)$$

que se estima através da seguinte fórmula recursiva,

$$E\hat{Q}M\left(\sum_{i=n-j}^n \hat{D}_{i,j+1}\right) = \sum_{i=n-j}^n \hat{D}_{i,j}^2 E\hat{Q}M\left(\frac{D_{i,j+1}}{D_{i,j}}\right) + \left(\sum_{i=n-j}^n \hat{D}_{i,j}\right)^2 E\hat{Q}M(\hat{f}_j) + E\hat{Q}M\left(\sum_{i=n-j+1}^n \hat{D}_{i,j}\right) \hat{f}_j^2.$$

Assumindo a hipótese de normalidade de \hat{R}_i e de \hat{R} , estimam-se os intervalos de confiança, com um nível de confiança $1 - \alpha$, para as estimativas das provisões, respetivamente,

$$\left[\hat{R}_i - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{E\hat{Q}M(\hat{R}_i)}; \hat{R}_i + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{E\hat{Q}M(\hat{R}_i)} \right]$$

e,

$$\left[\hat{R} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{E\hat{Q}M(\hat{R})}; \hat{R} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{E\hat{Q}M(\hat{R})} \right].$$

Assim sendo, este método tem como vantagem avaliar a variabilidade da estimativa pontual obtida pelo método de *Chain Ladder*, atribuindo um grau de confiança às estimativas apresentadas.

6

Modelos Lineares Generalizados

Neste capítulo é feita uma breve apresentação dos modelos lineares generalizados, referindo as suas componentes e apresentando a definição de Família Exponencial. Posteriormente, estuda-se a aplicação destes modelos na estimação da provisão para sinistros, nomeadamente, através do modelo de *Sobre-dispersão de Poisson*, que se trata de uma vertente estocástica do método *Chain Ladder*, ou seja, permite obter estimativas semelhantes às deste método determinístico. No Capítulo 7, é aplicada a técnica de *Bootstrap* recorrendo a este modelo.

6.1 Generalidades

Os modelos lineares generalizados, adiante designados por MLG, tiveram origem nos modelos de regressão linear e foram inicialmente propostos por Nelder e Wedderburn (1972).

6.1.1 Componentes

Os modelos lineares generalizados são constituídos por três componentes, a componente aleatória, a sistemática e a função de ligação.

A primeira componente identifica a variável resposta Y e assume que esta tem uma função densidade de probabilidade que pertence à Família Exponencial.

A componente sistemática especifica as variáveis explicativas do modelo, tomando uma combinação linear dessas mesmas variáveis, tal que:

$$\mathbf{H} = \mathbf{X}\vec{B}$$

onde \mathbf{X} é uma matriz com p variáveis explicativas e \vec{B} é o vetor de parâmetros a estimar de dimensão p .

Por último, o preditor linear, η , relaciona as duas primeiras componentes:

$$\eta = g(\mu) = g(E(Y)) = \mathbf{H}$$

onde $g(\cdot)$ é a função de ligação e μ é o vetor dos valores esperados da componente aleatória Y , tal que:

$$E(Y) = g^{-1}(\mathbf{H}) = g^{-1}(\mathbf{X}\vec{B}) = g^{-1}\left(\sum_{j=1}^p \beta_j x_j\right).$$

A função de ligação é escolhida consoante o tipo de resposta que se pretende obter. Na Tabela 6.1 encontram-se as funções de ligação mais utilizadas e respetivas designações e distribuições.

Tabela 6.1: Funções de ligação

Distribuição	Função de Ligação	Designação
Normal/ Gaussiana	μ	Identidade
Binomial	$\ln(\mu/(1-\mu))$	Logit
Poisson	$\ln(\mu)$	Logarítmica
Gama	$1/\mu$	Recíproca
Normal inversa	$1/\mu^2$	Quadrática inversa

6.1.2 Família Exponencial

Seja Y uma variável aleatória. Diz-se que Y tem distribuição pertencente à Família Exponencial se a sua função densidade de probabilidade se puder escrever da seguinte forma:

$$f(y|\theta, \phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right\} \quad (6.1)$$

onde θ e ϕ são parâmetros escalares e $a(\cdot)$, $b(\cdot)$ e $c(\cdot)$ são funções reais conhecidas. Os dois parâmetros são designados por, respetivamente, parâmetro canónico de localização e parâmetro de dispersão ou de escala. A função $a(\cdot)$ é da forma $a(\phi) = \frac{\phi}{w}$, onde w é um valor constante que depende da informação conhecida. Assim, a equação (6.1) pode ser escrita na forma:

$$f(y|\theta, \phi, w) = \exp\left\{\frac{w}{\phi}(y\theta - b(\theta)) + c(y, \phi, w)\right\} \quad (6.2)$$

Daqui resulta que,

$$E(Y) = \mu = b'(\theta) \quad (6.3)$$

e que,

$$V(Y) = b''(\theta)a(\phi) = \frac{b''(\theta)\phi}{w}. \quad (6.4)$$

O termo $b''(\theta)$ é designado por função de variância e denotado de $v(\mu)$.

6.1.2.1 Distribuição de Poisson

Supondo que Y segue uma distribuição de Poisson de valor médio μ , então a sua função probabilidade é dada por,

$$f(y|\mu) = \frac{\mu^y e^{-\mu}}{y!} = \exp\{y \ln(\mu) - \mu - \ln(y!)\}.$$

Pode verificar-se que esta função é da forma (6.2), ou seja, pertencente à Família Exponencial, com,

- $\theta = \ln(\mu)$;
- $a(\phi) = 1$;
- $b(\theta) = e^\theta = \mu$;
- $c(y, \phi) = -\ln(y!)$;
- $b'(\theta) = e^\theta = \mu = b''(\theta)$.

Deste modo, a média e a variância da variável aleatória Y são dadas por

$$E(Y) = b'(\theta) = \mu$$

e por

$$V(Y) = b''(\theta)a(\phi) = \mu$$

respetivamente, tal como é próprio desta distribuição.

6.2 MLG na estimação da Provisão para Sinistros

6.2.1 Pressupostos

Considerando o triângulo de montantes incrementais, $C_{i,j}$, $i = 0, \dots, n$, $j = 0, \dots, n$, em que os montantes $C_{i,j}$ são variáveis aleatórias independentes e identicamente distribuídas, com distribuição pertencente à Família Exponencial, supõe-se que $w_{i,j} = 1$, para todo o par (i, j) , visto que se consideram pagamentos anuais, ou seja, que $a_i(\phi) = \phi$. Assim, através das equações (6.3) e (6.4), tem-se, respetivamente, que,

$$E(C_{i,j}) = \mu_{i,j}$$

e que,

$$V(C_{i,j}) = \phi v(\mu_{i,j})$$

com $\phi > 1$, desconhecido e estimado a partir dos dados.

Cada observação corresponde a uma variável aleatória do modelo, ou seja, têm-se $\frac{1}{2}(n+1)(n+2)$ variáveis aleatórias e o objetivo é modelar o valor esperado destas variáveis, sendo as variáveis explicativas dadas pelo ano de ocorrência e pelo ano de desenvolvimento.

Neste contexto de estimação de reservas para sinistros utiliza-se o seguinte preditor linear e função de ligação:

$$\eta_{i,j} = \ln(\mu_{i,j}) = \mu + \alpha_i + \beta_j$$

onde o parâmetro α_i representa o efeito provocado pelo ano de ocorrência i , o parâmetro β_j o efeito provocado pelo ano de desenvolvimento j e o parâmetro μ é a média global.

Casos particulares deste modelo são, por exemplo, modelos cuja distribuição da variável resposta é a distribuição de Poisson, a Gama ou a Gaussiana Inversa, apresentadas na Tabela 6.1.

Com o intuito de evitar a sobre-parametrização introduzem-se as restrições $\alpha_0 = \beta_0 = 0$, ficando com um total de $2n+1$ parâmetros para estimar (número de anos de ocorrência e de anos de desenvolvimento estimados e a média global).

6.2.2 Modelo de Sobre-dispersão de Poisson

Com o objetivo de construir uma vertente estocástica do método *Chain Ladder*, em Renshaw e Verrall (1998) especifica-se uma distribuição de probabilidade para os dados em estudo, a distribuição de *Sobre-dispersão de Poisson*. Esta distribuição difere da distribuição de *Poisson* na medida em que a variância não é igual à média, mas sim proporcional, ou seja,

- $E(C_{i,j}) = \mu_{i,j}$;
- $V(C_{i,j}) = \phi v(\mu_{i,j}) = \phi \mu_{i,j}$, com $\phi > 1$;
- $\eta_{i,j} = g(\mu_{i,j}) = \ln(\mu_{i,j})$.

A este modelo adiciona-se a restrição de que a soma dos montantes incrementais em cada ano de desenvolvimento tem de ser positiva, ou seja,

$$\sum_{i=0}^{n-j} C_{i,j} \geq 0, \quad 0 \leq j \leq n. \quad (6.5)$$

Note-se que tal restrição permite que existam montantes incrementais negativos, desde que a soma por anos de desenvolvimento seja positiva.

Este modelo permite obter estimativas para as reservas idênticas às do método *Chain Ladder*, calculando o valor esperado das variáveis aleatórias $C_{i,j}$, $1 \leq i \leq n$, $n+1-i \leq j \leq n$, através de,

$$E(C_{i,j}) = \hat{\mu}_{i,j} = e^{\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j} \quad (6.6)$$

onde $\hat{\mu}$, $\hat{\alpha}_i$ e $\hat{\beta}_j$ são estimadores de máxima quasi-verosimilhança de difícil interpretação (consultar McCullagh e Nelder (1989)). Também se consegue, da mesma forma, estimar valores ajustados para $C_{i,j}$, $0 \leq i \leq n$, $0 \leq j \leq n-i$.

Outra parametrização possível para este modelo é uma parametrização multiplicativa, em que,

$$E(C_{i,j}) = \hat{\mu}_{i,j} = \hat{\alpha}'_i \hat{\beta}'_j$$

onde $\hat{\alpha}'_i$ representa o valor esperado do total de montantes pagos de sinistros ocorridos no ano i e $\hat{\beta}'_j$ representa a proporção destes pagamentos surgirem no ano de desenvolvimento j .

Sabendo que a reserva para cada ano de ocorrência é dada pela equação (3.5), tem-se o seguinte estimador para a mesma,

$$\hat{R}_i = \sum_{j=n+1-i}^n E(C_{i,j}) = \sum_{j=n+1-i}^n \hat{\mu}_{i,j}, \quad 1 \leq i \leq n \quad (6.7)$$

considerando que $R_0 = 0$. Consequentemente, o estimador para a reserva total é dado por,

$$\hat{R} = \sum_{i=0}^n \hat{R}_i = \sum_{i=1}^n \sum_{j=n+1-i}^n \hat{\mu}_{i,j} \quad (6.8)$$

A fim de analisar a qualidade do ajustamento do modelo, inicialmente deve-se estimar os resíduos de *Pearson*, $r_{i,j}$, para $i = 0, \dots, n$ e $j = 0, \dots, n-i$, tais que,

$$r_{i,j} = \frac{C_{i,j} - \hat{\mu}_{i,j}}{\sqrt{V(\hat{C}_{i,j})}} = \frac{C_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\phi}v(\hat{\mu}_{i,j})}} = \frac{C_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\phi}\hat{\mu}_{i,j}}} \quad (6.9)$$

onde o parâmetro de dispersão, ϕ , é estimado através da soma do quadrado dos resíduos de *Pearson*, conhecida por estatística de *Pearson*, dividida pelos graus de liberdade, ou seja,

$$\hat{\phi} = \frac{\sum_{i=0}^n \sum_{j=0}^{n-i} (r_{i,j})^2}{N - p} \quad (6.10)$$

onde N representa o número total de observações, ou seja, geralmente, $N = \frac{1}{2}(n+1)(n+2)$ e p o número de parâmetros estimados. No entanto, visto que ϕ é constante ao longo

dos dados, pode-se ignorar este parâmetro e fica-se com os seguintes resíduos,

$$r_{i,j}^* = \frac{C_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\mu}_{i,j}}} \quad (6.11)$$

De seguida prossegue-se com a estimação dos resíduos de *Pearson* corrigidos, $r_{i,j}^{**}$, da seguinte forma,

$$r_{i,j}^{**} = \sqrt{\frac{N}{N-p}} r_{i,j}^* \quad (6.12)$$

Caso a análise gráfica dos resíduos versus os anos de ocorrência e os anos de desenvolvimento não apresente nenhuma tendência, pode-se concluir que o modelo se ajusta aos dados em estudo.



Simulação *Bootstrap*

A simulação *Bootstrap* é uma técnica de reamostragem, desenvolvida por Efron (1979), e baseia-se na geração de amostras aleatórias com reposição que permitem obter estimativas de erros de previsão e intervalos de confiança para as estimativas.

Neste capítulo apresenta-se a metodologia de base da técnica *Bootstrap*, a sua adaptação à estimação da provisão para sinistros, as medidas de variabilidade inerentes e as suas limitações.

7.1 Metodologia

Este método tem por base as estimativas obtidas através de um modelo, estocástico ou determinístico, sendo as medidas de variabilidade determinadas através da técnica *Bootstrap*. Neste estudo a simulação terá por base o modelo de *Sobre-dispersão de Poisson*, visto que, através deste modelo, estimam-se estatísticas semelhantes às do método *Chain Ladder* e, assim, consegue-se obter medidas de variabilidade para as mesmas.

Inicialmente aplica-se o modelo escolhido aos dados em causa e, como a técnica *Bootstrap* pressupõe a utilização de dados independentes e identicamente distribuídos, a aplicação será feita sobre os resíduos do modelo. Assim, procede-se à determinação dos resíduos ajustados que serão estimados utilizando os resíduos de *Pearson* corrigidos da equação (6.12).

De seguida inicia-se a técnica *Bootstrap* propriamente dita, que será repetida k vezes, seguindo os seguintes passos:

1. Procede-se à reamostragem dos resíduos corrigidos (com reposição), ignorando os dois valores nulos dos extremos da diagonal do triângulo, de forma a construir um

novo triângulo baseado nestes resíduos, designados por $r_{i,j}^{**^{(b)}}$;

2. Com os resíduos estimados, $r_{i,j}^{**^{(b)}}$, e com os montantes incrementais ajustados, $\hat{\mu}_{i,j}$, consegue-se obter um novo conjunto de dados incrementais, $\hat{\mu}_{i,j}^{(b)}$, designados por pseudo-dados, da seguinte forma:

$$\hat{\mu}_{i,j}^{(b)} = r_{i,j}^{**^{(b)}} \sqrt{\hat{\mu}_{i,j}} + \hat{\mu}_{i,j}; \quad (7.1)$$

3. Aplica-se novamente o método de *Sobre-dispersão de Poisson* a esse conjunto de dados, obtendo estimativas para os montantes ainda por pagar;
4. Estimam-se as reservas para cada ano de ocorrência através da equação (3.5);
5. Guardam-se os resultados obtidos e volta-se ao passo 1, repetindo este processo até as k repetições serem completadas, terminando o ciclo de *Bootstrap*.

Em cada iteração do ciclo anterior obtêm-se estimativas para as reservas para cada ano de ocorrência, sendo que no final têm-se k amostras para estas estimativas, $\hat{R}_i^{(1)}, \dots, \hat{R}_i^{(k)}$, e têm-se as respectivas médias dos estimadores,

$$\hat{R}_i^{bs} = \frac{1}{k} \sum_{b=1}^k \hat{R}_i^{(b)}, \quad i = 0, \dots, n$$

$$\hat{R}^{bs} = \frac{1}{k} \sum_{b=1}^k \hat{R}^{(b)}$$

onde,

$$\hat{R}^{(b)} = \sum_{i=0}^n \hat{R}_i^{(b)}, \quad b = 1, \dots, k$$

7.2 Medidas de Variabilidade

A partir das amostras obtidas anteriormente pode-se estimar algumas estatísticas de interesse como o erro de previsão e intervalos de confiança.

Os desvios padrão associados às estimativas de *Bootstrap* são definidos por

$$\hat{\sigma}_{bs}(\hat{R}_i) = \sqrt{\frac{1}{k-1} \sum_{b=1}^k \left(\hat{R}_i^{(b)} - \hat{R}_i^{bs} \right)^2}, \quad i = 0, \dots, n \quad (7.2)$$

e

$$\hat{\sigma}_{bs}(\hat{R}) = \sqrt{\frac{1}{k-1} \sum_{b=1}^k \left(\hat{R}^{(b)} - \hat{R}^{bs} \right)^2}. \quad (7.3)$$

Segundo England e Verral (1998), têm-se as seguintes estimativas para os erros de previsão,

$$\hat{E}P_{bs}(\hat{R}_i) = \sqrt{\hat{\phi}\hat{R}_i + \frac{N}{N-p}\hat{\sigma}_{bs}^2(\hat{R}_i)}, \quad i = 0, \dots, n$$

e

$$\hat{E}P_{bs}(\hat{R}) = \sqrt{\hat{\phi}\hat{R} + \frac{N}{N-p}\hat{\sigma}_{bs}^2(\hat{R})}$$

onde $\hat{\phi}$ é estimado pela equação (6.10), \hat{R}_i^{bs} são as estimativas das reservas para o ano de ocorrência i estimadas através do método *Bootstrap* e \hat{R}_i são as estimativas das reservas para o mesmo ano pelo método de *Sobre-dispersão de Poisson*.

Tratando-se de amostras aleatórias de grande dimensão, e tendo em conta o Teorema Limite Central, pode-se aproximar a distribuição da média à distribuição assintótica normal. Assim, têm-se os seguintes intervalos de confiança (IC) para um nível de confiança α ,

$$IC_i = \left] \hat{R}_i - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \hat{E}P_{bs}(\hat{R}_i); \hat{R}_i + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \hat{E}P_{bs}(\hat{R}_i) \right[, \quad i = 0, \dots, n \quad (7.4)$$

$$IC = \left] \hat{R} - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \hat{E}P_{bs}(\hat{R}); \hat{R} + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \hat{E}P_{bs}(\hat{R}) \right[. \quad (7.5)$$



Método *Double Chain Ladder*

Este método foi estudado e desenvolvido em Miranda et al. (2012), onde os autores combinam o número de sinistros declarados com os montantes pagos, utilizando duas vezes o algoritmo *Chain Ladder* - daí a designação de *Double Chain Ladder*.

Neste capítulo apresentam-se os pressupostos necessários para a aplicação deste modelo e a metodologia que se segue para a estimação da provisão para sinistros. Posteriormente é aplicada a técnica de *Bootstrap* a este método, bem como é efetuada uma associação ao método *Bornhuetter-Ferguson*.

8.1 Pressupostos

Considere-se que $N_{i,j}^{pagos}$ corresponde ao número de sinistros pagos ocorridos no ano i e encerrados no ano de desenvolvimento j , tal que:

$$N_{i,j}^{pagos} = \sum_{l=0}^{\min(j,d)} N_{i,j-l,l}^{pagos} \quad (8.1)$$

onde $N_{i,j,l}^{pagos}$ é o número de sinistros pagos originados de $N_{i,j}$ (sinistros declarados) e pagos no ano de desenvolvimento $j+l$ e d é o máximo de anos de atraso para o pagamento, com $d \leq n$.

Assumindo que os sinistros estão encerrados com um único pagamento, tome-se como $Y_{i,j}^{(k)}$ os pagamentos individuais que surgem a partir de $N_{i,j}^{pagos}$, com $k = 1, \dots, N_{i,j}^{pagos}$,

$$C_{i,j} = \sum_{k=1}^{N_{i,j}^{pagos}} Y_{i,j}^{(k)}, \quad (i,j) \in I, \quad I = \{(i,j) : i = 0, \dots, n, j = 0, \dots, n, i+j \leq n\} \quad (8.2)$$

A equação (8.2) é uma representação estocástica dos montantes pagos utilizada neste método.

Este método tem por base alguns pressupostos, tais como:

1. Variáveis relativas a anos de ocorrência diferentes ou de anos de desenvolvimento diferentes são independentes;
2. As variáveis $Y_{i,j}^{(k)}$ são independentes de $N_{i,j}$ e do atraso de *RBNS* e de *IBNR*;
3. Dado $N_{i,j}$, o número de sinistros pagos tem a seguinte distribuição multinomial conjunta:

$$(N_{i,j,0}^{pagos}, \dots, N_{i,j,d}^{pagos}) \sim Multi(N_{i,j}; p_0, \dots, p_d), \quad (i, j) \in I$$

onde $p = (p_0, \dots, p_d)$ são as probabilidades de atraso, com $\sum_{l=0}^d p_l = 1$ e $0 < p_l < 1$ para $l = 0, \dots, d$. Caso $d = 0$ significa que não houve atraso no pagamento do sinistro. Assim, tem-se,

$$E \left[N_{i,j}^{pagos} \mid N_{i,j} : (i, j) \in I \right] = \sum_{l=0}^{\min(j,d)} N_{i,j-l} \times p_l$$

4. Os pagamentos individuais $Y_{i,j}^{(k)}$ são mutuamente independentes com distribuição f_i . Sejam μ_i e σ_i^2 a média e a variância de $Y_{i,j}^{(k)}$, respetivamente, para cada $i = 0, \dots, n$. Assume-se que

$$\mu_i = E \left[Y_{i,j}^{(k)} \right] = \mu \gamma_i$$

e que

$$\sigma_i^2 = V \left(Y_{i,j}^{(k)} \right) = \sigma^2 \gamma_i^2$$

onde μ e σ^2 são o fator média e o fator variância, respetivamente, definidos adiante, e γ_i é a influência do ano de ocorrência i , ou seja, assume-se que a distribuição depende do ano de ocorrência. Portanto,

$$E [C_{i,j} \mid N_{i,j} : (i, j) \in I] = E \left[N_{i,j}^{pagos} \mid N_{i,j} : (i, j) \in I \right] E \left[Y_{i,j}^{(k)} \right] = \sum_{l=0}^{\min(j,d)} N_{i,j-l} p_l \mu \gamma_i \quad (8.3)$$

5. $N_{i,j}$ são variáveis aleatórias independentes com distribuição de *Poisson* com uma parametrização multiplicativa, tal que,

$$E [N_{i,j}] = \alpha_i \beta_j$$

com $\sum_{j=0}^n \beta_j = 1$, onde α_i é o parâmetro que representa a influência do ano de ocorrência i e β_j é o parâmetro para o ano de desenvolvimento j . Do pressuposto

anterior conclui-se que,

$$E[C_{i,j}] = E[E[C_{i,j} | N_{i,j} : (i,j) \in I]] = \alpha_i \mu \gamma_i \sum_{l=0}^{\min(j,d)} \beta_{j-l} p_l \quad (8.4)$$

No entanto, do modelo de *Sobre-dispersão de Poisson*, estudado na Subsecção 6.2.2, sabe-se que,

$$E[C_{i,j}] = \hat{\alpha}'_i \hat{\beta}'_j.$$

Associando esta última equação à equação (8.4) e substituindo o parâmetro $\vec{p} = (p_0, \dots, p_d)$ que tem como restrição $\sum_{l=0}^d p_l = 1$, por $\vec{\pi} = (\pi_0, \dots, \pi_n)$, que não tem essa restrição de valores, assume-se que,

$$\begin{aligned} \alpha_i \mu \gamma_i &= \alpha'_i \\ \sum_{l=0}^j \beta_{j-l} \pi_l &= \beta'_j \end{aligned} \quad (8.5)$$

8.2 Estimação dos Parâmetros

Partindo de dois triângulos acumulados, um com o número de sinistros declarados e outro com os montantes pagos, aplica-se o método *Chain Ladder* aos montantes ajustados, obtidos através do sistema de equações (8.6),

$$\begin{aligned} D_{i,n-i}^a &= D_{i,n-i} \\ D_{i,j}^a &= D_{i,j-1}^a \hat{f}_{j-1} \end{aligned} \quad (8.6)$$

em que $D_{i,j}^a$ representa os montantes acumulados ajustados, para o ano de ocorrência $i = 0, \dots, n$ e ano de desenvolvimento $j = 0, \dots, n$, com $i + j \leq n$ e \hat{f}_j são os estimadores dos coeficientes de desenvolvimento do triângulo original. Note-se que este ajustamento estima valores ajustados idênticos aos do modelo de *Sobre-dispersão de Poisson*, referido na Subsecção 6.2.2. O mesmo raciocínio se segue para os dados com o número de sinistros.

De seguida, transformam-se os dados ajustados acumulados em dados incrementais e estimam-se os seguintes parâmetros já referidos na secção anterior:

- $\hat{\alpha}_i = \sum_{j=0}^n N_{i,j}$, que representa o total de sinistros ocorridos no ano i ;
- $\hat{\beta}_j = \frac{N_{n-j,j}}{\hat{\alpha}_{n-j}}$, que representa a proporção de sinistros declarados no ano j relativos ao ano de ocorrência $n - j$ sobre o total de sinistros ocorridos no ano $n - j$;
- $\hat{\alpha}'_i = \sum_{j=0}^n C_{i,j}$, que representa o total de montantes relativos ao ano de ocorrência i ;

- $\hat{\beta}'_j = \frac{C_{n-j,j}}{\hat{\alpha}_{n-j}}$, que representa a proporção de montantes pagos no ano j relativos ao ano de ocorrência $n - j$ sobre o total de montantes pagos no ano $n - j$;
- $\hat{\mu} = \frac{\hat{\alpha}'_0}{\hat{\alpha}_0}$, que representa a média de pagamentos individuais no primeiro ano de ocorrência;

onde $i = 0, \dots, n, j = 0, \dots, n$.

Utilizando as relações de (8.5) e substituindo os parâmetros pelos respectivos estimadores, estima-se:

- $\hat{\pi}_l$, que representa a proporção de sinistros encerrados l anos depois de serem declarados;
- $\hat{\gamma}_i$, que representa a inflação do ano de ocorrência i .

Para a estimação dos parâmetros da distribuição dos pagamentos individuais, considere-se a estimativa do fator média $\hat{\mu}$ já calculado e, segundo Verrall et al. (2010), o estimador do fator variância $\hat{\sigma}^2$, com $i = 0, \dots, n$, é dado por,

$$\hat{\sigma}^2 = \hat{\mu}\hat{\phi} - \hat{\mu}^2$$

onde,

$$\hat{\phi} = \frac{1}{N - (d + 1)} \sum_{(i,j) \in I} \frac{(C_{i,j} - \hat{C}_{i,j}^{DCL})^2}{\hat{C}_{i,j}^{DCL} \hat{\gamma}_i}$$

com $N = \frac{1}{2}(n + 1)(n + 2)$ e $\hat{C}_{i,j}^{DCL} = \sum_{l=0}^{\min(j,d)} N_{i,j-l} \hat{\pi}_l \hat{\mu} \hat{\gamma}_i$.

8.3 Estimação das Reservas

Com os parâmetros estimados anteriormente pode-se obter uma estimativa para a provisão de *RBNS* e de *IBNR*. Para o cálculo da primeira estimativa recorre-se ao número de sinistros que já foram declarados, enquanto que, para o cálculo da estimativa para *IBNR*, utiliza-se o número de sinistros que se esperam vir a declarar, estimado pelo método *Chain Ladder*. Considere-se a divisão da matriz de índices ilustrada na Figura 8.1, onde,

- $J = \{(i, j) : i = 1, \dots, n; j = 1, \dots, n; i + j = n + 1, \dots, 2n\}$
- $K = \{(i, j) : i = 0, \dots, n - 1; j = n + 1, \dots, n + d; i + j = n + 1, \dots, n + d + 1\}$
- $L = \{(i, j) : i = 1, \dots, n; j = n + 1, \dots, n + d; i + j = n + d + 1, \dots, 2n + d\}$

Note-se que o método *Chain Ladder* produz os resultados apenas dentro do conjunto de índices de J , enquanto que, no método *Double Chain Ladder* obtêm-se, adicionalmente, estimativas dentro dos índices $K \cup L$.

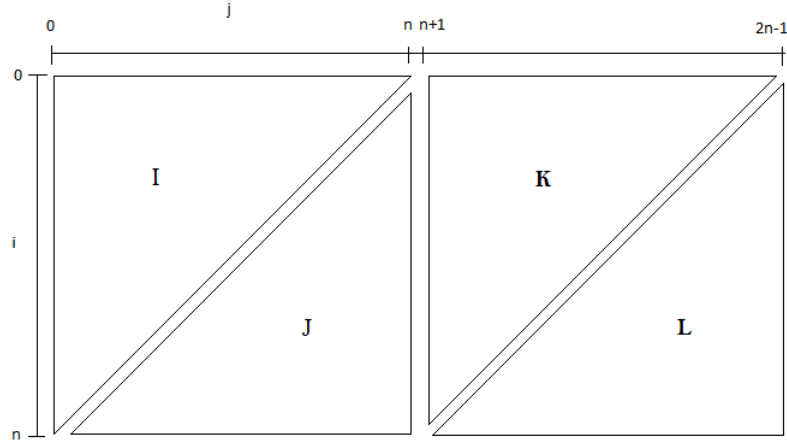


Figura 8.1: Conjunto de índices dos sinistros agregados, assumindo um atraso máximo de $d = n - 1$

Recorrendo à equação (8.3), a separação entre provisão para *IBNR* e para *RBNS* é dada por,

$$\hat{C}_{i,j}^{rbns} = \sum_{l=i-n+j}^j N_{i,j-l} \hat{\pi}_l \hat{\mu} \hat{\gamma}_i$$

e por,

$$\hat{C}_{i,j}^{ibnr} = \sum_{l=0}^{i-n+j-1} \hat{N}_{i,j-l} \hat{\pi}_l \hat{\mu} \hat{\gamma}_i$$

respetivamente, com $(i, j) \in J$. Portanto, o valor estimado de reservas para *RBNS* e para *IBNR*, para o ano de ocorrência i é dado por,

$$\hat{R}_i^{rbns} = \sum_{j=n+1-i}^n \hat{C}_{i,j}^{rbns}$$

e por,

$$\hat{R}_i^{ibnr} = \sum_{j=n+1-i}^n \hat{C}_{i,j}^{ibnr}$$

respetivamente. Ou seja, o total de reservas para o ano de ocorrência i é dado por,

$$\hat{R}_i = \hat{R}_i^{rbns} + \hat{R}_i^{ibnr}$$

Como se pode ver demonstrado em Miranda et al. (2012), para $(i, j) \in J$, tem-se,

$$\hat{C}_{i,j}^{CL} = \hat{C}_{i,j}^{rbns} + \hat{C}_{i,j}^{ibnr}$$

onde $\hat{C}_{i,j}^{CL}$ é a estimativa obtida pelo método *Chain Ladder* descrito na Secção 3.1. Ou seja, considerando as estimativas dentro do conjunto de índices de J , através do método *Double Chain Ladder* conseguem-se obter as mesmas estimativas que no método *Chain Ladder*, sendo esta mais uma especificação estocástica deste último método.

8.4 Fator Cauda

Por outro lado, algo que o método *Chain Ladder* não inclui por si só, ou seja, sem recorrer a métodos externos, é a incorporação do fator cauda na estimação dos montantes a pagar. Este fator cauda consiste em obter uma estimativa dentro do conjunto de índices de $K \cup L$, algo que se estima naturalmente com o método *DCL* através da seguinte fórmula,

$$\hat{R}^{cauda} = \sum_{(i,j) \in K \cup L} \sum_{l=0}^{\min(j,d)} \hat{N}_{i,j-l} \hat{\pi}_l \hat{\mu} \hat{\gamma}_i$$

e ainda separa os montantes que vão ser pagos de sinistros já declarados dos que vão ser pagos de sinistros ainda não declarados.

8.5 *Bootstrap*

A técnica de *Bootstrap* pode também ser aplicada com o método *Double Chain Ladder*. Os algoritmos para a implementação deste método à estimação da provisão para *RBNS* e para *IBNR* foram estudados em Miranda et al. (2011). O algoritmo que permite obter uma estimativa para a provisão para *RBNS* apresenta os seguintes passos:

1. *Estimação dos parâmetros e distribuições.* Inicialmente estimam-se os parâmetros através do método *DCL* referido na Secção 8.2 para os triângulos do número de sinistros e dos montantes pagos. Tal como referido na Secção 8.1, o número de sinistros pagos tem distribuição multinomial com parâmetro \hat{p} . Os pagamentos individuais são estimados pela distribuição gama com parâmetro de forma $\hat{\lambda} = \frac{\hat{\mu}^2}{\hat{\sigma}^2}$ e parâmetro de escala $\hat{\kappa} = \frac{\hat{\sigma}^2}{\hat{\mu}}$.
2. *Bootstrapping-triângulos.* Mantém-se o triângulo com o número de sinistros, mas gera-se um novo triângulo de pagamentos, $C^* = \{C_{i,j}^*, (i, j) \in I\}$, da seguinte forma:
 - Para cada $N_{i,j}$, gera-se o número de sinistros pagos, $N_{i,j}^{*pagos}$, através da equação (8.1) com a distribuição multinomial do passo 1.
 - Estima-se $C_{i,j}^*$ pela distribuição gama com parâmetro de forma $N_{i,j}^{*pagos} \hat{\lambda}$ e parâmetro de escala $\hat{\kappa}$, para cada $(i, j) \in I$, derivado da equação (8.2).

3. *Bootstrapping-parâmetros*. Estimam-se os parâmetros do método *DCL* como no passo 1, mas com o triângulo gerado no passo 2.
4. *Bootstrapping-RBNS*. De seguida,
 - Simula-se o número de sinistros pagos pela distribuição multinomial com parâmetro \hat{p}^* , estimado no passo 3. Estima-se o número de sinistros que originam provisão de *RBNS* através da equação (8.1), cuja denominação passa a ser $N_{i,j}^{*rbns}$, com $(i, j) \in J \cup K$.
 - Estima-se o valor de provisão para *RBNS* através da distribuição gama com parâmetro de forma $N_{i,j}^{*rbns} \hat{\lambda}^*$ e parâmetro de escala $\hat{\kappa}^*$, onde $\hat{\lambda}^* = \frac{\hat{\mu}^{*2}}{\hat{\sigma}^{*2}}$ e $\hat{\kappa}^* = \frac{\hat{\sigma}^{*2}}{\hat{\mu}^*}$, σ^{*2} e $\hat{\mu}^*$ estimados no passo 3.
5. *Técnica Bootstrap*. Repetem-se k vezes os passos de 2 a 4, obtendo k estimativas para a provisão de *RBNS*.

Relativamente ao algoritmo que nos dá uma estimativa média para a provisão de *IBNR*, este segue os seguintes passos:

1. *Estimação dos parâmetros e distribuições*. Mesmo procedimento do passo 1 do algoritmo para a provisão para *RBNS*.
2. *Bootstrapping-triângulos*. Gera-se um novo triângulo de pagamentos, $C^* = \{C_{i,j}^*, (i, j) \in I\}$, e um novo triângulo com o número de sinistros, $N^* = \{N_{i,j}^*, (i, j) \in I\}$, da seguinte forma:
 - O número de sinistros N^* é simulado através da distribuição de *Poisson* (tal como referido na Secção 8.1) com média obtida através dos parâmetros $(\hat{\alpha}, \hat{\beta})$ estimados no passo 1.
 - Estima-se C^* exatamente da mesma forma que no passo 2 do algoritmo anterior.
3. *Bootstrapping-parâmetros*. Com os triângulos (N, C^*) , através do método *DCL*, estimam-se os parâmetros $(\hat{\pi}^*, \hat{\mu}^*, \hat{\sigma}^{*2})$, e estimam-se os parâmetros $(\hat{\alpha}^*, \hat{\beta}^*)$ da mesma forma, mas com os dois triângulos gerados no passo 2, (N^*, C^*) . Obtêm-se as estimativas para o triângulo inferior, $N_{i,j}^*, (i, j) \in J$.
4. *Bootstrapping-IBNR*. De seguida,
 - Para cada $N_{i,j}^*, (i, j) \in J$, simula-se o número de sinistros pela distribuição multinomial de parâmetro \hat{p}^* , estimado no passo 3. Estima-se o número de sinistros que precisam de provisão para *IBNR* através da equação (8.1), cuja denominação passa a ser $N_{i,j}^{*ibnr}$, com $(i, j) \in J \cup K \cup L$.
 - Estima-se o valor de provisão para *IBNR* através da distribuição gama com parâmetro de forma $N_{i,j}^{*ibnr} \hat{\lambda}^*$ e parâmetro de escala $\hat{\kappa}^*$, da mesma forma do algoritmo para a provisão para *RBNS*.

5. *Técnica Bootstrap*. Repetem-se k vezes os passos de 2 a 4, obtendo k estimativas para a provisão para *IBNR*.

De forma similar ao procedimento final do Capítulo 7.1, em cada iteração dos dois ciclos anteriores obtêm-se estimativas de reservas para cada ano de ocorrência, sendo que no final têm-se k amostras para cada uma destas estimativas, $\hat{R}_i^{rbns(1)}, \dots, \hat{R}_i^{rbns(k)}$ e $\hat{R}_i^{ibnr(1)}, \dots, \hat{R}_i^{ibnr(k)}$, e têm-se as respectivas médias dos estimadores,

$$\hat{R}_i^{bs-rbns} = \frac{1}{k} \sum_{b=1}^k \hat{R}_i^{rbns(b)}, \quad i = 0, \dots, n$$

e,

$$\hat{R}_i^{bs-ibnr} = \frac{1}{k} \sum_{b=1}^k \hat{R}_i^{ibnr(b)}, \quad i = 0, \dots, n$$

portanto, tem-se a seguinte reserva média para *RBNS*,

$$\hat{R}^{bs-rbns} = \frac{1}{k} \sum_{b=1}^k \hat{R}^{rbns(b)}$$

e para *IBNR*,

$$\hat{R}^{bs-ibnr} = \frac{1}{k} \sum_{b=1}^k \hat{R}^{ibnr(b)}$$

onde,

$$\hat{R}^{rbns(b)} = \sum_{i=0}^n \hat{R}_i^{rbns(b)}, \quad b = 1, \dots, k$$

e,

$$\hat{R}^{ibnr(b)} = \sum_{i=0}^n \hat{R}_i^{ibnr(b)}, \quad b = 1, \dots, k$$

8.6 *Bornhuetter-Ferguson and Double Chain Ladder*

Denominado por método *BDCL* em Miranda et al. (2013), este método segue os mesmos passos do método *Double Chain Ladder* estudado na Secção 8.2, no entanto, o parâmetro de influência, γ , é ajustado através de informação extra, daí a designação do método associado ao método *Bornhuetter-Ferguson* apresentado na Secção 3.3, mas, neste caso, utiliza-se a matriz de custos totais (montantes pagos com reservas).

Este procedimento consiste em, essencialmente, duas etapas:

1. *Estimação dos parâmetros*. Inicialmente aplica-se o método *DCL* referido na Secção 8.2 com os triângulos do número de sinistros e dos montantes pagos, obtendo os

parâmetros estimados $(\hat{\pi}, \hat{\mu}, \hat{\gamma})$. De seguida, utiliza-se novamente o método *DCL*, mas substituindo o triângulo de montantes pagos pelo triângulo de custos totais, cujas componentes são $C_{i,j}^{BDCL}$, ou seja, os custos totais correspondentes ao ano de ocorrência i e ano de desenvolvimento j . Desta estimação apenas se guardam os valores do parâmetro influência, $\hat{\gamma}'$.

2. *Ajustamento Bornhuetter-Ferguson*. Substitui-se o estimador do parâmetro influência $\hat{\gamma}$ estimado com a matriz de montantes pagos por valores que representam melhor a realidade, tais como $\hat{\gamma}'$. O procedimento para a estimação da provisão para *IBNR*, da provisão para *RBNS* e da reserva final não se altera.

8.7 Comentários

Assim sendo, é de referir que o método *Double Chain Ladder* traz como vantagens, relativamente aos métodos apresentados anteriormente, a obtenção de estimativas semelhantes às do método *Chain Ladder*, a separação imediata entre a provisão que se destina a sinistros já declarados mas ainda não encerrados (*RBNS*) e a provisão para sinistros ocorridos mas não declarados (*IBNR*) e dá-nos ainda uma estimativa para o fator cauda, sem ter de recorrer a metodologias exteriores ao mesmo.

9

Aplicação Prática

Neste capítulo é feita a aplicação prática das metodologias apresentadas nos capítulos anteriores.

9.1 Dados Utilizados

Os dados disponibilizados são referentes à realidade de uma seguradora, no entanto aplicou-se uma percentagem sobre os dados originais, de modo a assegurar a confidencialidade da empresa.

A informação fornecida está organizada por períodos anuais, desde 2005 até 2013, e diz respeito a montantes pagos no ramo *Automóvel*. Na Tabela 9.1 os montantes estão apresentados sob a forma incremental, enquanto que na Tabela 9.2 encontram-se os dados sob forma acumulada.

Tabela 9.1: Montantes Incrementais Pagos (em u.m.)

	0	1	2	3	4	5	6	7	8
2005	419.086,40	825.900,80	661.004,00	669.390,40	486.258,40	557.706,40	79.404,00	106.346,38	30.488,98
2006	635.552,80	1.002.361,60	507.168,00	835.365,60	395.314,40	242.945,60	492.025,72	42.206,49	
2007	633.332,80	1.576.313,60	816.932,00	1.228.104,00	616.338,40	546.089,90	301.171,49		
2008	843.512,80	2.161.005,60	1.079.323,20	497.323,20	474.095,58	409.820,06			
2009	998.948,80	2.026.643,20	1.175.359,20	140.632,48	654.376,50				
2010	1.080.945,60	1.510.047,04	651.697,34	353.525,98					
2011	1.314.587,52	2.293.453,48	1.169.796,61						
2012	766.166,63	1.016.311,94							
2013	444.044,63								

Tabela 9.2: Montantes Acumulados Pagos (em u.m.)

	0	1	2	3	4	5	6	7	8
2005	419.086,40	1.244.987,20	1.905.991,20	2.575.381,60	3.061.640,00	3.619.346,40	3.698.750,40	3.805.096,78	3.835.585,77
2006	635.552,80	1.637.914,40	2.145.082,40	2.980.448,00	3.375.762,40	3.618.708,00	4.110.733,72	4.152.940,21	
2007	633.332,80	2.209.646,40	3.026.578,40	4.254.682,40	4.871.020,80	5.417.110,70	5.718.282,18		
2008	843.512,80	3.004.518,40	4.083.841,60	4.581.164,80	5.055.260,38	5.465.080,44			
2009	998.948,80	3.025.592,00	4.200.951,20	4.341.583,68	4.995.960,18				
2010	1.080.945,60	2.590.992,64	3.242.689,98	3.596.215,96					
2011	1.314.587,52	3.608.041,00	4.777.837,61						
2012	766.166,63	1.782.478,57							
2013	444.044,63								

Analisando a Tabela 9.1 verifica-se que, por exemplo, em 2013, foram pagos montantes no valor de 1.016.311,94 u.m. referentes a sinistros ocorridos em 2012.

O número de sinistros relativos aos custos apresentados na Tabela 9.1 tem a estrutura apresentada na Tabela 9.3.

Tabela 9.3: Número de Sinistros

	0	1	2	3	4	5	6	7	8
2005	12.622	832	422	133	50	22	8	3	2
2006	14.047	855	430	129	55	23	9	3	
2007	15.676	1.147	516	250	59	33	11		
2008	17.141	1.156	403	101	67	22			
2009	21.328	1.571	399	178	53				
2010	19.993	1.282	456	168					
2011	20.331	1.334	298						
2012	21.187	985							
2013	15.344								

9.2 Chain Ladder

Através da equação (3.1) estimam-se os fatores individuais que constam na Tabela 9.4.

Tabela 9.4: Fatores de desenvolvimento individuais

	0	1	2	3	4	5	6	7	8
2005	2,971	1,531	1,351	1,189	1,182	1,022	1,029	1,008	
2006	2,577	1,310	1,389	1,133	1,072	1,136	1,010		
2007	3,489	1,370	1,406	1,145	1,112	1,056			
2008	3,562	1,359	1,122	1,103	1,081				
2009	3,029	1,388	1,033	1,151					
2010	2,397	1,252	1,109						
2011	2,745	1,324							
2012	2,326								
2013									

Analisando a cadência dos fatores de desenvolvimento individuais verifica-se que apresentam um padrão aceitável, ou seja, dentro do mesmo ano de desenvolvimento não se observam discrepâncias e com o evoluir dos anos de desenvolvimento os coeficientes de desenvolvimento individuais vão diminuindo.

Procedendo à aplicação do método *Chain Ladder*, anteriormente descrito na Secção 3.1, aos montantes pagos e começando por assumir que não existem pagamentos para além

do 8º ano de desenvolvimento, ou seja, que $\hat{f}_8 = 1$, obtêm-se os resultados apresentados na Tabela 9.5. Este cenário torna-se irrealista para o ramo em causa, dada a sua grande dimensão, no entanto, na Secção 9.6 procede-se à inclusão do fator cauda.

Tabela 9.5: Aplicação do método *Chain Ladder*

	0	1	2	3	4	5	6	7	8	Provisão Esperada
2005										-
2006									4.186.216,34	33.276,14
2007							5.827.055,98		5.873.746,25	155.464,07
2008						5.841.909,64	5.953.035,10	6.000.734,80	6.074.490,38	535.654,36
2009					5.532.252,24	5.913.713,09	6.026.204,40	6.074.490,38	6.074.490,38	1.078.530,19
2010				4.100.401,63	4.540.559,83	4.853.641,32	4.945.967,83	4.985.598,23	4.985.598,23	1.389.382,27
2011			5.734.256,35	6.538.193,02	7.240.036,29	7.739.252,57	7.886.469,49	7.949.661,16	7.949.661,16	3.171.823,55
2012		2.406.211,07	2.887.881,98	3.292.759,97	3.646.221,76	3.897.636,70	3.971.777,97	4.003.602,52	4.003.602,52	2.221.123,95
2013	1.267.623,34	1.711.195,50	2.053.739,47	2.341.671,57	2.593.038,64	2.771.834,30	2.824.560,44	2.847.192,71	2.847.192,71	2.403.148,08
\hat{f}_j	2,855	1,350	1,200	1,140	1,107	1,069	1,019	1,008	1,000	
										10.988.402,60

A estimativa para o primeiro coeficiente de desenvolvimento $\hat{f}_0 = 2,855$ indica-nos que no ano seguinte a que ocorre o sinistro já se tem um total de montantes pagos 2,855 vezes superior ao que foi pago no ano em que ocorreu.

Assim, segundo esta metodologia, estima-se um valor de reservas para sinistros de 10.988.402,60 u.m..

9.3 Outros *Link Ratios*

Aplicando as metodologias de *Worst Factor* e de *Average Factor*, apresentadas na Subsecção 3.1.3, estimam-se os fatores de desenvolvimento pelas equações (3.3) e (3.2), respetivamente, que constam na Tabela 9.6, e encontram-se os resultados apresentados na Tabela 9.7.

Tabela 9.6: Coeficientes de desenvolvimento de outros *Link Ratios*

	0	1	2	3	4	5	6	7	8
<i>Worst factors</i>	3,5619	1,5309	1,4058	1,1888	1,1822	1,1360	1,0288	1,0080	1,0000
<i>Average factors</i>	2,8869	1,3620	1,2351	1,1441	1,1118	1,0712	1,0195	1,0080	1,0000

É de referir que, tal como esperado, no pior dos cenários a estimativa da provisão é superior, visto que os coeficientes de desenvolvimento são sempre superiores ou iguais aos obtidos pela média, o que significa que os montantes por pagar são os mais elevados possível.

Tabela 9.7: Aplicação de outros *Link Ratios*

Ano de ocorrência	<i>Worst Factor</i>	<i>Average Factor</i>
2005	-	-
2006	33.276,14	33.276,14
2007	211.548,00	158.274,47
2008	972.742,63	550.967,22
2009	1.961.287,64	1.118.675,69
2010	2.357.344,99	1.439.517,94
2011	6.341.467,40	3.485.501,44
2012	4.568.294,29	2.416.209,22
2013	5.191.190,98	2.575.588,57
Total Reservas	21.637.152,06	11.778.010,69

Segundo a metodologia *Worst Factor* estima-se um valor para a provisão de 21.637.152,06 u.m., enquanto que, aplicando *Average Factor*, estima-se que essa provisão seja de 11.778.010,69 u.m.. Note-se ainda que o primeiro valor apresentado é considerando o pior dos cenários, enquanto que o último já se aproxima mais do valor total estimado pelo método *Chain Ladder*, verificando-se também uma semelhança entre os montantes estimados por estes dois métodos em cada ano de ocorrência, uma vez que ambos se tratam de uma média dos *link ratios*.

9.4 Grossing Up

Segundo a metodologia descrita na Secção 3.2, este método baseia-se em percentagens que são apresentadas na Tabela 9.8. Ou seja, para o primeiro ano de ocorrência assume-se que 100% dos pagamentos foram feitos ao longo de 8 anos, sendo que no primeiro ano de desenvolvimento, ou seja, em 2005, foi pago 10,93% do montante que se espera vir a pagar no total, relativo a este mesmo ano de ocorrência.

Os resultados da aplicação de *Grossing Up* e de *Grossing Up - Worst Case* encontram-se na Tabela 9.9.

Tabela 9.8: Percentagens para aplicação de *Grossing Up*

	0	1	2	3	4	5	6	7	8
2005	10,93%	32,46%	49,69%	67,14%	79,82%	94,36%	96,43%	99,21%	100,00%
2006	15,18%	39,13%	51,24%	71,20%	80,64%	86,44%	98,20%	99,21%	
2007	10,78%	37,60%	51,51%	72,41%	82,90%	92,19%	97,31%		
2008	14,05%	50,03%	68,00%	76,28%	84,17%	91,00%			
2009	16,37%	49,59%	68,85%	71,16%	81,88%				
2010	21,53%	51,61%	64,59%	71,64%					
2011	16,23%	44,54%	58,98%						
2012	18,73%	43,57%							
2013	15,47%								

Tabela 9.9: Aplicação de *Grossing Up*

Ano de ocorrência	<i>Grossing Up</i>	<i>Worst Case</i>
2005	-	-
2006	33.276,14	33.276,14
2007	157.791,53	211.548,00
2008	540.612,83	857.067,98
2009	1.105.375,68	1.262.918,64
2010	1.423.819,71	1.759.726,14
2011	3.322.752,43	4.837.005,27
2012	2.308.996,44	3.709.023,21
2013	2.425.592,51	3.713.499,81
Total Reservas	11.318.217,26	16.384.065,19

Segundo a metodologia *Grossing Up* estima-se um valor para a provisão de 11.318.217,26 u.m., enquanto que, aplicando *Grossing Up - Worst Case*, estima-se que essa provisão seja de 16.384.065,19 u.m..

Note-se que o segundo valor apresentado é, mais uma vez, considerando o pior dos cenários, enquanto que o primeiro já se aproxima mais do valor estimado pelo método *Chain Ladder*, tal como aplicando *Average Factors* nos *Link Ratios*.

9.5 *Bornhuetter-Ferguson*

Para a aplicação deste método foram considerados os prêmios emitidos que constam na Tabela 9.10.

Tabela 9.10: Prêmios Emitidos

Ano de ocorrência	Prêmios Emitidos
2005	5.000.000,00
2006	7.500.000,00
2007	6.000.000,00
2008	7.104.604,57
2009	5.000.000,00
2010	4.675.080,75
2011	6.211.188,89
2012	7.000.000,00
2013	5.300.000,00

Tabela 9.11: Taxas de sinistralidade no método *Bornhuetter-Ferguson*

Ano de desenvolvimento	0	1	2	3	4	5	6	7	8
$\frac{\hat{S}_j}{\hat{p}_j}$	0,133	0,256	0,146	0,106	0,086	0,069	0,047	0,012	0,006

A partir da equação (3.4) e dos rácios $\frac{\hat{S}_j}{\hat{p}_j}$, $j = 0, \dots, 8$, que constam na Tabela 9.11, estima-se uma taxa de sinistralidade global $\theta = 85,99\%$.

Aplicando o método *Bornhuetter-Ferguson*, anteriormente apresentado na Secção 3.3, obtêm-se os resultados apresentados na Tabela 9.12.

Tabela 9.12: Aplicação do método *Bornhuetter-Ferguson*

Ano de ocorrência	Reserva
2005	-
2006	45.733,47
2007	107.892,16
2008	462.862,57
2009	668.765,10
2010	1.026.505,73
2011	2.019.479,92
2012	3.298.560,16
2013	3.854.103,49
Total	11.483.902,61

Assim, segundo esta metodologia, estima-se um valor de reservas para sinistros de 11.483.902,61 u.m.. Note-se que, pelo facto de se ter informação extra sobre os prémios emitidos, este método torna-se mais credível do que os anteriormente apresentados. Por exemplo, nos métodos anteriores, a reserva final para o ano de ocorrência de 2011 apresentava um valor muito elevado, derivado do histórico dos montantes pagos, que, agora, aplicando este método, se diluiu devido ao suplemento de informação.

9.6 Fator Cauda

Considerando que o triângulo fornecido não está encerrado, ou seja, que o número de anos de desenvolvimento considerado não é o suficiente para os sinistros serem todos encerrados, surge a necessidade de se projetar os coeficientes de desenvolvimento, como é o caso do ramo em estudo.

Aplicando a metodologia descrita na Secção 3.4 e considerando várias hipóteses para o parâmetro δ , como $\hat{\delta} = \{0, 50; 0, 70; 0, 90\}$, concluí-se que a curva gerada por $\hat{\delta} = 0, 50$ é a que melhor se ajusta à curva dos coeficientes já estimados, tal como se pode observar na Figura 9.1.

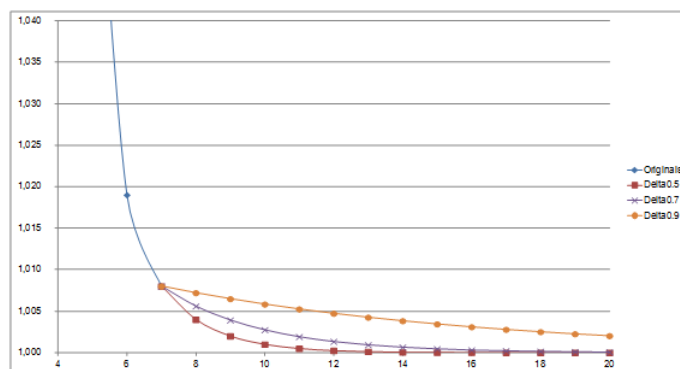


Figura 9.1: Projeção dos Coeficientes de desenvolvimento no método *Chain Ladder*

Na Tabela 9.13 encontram-se as projeções dos coeficientes de desenvolvimento onde o parâmetro é $\hat{\delta} = 0,50$.

Tabela 9.13: Coeficientes de desenvolvimento projetados no método *Chain Ladder*

Ano de desenvolvimento							
0	1	2	3	4	5	6	7
2,855	1,350	1,200	1,140	1,107	1,069	1,019	1,008
8	9	10	11	12	13	14	15
1,004	1,002	1,001	1,001	1,000	1,000	1,000	1,000

Note-se que a partir do 12º ano de desenvolvimento os coeficientes já tomam um valor muito próximo de 1,000, ou seja, a partir deste ano já se pode considerar que os sinistros estão todos encerrados.

Aplicando o método *Chain Ladder* com os coeficientes projetados, ou seja, com $\hat{f}_8 = 1,008$ (derivado da multiplicação dos coeficientes projetados estimados a partir do 8º ano de desenvolvimento), estimam-se as reservas finais apresentadas na Tabela 9.14.

Tabela 9.14: Resultados com os coeficientes projetados no método *Chain Ladder*

Ano de ocorrência	Reservas
2005	30.754,95
2006	66.842,56
2007	202.561,64
2008	583.770,17
2009	1.127.237,40
2010	1.429.358,39
2011	3.235.566,47
2012	2.253.226,11
2013	2.425.977,78
Total	11.355.295,47

Assim, estima-se um valor de provisões para sinistros de 11.355.295,47 u.m., 366.892,86 u.m. acima do estimado não considerando a projeção do fator cauda na Secção

9.2.

No caso de se aplicar o método *Bornhuetter-Ferguson*, a projeção das taxas de sinistralidade é feita com base na função de regressão linear $y = -0,057 \ln(j) + 0,1172$, como se pode observar na Figura 9.2.

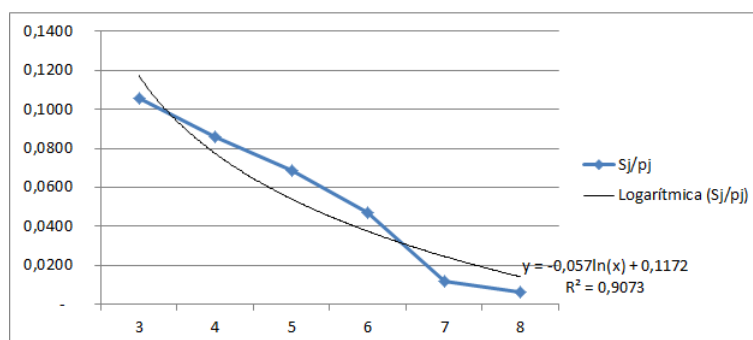


Figura 9.2: Projeção das taxas de sinistralidade no método *Bornhuetter-Ferguson*

Na Tabela 9.15 encontram-se as projeções das taxas de sinistralidade, onde se vê que a partir do 10º ano de desenvolvimento as taxas são negativas, não devendo ser consideradas para o cálculo da projeção e, portanto, $\frac{\hat{S}_9}{\hat{p}_9} = 0,006$.

Tabela 9.15: Taxas de sinistralidade projetadas no método *Bornhuetter-Ferguson*

Ano de desenvolvimento							
0	1	2	3	4	5	6	7
0,133	0,256	0,146	0,117	0,078	0,055	0,038	0,025
8	9	10	11	12	13	14	15
0,015	0,006	-0,001	-0,008	-0,014	-0,019	-0,024	-0,029

Aplicando o método *Bornhuetter-Ferguson* com as taxas de sinistralidade projetadas, obtêm-se os resultados apresentados na Tabela 9.16.

Tabela 9.16: Resultados com as taxas projetados no método *Bornhuetter-Ferguson*

Ano de ocorrência	Reservas
2005	31.415,61
2006	160.146,24
2007	280.889,22
2008	603.863,62
2009	697.875,96
2010	1.015.735,17
2011	2.077.430,28
2012	3.363.870,12
2013	3.903.552,46
Total	12.134.778,68

Assim, estima-se um valor de provisões para sinistros de 12.134.778,68 u.m.,

650.876,07 u.m. acima do estimado na Secção 9.5 quando se considera que o triângulo de montantes pagos está encerrado ao fim de 8 anos de desenvolvimento.

9.7 Teoria da Credibilidade

Para a aplicação da teoria da credibilidade, tal como referido no Capítulo 4, utiliza-se o volume de prémios da Tabela 9.10 e estimam-se as taxas de sinistralidade que constam na Tabela 9.17, de onde se estima que $\hat{\theta} = 85,99\%$, tal como na Secção 9.5.

Tabela 9.17: Taxas de sinistralidade

Ano de desenvolvimento	0	1	2	3	4	5	6	7	8
$\hat{\theta}_j$	13,27%	25,60%	14,61%	10,56%	8,58%	6,86%	4,72%	1,19%	0,61%

O fator de credibilidade, Z_i , pode ser estimado de diferentes formas, como descrito na Tabela 4.1. Os valores estimados para os dados em estudo são apresentados na Tabela 9.18.

Tabela 9.18: Fator de Credibilidade

Ano de ocorrência	<i>Benktander</i>	<i>Optimal</i>	<i>Neuhaus</i>
	z_i	$\frac{z_i}{z_i + \sqrt{z_i}}$	$z_i \times \theta$
2005	1,000	0,500	0,860
2006	0,993	0,499	0,854
2007	0,979	0,497	0,842
2008	0,924	0,490	0,795
2009	0,844	0,479	0,726
2010	0,745	0,463	0,640
2011	0,622	0,441	0,535
2012	0,452	0,402	0,389
2013	0,154	0,282	0,133

Tabela 9.19: Resultados da aplicação de Teoria da Credibilidade

Ano de ocorrência	<i>Coletiva</i>	<i>Individual</i>	<i>Benktander</i>	<i>Optimal</i>	<i>Neuhaus</i>
2005	-	-	-	-	-
2006	45.733,47	29.661,59	29.775,57	37.711,83	32.012,00
2007	107.892,16	122.139,99	121.842,02	114.978,43	119.887,01
2008	462.862,57	448.025,59	449.149,76	455.590,19	451.071,55
2009	668.765,10	920.291,08	881.165,44	789.215,46	851.398,50
2010	1.026.505,73	1.233.231,90	1.180.442,99	1.122.263,44	1.158.869,38
2011	2.019.479,92	2.905.164,83	2.570.261,82	2.409.978,19	2.493.072,24
2012	3.298.560,16	2.161.287,26	2.784.542,68	2.841.356,48	2.856.579,89
2013	3.854.103,49	2.433.980,28	3.634.995,56	3.453.601,35	3.665.702,54
Total Reservas	11.483.902,61	10.253.782,52	11.652.175,84	11.224.695,38	11.628.593,11

Os resultados obtidos, utilizando as diferentes propostas para o fator de credibilidade e recorrendo à equação (4.3) para estimar a reserva final, encontram-se na Tabela

9.19 e mostram que o valor estimado para a provisão para sinistros deve situar-se entre 10.253.782,52 u.m. e 11.652.175,84 u.m., sendo o menor valor estimado pela reserva *Individual* e o valor mais elevado pela proposta de *Benktander*.

9.8 Thomas Mack

A aplicação deste modelo aos dados da sinistralidade observada numa carteira implica a verificação de alguns pressupostos já referidos no Capítulo 5. Este método permite-nos obter estimativas para o erro padrão e intervalos de confiança dos resultados estimados pelo método *Chain Ladder* na Secção 9.2.

9.8.1 Validação dos Pressupostos

O primeiro pressuposto, descrito na Subsecção 5.1.1, é verificado se existir proporcionalidade entre os anos de desenvolvimento. Assim, observando os gráficos da Figura 9.3 que nos dão os pares ordenados $(D_{i,j}, D_{i,j+1})$, conclui-se que as retas de declive \hat{f}_j estão bem ajustadas face aos dados observados, ou seja, aceita-se a hipótese de existência de proporcionalidade entre os anos de desenvolvimento.

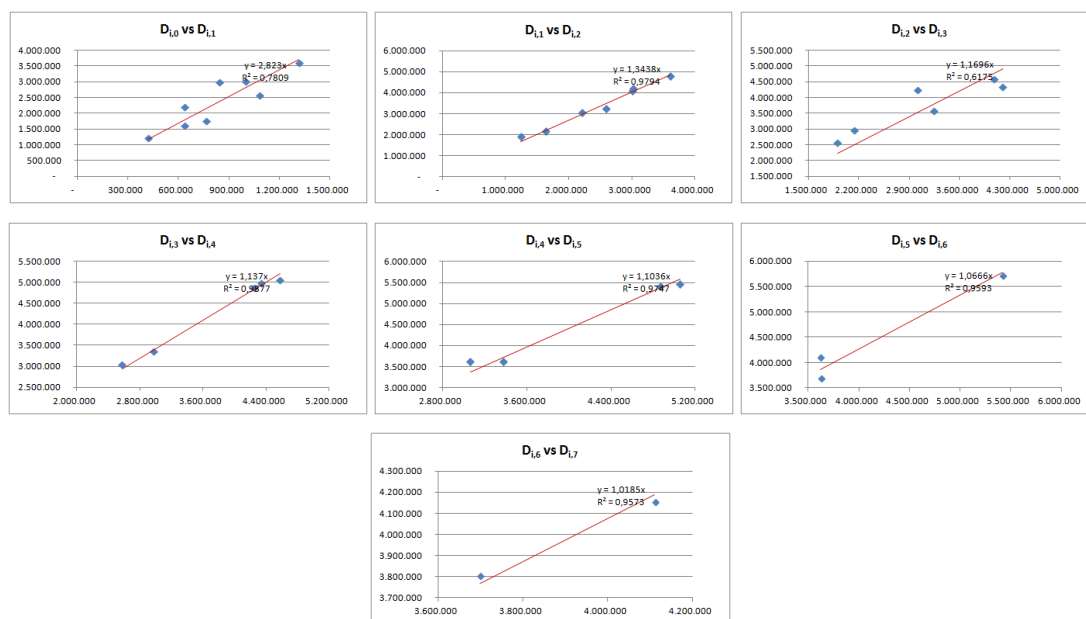


Figura 9.3: Ajustamento dos dados aos fatores de desenvolvimento

Ainda dentro do primeiro pressuposto é necessário verificar a hipótese de inexistência de correlação entre os fatores de desenvolvimento. Para tal recorre-se ao teste de *Spearman*.

Nas Tabelas 9.20 e 9.21 apresentam-se os números de ordem $r_{i,j}$ e $s_{i,j}$, respetivamente.

Tabela 9.20: Matriz **R**

	0	1	2	3	4	5	6	7	8
2005		7	4	5	4	1	2	1	
2006		2	5	2	1	3	1		
2007		5	6	3	3	2			
2008		4	3	1	2				
2009		6	1	4					
2010		1	2						
2011		3							
2012									
2013									

Tabela 9.21: Matriz **S**

	0	1	2	3	4	5	6	7	8
2005		4	6	3	4	3	1	1	
2006		2	2	4	2	1	2		
2007		6	4	5	3	2			
2008		7	3	2	1				
2009		5	5	1					
2010		1	1						
2011		3							
2012									
2013									

Os coeficientes de correlação de *Spearman*, T_j , são estimados com recurso à equação (5.1) e encontram-se na Tabela 9.22, bem como a respetiva variância, $V(T_j)$.

Tabela 9.22: Coeficientes de correlação de *Spearman*

j	1	2	3	4	5	6
T_j	0,6429	0,0286	-0,1000	0,8000	-1,0000	-1,0000
$V(T_j)$	0,1667	0,2000	0,2500	0,3333	0,5000	1,0000

De acordo com a equação (5.2) obtém-se a estimativa para T no valor de 0,1429. Na hipótese de ausência de correlação tem-se que $E(T) = 0$ e $V(T) = 0,0476$. O intervalo de confiança a 95% tem limite inferior de $-0,4277$ e limite superior de $0,4277$. Uma vez que o valor observado para T se encontra dentro do intervalo de confiança, aceita-se a hipótese de não correlação entre os fatores de desenvolvimento.

Para validar o segundo pressuposto é necessário provar que existe independência entre os diferentes anos de ocorrência. A partir dos fatores de desenvolvimento da Tabela 9.4 separam-se os coeficientes em dois conjuntos distintos, S e L , tal como descrito na Subsecção 5.1.2, como se pode ver na Tabela 9.23.

Tabela 9.23: Conjuntos S e L

	0	1	2	3	4	5	6	7	8
2005	L	L	L	L	L	S	L	*	
2006	S	S	L	S	S	L	S		
2007	L	L	L	*	L	*			
2008	L	*	S	S	S				
2009	L	L	S	L					
2010	S	S	S						
2011	S	S							
2012	S								
2013									

Nos anos de desenvolvimento com um número ímpar de fatores de desenvolvimento ignora-se o valor mediano, denotado por *.

A Tabela 9.24 apresenta os momentos das variáveis Z_j e Z , estimados a partir das equações (5.4) a (5.7).

Tabela 9.24: Momentos das variáveis Z_j e Z

k	S_j	L_j	Z_j	z	m	$E(Z_j)$	$V(Z_j)$
0	0	1	0	1	0	0,000	0,000
1	1	1	1	2	0	0,500	0,250
2	1	2	1	3	1	0,750	0,188
3	0	4	0	4	1	1,250	0,438
4	1	3	1	4	1	1,250	0,438
5	4	1	1	5	2	1,563	0,371
6	4	3	3	7	3	2,406	0,554
7	5	1	1	6	2	2,063	0,621
Total			8			9,78	2,858

Uma vez que $Z = 8$ não se situa dentro do intervalo de confiança a 95%, [8, 622; 10, 940], nem a 99%, com o intervalo [8, 258; 11, 305], rejeita-se a hipótese de independência entre os anos de ocorrência. No entanto, o valor da estimativa encontra-se próximo do limite inferior dos dois intervalos, logo, prossegue-se com a aplicação deste modelo. Caso o valor da estimativa estivesse muito distante dos limites do intervalo de confiança, recomendaria-se a estimação de outros coeficientes de desenvolvimento ou até mesmo rejeitar a aplicação deste modelo.

Para se verificar o terceiro pressuposto exige-se que os estimadores dos fatores de desenvolvimento estimados na Tabela 9.5 correspondam aos que têm menor variância. Tal como descrito na Subsecção 5.1.3, para validar este pressuposto é necessário analisar cuidadosamente os gráficos dos resíduos versus os próprios dados acumulados, para cada ano de desenvolvimento. Na Figura 9.4, observa-se que não existe nenhuma tendência nos resíduos, ou seja, aparentemente são aleatórios, tal como pretendido.

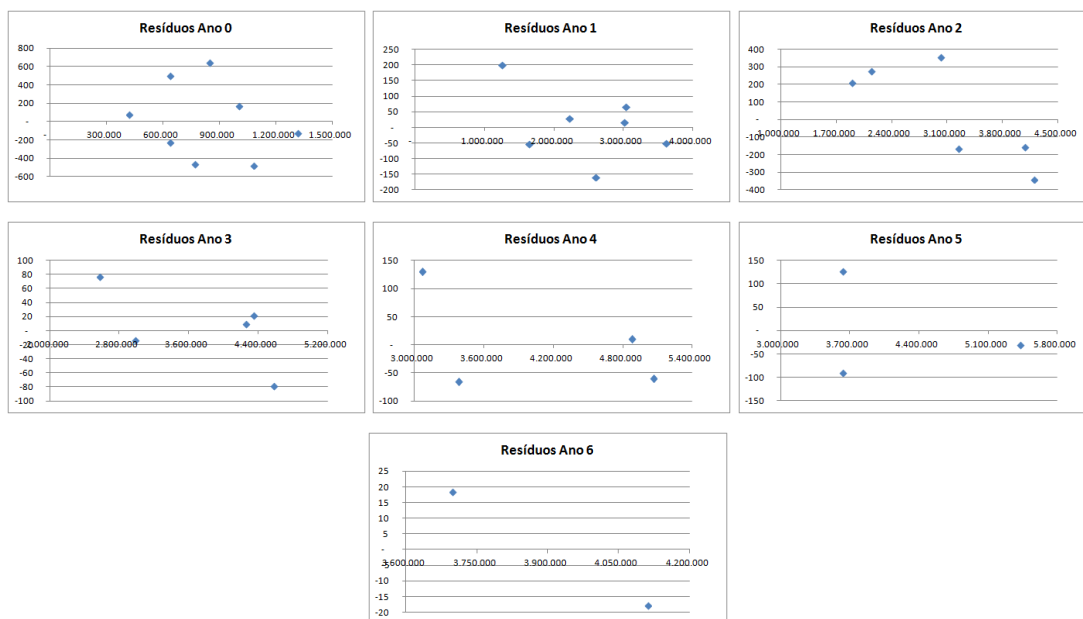


Figura 9.4: Resíduos ponderados

9.8.2 Medidas de Variabilidade

Depois de analisados os pressupostos, utilizam-se as estimativas das reservas apresentadas na Secção 9.2 para obter estatísticas de interesse.

A partir das equações (5.10) e (5.11) estima-se $\hat{\sigma}_j^2$, com $j = 0, \dots, 8$, e a partir de (5.13) estima-se $E\hat{Q}M(\hat{f}_j)$, $j = 0, \dots, 8$, valores esses que se encontram na Tabela 9.25.

Tabela 9.25: Estimativas para a variância e erro padrão

Ano de desenvolvimento	0	1	2	3	4	5	6	7	8
\hat{f}_j	2,855	1,350	1,200	1,140	1,107	1,069	1,019	1,008	1,008
$\hat{\sigma}_j^2$	173953,33	12756,88	83404,78	3250,90	8320,90	12608,84	665,23	35,10	35,42
$E\hat{Q}M(\hat{f}_j)$	2,599E-02	7,365E-04	4,483E-03	1,735E-04	5,085E-04	9,963E-04	8,518E-05	9,224E-06	9,263E-06
$EP(\hat{f}_j)$	0,16123	0,02714	0,06695	0,01317	0,02255	0,03156	0,00923	0,00304	0,00304

Considerando para \hat{f}_8 o valor estimado na Secção 9.6 de 1,008, este situa-se entre \hat{f}_6 e \hat{f}_7 . Com base nesta relação e nas equações (5.12) e (5.14) estimam-se $\hat{\sigma}_8^2$ e $E\hat{Q}M(\hat{f}_8)$, respetivamente.

De seguida estimam-se os erros quadráticos médios de $\hat{D}_{i,j+1}$, \hat{R}_i e \hat{R} , para $i = 0, \dots, 8$ e $j \geq n - i + 1$, através das equações (5.9), (5.8) e (5.15), respetivamente.

Os erros padrão e os intervalos de confiança para as estimativas intervalares apresentam-se na Tabela 9.26, de onde se estima um intervalo de confiança a 95% para o valor da reserva: $IC_{95\%}(R) = [7.066.503, 04; 15.644.087, 89]$, e uma percentagem de erro padrão sobre a provisão esperada total de 19%.

Tabela 9.26: Estimativas para o intervalo de confiança

Ano de ocorrência	Provisão Esperada	Erro Padrão	% erro	Limite Inferior	Limite Superior
2005	30.755	16.496,26	54%	-	63.087,03
2006	66.843	24.906,66	37%	18.026,40	115.658,71
2007	202.562	88.638,58	44%	28.833,22	376.290,06
2008	583.770	337.467,57	58%	-	1.245.194,45
2009	1.127.237	426.850,24	38%	290.626,31	1.963.848,49
2010	1.429.358	403.513,65	28%	638.486,18	2.220.230,60
2011	3.235.566	1.128.758,19	35%	1.023.241,07	5.447.891,87
2012	2.253.226	798.649,11	35%	687.902,62	3.818.549,61
2013	2.425.978	926.815,06	38%	609.453,64	4.242.501,91
Total	11.355.295	2.188.199,61	19%	7.066.503,04	15.644.087,89

9.9 Aplicação dos Modelos Lineares Generalizados

Com o intuito de aplicar o modelo de *Sobre-dispersão de Poisson*, descrito na Subsecção 6.2.2, à estimação da provisão para sinistros, vai-se assumir que as variáveis $C_{i,j}$, $0 \leq i \leq 8$, $0 \leq j \leq 8 - i$ têm distribuição de *Poisson* com sobre-dispersão. Note-se que a restrição implícita neste modelo, referida em (6.5), é admissível, uma vez que todos os montantes incrementais que constam da Tabela 9.1 são estritamente positivos.

De forma a verificar se este modelo se ajusta corretamente aos dados em estudo, apresentam-se na Tabela 9.27 os montantes incrementais ajustados e na Tabela 9.28 os resíduos de *Pearson* estimados pela fórmula (6.11). De seguida, estimam-se os resíduos de *Pearson* corrigidos pela equação (6.12) que constam na Tabela 9.29 e comparam-se esses mesmos resíduos por ano de ocorrência e de desenvolvimento, como na Figura 9.5, de onde se conclui que não existe nenhuma tendência, ou seja, o modelo ajusta-se adequadamente aos dados.

Tabela 9.27: Montantes Incrementais Pagos Ajustados (em u.m.)

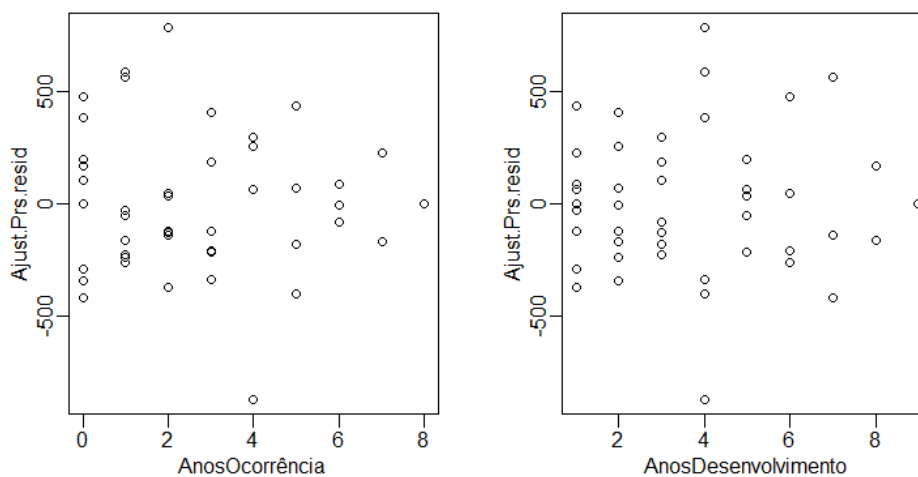
	0	1	2	3	4	5	6	7	8
2005	598.193	1.109.481	597.557	461.457	387.887	338.628	240.864	71.030	30.489
2006	652.877	1.210.905	652.182	503.641	423.345	369.584	262.883	77.523	
2007	916.062	1.699.039	915.087	706.667	594.003	518.569	368.855		
2008	935.867	1.735.772	934.871	721.945	606.845	529.780			
2009	947.370	1.757.107	946.362	730.818	614.304				
2010	777.548	1.442.134	776.720	599.814					
2011	1.239.819	2.299.518	1.238.500						
2012	624.397	1.158.082							
2013	444.045								

Tabela 9.28: Resíduos de *Pearson*

	0	1	2	3	4	5	6	7	8
2005	-231,57	-269,23	82,08	306,10	157,95	376,48	-328,99	132,51	-0,00
2006	-21,44	-189,51	-179,57	467,43	-43,08	-208,31	446,92	-126,84	
2007	-295,40	-94,15	-102,61	620,29	28,98	38,22	-111,44		
2008	-95,47	322,76	149,40	-264,36	-170,41	-164,81			
2009	52,99	203,34	235,40	-690,37	51,13				
2010	344,07	56,55	-141,86	-318,01					
2011	67,15	-4,00	-61,73						
2012	179,41	-131,74							
2013	-								

Tabela 9.29: Resíduos de *Pearson* corrigidos

	0	1	2	3	4	5	6	7	8
2005	-293,57	-341,31	104,05	388,05	200,24	477,27	-417,07	167,99	-0,00
2006	-27,18	-240,25	-227,64	592,58	-54,62	-264,08	566,57	-160,80	
2007	-374,49	-119,36	-130,08	786,36	36,74	48,45	-141,28		
2008	-121,03	409,17	189,40	-335,14	-216,03	-208,94			
2009	67,18	257,78	298,42	-875,21	64,82				
2010	436,19	71,69	-179,84	-403,15					
2011	85,13	-5,07	-78,26						
2012	227,45	-167,01							
2013	-								

Figura 9.5: Gráfico dos Resíduos de *Pearson* corrigidos

Procedendo à aplicação deste modelo, utiliza-se o preditor linear definido em (6.2.1) e obtêm-se as estimativas para os parâmetros do modelo, apresentadas na Tabela 9.30.

Através destas estimativas para os parâmetros e das equações (6.6), (6.7) e (6.8), calculam-se, respetivamente, o valor esperado das variáveis aleatórias $C_{i,j}$, $0 \leq i \leq 8$, $0 \leq j \leq 8$, a estimativa para as reservas por ano de ocorrência e a reserva total. Estas duas últimas estimativas encontram-se na Tabela 9.31. Note-se que o valor da provisão é semelhante ao estimado pelo método *Chain Ladder* na Secção 9.2, tal como esperado.

Tabela 9.30: Estimativas dos parâmetros para o MLG - *Sobre-dispersão de Poisson*

Parâmetros	Estimativa
$\hat{\mu}$	13,3017
$\hat{\alpha}_0$	0,0000
$\hat{\alpha}_1$	0,0875
$\hat{\alpha}_2$	0,4262
$\hat{\alpha}_3$	0,4476
$\hat{\alpha}_4$	0,4598
$\hat{\alpha}_5$	0,2622
$\hat{\alpha}_6$	0,7288
$\hat{\alpha}_7$	0,0429
$\hat{\alpha}_8$	-0,2980
$\hat{\beta}_0$	0,0000
$\hat{\beta}_1$	0,6177
$\hat{\beta}_2$	-0,0011
$\hat{\beta}_3$	-0,2595
$\hat{\beta}_4$	-0,4332
$\hat{\beta}_5$	-0,5690
$\hat{\beta}_6$	-0,9097
$\hat{\beta}_7$	-2,1308
$\hat{\beta}_8$	-2,9765

Tabela 9.31: Estimativas das reservas pelo MLG - *Sobre-dispersão de Poisson*

Ano de ocorrência	Provisão
2005	-
2006	33.276,18
2007	155.463,47
2008	535.653,72
2009	1.078.529,45
2010	1.389.381,63
2011	3.171.822,43
2012	2.221.123,62
2013	2.403.149,10
Total	10.988.399,60

9.10 Simulação *Bootstrap*

Atendendo ao que foi exposto no Capítulo 7, pretende-se aplicar a técnica *Bootstrap* de modo a obter uma estimativa para a provisão média para sinistros e a sua variabilidade, através do método base *Sobre-dispersão de Poisson*, anteriormente aplicado na Secção 9.9.

Com o objetivo de utilizar resíduos independentes e identicamente distribuídos, recorreremos aos resíduos corrigidos que constam na Tabela 9.29, estimados a partir dos dados incrementais ajustados da Tabela 9.27. Antes de se proceder à técnica de *Bootstrap* propriamente dita, é necessário verificar se esses resíduos satisfazem a hipótese de independência. Através da Figura 9.5 verifica-se que os resíduos de *Pearson* corrigidos apresentam um padrão satisfatório, não apresentando tendências sistemáticas que poderiam sugerir a dependência entre si.

Através de um processo de extração, com reposição, obtém-se uma nova amostra de resíduos simulados com o objetivo de construir um conjunto de pseudo-dados. Note-se que os valores nulos dos resíduos são desprezados neste processo. Na Tabela 9.32 tem-se um exemplo dessa reamostragem.

Tabela 9.32: Resíduos de *Pearson* corrigidos simulados (exemplo)

	0	1	2	3	4	5	6	7	8
2005	257,78	477,27	786,36	-417,07	-5,07	-208,94	-121,03	-130,08	67,18
2006	-54,62	-335,14	200,24	-293,57	227,45	104,05	-208,94	-240,25	
2007	-264,08	-141,28	-179,84	36,74	-374,49	-341,31	48,45		
2008	-208,94	257,78	409,17	388,05	167,99	36,74			
2009	48,45	-403,15	298,42	-130,08	36,74				
2010	-121,03	-227,64	-216,03	-160,80					
2011	-227,64	-167,01	67,18						
2012	-27,18	592,58							
2013	-119,36								

Através da equação (7.1) constrói-se um triângulo de pseudo-dados com montantes incrementais, tal como apresentado na Tabela 9.33, tendo por base o exemplo simulado anteriormente.

Tabela 9.33: Triângulo de pseudo-dados (exemplo)

	0	1	2	3	4	5	6	7	8
2005	797.566,38	1.612.199,41	1.205.427,88	178.140,50	384.728,99	217.043,59	181.467,12	36.361,68	42.219,32
2006	608.747,01	842.112,59	813.889,63	295.297,97	571.333,94	432.840,85	155.755,76	10.629,65	
2007	663.307,82	1.514.884,89	743.052,68	737.551,21	305.380,38	272.789,02	398.279,33		
2008	733.740,12	2.075.391,32	1.330.496,97	1.051.659,13	737.709,95	556.521,69			
2009	994.526,67	1.222.712,57	1.236.668,95	619.615,50	643.099,11				
2010	670.828,90	1.168.760,04	586.326,57	475.277,17					
2011	986.345,46	2.046.262,29	1.313.263,13						
2012	602.918,69	1.795.777,97							
2013	364.506,68								

Aplicando o método de *Sobre-dispersão de Poisson*, calculam-se as estimativas para a provisão para esta matriz de pseudo-dados, a seguir apresentadas na Tabela 9.34.

Tabela 9.34: Estimativas da provisão (exemplo)

Ano de ocorrência	Provisão
2005	-
2006	29.892
2007	126.019
2008	635.672
2009	1.018.227
2010	1.120.863
2011	2.885.057
2012	2.988.986
2013	1.972.693
Total	10.777.408

Repetindo este processo 10.000 vezes, obtém-se 10.000 estimativas para as provisões

para cada ano de ocorrência, bem como para a provisão total, representadas no histograma da Figura 9.6.

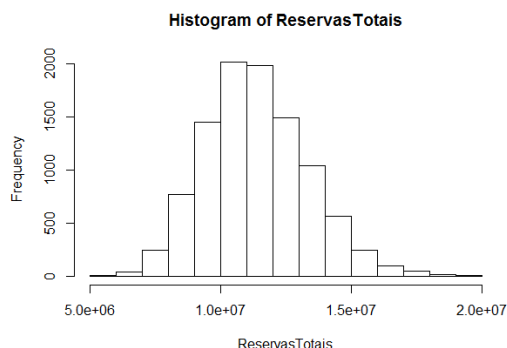


Figura 9.6: Histograma das reservas totais aplicando a técnica *Bootstrap*

Na Tabela 9.35 apresentam-se os resultados obtidos após 10.000 simulações. Os limites estimados para as provisões \hat{R}_i e \hat{R} são obtidos através das equações (7.4) e (7.5), respectivamente. Nos casos em que o limite inferior estimado se situava abaixo de 0, considerou-se limite nulo.

Tabela 9.35: Resultados da aplicação da técnica *bootstrap*

Ano de ocorrência	Média	Desvio Padrão	Variabilidade	Erro de Bootstrap	Limite Inferior	Limite Superior
2005	-	-	-	-	-	-
2006	43.550,33	50.474,18	81.119,22	75.984,87	-	182.203,79
2007	182.781,31	110.530,04	177.637,57	165.193,09	-	479.235,97
2008	573.884,24	185.596,81	298.280,59	293.899,15	-	1.111.685,46
2009	1.126.471,40	276.453,15	444.299,70	425.426,33	244.709,16	1.912.349,75
2010	1.436.544,58	324.613,98	521.701,04	489.971,37	429.055,39	2.349.707,86
2011	3.259.103,88	634.708,85	1.020.067,80	842.826,80	1.519.912,25	4.823.732,61
2012	2.274.989,26	595.112,37	956.430,60	754.647,02	742.042,64	3.700.204,60
2013	2.477.518,23	1.173.236,23	1.885.558,22	1.268.646,38	-	4.889.650,31
Total	11.374.843,22	3.163.231,80	5.385.094,72	2.222.441,10	6.632.495,10	15.344.304,00

Assim, aplicando a técnica de *Bootstrap* estima-se uma provisão média total de 11.374.843,22 u.m., com um limite superior de 15.344.304,00 u.m. para um nível de significância de 5%.

9.11 *Double Chain Ladder*

Tal como descrito no Capítulo 8, inicialmente aplica-se o método *Chain Ladder* aos montantes acumulados ajustados obtidos através da Tabela 9.27 e ao número de sinistros acumulado ajustado obtido através da Tabela 9.36.

Tabela 9.36: Triângulo incremental do número de sinistros ajustado

	0	1	2	3	4	5	6	7	8
2005	12.765	822	307	121	45	21	8	3	2
2006	14.087	907	339	133	49	23	9	3	
2007	16.029	1.032	386	152	56	27	10		
2008	17.125	1.102	412	162	60	29			
2009	21.363	1.375	514	202	75				
2010	19.946	1.284	480	189					
2011	20.178	1.299	486						
2012	20.831	1.341							
2013	15.344								

De seguida, para a aplicação deste método é necessário estimar alguns parâmetros que são apresentados nas Tabelas 9.37 e 9.38.

Tabela 9.37: Estimativas dos parâmetros do método *Double Chain Ladder*

Ano de ocorrência	$\hat{\alpha}$	$\hat{\alpha}'$	$\hat{\gamma}$
2005	14.094,00	3.835.585,77	1,000
2006	15.553,00	4.165.040,60	0,984
2007	17.698,47	5.819.023,52	1,208
2008	18.902,54	5.819.438,78	1,131
2009	23.565,02	6.082.846,98	0,949
2010	22.000,32	4.329.177,50	0,723
2011	22.137,46	7.515.139,29	1,247
2012	22.730,43	3.587.961,57	0,580
2013	16.891,74	2.730.714,83	0,594

Tabela 9.38: Estimativas dos parâmetros do método *Double Chain Ladder*

Ano de desenvolvimento	0	1	2	3	4	5	6	7	8
$\hat{\beta}$	0,908	0,043	0,013	0,008	0,002	0,001	0,001	0,000	0,000
$\hat{\beta}'$	0,163	0,283	0,156	0,082	0,108	0,070	0,052	0,010	0,008
$\hat{\pi}$	0,179	0,303	0,154	0,077	0,109	0,069	0,051	0,006	0,007

Assim, através da equação $\hat{\mu} = \frac{\hat{\alpha}'_0}{\hat{\alpha}_0}$ estima-se uma média de pagamentos individuais no ano de 2005 de 272,14 u.m..

Na Tabela 9.39 encontram-se os resultados utilizando esta metodologia, considerando que o triângulo encerra ao fim dos 8 anos de desenvolvimento, notando-se que o montante total a provisionar estimado é semelhante ao estimado pelo método *Chain Ladder* na Secção 9.2, no entanto é feita a distinção entre provisão para *RBNS* e para *IBNR*.

Tabela 9.39: Resultados pelo método *Double Chain Ladder* sem fator cauda

Ano de ocorrência	RBNS	IBNR	Total
2005	-	-	-
2006	33.173,89	102,29	33.276,18
2007	154.748,79	714,68	155.463,47
2008	532.693,16	2.960,58	535.653,74
2009	1.068.722,43	9.807,01	1.078.529,44
2010	1.368.094,25	21.287,37	1.389.381,62
2011	3.074.791,57	97.030,84	3.171.822,41
2012	2.086.301,76	134.821,89	2.221.123,65
2013	2.141.647,85	261.501,28	2.403.149,13
Total Reservas	10.460.173,71	528.225,94	10.988.399,65

9.11.1 Fator Cauda

Na Tabela 9.40 encontram-se os resultados utilizando esta metodologia com a inclusão do fator cauda, notando-se que o valor estimado para as reservas aumenta em 123.091,91 u.m. do que quando não se considera este fator, mas, ainda assim, é inferior ao estimado quando se considera fator cauda no método *Chain Ladder* na Secção 9.6.

Tabela 9.40: Resultados pelo método *Double Chain Ladder* com fator cauda

Ano de ocorrência	RBNS	IBNR	Total
2005	10.318,23	-	10.318,23
2006	43.942,02	595,64	44.537,66
2007	169.236,94	2.027,68	171.264,62
2008	546.166,61	5.629,90	551.796,51
2009	1.079.978,24	14.892,38	1.094.870,62
2010	1.374.767,91	28.025,63	1.402.793,54
2011	3.080.184,29	113.023,75	3.193.208,04
2012	2.087.467,80	144.426,07	2.231.893,87
2013	2.141.647,85	269.160,60	2.410.808,46
Total Reservas	10.533.709,90	577.781,66	11.111.491,56

Pode-se concluir que, pelo método *Double Chain Ladder*, são necessárias provisões para sinistros no valor de 11.111.491,56 u.m., considerando o fator cauda por questões de prudência. Sendo que, deste valor, apenas 577.781,66 u.m. são estimadas para montantes a pagar respeitantes a sinistros já ocorridos mas ainda não declarados e 10.533.709,90 u.m. é o que se espera vir a pagar por sinistros já declarados mas ainda não encerrados.

9.11.2 *Bootstrap*

A técnica de *Bootstrap* aplicada ao método *Double Chain Ladder* encontra-se descrita na Secção 8.5 e, seguindo esse procedimento, após 10.000 simulações e considerando que os sinistros estão todos encerrados ao fim de 8 anos de desenvolvimento, estimam-se os valores que constam na Tabela 9.41. Sendo que, em média, o total da provisão é de 11.046.208,56 u.m., dos quais 10.511.569,85 u.m. são para montantes a pagar respeitantes

a sinistros já declarados mas ainda não encerrados e 534.638,71 u.m. são estimados para montantes a pagar de sinistros já ocorridos mas ainda não declarados, com um desvio padrão total de 1.823.185,41 u.m..

Tabela 9.41: Resultados pelo método *Bootstrap - Double Chain Ladder* sem fator cauda

Ano de desenvolvimento	Média			Desvio Padrão		
	RBNS	IBNR	Total	RBNS	IBNR	Total
2014	3.936.441,13	60.209,57	3.996.650,70	638.924,26	53.358,11	644.679,04
2015	2.611.550,03	131.678,60	2.743.228,63	517.915,10	81.026,49	530.870,19
2016	1.836.138,19	104.406,21	1.940.544,40	447.459,28	71.803,70	457.564,15
2017	1.213.013,24	80.611,14	1.293.624,38	364.677,62	64.115,82	374.113,44
2018	616.967,25	64.027,06	680.994,31	250.710,37	56.675,99	261.894,70
2019	245.219,56	52.466,83	297.686,39	161.326,97	50.576,50	172.554,63
2020	44.197,04	31.371,46	75.568,50	68.285,05	35.873,36	77.468,10
2021	8.043,41	9.867,84	17.911,25	32.278,78	20.794,90	39.110,76
2022	-	-	-	-	-	-
Total	10.511.569,85	534.638,71	11.046.208,56	1.753.170,85	183.387,71	1.823.185,41

Caso se considere que os 8 anos de desenvolvimento não são suficientes para que todos os sinistros sejam encerrados, inclui-se um fator cauda, estimando os resultados da Tabela 9.42 após 10.000 simulações. Agora, em média, espera-se que o total da provisão seja de 11.164.511,04 u.m., 118.302,48 u.m. acima do resultado anterior, dos quais 10.584.915,75 u.m. são para montantes a pagar respeitantes a sinistros já declarados mas ainda não encerrados e 579.595,29 u.m. são estimados para montantes a pagar de sinistros já ocorridos mas ainda não declarados, com um desvio padrão total de 1.808.867,65 u.m., ligeiramente superior ao que se estima não considerando o fator cauda.

Tabela 9.42: Resultados pelo método *Bootstrap - Double Chain Ladder* com fator cauda

Ano de desenvolvimento	Média			Desvio Padrão		
	RBNS	IBNR	Total	RBNS	IBNR	Total
2014	3.952.080,07	60.187,09	4.012.267,16	639.371,92	52.985,16	645.742,66
2015	2.625.232,56	132.591,64	2.757.824,20	514.113,38	81.490,24	526.741,63
2016	1.850.131,91	104.011,81	1.954.143,72	439.982,44	72.724,08	451.345,94
2017	1.221.172,49	80.964,13	1.302.136,62	359.172,38	63.276,71	368.869,68
2018	629.399,09	66.804,16	696.203,25	253.901,04	57.474,41	263.450,31
2019	250.325,70	56.747,82	307.073,52	164.769,89	53.712,77	176.469,53
2020	47.680,44	43.415,42	91.095,86	74.545,08	46.753,13	89.472,58
2021	8.893,49	20.158,76	29.052,25	34.742,43	30.013,05	46.224,03
2022	-	9.039,67	9.039,67	-	21.748,81	21.748,81
2023	-	3.601,41	3.601,41	-	13.699,12	13.699,12
2024	-	1.319,47	1.319,47	-	8.335,76	8.335,76
2025	-	545,36	545,36	-	5.684,69	5.684,69
2026	-	127,42	127,42	-	1.867,84	1.867,84
2027	-	69,13	69,13	-	2.505,06	2.505,06
2028	-	11,94	11,94	-	857,88	857,88
2029	-	0,06	0,06	-	5,56	5,56
Total	10.584.915,75	579.595,29	11.164.511,04	1.731.747,23	192.793,63	1.808.867,65

9.11.3 *Bornhuetter-Ferguson and Double Chain Ladder*

Tal como explicitado na Secção 8.6, este método difere do modelo *Double Chain Ladder* na medida em que o parâmetro γ passa a ser estimado através do triângulo de custos totais que consta na Tabela 9.43 em vez do triângulo de montantes pagos, denominado agora por γ' na Tabela 9.44.

Tabela 9.43: Custos totais incrementais(em u.m.)

	0	1	2	3	4	5	6	7	8
2005	1.926.646,05	1.456.158,43	1.124.228,85	899.221,46	669.329,30	460.778,34	342.434,83	240.818,04	145.904,49
2006	2.253.246,43	1.703.002,89	1.314.805,40	1.051.655,30	782.792,38	538.888,38	400.483,56	281.640,93	
2007	1.863.408,71	1.408.363,67	1.087.328,85	869.706,76	647.360,24	445.654,45	331.195,27		
2008	1.720.062,99	1.300.023,03	1.003.684,32	802.803,17	597.561,01	411.371,76			
2009	1.683.202,47	1.272.163,86	982.175,62	785.599,30	584.755,42				
2010	1.383.130,30	1.045.369,41	807.078,70	645.546,93					
2011	1.170.601,85	884.740,48	683.064,94						
2012	1.162.068,68	878.291,12							
2013	851.551,72								

Tabela 9.44: Parâmetro γ' do método *Bornhuetter-Ferguson and Double Chain Ladder*

Ano de ocorrência	$\hat{\gamma}'$
2005	1,894
2006	2,008
2007	1,459
2008	1,261
2009	0,989
2010	0,870
2011	0,728
2012	0,700
2013	0,697

Na Tabela 9.45 encontram-se os resultados utilizando esta metodologia sem a inclusão do fator cauda, notando-se que o valor obtido para a estimativa da provisão para sinistros total é inferior ao estimado pelo método *Double Chain Ladder*.

Tabela 9.45: Resultados pelo método *Bornhuetter-Ferguson and Double Chain Ladder* sem fator cauda

Ano de ocorrência	RBNS	IBNR	Total
2005	-	-	-
2006	67.336,13	207,63	67.543,75
2007	185.133,50	855,01	185.988,51
2008	575.813,46	3.200,23	579.013,69
2009	1.116.750,78	10.247,74	1.126.998,51
2010	1.431.286,72	22.270,64	1.453.557,36
2011	1.707.422,76	53.880,94	1.761.303,70
2012	2.283.614,07	147.572,69	2.431.186,75
2013	2.415.498,47	294.939,21	2.710.437,69
Total Reservas	9.782.855,88	533.174,09	10.316.029,97

Na Tabela 9.46 encontram-se os resultados utilizando esta metodologia com a inclusão do fator cauda, notando-se que o valor estimado para as reservas aumenta em 142.166,74 u.m. do que quando não se considera este fator, sendo também inferior ao estimado pelo método *Double Chain Ladder* na Subsecção 9.11.1.

Tabela 9.46: Resultados pelo método *Bornhuetter-Ferguson and Double Chain Ladder* com fator cauda

Ano de ocorrência	RBNS	IBNR	Total
2005	19.545,21	-	19.545,21
2006	89.193,20	1.209,03	90.402,23
2007	202.466,38	2.425,81	204.892,19
2008	590.377,55	6.085,62	596.463,18
2009	1.128.512,42	15.561,65	1.144.074,06
2010	1.438.268,64	29.320,14	1.467.588,78
2011	1.710.417,33	62.761,76	1.773.179,08
2012	2.284.890,38	158.085,19	2.442.975,57
2013	2.415.498,47	303.577,93	2.719.076,41
Total Reservas	9.879.169,58	579.027,12	10.458.196,71

Pode-se concluir que, segundo o método *Bornhuetter-Ferguson and Double Chain Ladder*, são necessárias provisões para sinistros no valor de 10.458.196,71 u.m., considerando o fator cauda por questões de prudência. Sendo que, deste valor, apenas 579.027,12 u.m. são estimadas para montantes a pagar respeitantes a sinistros já ocorridos mas ainda não declarados e 9.879.169,58 u.m. é o que se espera vir a pagar por sinistros já declarados mas ainda não encerrados.

9.12 Comentários sobre os Resultados

Com o intuito de comparar os resultados obtidos nas secções precedentes, na Tabela 9.47 apresentam-se os montantes estimados para a provisão para sinistros do ramo Automóvel por ano de ocorrência, de 2005 a 2013, inicialmente considerando que todos os sinistros estão encerrados ao fim de 8 anos de desenvolvimento.

Numa primeira análise, note-se que as estimativas obtidas pelos métodos *Chain Ladder*, *Sobre-dispersão de Poisson* e pelo *Double Chain Ladder* são semelhantes, tal como esperado, visto que os modelos *Sobre-dispersão de Poisson* e *Double Chain Ladder* são versões estocásticas do modelo *Chain Ladder*, cujos resultados são idênticos aos do modelo determinístico.

Relativamente aos métodos *Average Factors* e *Grossing Up* e à técnica *Bootstrap*, estes apresentam semelhanças entre os montantes estimados com o método *Chain Ladder* em cada ano de ocorrência, visto que estes se baseiam somente na matriz de montantes pagos. No entanto, com os métodos em que se insere informação extra, como no método *Bornhuetter-Ferguson* e *Bornhuetter-Ferguson and Double Chain Ladder*, onde se consideram

adicionalmente os prêmios emitidos e a matriz de custos totais, respetivamente, estimam-se valores mais credíveis do que nos modelos anteriormente referidos. Por exemplo, nestes últimos modelos a reserva final para o ano de ocorrência de 2011 apresenta um valor muito elevado, derivado do histórico dos montantes pagos, que, com suplemento de informação se dilui.

De um modo geral, não considerando o fator cauda, verifica-se que o montante total da provisão se deve situar entre os 10,3 milhões e os 11,8 milhões de u.m..

Com o objetivo de determinar uma estimativa final da provisão, tem-se em conta a inclusão do fator cauda, ou seja, considera-se que os sinistros não estão encerrados ao fim de apenas 8 anos de desenvolvimento, visto se tratar de um ramo com uma grande dimensão e com montantes significativos mesmo nos anos de desenvolvimento mais remotos, obtendo os resultados apresentados na Tabela 9.48.

Com base nestes resultados, conclui-se que a provisão se deve situar entre os 10,4 milhões e os 12,1 milhões de u.m., sendo a menor estimativa obtida pelo método *Bornhuetter-Ferguson and Double Chain Ladder*.

Note-se que todos os cálculos possíveis foram programados no *Software Excel*, sendo que no Anexo A apresentam-se algumas funções do *Software R* que confirmam os valores obtidos ao longo deste capítulo.

Tabela 9.47: Resultados sem fator cauda

Ano de ocorrência	Chain Ladder	Average Factors	Grossing Up	Bornhuetter-Ferguson	Neuhaus	Banktander	Optimal Credibility	Sobre-dispersão de Poisson	Boots-strap	Double Chain Ladder	Bornhuetter-Ferguson and Double Chain Ladder
2005	-	-	-	-	-	-	-	-	-	-	-
2006	33.276,14	33.276,14	33.276,14	45.733,47	32.012,00	29.775,57	37.711,83	33.276,18	43.550,33	33.276,18	67.543,75
2007	155.464,07	158.274,47	157.791,53	107.892,16	119.887,01	121.842,02	114.978,43	155.463,47	182.781,31	155.463,47	185.988,51
2008	535.654,36	550.967,22	540.612,83	462.862,57	451.071,55	449.149,76	455.590,19	535.653,72	573.884,24	535.653,74	579.013,69
2009	1.078.530,19	1.118.675,69	1.105.375,68	668.765,10	851.398,50	881.165,44	789.215,46	1.078.529,45	1.126.471,40	1.078.529,44	1.126.998,51
2010	1.389.382,27	1.439.517,94	1.423.819,71	1.026.505,73	1.158.869,38	1.180.442,99	1.122.263,44	1.389.381,63	1.436.544,58	1.389.381,62	1.453.557,36
2011	3.171.823,55	3.485.501,44	3.322.752,43	2.019.479,92	2.493.072,24	2.570.261,82	2.409.978,19	3.171.822,43	3.259.103,88	3.171.822,41	1.761.303,70
2012	2.221.123,95	2.416.209,22	2.308.996,44	3.298.560,16	2.856.579,89	2.784.542,68	2.841.356,48	2.221.123,62	2.274.989,26	2.221.123,65	2.431.186,75
2013	2.403.148,08	2.575.588,57	2.425.592,51	3.854.103,49	3.665.702,54	3.634.995,56	3.453.601,35	2.403.149,10	2.477.518,23	2.403.149,13	2.710.437,69
Total	10.988.402,60	11.778.010,69	11.318.217,26	11.483.902,61	11.628.593,11	11.652.175,84	11.224.695,38	10.988.399,60	11.374.843,22	10.988.399,65	10.316.029,97

Tabela 9.48: Resultados com fator cauda

Ano de ocorrência	Chain Ladder	Bornhuetter-Ferguson	Double Chain Ladder	Bornhuetter-Ferguson and Double Chain Ladder
2005	30.754,95	31.415,61	10.318,23	19.545,21
2006	66.842,56	160.146,24	44.537,66	90.402,23
2007	202.561,64	280.889,22	171.264,62	204.892,19
2008	583.770,17	603.863,62	551.796,51	596.463,18
2009	1.127.237,40	697.875,96	1.094.870,62	1.144.074,06
2010	1.429.358,39	1.015.735,17	1.402.793,54	1.467.588,78
2011	3.235.566,47	2.077.430,28	3.193.208,04	1.773.179,08
2012	2.253.226,11	3.363.870,12	2.231.893,87	2.442.975,57
2013	2.425.977,78	3.903.552,46	2.410.808,46	2.719.076,41
Total	11.355.295,47	12.134.778,68	11.111.491,56	10.458.196,71

10

Conclusão

Uma das funções do atuário consiste em obter estimativas para a provisão para sinistros, provisão essa de extrema importância para a solvabilidade de uma companhia do ramo Não Vida, visto que se trata de uma rubrica com elevado peso nas contas da seguradora. Esta provisão tem de ser suficiente para cobrir qualquer custo com sinistros que ocorra dentro do contrato estabelecido, mas, simultaneamente, não pode ser excessiva para não comprometer a rentabilidade da companhia.

Para a estimativa desta provisão são utilizados modelos determinísticos, de fácil aplicação, que permitem obter estimativas pontuais, enquanto que a aplicação de modelos estocásticos informam sobre a variabilidade e sobre o erro da estimativa, bem como nos permitem obter intervalos de confiança para a mesma.

Outro aspeto importante na mensuração das provisões é o facto de nem sempre se poder considerar que, num determinado triângulo, todos os sinistros encerram ao fim de n anos de desenvolvimento. Esta conclusão é demonstrada na aplicação prática que se estudou. Tratando-se de um ramo muito significativo, como é o caso do ramo automóvel, e com montantes elevados em todos os anos de desenvolvimento, deve-se incluir um fator cauda, por questões de prudência, fator este que representa todos os anos posteriores ao último ano de desenvolvimento considerado. Este fator pode ser um valor que o atuário considere adequado devido à sua experiência com este ramo nesta seguradora, ou pode ser estimado através de metodologias próprias.

É de se destacar um dos métodos estocásticos mais recentes, o método *Double Chain Ladder*, que, por um lado, permite obter estimativas semelhantes às do método *Chain Ladder* e, por outro lado, dá-nos automaticamente a separação entre a provisão que se destina a sinistros já declarados mas ainda não encerrados (*RBNS*) e a provisão para sinistros ocorridos mas não declarados (*IBNR*). Este modelo estima ainda, instantaneamente, a

inclusão do fator cauda, sem ter de recorrer a metodologias exteriores ao mesmo.

Salienta-se ainda que nos métodos em que se considera informação extra, como os prémios emitidos ou a matriz de custos totais, as provisões estimadas são mais credíveis, como no método *Bornhuetter-Ferguson* e no método *Bornhuetter-Ferguson and Double Chain Ladder*. Ao contrário de outros métodos, estes não falham quando, para um determinado ano de ocorrência, ainda não se verificaram pagamentos, desde que existam prémios emitidos ou custos totais contabilizados para esse mesmo ano.

Por último, relativamente à aplicação prática considerada, concluiu-se que a provisão devia situar-se entre os 10,4 milhões e os 12,1 milhões de u.m., sendo a menor estimativa obtida pelo método *Bornhuetter-Ferguson and Double Chain Ladder*, considerando a inclusão do fator cauda.

Quanto a estudos posteriores, um acompanhamento do desenvolvimento do método *Double Chain Ladder* e de metodologias associadas ao mesmo será de grande interesse, visto se tratar de um método recente que apresentou resultados muito satisfatórios. Propõe-se também refazer todos os cálculos colocando os valores dos triângulos a preços correntes.

Bibliografia

- Benktander, G. (1976). "An Approach to Credibility in Calculating IBNR". *The Actuarial Review*, p. 7.
- Borginho, H. M. M. (2001). "Metodologias Estocásticas de Estimação de Provisões para Sinistros". Tese de mestrado. Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Bornhuetter, R. e R. Ferguson (1972). "The Actuary and IBNR". *Proceedings of the Casualty Actuarial Society* LIX, pp. 181–195.
- Bühlmann, H. (1967). "Experience rating and credibility". *ASTIN Bulletin* 4(3), pp. 199–207.
- Bühlmann, H. e E. Straub (1970). "Glaubwürdigkeit für Schadensätze". *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, pp. 111–133.
- Efron, B. (1979). "Bootstrap methods: another look at the jackknife". *The Annals of Statistics*, pp. 1–26.
- England, P. e R. Verrall (1998). "Standard errors of prediction in claims reserving: a comparison of methods". *Institute of Actuaries, London*.
- England, P. D. e R. J. Verrall (2002). "Stochastic claims reserving in general insurance". *British Actuarial Journal* 8(3), pp. 443–518.
- Hürlimann, W. (2009). "Credible loss ratio claims reserves: the Benktander, Neuhaus and Mack methods revisited". *ASTIN Bulletin* 39(1), pp. 81–99.
- Institute and Faculty of Actuaries (1997). *Claims Reserving Manual*. URL: <http://www.actuaries.org.uk/research-and-resources/pages/claims-reserving-manual>.
- Instituto de Seguros de Portugal. *Autoridade de Supervisão de Seguros e Fundos de Pensões*. URL: <http://www.isp.pt/>.
- Klaus, D. (1996). "An extension of Mack's model for the Chain Ladder Method". *ASTIN Bulletin* 26(2), pp. 247–262.
- Mack, T. (1993). "Measuring the variability of Chain Ladder reserve estimates". *Casualty Actuarial Society*, pp. 101–182.

- Mack, T. (1999). "The standard error of chain ladder reserve estimates: Recursive calculation and inclusion of a tail factor". *ASTIN Bulletin* 29(2), pp. 361–366.
- Mack, T. (2000). "Credible Claims Reserve: The Benktander Method". *ASTIN Bulletin* 30(2), pp. 333–348.
- McCullagh, P. e J. A. Nelder (1989). *Generalized Linear Models*. Chapman and Hall.
- Miranda, M., J. Nielsen e R. Verrall (2013). "Double Chain Ladder and Bornhuetter-Ferguson". *North American Actuarial Journal* 17(2), pp. 101–113.
- Miranda, M. D. M., B. Nielsen, J. P. Nielsen e R. Verrall (2011). "Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers". *ASTIN Bulletin* 41(1), pp. 107–129.
- Miranda, M. D. M., J. P. Nielsen e R. Verrall (2012). "Double chain ladder". *ASTIN Bulletin* 42(1), pp. 59–76.
- Nelder, J. A. e R. W. M. Wedderburn (1972). "Generalized Linear Models". *Journal of the Royal Statistic Society A*. 135, pp. 370–384.
- Neuhaus, W. (1992). "Another pragmatic loss reserving method or Bornhuetter-Ferguson revisited". *Scandinavian Actuarial Journal* 1992(2), pp. 151–162.
- Neuhaus, W. (2004). "On the estimation of outstanding claims". *Australian Actuarial Journal* 10, pp. 485–518.
- Norberg, R. (2004). *Credibility theory*.
- Renshaw, A. E. e R. J. Verrall (1998). "A stochastic model underlying the Chain Ladder technique". *British Actuarial Journal* 4, pp. 903–923.
- Salzmann, R. E. (1984). *Estimated Liabilities for losses & loss adjustment expenses*. Prentice-Hall.
- Verrall, R., J. Nielsen e A. Jessen (2010). "Prediction of RBNS and IBNR claims using claim amounts and claims counts". *ASTIN Bulletin* 40(2), pp. 871–887.
- Waters, H. (1987). "An Introduction to Credibility Theory". *Institute and Faculty of Actuaries*.



Resultados do R

O software aplicado ao exemplo prático foi o R, que vem confirmar os valores já obtidos no capítulo 9.

A.1 Dados Utilizados

A matriz de montantes pagos acumulados é denominada por *MontantesAcumulados*, a matriz de montantes pagos incrementais, do número de sinistros incremental e dos custos totais incrementais são identificados por *MontantesIncrementais*, *NrSinistros* e por *CustosTotais*, respetivamente.

Na listagem A.1 são obtidas algumas estatísticas referentes aos montantes pagos acumulados por anos de desenvolvimento, tais como o montante mínimo, o médio e o máximo.

Listing A.1: Estatísticas dos montantes acumulados

```
1 > summary(MontantesAcumulados)
2           V1           V2           V3           V4
3 Min.      : 419086   Min.      :1244987   Min.      :1905991   Min.      :2575381
4 1st Qu.: 633333   1st Qu.:1746338   1st Qu.:2585831   1st Qu.:3134391
5 Median : 766167   Median :2400320   Median :3242690   Median :3925450
6 Mean    : 792909   Mean    :2388022   Mean    :3340425   Mean    :3721580
7 3rd Qu.: 998949   3rd Qu.:3009787   3rd Qu.:4142396   3rd Qu.:4319858
8 Max.    :1314588   Max.    :3608041   Max.    :4777838   Max.    :4581165
9           V5           V6           V7           V8
10 Min.     :3061639   Min.     :3618709   Min.     :3698749   Min.     :3805095
11 1st Qu.:3375763   1st Qu.:3619186   1st Qu.:3904742   1st Qu.:3892056
12 Median :4871021   Median :4518228   Median :4110735   Median :3979018
```

A. RESULTADOS DO R

```
13 Mean :4271929 Mean :4530062 Mean :4509255 Mean :3979018
14 3rd Qu.:4995960 3rd Qu.:5429104 3rd Qu.:4914508 3rd Qu.:4065980
15 Max. :5055261 Max. :5465081 Max. :5718282 Max. :4152941
16 V9
17 Min. :3835584
18 1st Qu.:3835584
19 Median :3835584
20 Mean :3835584
21 3rd Qu.:3835584
22 Max. :3835584
```

A.2 Chain Ladder e Thomas Mack

Na listagem A.2 obtêm-se valores para a provisão aplicando o método *Chain Ladder* e o erro padrão aplicando o modelo proposto por *Thomas Mack*, não considerando nenhum valor para a cauda do triângulo, ou seja, considerando que os sinistros estão encerrados ao fim de 8 anos de desenvolvimento.

Listing A.2: Resultados da aplicação do modelo proposto por Mack

```
1 MackChainLadder(Triangle = MontantesAcumulados, est.sigma = "Mack")
2
3 Latest Dev.To.Date Ultimate IBNR Mack.S.E CV(IBNR)
4 1 3,835,584 1.000 3,835,584 0 0 NaN
5 2 4,152,941 0.992 4,186,217 33,276 17,460 0.525
6 3 5,718,282 0.974 5,873,745 155,463 84,930 0.546
7 4 5,465,081 0.911 6,000,735 535,654 333,980 0.623
8 5 4,995,960 0.822 6,074,489 1,078,529 422,807 0.392
9 6 3,596,216 0.721 4,985,598 1,389,382 399,803 0.288
10 7 4,777,838 0.601 7,949,660 3,171,822 1,119,398 0.353
11 8 1,782,479 0.445 4,003,603 2,221,124 792,116 0.357
12 9 444,045 0.156 2,847,194 2,403,149 919,349 0.383
13
14 Totals
15 Latest: 34,768,426.00
16 Dev: 0.76
17 Ultimate: 45,756,825.60
18 IBNR: 10,988,399.60
19 Mack S.E.: 2,166,025.27
20 CV (IBNR): 0.20
```

Assim, segundo este método tem-se uma provisão esperada para os montantes a pagar de 10.988.399,60 u.m., com um erro padrão de 2.166.025,27 u.m., tal como já obtido em 9.2 e em 9.8.2, respetivamente.

A.3 Bootstrap

Na listagem A.3 aplica-se a técnica *Bootstrap* com base no modelo *Chain Ladder*, com um ciclo de 10.000 iterações.

Listing A.3: Resultados da aplicação da técnica *Bootstrap*

```

1 > BootChainLadder(Triangle = MontantesAcumulados, R = 10000, process.distr
2 = c("gamma", "od.pois"))
3
4 Latest Mean Ultimate Mean IBNR SD IBNR IBNR 75% IBNR 95%
5 1 3,835,584 3,835,584 0 0 0 0
6 2 4,152,941 4,188,618 35,677 96,010 53,623 223,732
7 3 5,718,282 5,876,282 158,000 186,591 249,797 511,451
8 4 5,465,081 6,002,088 537,007 306,130 711,014 1,096,844
9 5 4,995,960 6,085,738 1,089,778 438,928 1,346,106 1,897,615
10 6 3,596,216 4,996,698 1,400,482 496,190 1,700,204 2,314,692
11 7 4,777,838 7,994,118 3,216,280 853,276 3,735,799 4,733,581
12 8 1,782,479 4,032,104 2,249,625 756,791 2,710,485 3,605,674
13 9 444,045 2,899,173 2,455,128 1,302,098 3,205,878 4,840,527
14
15 Totals
16 Latest: 34,768,426
17 Mean Ultimate: 45,910,404
18 Mean IBNR: 11,141,978
19 SD IBNR: 2,300,700
20 Total IBNR 75%: 12,571,235
21 Total IBNR 95%: 15,126,810

```

Segundo esta técnica, obtém-se uma média para a provisão esperada de 11.141.978 u.m., com um erro padrão de 2.300.700 u.m., semelhante ao obtido na Secção 9.10.

A.4 Double Chain Ladder

Nas listagens A.4, A.5 e A.6 têm-se os parâmetros dos montantes pagos, do número de sinistros e os parâmetros comuns, respetivamente, necessários para a aplicação do método *Double Chain Ladder*.

Listing A.4: Parâmetros dos montantes pagos aplicando DCL

```

1 > clm(MontantesIncrementais)
2 $triangle.hat
3 V1 V2 V3 V4 V5 V6 V7
4 [1,] 598193.1 1109480.8 597556.5 461456.7 387886.5 338628.2 240863.8
5 [2,] 652877.4 1210905.0 652182.6 503641.1 423345.5 369584.1 262882.6
6 [3,] 916062.3 1699039.3 915087.4 706666.6 594002.6 518569.2 368854.6
7 [4,] 935867.4 1735772.2 934871.4 721944.6 606844.8 529780.5 376829.2
8 [5,] 947370.1 1757106.5 946361.8 730818.0 614303.5 536292.0 381460.7
9 [6,] 777547.9 1442133.7 776720.3 599814.1 504185.6 440158.2 313081.4
10 [7,] 1239819.6 2299518.4 1238500.0 956418.6 803936.6 701843.3 499216.2

```

A. RESULTADOS DO R

```

11 [8,] 624397.1 1158081.9 623732.5 481670.9 404878.0 353461.8 251414.9
12 [9,] 444045.0 823579.2 443572.4 342544.1 287932.2 251367.2 178795.7
13          V8          V9
14 [1,] 71029.39 30489.00
15 [2,] 77522.61 33276.18
16 [3,] 108773.16 46690.32
17 [4,] 111124.81 47699.75
18 [5,] 112490.64 48286.03
19 [6,] 92325.96 39630.44
20 [7,] 147216.06 63191.73
21 [8,] 74140.85 31824.58
22 [9,] 52725.86 22632.30
23
24 $alpha
25 [1] 3835584 4186217 5873745 6000735 6074489 4985598 7949660 4003603 2847194
26
27 $beta
28 [1] 0.155958808 0.289259953 0.155792821 0.120309364 0.101128419 0.088285945
29 [7] 0.062797170 0.018518534 0.007948985
30
31 $Fj
32          V2          V3          V4          V5          V6          V7          V8          V9
33 2.854720 1.349924 1.200178 1.140199 1.107345 1.068952 1.019022 1.008013

```

Listing A.5: Parâmetros do número de sinistros aplicando DCL

```

1 > clm(NrSinistros)
2 $triangle.hat
3          V1          V2          V3          V4          V5          V6          V7          V8
4 [1,] 12765.19 821.7439 307.2204 120.8286 44.55377 21.27300 8.335188 2.852343
5 [2,] 14086.82 906.8222 339.0281 133.3385 49.16660 23.47548 9.198163 3.147657
6 [3,] 16029.49 1031.8788 385.7822 151.7267 55.94699 26.71290 10.466648 3.581739
7 [4,] 17125.04 1102.4038 412.1490 162.0967 59.77075 28.53862 11.182005 3.826538
8 [5,] 21362.88 1375.2095 514.1412 202.2098 74.56188 35.60091 13.949152 4.773469
9 [6,] 19946.15 1284.0091 480.0447 188.7998 69.61713 33.23995 13.024080 4.456905
10 [7,] 20178.41 1298.9605 485.6344 190.9982 70.42777 33.62701 13.175736 4.508802
11 [8,] 20831.03 1340.9723 501.3411 197.1756 72.70559 34.71459 13.601874 4.654629
12 [9,] 15344.00 987.7515 369.2846 145.2383 53.55447 25.57054 10.019052 3.428569
13          V9
14 [1,] 2.000000
15 [2,] 2.207068
16 [3,] 2.511437
17 [4,] 2.683084
18 [5,] 3.347051
19 [6,] 3.125083
20 [7,] 3.161473
21 [8,] 3.263723
22 [9,] 2.404037
23
24 $alpha
25 [1] 14094.00 15553.21 17698.09 18907.69 23586.67 22022.46 22278.90 22999.46

```

A. RESULTADOS DO R

```

26 [9] 16941.25
27
28 $beta
29 [1] 0.9057182366 0.0583045177 0.0217979549 0.0085730536 0.0031611873
30 0.0015093658
31 [7] 0.0005913998 0.0002023799 0.0001419044
32
33 $Fj
34      V2      V3      V4      V5      V6      V7      V8      V9
35 1.064374 1.022611 1.008696 1.003179 1.001513 1.000592 1.000202 1.000142

```

Listing A.6: Parâmetros do modelo DCL

```

1 > Parametros<-dcl.estimation(MontantesIncrementais,NrSinistros)
2 delay.par delay.prob inflation severity.mean severity.var
3 1 0.1722 0.1722 1.0000 272.8077 19721220
4 2 0.3083 0.3083 0.9890 269.8120 19290475
5 3 0.1480 0.1480 1.2195 332.6964 29330307
6 4 0.1143 0.1143 1.1662 318.1452 26820764
7 5 0.0972 0.0972 0.9463 258.1681 17661419
8 6 0.0857 0.0857 0.8319 226.9399 13647150
9 7 0.0593 0.0593 1.3112 357.6962 33903850
10 8 0.0128 0.0128 0.6396 174.4990 8068746
11 9 0.0050 0.0023 0.6176 168.4733 7521120
12 mean.factor mean.factor.adj variance.factor
13 1 272.143 272.8077 19721220

```

Listing A.7: Resultados do modelo DCL sem fator cauda

```

1 > dcl.predict(Parametros,Model=1,Tail=FALSE,num.dec=0)
2 Future.years rbns ibnr total clm
3 1 1 3902528 60558 3963087 3963087
4 2 2 2577185 131429 2708614 2708614
5 3 3 1820166 102371 1922537 1922537
6 4 4 1201881 79341 1281222 1281222
7 5 5 626460 63168 689629 689629
8 6 6 265064 51064 316128 316128
9 7 7 54008 30543 84550 84550
10 8 8 12882 9751 22632 22632
11 9 9 0 0 0 NA
12 10 10 0 0 0 NA
13 11 11 0 0 0 NA
14 12 12 0 0 0 NA
15 13 13 0 0 0 NA
16 14 14 0 0 0 NA
17 15 15 0 0 0 NA
18 16 16 0 0 0 NA
19 17 Tot. 10460174 528226 10988400 10988400

```

Pela listagem [A.7](#) retira-se que o valor esperado para a provisão de *RBNS* é de 10.460.174 u.m. e para a provisão de *IBNR* é de 528.226 u.m., obtendo um total de provisão sem fator

cauda de 10.988.400 u.m., tal como na Tabela 9.39.

A.4.1 Fator Cauda

Listing A.8: Resultados do modelo DCL com fator cauda

```

1 > dcl.predict(Parametros,Model=1,Tail=TRUE,num.dec=0)
2   Future.years      rbns      ibnr      total      clm
3 1             1 3908805  60558 3969363 3963087
4 2             2 2586294 131612 2717906 2708614
5 3             3 1832916 102892 1935808 1922537
6 4             4 1215738  80516 1296254 1281222
7 5             5  639689  65796  705486  689629
8 6             6  274514  56007  330521  316128
9 7             7   60461  42753  103215   84550
10 8            8   15292  21191   36484   22632
11 9            9           0  10115   10115      NA
12 10           10           0   3956   3956      NA
13 11           11           0   1546   1546      NA
14 12           12           0    575    575      NA
15 13           13           0    194    194      NA
16 14           14           0     57     57      NA
17 15           15           0     11     11      NA
18 16           16           0      2      2      NA
19 17           Tot. 10533710 577782 11111492 10988400

```

Pela listagem A.8 retira-se que o valor esperado para a provisão de *RBNS* é de 10.533.710 u.m. e para a de *IBNR* é de 577.782 u.m., obtendo um total de provisão com a inclusão do fator cauda de 11.111.492 u.m., tal como na Tabela 9.40.

A.4.2 Bootstrap

Na listagem A.9 é aplicada a técnica *Bootstrap* ao método *Double Chain Ladder*, com um ciclo de 10.000 iterações e não considerando a inclusão do fator cauda.

Listing A.9: Resultados do modelo *Bootsatrap-DCL* sem fator cauda

```

1 > boot<-dcl.boot(dcl.par=Parametros,Ntriangle=NrSinistros,boot.type=2,B=10000,
2 Tail=FALSE)
3 [1] "Please_wait,_simulating_the_distribution..."
4 [1] "Done!"
5   period      rbns   mean.rbns   sd.rbns
6 1         1 3884470.93 3936441.13 638924.26
7 2         2 2567597.08 2611550.03 517915.10
8 3         3 1804111.80 1836138.19 447459.28
9 4         4 1184016.18 1213013.24 364677.62
10 5         5  609148.29  616967.25 250710.37
11 6         6  239715.97  245219.56 161326.97
12 7         7   43711.38   44197.04  68285.05
13 8         8    5958.93    8043.41  32278.78

```

A. RESULTADOS DO R

14	9	9	0.00	0.00	0.00
15	10	10	0.00	0.00	0.00
16	11	11	0.00	0.00	0.00
17	12	12	0.00	0.00	0.00
18	13	13	0.00	0.00	0.00
19	14	14	0.00	0.00	0.00
20	15	15	0.00	0.00	0.00
21	16	16	0.00	0.00	0.00
22	17	Tot.	10338730.56	10511569.85	1753170.85
23		period	ibnr	mean.ibnr	sd.ibnr
24	1	1	60706.33	60209.57	53358.11
25	2	2	131749.78	131678.60	81026.49
26	3	3	102621.39	104406.21	71803.70
27	4	4	79535.21	80611.14	64115.82
28	5	5	63322.62	64027.06	56675.99
29	6	6	51189.01	52466.83	50576.50
30	7	7	30617.30	31371.46	35873.36
31	8	8	9774.52	9867.84	20794.90
32	9	9	0.00	0.00	0.00
33	10	10	0.00	0.00	0.00
34	11	11	0.00	0.00	0.00
35	12	12	0.00	0.00	0.00
36	13	13	0.00	0.00	0.00
37	14	14	0.00	0.00	0.00
38	15	15	0.00	0.00	0.00
39	16	16	0.00	0.00	0.00
40	17	Tot.	529516.15	534638.71	183387.71
41		period	total	mean.total	sd.total
42	1	1	3945177.26	3996650.70	644679.04
43	2	2	2699346.86	2743228.63	530870.19
44	3	3	1906733.19	1940544.40	457564.15
45	4	4	1263551.40	1293624.39	374113.44
46	5	5	672470.91	680994.31	261894.70
47	6	6	290904.99	297686.39	172554.63
48	7	7	74328.68	75568.50	77468.10
49	8	8	15733.44	17911.25	39110.76
50	9	9	0.00	0.00	0.00
51	10	10	0.00	0.00	0.00
52	11	11	0.00	0.00	0.00
53	12	12	0.00	0.00	0.00
54	13	13	0.00	0.00	0.00
55	14	14	0.00	0.00	0.00
56	15	15	0.00	0.00	0.00
57	16	16	0.00	0.00	0.00
58	17	Tot.	10868246.72	11046208.56	1823185.41

Segundo esta técnica, estima-se uma média para a provisão de *RBNS* de 10.511.569,85 u.m. e para a provisão de *IBNR* de 534.638,71 u.m., fazendo um total de provisão de 11.046.208,56 u.m..

Na listagem A.10 é aplicada a técnica de *Bootstrap* ao método *Double Chain Ladder*, com um ciclo de 10.000 iterações, mas considerando a inclusão do fator cauda.

Listing A.10: Resultados do modelo *Bootsatrap-DCL* com fator cauda

```

1 > boot<-dcl.boot(dcl.par=Parametros,Ntriangle=NrSinistros,boot.type=2,B=10000,
2 Tail=TRUE)
3 [1] "Please_wait,_simulating_the_distribution..."
4 [1] "Done!"
5   period      rbns  mean.rbns  sd.rbns
6 1      1 3890904.70 3952080.07 639371.92
7 2      2 2576381.98 2625232.56 514113.38
8 3      3 1818672.45 1850131.91 439982.44
9 4      4 1196113.56 1221172.49 359172.38
10 5     5  619732.22  629399.09 253901.04
11 6     6  246905.46  250325.70 164769.89
12 7     7   47412.88   47680.44  74545.08
13 8     8    6747.97    8893.49  34742.43
14 9     9     0.00     0.00    0.00
15 10    10     0.00     0.00    0.00
16 11    11     0.00     0.00    0.00
17 12    12     0.00     0.00    0.00
18 13    13     0.00     0.00    0.00
19 14    14     0.00     0.00    0.00
20 15    15     0.00     0.00    0.00
21 16    16     0.00     0.00    0.00
22 17  Tot. 10402871.24 10584915.74 1731747.23
23   period      ibnr  mean.ibnr  sd.ibnr
24 1      1  60706.33  60187.09  52985.16
25 2      2 131933.36 132591.64  81490.24
26 3      3 103143.50 104011.81  72724.08
27 4      4  80712.90  80964.13  63276.71
28 5      5  65956.93  66804.16  57474.41
29 6      6  56144.16  56747.82  53712.77
30 7      7  42857.57  43415.42  46753.13
31 8      8  21243.14  20158.76  30013.05
32 9      9   9190.88   9039.67  21748.81
33 10     10   3602.93   3601.41  13699.12
34 11     11   1404.04   1319.47   8335.76
35 12     12    517.71    545.36   5684.69
36 13     13    169.78    127.42   1867.84
37 14     14     47.01     69.13   2505.06
38 15     15     7.81     11.94    857.88
39 16     16     0.93     0.06     5.56
40 17  Tot. 577638.99 579595.30 192793.63
41   period      total  mean.total  sd.total
42 1      1 3951611.03 4012267.16 645742.66
43 2      2 2708315.34 2757824.20 526741.63
44 3      3 1921815.95 1954143.72 451345.94
45 4      4 1276826.46 1302136.61 368869.68
46 5      5  685689.16  696203.25 263450.31

```

A. RESULTADOS DO R

47	6	6	303049.62	307073.52	176469.53
48	7	7	90270.45	91095.86	89472.58
49	8	8	27991.12	29052.25	46224.03
50	9	9	9190.88	9039.67	21748.81
51	10	10	3602.93	3601.41	13699.12
52	11	11	1404.04	1319.47	8335.76
53	12	12	517.71	545.36	5684.69
54	13	13	169.78	127.42	1867.84
55	14	14	47.01	69.13	2505.06
56	15	15	7.81	11.94	857.88
57	16	16	0.93	0.06	5.56
58	17	Tot.	10980510.22	11164511.03	1808867.65

Segundo esta técnica, estima-se uma média para a provisão de *RBNS* de 10.584.915,74 u.m. e para a de *IBNR* de 579.595,30 u.m., fazendo um total de provisão de 11.164.511,03 u.m., ligeiramente acima do obtido não considerando o fator cauda, tal como na Subsecção 9.11.2.

A.4.3 Bornhuetter-Ferguson and Double Chain Ladder

Na listagem A.11 encontram-se os parâmetros necessários para a aplicação deste método. Os resultados, não incluindo o fator cauda, são obtidos pela função que consta na listagem A.12, verificando-se que a provisão para *RBNS* é de 9.782.856 u.m. e para *IBNR* é de 533.174 u.m., fazendo um total de provisão esperada de 10.316.030 u.m., tal como obtido na Tabela 9.45.

Listing A.11: Parâmetros do modelo *BF-DCL*

```

1 > bdcl.parametros<-bdcl.estimation(MontantesIncrementais,NrSinistros,
2 CustosTotais)
3 delay.par delay.prob inflat.DCL inflat.BDCL severity.mean severity.var
4 1 0.1722 0.1722 1.0000 1.8942 516.7635 163967376
5 2 0.3083 0.3083 0.9890 2.0075 547.6624 184161814
6 3 0.1480 0.1480 1.2195 1.4590 398.0209 97271452
7 4 0.1143 0.1143 1.1662 1.2606 343.8983 72616246
8 5 0.0972 0.0972 0.9463 0.9889 269.7702 44684978
9 6 0.0857 0.0857 0.8319 0.8703 237.4222 34611189
10 7 0.0593 0.0593 1.3112 0.7281 198.6276 24224402
11 8 0.0128 0.0128 0.6396 0.7001 191.0022 22400130
12 9 0.0050 0.0023 0.6176 0.6965 190.0158 22169363
13 mean.factor mean.factor.adj variance.factor
14 1 272.143 272.8077 45696934

```

Listing A.12: Resultados do modelo *BF-DCL* sem fator cauda

```

1 > dcl.predict(bdcl.parametros,Model=1,Tail=FALSE,num.dec=0)
2 Future.years rbns ibnr total clm
3 1 1 3774775 60807 3835582 3963087
4 2 2 2377191 131820 2509011 2708614
5 3 3 1613376 102519 1715895 1922537

```

A. RESULTADOS DO R

6	4	4	1056750	79150	1135900	1281222
7	5	5	618926	62985	681911	689629
8	6	6	267096	50805	317901	316128
9	7	7	60211	34091	94302	84550
10	8	8	14529	10998	25526	22632
11	9	9	0	0	0	NA
12	10	10	0	0	0	NA
13	11	11	0	0	0	NA
14	12	12	0	0	0	NA
15	13	13	0	0	0	NA
16	14	14	0	0	0	NA
17	15	15	0	0	0	NA
18	16	16	0	0	0	NA
19	17	Tot. 9782856 533174 10316030 10988400				

Na listagem A.13 tem-se os resultados obtidos por este método considerando o fator cauda, de onde se concluí que a provisão para *RBNS* é de 9.879.169 u.m. e para a de *IBNR* é de 579.027 u.m., fazendo um total de provisão esperada de 10.458.197 u.m., tal como obtido na Tabela 9.46.

Listing A.13: Resultados do modelo *BF-DCL* com fator cauda

1	> dcl.predict(bdcl.parametros, Model=1, Tail=TRUE, num.dec=0)					
2	Future.years	rbns	ibnr	total	clm	
3	1	3786664	60807	3847471	3963087	
4	2	2395350	132191	2527541	2708614	
5	3	1631471	103215	1734687	1922537	
6	4	1073412	80514	1153926	1281222	
7	5	633553	65842	699394	689629	
8	6	277234	56070	333304	316128	
9	7	64793	42770	107563	84550	
10	8	16693	21191	37884	22632	
11	9	0	10082	10082	NA	
12	10	0	3943	3943	NA	
13	11	0	1544	1544	NA	
14	12	0	578	578	NA	
15	13	0	203	203	NA	
16	14	0	60	60	NA	
17	15	0	12	12	NA	
18	16	0	2	2	NA	
19	17	Tot. 9879169 579027 10458197 10988400				