



Leonardo Henrique Costa

**Modelagem de sinistros IBNR com cauda:
chainladder estendido, análise de regressão com
heterocedasticidade e modelagem em espaço de
estado linear**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Ciências Atuariais do Instituto de Gestão de Riscos Financeiros e Atuariais da PUC-Rio.

Orientador: Prof. Cristiano Augusto Coelho Fernandes
Co-orientador: Prof. Adrian Heringer Pizzinga

Rio de Janeiro, março de 2010



Leonardo Henrique Costa

**Modelagem de sinistros IBNR com cauda:
chainladder estendido, análise de regressão com
heterocedasticidade e modelagem em espaço de
estado linear**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Ciências Atuariais do Instituto de Gestão de Riscos Financeiros e Atuariais da PUC-Rio.

Prof. Cristiano Augusto Coelho Fernandes
Orientador
DEE – PUC-Rio

Prof. Adrian Heringer Pizzinga
Co-orientador
IAPUC – PUC-Rio

Prof. Marcelo Cunha Medeiros
DEC – PUC-Rio

Prof. Rodrigo Simões Atherino
DEE – PUC-Rio

Prof.^a Mônica Herz
Coordenadora Setorial do Centro de Ciências Sociais – PUC-Rio

Rio de Janeiro, 13 de março 2010

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Leonardo Henrique Costa

Graduado em Ciências Atuariais pela Universidade Federal de Minas Gerais.

Ficha Catalográfica

Costa, Leonardo Henrique

Modelagem de sinistros IBNR com cauda chainladder estendido, análise de regressão com heterocedasticidade e modelagem em espaço de estado linear / Leonardo Henrique Costa ; orientador: Cristiano Augusto Coelho Fernandes ; co-orientador: Adrian Heringer Pizzinga. – 2010.

123 f. ; 30 cm

Dissertação (Mestrado em Gestão de Riscos Financeiros e Atuariais) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.

Inclui bibliografia

1. Gestão de Riscos Financeiros e Atuariais - Teses. 2. Chain ladder. 3. Filtro de Kalman. 4. Erro médio quadrático. 5. IBNR. 6. Mínimos quadráticos generalizados. 7. Modelos de regressão. 8. Modelos em espaço de estado. 9. Triângulo de runoff. I. Fernandes, Cristiano Augusto Coelho. II. Pizzinga, Adrian Heringer. III. Pontifícia Universidade Católica do Rio de Janeiro. Instituto de Gestão de Riscos Financeiros e Atuariais. IV. Título.

CDD: 368.01

Dedico à minha mãe, ao meu pai, aos meus irmãos e à Gabriela.

Agradecimentos

Agradeço primeiramente a Deus, pelo dom da vida e por ter me dado força neste período do mestrado. Gostaria de agradecer também ao professor Adrian, pelas aulas que foram fundamentais na pesquisa, pela impecável orientação e pela agradável companhia durante este período do mestrado.

Ao Rodrigo Atherino, pela disponibilidade em tirar dúvidas e pela grande ajuda na obtenção de referências bibliográficas em assuntos dentro e fora do tema da pesquisa.

Ao professor Luiz Roberto Cunha, pela dedicação ao Instituto Financeiro e Atuarial da PUC-Rio (IAPUC) e pela ajuda prestada aos alunos em seus mais diversos aspectos, disponibilizando inclusive um curso de Probabilidade no período de férias aos alunos que foi fundamental para que este trabalho se concretizasse.

Agradeço também ao IAPUC e a todos os seus funcionários, pela dedicação com os alunos e a coragem em implementar o primeiro e único mestrado em atuária do Brasil. Em especial agradeço ao professor Cristiano Fernandes, pela competência enquanto professor e coordenador, e pela paciência e interesse pelos alunos.

Agradeço aos professores Ponce de Leon e Washington Junger, pela oportunidade de me desenvolver computacionalmente bem como agradeço à Elizabeth Albuquerque e à Gabriela de Queiroz pela amizade durante todo o período que estive presente na UERJ.

Gostaria de fazer um agradecimento especial à Roberta Chun e a toda Mongeral Seguros e Previdência, pela disponibilidade em ajudar na dissertação sempre que possível.

Gostaria de agradecer a todos os meus colegas do IAPUC que me acompanharam durante todo este período, em especial a Bruna Casotti, pela raça e companhia nas madrugadas de estudo, e ao Caio Azevedo, pelo exemplo de pessoa e pela ajuda dada desde o curso de verão, o qual tinha como objetivo selecionar alunos para o mestrado. Agradeço também ao Eduardo Lourenço, pela amizade e pela companhia nas horas de estudo da disciplina Modelos em Espaço de Estado.

Ao Reinaldo Marques e à Máris Caroline, pela companhia nas batalhas da vida no Rio de Janeiro bem como pela ajuda prestada nos estudos e trabalhos do mestrado.

À minha mãe, meu pai e meus irmãos, bem como à Gabriela (minha namorada e fonte de inspiração), à minha madrinha Imaculada e ao meu primo e grande amigo Fabrício. Um agradecimento especial ao Alexandre, por incentivar-me e orientar-me na parte de programação desde a graduação e pela paciência em tirar minhas dúvidas nesta área.

Gostaria de agradecer a toda minha família, tios, primos e amigos que, de uma forma ou outra, me ajudaram a conquistar mais uma etapa em minha vida.

Finalmente um agradecimento ao meu grande amigo Rodrigo Sânzio, que em vida ofereceu sua amizade sem nunca pedir nada em troca e foi e sempre será um exemplo de vida para mim e para todos os que o conheceram.

Resumo

Costa, Leonardo Henrique; Fernandes, Cristiano Augusto Coelho (Orientador); Pizzinga, Adrian Henrique (Co-orientador). **Modelagem de sinistros IBNR com cauda chainladder estendido, análise de regressão com heterocedasticidade e modelagem em espaço de estado linear**. Rio de Janeiro, 2010. 123p. Dissertação de Mestrado - Departamento de Ciências Atuariais, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho utiliza três metodologias para modelagem de sinistros IBNR apresentados no formato do triângulo de *runoff* com cauda, e verifica, por meio de quatro exercícios empíricos com dados reais, se existe uma abordagem estatisticamente mais eficaz. A primeira metodologia se baseia no método do chain ladder clássico, com uma extensão de cálculo de reserva para ano de calendário. A segunda metodologia baseia-se em modelos de regressão linear com heterocedasticidade, sob o arranjo usual do triângulo via duplo-índice. A terceira insere-se no arcabouço dos modelos de espaço de estado lineares e do filtro de Kalman, considerando, desta vez, a ordenação por linhas do triângulo de Atherino et al. (2010). Para todas as abordagens, efetivam-se derivações teóricas e implementações computacionais tanto dos cálculos de reservas IBNR totais e parciais, resultantes dos modelos estimados, quanto dos correspondentes erros médios quadráticos teóricos. Como conclusões desta Dissertação, citam-se: (i) apesar de superiores ao chain ladder, nenhuma das outras duas abordagens se destaca sistematicamente em relação à outra; (ii) a adoção do efeito cauda se mostrou computacional e tecnicamente viável; e (iii) há fatos estilizados nos dados, modelados sob as três abordagens, que possibilitariam a confecção de *softwares* de estimação de reserva.

Palavras-chave

Chain ladder; filtro de Kalman; erro médio quadrático; IBNR; mínimos quadrados generalizados; modelos de regressão; modelos em espaço de estado; triângulo de *runoff*.

Abstract

Costa, Leonardo Henrique; Fernandes, Cristiano Augusto Coelho (Advisor); Pizzinga, Adrian Henrique (Co-advisor). **Modeling IBNR claims with tail effect: extended chain ladder, heterocedastic linear regression models and linear state space models**. Rio de Janeiro, 2010. 123p. MSc Dissertation – Departamento de Ciências Atuariais, Pontifícia Universidade Católica do Rio de Janeiro.

This work makes use of three methodologies for modeling IBNR data arranged in the runoff triangle with a tail effect, and evaluates their performances in four empirical examples. The first methodology is the traditional chain ladder, duly extended to calculate a reserve corresponding to the calendar year. The second methodology remains on linear regression models with heteroscedastic errors, under the well-established double index notation of the triangle. The third methodology uses the linear state space modeling and the theory of the Kalman filter, adopting, this time, the row-wise ordering proposed by Atherino et al. (2010). For each approach, theoretical results and numerical implementations are obtained, where both the punctual IBNR reserve estimators and their corresponding theoretical mean square errors are considered. The main conclusions from this Dissertation are: (i) even though proving to be superior to the chain ladder, none of the remaining two approaches seems to outperform the other; (ii) the adding of a tail effect does not entail major theoretical and/or computational problems; and (iii) the approaches have uncovered stylized facts that would enable the planning of softwares for IBNR reserve estimation.

Keywords

Chain ladder; generalized least squares; IBNR; Kalman filter; mean square error; regression models; runoff triangle; state space models.

Sumário

1	Introdução	15
1.1.	Motivação e conceitos básicos	15
1.2.	Revisão da Literatura	16
1.2.1.	O chain ladder	16
1.2.2.	Modelos de regressão	17
1.2.3.	Modelos em espaço de estado	17
1.3.	Contribuições desta Dissertação	18
1.4.	Organização do texto	19
2	Triângulo de runoff	21
2.1.	Conceitos básicos	21
2.2.	Adição do efeito cauda	22
2.3.	Extensões da forma básica do triângulo de runoff	24
2.4.	Diferentes ordenações do triângulo de runoff	25
2.5.	Tipos de reserva calculados	27
3	Chain ladder estendido	28
3.1.	Descrição do método chain ladder tradicional	28
3.2.	Extensão do chain ladder para a reserva por ano de calendário	30
3.3.	Apêndice: Provas	31
3.3.1.	Prova do Lema 1	31
3.3.2.	Prova do Teorema 1	32
4	Modelos de regressão linear com heterocedasticidade	33
4.1.	Abordagem via mínimos quadrados ordinários	33
4.2.	Extensão do modelo de regressão: mínimos quadrados generalizados e (ainda) mínimos quadrados ordinários	34
4.3.	Versões factíveis dos EQMs teóricos	37
4.4.	Algoritmo de modelagem e cálculo da reserva IBNR: análise de regressão	42
4.5.	Apêndice: Provas	45
4.5.1.	Prova do Teorema 2.	45
4.5.2.	Prova do Teorema 3	46
4.5.3.	Prova do Teorema 4	46
5	Modelos em espaço de estado	50
5.1.	O modelo em espaço de estado linear e o filtro de Kalman	50
5.2.	O método do acumulador	55
5.3.	Modelos de espaço de estado para reservas IBNR	59
5.4.	Algoritmo de modelagem e cálculo da reserva IBNR: modelos em espaço de estado	63
5.5.	Apêndice: Provas	65
5.5.1.	Prova do Teorema 5	65
5.5.2.	Primeira Prova do Teorema 6	68

5.5.3. Segunda prova do Teorema 6.	71
6 Aplicações	72
6.1. Introdução	72
6.2. O triângulo T1	72
6.2.1. O chain ladder	74
6.2.2. Modelagem utilizando análise de regressão	76
6.2.3. Modelagem utilizando espaço de estado	82
6.3. O trapézio Tp1	87
6.3.1. O chain ladder	89
6.3.2. Modelagem via análise de regressão	90
6.3.3. Modelagem via espaço de estado	96
6.4. Resultados para os demais triângulos	102
6.4.1. O triângulo AFG	102
6.4.2. O triângulo T2	107
6.4.3. O trapézio Tp2	111
6.4.4. O trapézio Tp3	115
7 Conclusões	119
Referências bibliográficas	121

Lista de figuras

Figura 1. O triângulo de runoff na ordenação usual de duplo índice.	21
Figura 2. Triângulo de runoff com valores observados e os valores a serem estimados (em sombreado).	22
Figura 3. Reserva total formada pela seguradora por ano de calendário.	22
Figura 4. O triângulo de runoff adicionando um efeito cauda genérico com até $n+1$ instantes de tempo de atraso.	23
Figura 5. O triângulo de runoff com o efeitos cauda usual e originado da extensão do efeito linha.	24
Figura 6. Formato de trapézio do triângulo de runoff na ordenação usual de duplo índice.	25
Figura 7. Triângulo de runoff na ordenação usual de duplo índice e valores acumulados por linha.	25
Figura 8. Triângulo de runoff na ordenação por tempo de calendário.	26
Figura 9. Triângulo de runoff ordenado por linha.	26
Figura 10. Fluxograma para ajuste do modelo de regressão.	44
Figura 11. Triângulo IBNR com notação diferenciada para valores na cauda (em sombreado).	60
Figura 12. Fluxograma para ajuste do modelo via EE.	64
Figura 13. Triângulo T1: Valores de sinistros IBNR.	73
Figura 14. Diagrama de dispersão das observações do triângulo T1 pelos fatores linha e coluna.	73
Figura 15. Gráfico dos dados do triângulo T1 via ordenação por linhas.	74
Figura 16. Gráfico de resíduos do modelo MQO por fatores linha, coluna ajustados ao T1.	77
Figura 17. Gráfico de resíduos padronizados do modelo MQG por fatores linha, coluna ajustados ao T1.	78
Figura 18. Gráfico QQ de normalidade dos resíduos padronizados para o modelo via MQG ajustado ao T1.	78
Figura 19. Suavização da série, nível e periodicidade de T1 utilizando o modelo M1.	83

Figura 20. Suavização da série, nível e periodicidade de T1 utilizando o modelo M2.	84
Figura 21. Diagnósticos com as inovações padronizadas do modelo M1 (ou M2) ajustado ao triângulo T1.	84
Figura 22. Trapézio Tp1: valores de sinistros IBNR (valores em milhares de dólares australianos).	87
Figura 23. Diagrama de dispersão das observações do trapézio Tp1 pelos fatores linha e coluna.	88
Figura 24. Gráfico dos dados do trapézio Tp1 via ordenação por linhas	88
Figura 25. Gráfico de resíduo do modelo MQO por fatores linha e coluna ajustados ao Tp1.	91
Figura 26. Gráfico de resíduos padronizados do modelo MQG por fatores linha e coluna ajustado ao Tp1.	92
Figura 27. Gráfico QQ de normalidade dos resíduos padronizados para o modelo via MQG ajustado ao Tp1.	92
Figura 28. Diagnósticos com as inovações padronizadas do modelo M1 ajustado ao trapézio Tp1.	97
Figura 29. Diagnósticos com as inovações padronizadas do modelo M2 ajustado ao trapézio Tp1.	97
Figura 30. Diagnósticos com as inovações padronizadas do modelo M1* (ou M2*) ajustado ao Tp1.	98
Figura 31. Estimação da série, nível e periodicidade de Tp1 utilizando M1*	99
Figura 32. Estimação da série, nível e periodicidade de Tp1 utilizando M2*	99
Figura 33. Triângulo AFG: valores de sinistros IBNR (escala em milhares de dólares).	102
Figura 34. O triângulo T2. Valor de pagamento de sinistro por número de corrências.	107
Figura 35. Trapézio Tp2: valores de sinistros IBNR (valores em milhares de dólares).	111
Figura 36. Trapézio Tp3: valores de sinistros IBNR (valores em milhares de dólares).	115

Lista de tabelas

Tabela 1. Reservas calculadas para T1 via chain ladder e correspondentes EQMs teóricos.	75
Tabela 2. Informações analíticas sobre a estimação das reservas para T1 por chain ladder.	75
Tabela 3. Informações analíticas sobre a estimação de reservas em T1 via modelo de regressão heterocedástico (p-valores em parêntesis)	79
Tabela 4. Medidas de dependência no triângulo T1 para os resíduos padronizados referentes à estimação via MQG sob diferentes ordenações das observações (p-valores em parêntesis)	80
Tabela 5. Reservas calculadas para T1 via análise de regressão (com e sem heterocedasticidade) e correspondentes EQMs teóricos.	80
Tabela 6. Resultados das estimções dos parâmetros estimados do modelo em EE para o triângulo T1.	83
Tabela 7. Informações analíticas sobre a estimação de reservas em T1 via modelo em espaço de estado (p-valores em parêntesis).	85
Tabela 8. Reservas calculadas para T1 via modelo em EE M1 (ou M2) e correspondentes EQMs teóricos.	86
Tabela 9. Reservas calculadas para Tp1 via chain ladder e correspondentes EQMs teóricos.	89
Tabela 10. Informações analíticas sobre a estimação das reservas em Tp1 por chain ladder.	90
Tabela 11. Informações analíticas sobre a estimação de reservas em Tp1 via modelo de regressão heterocedástico (p-valores em parêntesis).	93
Tabela 12. Medidas de dependência no trapézio Tp1 para os resíduos padronizados referentes à estimação via MQG sob diferentes ordenações das observações (p-valores em parêntesis)	94
Tabela 13. Reservas calculadas para Tp1 via análise de regressão (com e sem heterocedasticidade) e correspondentes EQMs teóricos.	95
Tabela 14. Parâmetros estimados do modelo em EE para o trapézio Tp1.	99

Tabela 15. Informações analíticas sobre a estimação de reservas em Tp1 via modelo em espaço de estado (p-valores em parêntesis).	100
Tabela 16. Reservas calculadas para Tp1 via modelo em EE M1* (ou M2*) e correspondentes EQMs teóricos.	101
Tabela 17. Informação analítica sobre a estimação do AFG via o chain ladder, análise de regressão e EE (p-valores em parêntesis)	104
Tabela 18. Reservas estimadas de AFG para cada método (e raiz dos EQMS entre parêntesis).	105
Tabela 19. Medidas de dependência no triângulo AFG para os resíduos padronizados referentes à estimação via MQG e EE sob diferentes ordenações das observações (p-valores em parêntesis)	106
Tabela 20. Informação analítica sobre a estimação do T2 via chain ladder, análise de regressão e EE (p-valores em parêntesis)	108
Tabela 21. Reservas estimadas do T2 para cada método (e raiz dos EQMs entre parêntesis)	109
Tabela 22. Medidas de dependência no triângulo T2 para os resíduos padronizados referentes à estimação via MQG e EE sob diferentes ordenações das observações (p-valores entre parêntesis)	110
Tabela 23. Informação analítica sobre a estimação do Tp2 via o chain ladder, análise de regressão e EE (p-valores em parêntesis).	112
Tabela 24. Reservas estimadas de Tp2 para cada método (e raiz dos EQMS entre parêntesis).	113
Tabela 1. Medidas de dependência no trapézio Tp2 para os resíduos padronizados referentes à estimação via MQO e EE sob diferentes ordenações das observações (p-valores em parêntesis)	114
Tabela 2. Informação analítica sobre a estimação do Tp3 via o chain ladder, análise de regressão e EE (p-valores em parêntesis)	116
Tabela 27. Reservas estimadas do Tp3 para cada método (e raiz dos EQMs entre parêntesis)	117
Tabela 28. Medidas de dependência no trapézio Tp3 para os resíduos padronizados referentes à estimação via MQG e EE sob diferentes ordenações das observações (p-valores entre parêntesis)	118