FUNDAÇÃO GETULIO VARGAS

ESCOLA DE ECONOMIA DE SÃO PAULO

CARLOS EDUARDO DA SILVA MATOS

MODELO DE REGRESSÃO BASEADO NO LOGARITMO DAS INDENIZAÇÕES INCREMENTAIS – UMA ABORDAGEM ALTERNATIVA AO MODELO DE CHAIN LADDER PADRÃO

SÃO PAULO
2019

CARLOS EDUARDO DA SILVA MATOS

MODELO DE REGRESSÃO BASEADO NO LOGARITMO DAS INDENIZAÇÕES INCREMENTAIS – UMA ABORDAGEM ALTERNATIVA AO MODELO DE CHAIN LADDER PADRÃO

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas, como requisito para obtenção do título de Mestre em Economia.

Área de Concentração:
Engenharia Financeira

Orientador:
Prof. Dr. João Luiz Chela

SÃO PAULO
2019

Matos, Carlos Eduardo da Silva.

Modelo de regressão baseado no logaritmo das indenizações incrementais: uma abordagem alternativa ao modelo de *Chain Ladder* padrão / Carlos Eduardo da Silva Matos. - 2019.

61 f.

Orientador: João Luiz Chela.

Dissertação (mestrado profissional MPFE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Seguros - Brasil. 2. Seguros - Métodos estatísticos. 3. Análise de regressão. 4. Indenização. 5. Atuária. I. Chela, João Luiz. II. Dissertação (mestrado profissional MPFE) – Escola de Economia de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 368(81)

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O

Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

CARLOS EDUARDO DA SILVA MATOS

MODELO DE REGRESSÃO BASEADO NO LOGARITMO DAS INDENIZAÇÕES INCREMENTAIS – UMA ABORDAGEM ALTERNATIVA AO MODELO DE CHAIN LADDER PADRÃO

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas, como requisito para obtenção do título de Mestre em Economia.

Área de Concentração:
Engenharia Financeira

Data da aprovação:
___/___/_____

Banca examinadora:

____
Prof. Dr. João Luiz Chela
(Orientador)
Fundação Getulio Vargas

____
Prof. Dr. Alexandre de Oliveira
Fundação Getulio Vargas

____
Prof. Dr. Oswaldo Luiz do Valle Costa
Universidade de São Paulo

RESUMO

Com o objetivo de se adicionar uma estrutura estatística ao atual modelo de *Chain Ladder* para se obter não somente uma estimativa pontual para a projeção de indenizações futuras e para a provisão de sinistros ocorridos e não avisados (IBNR) mas sim uma estimativa para a variabilidade dessa projeção além de se fazer uma análise residual das estimativas obtidas. Este trabalho foca na implementação do modelo de regressão baseado no logaritmo natural das indenizações incrementais apresentada por Cristofides (1997) através de uma abordagem de regressão linear múltipla, onde os fatores de desenvolvimento do atual modelo de *Chain Ladder* passariam a ser interpretados como os coeficientes dessa regressão linear múltipla. A variável dependente “Y” seria o logaritmo natural das indenizações incrementais e cada coluna da matriz *design* passariam a ser a variável independente. O modelo foi implementado com dados de uma carteira de seguros patrimoniais considerando o período de 2012 a 2018. Os resultados obtidos por este modelo foram comparados com o modelo de *Chain Ladder.*

Palavras-chave: Seguros - Brasil, Seguros - Métodos estatísticos, Análise de regressão, Indenização e Atuária.

ABSTRACT

The main goal of this paper is to add a statistical structure on pure Chain Ladder method to obtain the variability of ultimate claims and IBNR provision besides others analysis like residual analysis of claims estimation. This paper focuses on regression implementation based on Log of incremental payments of claims as described by Cristofides (1997). This considers the link ratio factor from Chain Ladder method as coefficient of this multiple linear regression. The dependent variable "Y" is the Log of incremental payments and all columns from design matrix is the independent variable. The model was implemented using information from property insurance database considering information of claims between 2012 – 2018. The results obtained was compared with pure Chain Ladder method.

Keywords: Insurance - Brazil, Insurance - Statistical Methods, Regression Analysis, Claims and Actuarial.

DEDICATÓRIA

Ao meu avô José de Matos (In Memoriam)

AGRADECIMENTOS

À Deus pela oportunidade de poder continuar os meus estudos.

À minha esposa, Fernanda Fujimoto, por toda paciência, incentivo e principalmente por não me deixar desistir nunca dos meus sonhos.

Aos meus pais Euzébio e Alzira e minha irmã Carolina por sempre me apoiarem em todas as minhas escolhas.

Aos professores presentes nesta banca pela disponibilidade em poderem participar da minha defesa.

LISTA DE FIGURAS

LISTA DE TABELAS

Sumário

1. Introdução

1.1. O Mercado de Seguros Brasileiro

O mercado segurador brasileiro teve seu início em 1808. Com vinda da família real portuguesa para o Brasil houve a abertura dos portos ao comércio internacional. Naquele ano fundou-se a primeira companhia seguradora do Brasil, a então conhecida Companhia de Seguros Boa Fé cujo objetivo era atuar com seguros marítimos.

Com a intensificação das atividades econômicas muitas companhias seguradoras nacionais apareceram oferecendo outros tipos de seguros. Somente 1862 as primeiras filiais de companhias estrangeiras apareceram oferecendo produtos mais sofisticados e em linha com os mercados internacionais.

Atualmente o mercado de seguros brasileiro está estruturado com base nos seguintes agentes:

- Conselho Nacional de Seguros Privados (CNSP);
- Superintendência de Seguros Privados (SUSEP);
- Resseguradores;
- Sociedades Seguradoras; e
- Corretores de Seguros.

O Conselho Nacional de Seguros Privados (CNSP) é o órgão máximo do setor de seguros sendo responsável por fixar as diretrizes, normas da política de seguros, resseguros, previdência privada aberta e capitalização. O órgão é composto por representantes do ministério da fazenda, ministério da justiça, ministério da previdência social, da SUSEP, banco central e representante da comissão de valores mobiliários.

A Superintendência de Seguros Privados (SUSEP) é uma autarquia vinculada ao ministério da fazenda responsável pela fiscalização e controle do mercado de seguros, resseguradores, previdência privada aberta e capitalização.

Companhia Seguradoras são empresas subordinadas à fiscalização da SUSEP e as diretrizes do CNSP sendo o objeto central de seus negócios a venda de contratos de seguros.

Resseguradores são empresas também subordinadas a fiscalização da SUSEP e as diretrizes do CNSP, porém atuam junto as seguradoras, aceitando riscos em sua totalidade ou não sendo agentes de pulverização dos riscos das companhias seguradoras.

O corretor de seguros pode ser uma empresa ou pessoa física que intermedia a aquisição de um seguro entre a companhia seguradora e o seu cliente. De acordo com a legislação brasileira toda celebração de um contrato de seguros deve ser intermediada por um corretor de seguros.

Nos últimos anos o setor de seguro brasileiro experimentou uma fase de crescimento em relação aos prêmios emitidos impulsionado pelas obras realizadas por governos anteriores, novos produtos passaram a ser oferecidos, companhias seguradoras internacionais passaram a ter escritórios de representação no Brasil mudando toda a dinâmica do setor.

A SUSEP através da circular n°535/2016 estabeleceu que os produtos de seguros devem ser compreendidos dentro dos seguintes grupos:

Tabela 1 - Grupos de Seguros

Grupos	**Características**
01 - Patrimonial	Seguros contra incêndio, desmoronamento, roubo de imóveis residenciais, empresarial e condominiais etc.

03 - Responsabilidades	Seguro contra danos causados a terceiro causado de forma involuntária pelo segurado.
05 - Automóvel	Seguros roubos e acidentes de carros, além do seguro obrigatório DPVAT.
06 - Transportes	Seguro de transportes nacional e internacional de cargas além da responsabilidade civil do transportador, e operador.
07 - Riscos Financeiros	Envolve os seguros de garantias (Fiança locatícia, execução de obra pública etc.).
09 - Pessoas Coletivo	Seguro na forma coletiva para vida, acidentes pessoais, prestamistas etc.
10 - Habitacional	Seguro com cobertura de morte, invalidez, danos ao imóvel financiado etc.
11 – Rural	Seguros agrícolas, pecuário etc
13 - Pessoas Individual	Seguro na forma individual para vida, acidentes pessoais, prestamistas etc.
14 – Marítimos	Seguros para operadores portuários, responsabilidade civil para embarcações etc.
15 - Aeronáuticos	Seguros para responsabilidade civil para aeronaves, hangar, transportador aéreo, satélites etc.

16 - Microsseguros	Microsseguros de seguros de vida, previdência e de danos gerais.
17 – Petróleo	Seguro contra riscos de petróleo.
18 – Nucleares	Seguro contra riscos nucleares
19 – Saúde	Seguro Saúde.
20 - Aceitações do Exterior	Seguros contratados no exterior.
21 - Sucursais no Exterior	Seguro de sucursais de seguradoras no exterior.
22 - Pessoas EFPC	Seguro para a transferência de riscos biométricos, sobrevivência do assistido e de vida de entidades de previdência complementar fechada (Fundos de pensão).

Conhecido todos os grupos onde a SUSEP permite as seguradoras comercializar os seus produtos, uma outra divisão bastante comum feita no setor é a divisão em produtos de vida e de produtos de não vida (seguros gerais) formando assim dois grandes grupos de análise.

Abaixo apresentamos a evolução de alguns indicadores do setor nos últimos anos. Vale ressaltar que todos os dados forma obtidos do sistema de estatística da SUSEP (SES).

Em relação ao montante de prêmio emitido do setor observamos nas figuras abaixo uma tendência de aumento na arrecadação do setor. Quando comparamos o ano de 2018 observamos um aumento de 6,2% (6,1 bilhão de reais) quando comparado com 2017, sendo que os produtos de vida cresceram 9,7% e os produtos de seguros gerais cresceram 4,7%.

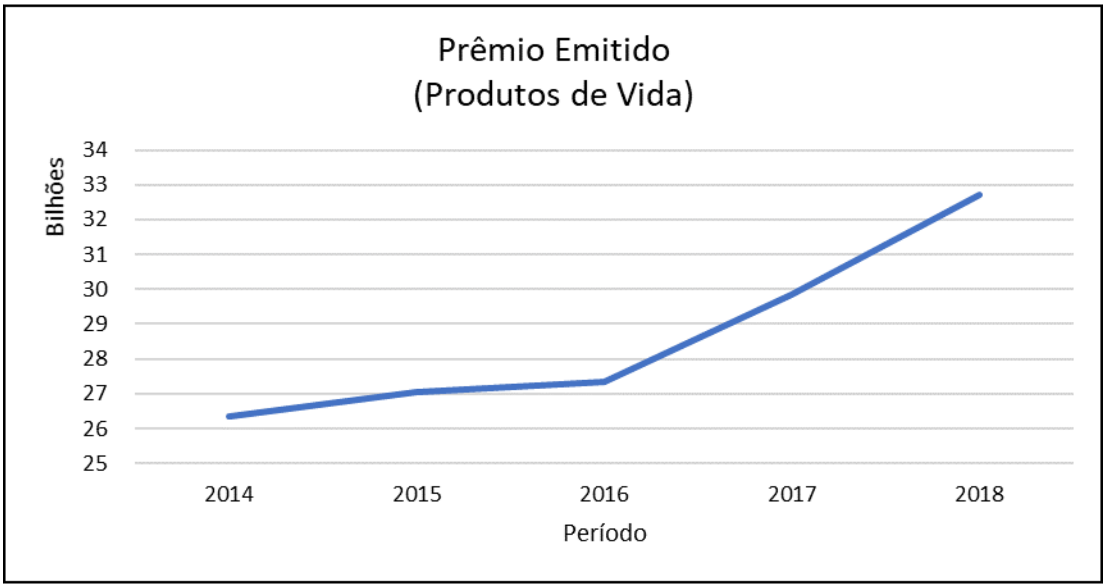

Figura 1 - Prêmio Emitido (Produtos de Vida)

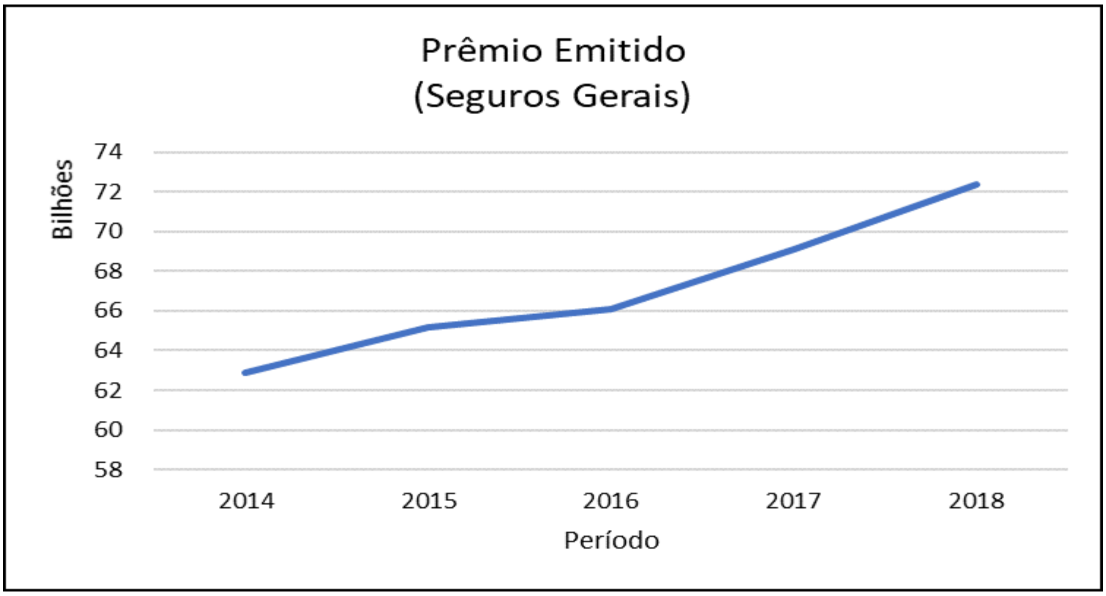

Figura 2 - Prêmio Emitido (Seguros Gerais)

Quando analisamos os resultados dos anos de 2018 por empresas, observamos alguns pontos interessantes.

Para os produtos de vida, as dez maiores empresas, concentram cerca de 70% do prêmio emitido do ano sendo que as cinco primeiras empresas são grupos seguradores vinculados aos maiores bancos do país. Somente essas cinco empresas concentram aproximadamente 54% do prêmio emitido total. Essa

concentração em produtos de vida reflete uma tendência dos últimos 3 anos onde os grandes bancos venderam ou procuraram seguradoras parceiras para administrar suas carteiras de seguros de grandes riscos.

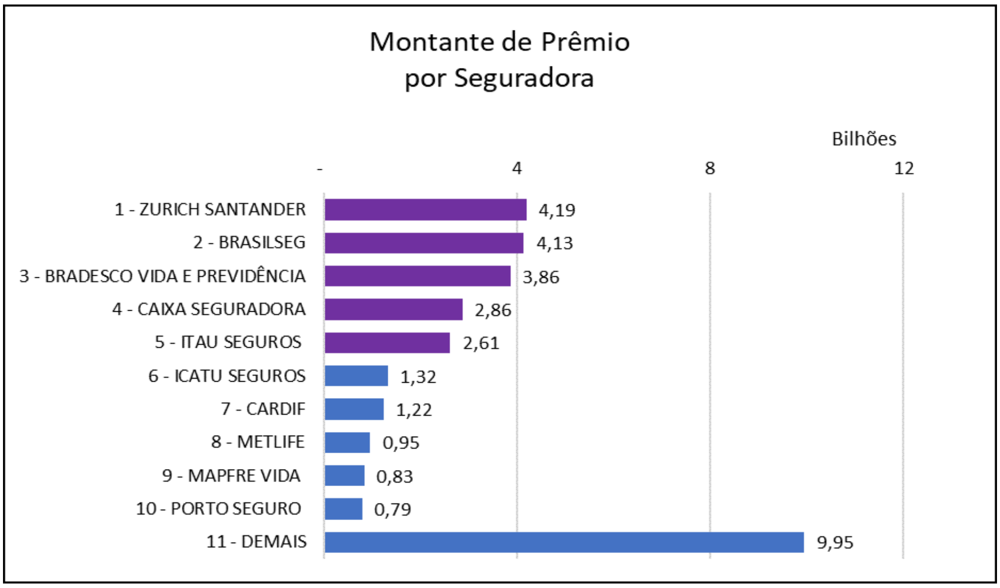

Figura 3 - Montante de Prêmio por Seguradora (Seguros de Vida)

Quando analisamos os resultados do grupo de seguros gerais, observamos a presença maior de seguradoras independentes, isto é, seguradoras que não estão vinculadas a instituições bancárias.

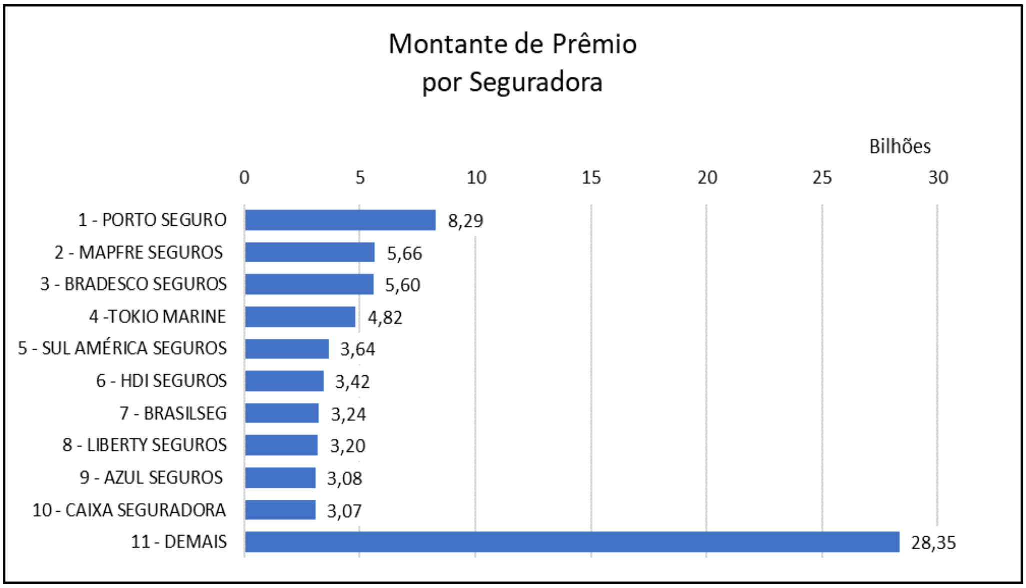

Figura 4 - Montante de Prêmio por Seguradora (Seguros Gerais)

Nesse contexto as dez maiores seguradoras concentram aproximadamente 60,8% do prêmio emitido total.

Um outro bom indicador utilizado para medir a performance das carteiras de seguros é a sinistralidade. A sinistralidade nada mais é do que a razão dos sinistros ocorridos pelo prêmio ganho da seguradora. Em outras palavras esse indicador fornece o quanto da receita (prêmio) está sendo consumida pela despesa (sinistro).

Dentro desse contexto observamos uma melhora ao longo dos últimos cinco anos. Entre os anos de 2017 e 2018 a sinistralidade do mercado reduziu 2,2 pontos percentuais sendo maior nos produtos de seguros gerais.

O volume de sinistros ocorridos vem diminuindo ao longo dos anos apesar de ter ensaiado uma leve alta entre os anos de 2017 e 2018. Mas a dinâmica da ocorrência de sinistros é muito aleatória e depende para alguns produtos de fenômenos naturais, dinâmica populacional entre outros.

Conforme podemos observar no gráfico abaixo, os seguros de vida possuem uma tendência de queda na sinistralidade reduzindo 1,3 pontos percentuais em relação ao último ano.

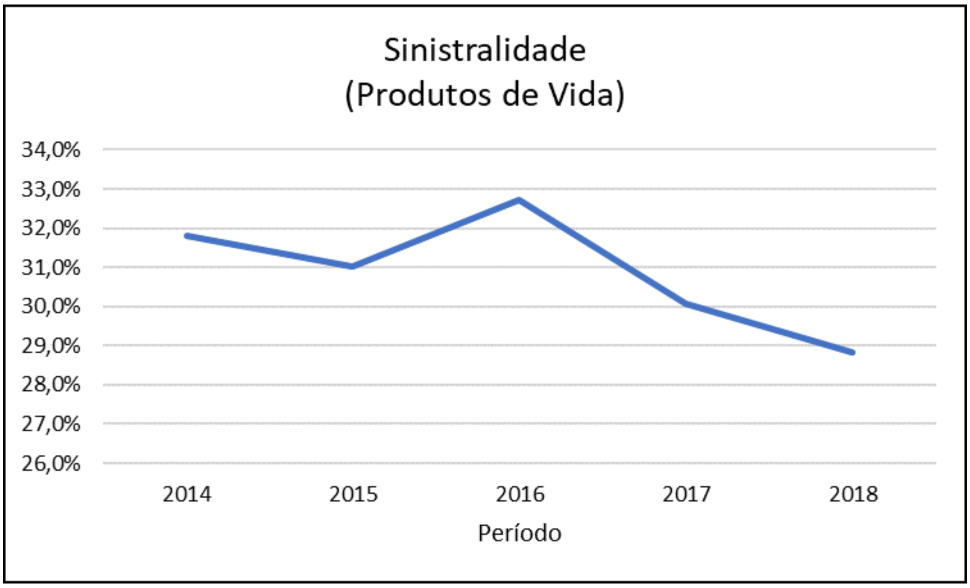

Figura 5 - Sinistralidade (Produtos de Vida)

Já os produtos de seguros gerais também apresentam uma tendência de queda na sinistralidade observada apresentando uma redução de 2,4 pontos percentuais em relação ao ano de 2017.

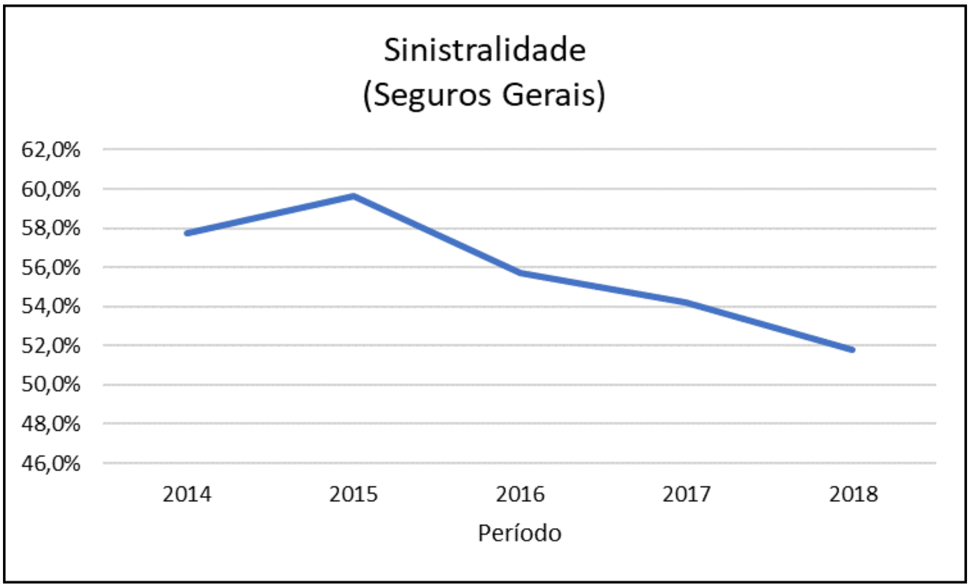

Figura 6 - Sinistralidade (Seguros Gerais)

1.2. Dinâmica do Processo de Sinistros em uma Seguradora

A figura abaixo é uma adaptação da apresentada por Wüthrich e Merz (2008) em seu livro e ilustra de forma básica a dinâmica dentro de uma companhia seguradora do processo de ocorrência de um sinistro até a sua liquidação final. Não é o objetivo tratar com detalhes as particularidades que envolvem o processo de regulação de sinistros mesmo porque cada produto tem as suas particularidades.

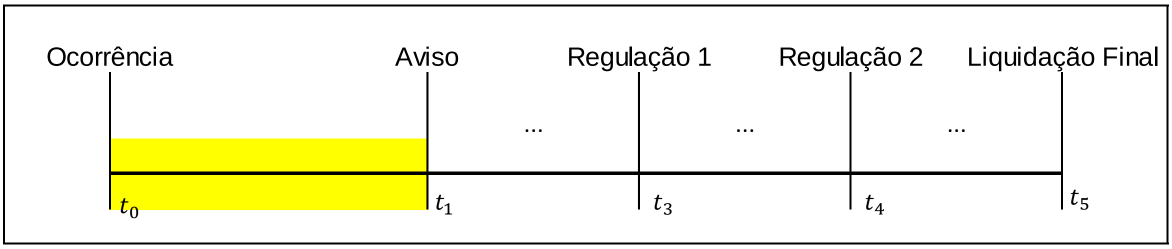

Figura 7 - Dinâmica da evolução do sinistro

A figura acima ilustra que um sinistro ocorreu em um determinado instante t_0 e após algum tempo de atraso foi avisado a companhia seguradora no instante t_1. Após o aviso, seguradora inicia o processo de regulação do sinistro até a sua liquidação final.

Pelo o disposto na regulamentação vigente, Circular SUSEP n° 517 de 2015 a companhia seguradora deve constituir mensalmente uma provisão para a cobertura dos valores esperados relativos a sinistros ocorridos e não avisados (IBNR). Observando a figura 7 acima, essa necessidade vem do fato de ocorrer esse atraso entre a ocorrência do sinistro e o aviso a companhia seguradora, no espaço temporal realçado em amarelo na figura acima a companhia seguradora já tem uma obrigação com o segurado sendo necessário constituir um passivo em seu balanço contábil que seja adequado para cobrir essa obrigação financeira.

Para estimar o valor da provisão de sinistros ocorridos e não avisados o método mais comum utilizados pelos atuários segundo uma pesquisa realizada pela instituição *Casualty Actuarial Society*, organização norte-americana credenciadora de atuários para trabalharem com riscos de seguros gerais, é o modelo de *Chain Ladder.*

Nos capítulos posteriores analisaremos esse método, mostrando suas premissas e seus pontos fracos. Assim proporemos a introdução de uma modelagem estatística ao modelo de Chain Ladder para obter estimativas mais robustas e com maior acuracidade.

Este trabalho está dividido em seis capítulos. O capítulo dois trata da revisão bibliográfica contextualizando o trabalho em relação aos estudos e técnicas já realizadas ao longo do tempo. No terceiro capítulo é apresentado o modelo de *Chain Ladder* e o modelo de regressão baseado no logaritmo das indenizações incrementais. O capítulo quatro apresenta a metodologia para a implementação dos modelos apresentados. No capítulo cinco são demostrados os resultados da aplicação dos modelos em uma carteira de seguros patrimoniais. Por fim, o capítulo seis encerra com a conclusão do trabalho e possíveis extensões.

2. Revisão Bibliográfica

Como apresentado por Friedland (2009) o modelo de *Chain ladder* é amplamente utilizado pela comunidade atuarial para estimar a provisão de sinistros ocorridos, mas não avisados à companhia seguradora além de projetar o desenvolvimento dos sinistros futuros. O modelo produz a partir dos montantes de indenizações avisadas uma sequência de fatores de desenvolvimentos para cada período de ocorrência. Assim a técnica consiste em identificar um padrão para o desenvolvimento das indenizações futuras. Segundo, Da Rosa Pinto (2013) este modelo tem como premissa que o montante de indenizações em cada período de ocorrência são independentes entre si e que os padrões de desenvolvimentos das indenizações serão os mesmos para cada período de desenvolvimento futuro, em outras palavras, as indenizações futuras irão se desenvolver da mesma forma que as indenizações já avisadas.

Segundo Da Rosa Pinto (2013) os modelos de estimativas da provisão de sinistros ocorridos e não avisados, podem ser divididos em modelos determinísticos e estocásticos. Modelos determinísticos são aqueles em que a suposição do modelo está relacionada apenas ao valor esperado dos sinistros finais projetados, sem qualquer quantificação do grau de incerteza e da variabilidade das estimativas realizadas. Os modelos estocásticos, por sua vez, produzem estimativas das variações do valor final projetado de sinistros ocorridos além de permitir associar um erro as estimativas feitas. Dentre os modelos determinísticos o autor apresenta o modelo de *Chain ladder, Gross up e Link Ratio.* Já dentro do universo dos modelos estocásticos o modelo de Mack é apresentado bem como a abordagem de bootstrap dentro da modelagem de *Munich chain Ladder.* Wüthrich, M. V (2008) apresenta alguns outros modelos como os modelos multivariados de cálculo de reserva além da aplicação da técnica de bootstrap em modelos Log-Normal. Já Carvalho, A (2010) apresenta modelagem Bayesiana no contexto de modelos Binomial Negativas, Sobredispersão de Poisson e modelos Log-Normal de três parâmetros para estimar o valor da provisão de sinistros ocorridos e não avisados.

Zenhnwirth (1994) desenvolveu estudos considerando o fato que poderia haver uma tendência no desenvolvimento das indenizações avisadas. Quando o atuário realiza

a estimativa dos sinistros futuros de uma carteira de seguros que possui uma certa tendência no desenvolvimento dessas indenizações o modelo de *Chain Ladder* pode não ser capaz de separar a tendência da flutuação aleatória normal dos avisos dos sinistros levando a estimativas erradas e imprecisas para o montante de sinistros futuros. Assim em busca de um modelo que seja capaz de incorporar a tendência presente nos dados, Zenhnwirth (1994) passou a analisar os dados na escala logarítmica associando uma dependência linear entre os períodos de desenvolvimentos das indenizações. Esta relação ficou conhecida como Família de Tendência Probabilística. Esta família se resume a um modelo de regressão linear que explicar o desenvolvimento o logaritmo das indenizações futuras através de um modelo que segrega a tendência dos dados da flutuação aleatória.

Cristofides (1997) em seu trabalho apresenta que em muitas situações o atual modelo de *Chain Ladder* utilizado por grande parte das companhias seguradoras para projetar o desenvolvimento dos sinistros futuros pode estar superparametrizado sendo potencialmente instável em suas estimativas, isto é, diferentes valores de sinistros projetados podem ser gerados com pequenas mudanças nos dados observados. Um outro ponto levantado por Cristofides é que o modelo de *Chain Ledder* produz uma estimativa pontual para o valor desses sinistros projetados não tendo nenhuma medida de variabilidade ou de erro associada a essa estimativa. Com a finalidade de trazer mais robustez para essa estimativa Cristofides apresenta em seu artigo uma aplicação do modelo de regressão múltipla ao já existente modelo de *Chain Ladder*. Ao incorporar um modelo estatístico as estimativas do modelo padrão, as estimativas passaram a ser mais robustas tendo em vista que o modelo passou a ter menos parâmetros, um erro padrão e um coeficiente de variação foi associado a cada estimativa. Na prática deixou-se de ter apenas uma estimativa pontual e passou-se a ter a percepção da variação e os erros dessa estimativa. Com isso ganha-se uma ferramenta comparativa entre as diversas linhas de negócios da companhia e entre empresas do mesmo grupo.

Este trabalho, implementa o modelo proposto por Cristofides (1997), onde uma estrutura de regressão linear múltipla é incorporada ao atual modelo de *Chain Ladder* descrito por Friedland (2009). Ainda que o modelo de regressão não constrói uma distribuição de probabilidade para os valores de IBNR conforme alguns modelos estocásticos apresentados por Wüthrich, M. V (2008), o conhecimento do

erro padrão e do coeficiente de variação fornecido pelo modelo de regressão já constitui um excelente indicador das estimativas produzidas e servem como balizador entre as estimativas das linhas de negócios. Pela primeira vez na literatura brasileira, o modelo apresentado por Cristofides (1997) é aplicado a dados do mercado de seguros brasileiros mostrando um comparativo entre as estimativas deste modelo com a do modelo de *Chain Ladder.*

3. Arcabouço Teórico

Apresentaremos a seguir uma breve descrição do modelo de *Chain Ladder,* indicando suas premissas básicas e toda mecânica de cálculo para se obter os sinistros finais projetados e a estimativa do IBNR. Em seguida, descreveremos o modelo Log linear proposto por Cristofides (1997) como uma adaptação do modelo de *Chain Ladder* para estimar os sinistros futuros.

3.1. O Modelo de *Chain Ladder*

O modelo de *Chain Ladder* é o método mais comum utilizado para realizar as projeções de sinistros futuros e consequentemente estimar a reserva de IBNR. Este modelo não faz nenhuma suposição quanto ao padrão de aviso dos sinistros, isto é, os sinistros ocorrem de forma independente entre os anos de análise.

Para modelar o desenvolvimento das indenizações futuras o método de *Chain Ladder* propõe que as indenizações já observadas sejam alocadas pelo seu respectivo período de ocorrência “i” e período de desenvolvimento “k”. Assim o portfólio passa ser interpretado como uma sequência de variáveis aleatórias ($Z_{i,k}$) com ano de ocorrência “i” e desenvolvimento “k” anos após a sua ocorrência. A variável ($Z_{i,k}$) é comumente interpretada como indenização incremental do ano de ocorrência “i” e ano de desenvolvimento “k”.

Segundo Friedland (2009) o método pode ser executado nas seguintes etapas. São elas:

- Compilação das indenizações em triângulos de desenvolvimentos;
- Cálculo dos fatores de desenvolvimento dos sinistros;
- Análise de diversas médias dos fatores de desenvolvimento;
- Seleção do melhor padrão de desenvolvimento das indenizações; e
- Projeção futura das indenizações.

Desse modo, a tabela abaixo ilustra como o modelo de *Chain Ladder* agrupa os valores dessas indenizações na forma de matricial ou triangular onde as linhas são os anos de ocorrência e as colunas representam o tempo até o aviso. A esta forma peculiar de agrupamento das indenizações damos o nome de triângulos de desenvolvimento ou triângulos de *run-off*.

Nas tabelas abaixo, apresenta-se, respectivamente, o triângulo de desenvolvimento das indenizações já avisadas e o triângulo das indenizações acumuladas para cada período de ocorrência.

Tabela 2 - Triângulo de Desenvolvimento - Indenizações Incrementais

Período	Período de desenvolvimento					
Ocorrência	0	1	2	(...)	K-1	K
0	$Z_{0,0}$	$Z_{0,1}$	$Z_{0,2}$	...	$Z_{0,k-1}$	$Z_{0,k}$
1	$Z_{1,0}$	$Z_{1,1}$	$Z_{1,2}$	...	$Z_{1,k-1}$	
2	$Z_{2,0}$	$Z_{2,1}$	$Z_{2,2}$	...		
3	$Z_{3,0}$	$Z_{3,1}$	$Z_{3,2}$	...		
4	$Z_{4,0}$	$Z_{4,1}$				
(...)	.	.	.			
i	$Z_{i,0}$					

Tabela 3 - Triângulo de Desenvolvimento - Indenizações Acumuladas

Período	Período de desenvolvimento					
Ocorrência	0	1	2	(...)	K-1	K
0	$S_{0,0}$	$S_{0,1}$	$S_{0,2}$	...	$S_{0,k-1}$	$S_{0,k}$
1	$S_{1,0}$	$S_{1,1}$	$S_{1,2}$	...	$S_{1,k-1}$	
2	$S_{2,0}$	$S_{2,1}$	$S_{2,2}$	...		
3	$S_{3,0}$	$S_{3,1}$	$S_{3,2}$	...		
4	$S_{4,0}$	$S_{4,1}$				
(...)	.	.	.			
i	$S_{i,0}$					

Posto isto, definimos a variável $S_{i,k}$ apresentada acima como sendo a soma das indenizações ocorridas no período "i" até o ano de desenvolvimento "k". A formulação matemática da variável $S_{i,k}$ é apresentada abaixo.

$$S_{i,k} = \sum_{l=0}^{k} (Z_{i,l}), \quad (1)$$

Após a elaboração do triângulo de indenizações acumuladas e partindo da premissa do modelo em que o padrão de desenvolvimento das indenizações se manterá o mesmo para cada período de ocorrência das indenizações futuras, um importante elemento aparece no modelo.

Os fatores de desenvolvimento, representam a proporção de mudança entre um período e outro, isto é, o quanto as indenizações estão evoluindo entre os períodos do triângulo de desenvolvimento. A definição desses fatores de desenvolvimento é essencial para se projetar os sinistros futuros.

Podemos definir os fatores de desenvolvimentos (f_k) como a razão entre o montante de indenizações avisadas acumuladas no final do ano de desenvolvimento "K" e o montante de indenizações avisadas acumuladas até o ano de desenvolvimento "K-1". Assim temos que:

$$f_k = \frac{\sum_{l=1}^{n-k}(S_{l,k+1})}{\sum_{l=1}^{n-k}(S_{l,k})}, \quad \forall\ 1 \leq k \leq n-1, \quad (2)$$

O modelo de *Chain Ladder* assume que os montantes de indenizações de cada período de ocorrência são independentes entre si, além disso, uma outra suposição feita pelo modelo é que as indenizações avisadas até o momento continuarão a se desenvolver de forma semelhante no futuro. Em outras palavras a atividade futura de avisos de indenizações podem ser projetadas com base na experiência histórica da companhia seguradora.

Dessa maneira o objetivo central do método é projetar a evolução das indenizações futuras ao longo dos períodos do triângulo de desenvolvimento. Esta evolução é baseada nos fatores de desenvolvimento calculados anteriormente.

Assim os valores finais estimados de sinistros de cada i-ésimo ano de ocorrência ($\hat{S}_i$) são obtidos da seguinte forma:

$$\hat{S}_i = S_{i,k-i+1} * f_{k-i+1} \dots * f_{k-1} \qquad 2 \leq i \leq k, \tag{3}$$

O montante a ser provisionado para cobrir o volume de sinistros ocorridos e não avisados à companhia seguradora para cada período de ocorrência “i” é obtido conforme demostrado abaixo:

$$IBNR_i = \hat{S}_i - S_{i,k} \tag{4}$$

$$IBNR_{total} = \sum_{i=0}^{m} IBNR_i \tag{5}$$

O modelo *Chain Ladder* é de simples compreensão e de fácil aplicabilidade, porém o seu uso deve ser feito com certa cautela quando envolve algumas situações ou até mesmo carteiras de seguros. Situações que envolvam mudanças no processo de aviso de sinistros, podendo a levar a uma diminuição ou aumento na frequência dos sinistros quer seja por questões operacionais da companhia ou por decisão de parar de operar em uma determinada linha de negócio, o atuário deve levar em consideração esses fatores pois haverá uma mudança na dinâmica das indenizações no triângulo de desenvolvimento. O descuido desses fatores pode produzir estimativas superestimadas ou subestimadas.

Segundo Cristofides (1997) o modelo de *Chain Ladder* produz estimativas baseadas em todas as observações de um determinada o período de análise podendo ser um modelo com mais parâmetros do que o necessário podendo também apresentar uma potencial instabilidade nas estimativas, isto porque, eventuais indenizações com valores elevados que tenham um atraso maior no aviso à seguradora podem

distorcer a estimativa de indenizações futuras elevando de forma equivocada o IBNR da companhia seguradora.

Conforme demonstrado por Bornhuetter e Ferguson (1972) o modelo de *Chain Ladder* é um método determinístico, isto é, produz estimativas pontuais para os sinistros futuros projetados e para a reserva de IBNR sem trazer qualquer sensibilidade quanto a variação das estimativas obtidas.

3.2. Modelo de Regressão Baseado no Logaritmo das Indenizações Incrementais

A busca por incorporar uma estrutura estatística ao modelo de *Chain Ladder* surgiu nos anos de 1990. Com o desenvolvimento e acessibilidade a softwares capazes de performar modelos estatísticos com mais acuracidade e qualidade permitiu-se aos atuários adicionarem uma abordagem mais estatística a técnica tradicional.

Alguns autores passaram a se dedicar a encontrar modelos estatísticos que se ajustassem bem ao modelo de *Chain Ladder*. As primeiras simulações apareceram Murphy (1994) e Zenhnwirth (1994). Os primeiros ensaios feitos por Murphy (1994) sugerem que se o modelo de *Chain Ladder* propõe, por exemplo, que o valor das indenizações futuras é baseado nas indenizações já incorridas multiplicada por um fator de desenvolvimento, ou seja, o valor da indenização futura passa a ser o valor da indenização incorrida multiplicada por uma constante mais um erro aleatório. Essa relação descreve um modelo de regressão onde os fatores de desenvolvimentos de um período para outro são interpretados como sendo os coeficientes de uma regressão linear múltipla.

A vantagem de se utilizar um modelo baseado em uma abordagem estatística está no fato de termos mais robustez para as estimativas, visto que passamos a ter uma quantidade ótima de variáveis sendo utilizada pelo modelo. Além disso, o modelo permite conhecermos a variabilidade da projeção das indenizações futuras e da provisão de IBNR.

Um desafio muito comum que atuários encontram quando estão analisando triângulos de desenvolvimentos é a heterogeneidade das indenizações ao longo dos

períodos de ocorrência, dificultando a estimativa final dos sinistros. Isso ocorre porque muitas coberturas vendidas pelas seguradoras possuem particularidades no aviso do sinistro e até mesmo no processo de regulação e liquidação desse sinistro. A variabilidade dessas indenizações ou até mesmo a severidade ao longo dos anos contribuem para deixar os dados menos homogêneos.

Realizar um tratamento a priori nos dados retirar a heterogeneidade nem sempre pode ser realizado isso porque pode prejudicar a interpretação do desenvolvimento das indenizações. Uma solução comentada por Zenhnwirth (1997) é utilizar a transformação logarítmica para a remoção da heterogeneidade dos dados. na escala logarítmica

O modelo descrito por Cristofides (1997) considera que as indenizações incrementais $(Z_{i,k})$ são independentes entre si e seguem uma distribuição Log-Normal.

A formulação do modelo partiu da decomposição dos sinistros finais projetados pelo modelo de *Chain Ladder* pelo percentual de cada período de desenvolvimento, chegando novamente nas indenizações incrementais conforme demonstrado abaixo:

$$Z_{i,k} = S_i * \mathrm{B}_k \tag{6}$$

Onde:

S_i = Montante acumulado de indenizações avisadas referente ao ano de ocorrência "i".

B_k = Percentual do montante final estimado de sinistro referente ao período de desenvolvimento "k". Sendo que $\sum_{i=0}^{k} B_i = 1$.

Para estimar os parâmetros do novo modelo primeiramente devemos fazer a relação descrita no item (6) ser linear. Para isto, iremos reescrever a relação aplicando o logaritmo em ambos os lados da equação. Lembrando que deveremos voltar ao espaço original quando for comparar os resultados desse modelo com o do *Chain Ladder.*

$$\ln (Z_{i,k}) = \ln (S_i * \mathrm{B}_k) \tag{7}$$

$$\ln (Z_{i,k}) = \ln (S_i) + \ln (\mathrm{B}_k) \tag{8}$$

Segundo mencionado por Cristofides (1997) é possível obter uma estimativa para o segundo termo da equação (8), porém é mais conveniente deixar esse termo de lado e se concentrar em obter estimativas dos parâmetros do conjunto de equações do agora linear conjunto de dados. A retirada desse segundo termo da equação não afeta os resultados que serão produzidos.

Da relação (8) descrita acima surge a relação entre duas variáveis aleatórias sendo $\ln (Z_{i,k})$ a variável a ser estimada.

Redefinindo o valor do $\ln (S_i) = a_i$ e introduzindo um erro $e_{i,k}$. Com isso temos um modelo de regressão descrito como:

$$\ln\left(Z_{i,k}\right) = Y_{i,k} = a_i + b_k + e_{i,k} \tag{9}$$

Ou de forma mais genérica, esta regressão linear múltipla pode ser escrita da seguinte maneira:

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \varepsilon \tag{10}$$

Onde:

Y: Vetor com os dados do logaritmo das indenizações incrementais,($Z_{i,k}$);

X: Matriz design;

β: Vetor dos parâmetros da regressão; e

ε: Vetor de erros aleatórios.

Com isso o modelo a ser ajustado é o demonstrado acima, com a suposição dos erros aleatórios, $e_{i,k}$, serem independentes e identicamente distribuídos com distribuição normal de média zero e variância constante δ^2. Para facilitar o entendimento e o uso dos dados no modelo de regressão múltipla, a seguinte matriz é definida como *input* do modelo.

Tabela 4 - Matriz *Design* X

Variável - Y				Matriz *Design* - X					
I	k	$Z_{i,k}$	$Y_{i,k}$	a_0	a_1	(...)	a_i	d	s
0	0	$Z_{0,0}$	$\ln(Z_{0,0})$	1	0	(...)	0	1	0
0	1	$Z_{0,1}$	$\ln(Z_{0,1})$	1	0	(...)	0	0	1
0	2	$Z_{0,2}$	$\ln(Z_{0,2})$	1	0	(...)	0	0	2
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
0	k	$Z_{0,k}$	$\ln(Z_{0,k})$	1	1	(...)	0	0	k
1	0	$Z_{1,0}$	$\ln(Z_{1,0})$	0	1	(...)	0	1	0
1	1	$Z_{1,1}$	$\ln(Z_{1,1})$	0	1	(...)	0	0	1
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
1	k-1	$Z_{1,k-1}$	$\ln(Z_{1,k-1})$	0	1	(...)	0	0	k-1
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Pela definição do modelo dada acima, o vetor $\boldsymbol{Y}$ passa a ser a variável dependente e cada uma das colunas da matriz **X** como as variáveis independentes. Os

coeficientes a_i assumem valores 0 ou 1, indicando os valores de $Y_{i,k}$ que são referentes ao ano de ocorrência “i”. O coeficiente “d” indica a origem do ano de ocorrência “i” e a variável “s” indica qual período de desenvolvimento está relacionado o valor de $Y_{i,k}$.

Conforme apresentado por Montgomery (2012), os coeficientes do modelo de regressão linear múltipla descrito no item (10) podem ser estimados através do método de mínimos quadrados. Este método é uma técnica de otimização matemática que visa estimar os coeficientes que minimizam a soma dos quadrados das diferenças entre o valor estimado e os dados observados.

Para encontrar o vetor de coeficientes da regressão linear múltipla por este método, a seguinte relação matricial deve ser calculada:

$$\boldsymbol{\beta} = (\boldsymbol{X^T X})^{-1} \boldsymbol{X^T Y} \tag{11}$$

Onde **X** é a matriz *design*, $\boldsymbol{X^T}$ é a matriz transposta de **X**, **Y** é o vetor de dados e $\boldsymbol{\beta}$ é a matriz contendo as estimativas para cada coeficiente do modelo de regressão.

Uma vez estimados os coeficientes da regressão linear múltipla, o procedimento seguinte é ajustar os dados ao modelo estabelecido. Para obter uma estimativa para cada valor de $Y_{i,k}$, usaremos a relação matricial descrita abaixo.

$$\boldsymbol{\hat{Y}} = \mathbf{X}\boldsymbol{\beta} \tag{12}$$

Assim, a matriz coluna calculada acima ($\boldsymbol{\hat{Y}}$) representa o valor ajustado, $\widehat{Y_{i,k}}$, de cada dado do vetor de dados (**Y**).

Com base nos valores ajustados para cada elemento,$Y_{i,k}$, do vetor de dados (**Y**) é possível calcular o resíduo de cada observação do modelo, o erro padrão dos resíduos além do erro padrão para o modelo. Não é o objetivo central desse trabalho mostrar a formulação de todas as medidas que evolvem o processo de regressão linear sendo assim para maiores detalhes dessas medidas ver Montgomery (2012).

Para estimar os valores de indenizações futuras iremos recorrer mais uma vez a elaboração de uma matriz design no mesmo formato da apresentada na tabela 4 acima. A diferença é que esta nova matriz design irá conter apenas os valores de quantos períodos de desenvolvimentos futuros será estimado, em outras palavras, quantos passos à frente a previsão será realizada.

O vetor de indenizações futuras estimadas, $\widehat{Y_{i,k}}$, será denotado por $\mathbf{Y_{futuro}}$ sendo obtido pelo produto matricial da matriz *design* futura pela matriz de coeficientes estimados do modelo, conforme demostrado abaixo:

$$\boldsymbol{Y_{futuro}} = \mathbf{X}_{futuro}\boldsymbol{\beta} \tag{13}$$

A ideia central para definir o valor das indenizações incrementais futuras vem da premissa do modelo mencionada no início deste capítulo onde considera-se que as indenizações seguem distribuição Log-Normal. Com base na distribuição Log-Normal, o valor esperado pode ser calculado da seguinte maneira:

$$E[X] = exp\{E[Y] + 0{,}5 * Var(Y)\} \tag{14}$$

$$\mathrm{Var(X)} = exp\{2 * E[Y] + Var(Y)\} * \ exp\{Var(Y) - 1\} \tag{15}$$

Podemos fazer um paralelo entre o valor esperado das indenizações provenientes da distribuição Log-Normal apresentado no item (14) com a formulação apresentada por Cristofides (1997) da estimativa dos valores futuros, conforme definido abaixo:

$$\hat{Z}_{i,k} = \exp\,(\widehat{Y_{i,k}} + {}^{1}\!/_{2} * Var(\widehat{Y_{i,k}})) \tag{16}$$

O valor de $E[Y]$ é caracterizado como o valor esperado das indenizações pelo modelo de regressão linear múltipla ($\widehat{Y_{i,k}}$).

Já para calcular a $Var(\widehat{Y_{i,k}})$ devemos recorrer a seguinte formulação apresentada por Cristofides (1997).

$$\boldsymbol{\sigma}^2 \mathbf{X}_{futuro} (\boldsymbol{X}^T \boldsymbol{X})^{-1} \mathbf{X}_{futuro}{}^{\boldsymbol{T}} \tag{17}$$

Onde:

$\boldsymbol{\sigma}^2$**:** É a variância dos dados; e

$\mathbf{X}_{futuro}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\,\mathbf{X}_{futuro}{}^{\boldsymbol{T}}$**:** O produto matricial representa a matriz de variância – covariância do modelo.

Da mesma forma que foi feito anteriormente para cálculo do valor esperado das indenizações futuras podemos fazer para o cálculo do erro padrão da estimativa. Partindo do cálculo da variância da distribuição Log-Normal apresentada no item (15), reescreveremos a fórmula da seguinte maneira:

$$\mathrm{Var(X)} = exp\{E[Y] + 0{,}5 * Var(Y)\}^2 * \; exp\{Var(Y) - 1\} \tag{18}$$

A primeira parte da equação, $exp\{E[Y] + 0{,}5 * Var(Y)\}^2$, é definido como sendo o valor da equação (16) elevado ao quadrado ($\hat{Z}_{i,k}{}^2$). Já a segunda parte da equação $exp\{Var(Y) - 1\}$ pode ser obtida considerando o valor da $Var(Y)$ obtida no item (17).

Assim o erro padrão do modelo é dado pela raiz quadrada dos valores calculados acima, em outras palavras, podemos definir como:

$$\mathrm{s.e}(\,\hat{Z}_{i,k}) = \;\hat{Z}_{i,k} * \sqrt{\exp\left(Var\big(\widehat{Y_{i,k}}\big)\right) - 1} \tag{19}$$

Com a incorporação de um modelo de regressão ao método padrão de *Chain Ladder* passamos a ter estimativas embasadas por teoria estatística sendo o modelo capaz de fornecer não somente uma estimativa pontual para as indenizações futuras mais uma gama de outras informações, como análise residual das estimativas, o erro padrão do modelo além de termos a possibilidade de extrapolarmos as estimativas para períodos de desenvolvimento maiores que “k”.

No capítulo 5 a metodologia apresentada será implementada a dados de uma carteira de seguros hipotética.

4. Metodologia

Este capítulo descreve os procedimentos necessários para realizar a estimativa dos sinistros finais projetados e do valor da provisão de sinistros ocorridos e não avisados (IBNR) tanto pelo modelo de *Chain Ladder* quanto pelo modelo de regressão baseado no logaritmo das indenizações incrementais.

4.1. Base de Dados

Para a realização desse estudo, foram utilizadas informações relativas a uma carteira hipotética de seguros gerais referente ao grupo patrimonial.

A escolha por essa carteira vem do fato de no último anos, 2018, esse grupo ter sido o segundo com maior representatividade em termos de prêmio emitido no grupo dos seguros de não vida, ficando apenas atrás do seguro de automóveis segundo dados do sistema de estatísticas da SUSEP. Além disso, esse grupo abrange coberturas de grandes riscos, fazendo com que atuários procurem a testar novas metodologias que sejam capazes de prever com mais acuracidade o valor final das indenizações futuras.

Como podemos observar na figura abaixo, cerca de 42% das indenizações são avisadas dentro do mesmo ano de ocorrência (ano de atraso igual a zero) já 58% são avisados em anos posteriores ao da ocorrência (anos igual ou maior que um). O desafio aqui é estabelecer um modelo que seja capaz de prever o valor final dessas indenizações capturando a essência dos avisos de sinistros dessa carteira de seguros.

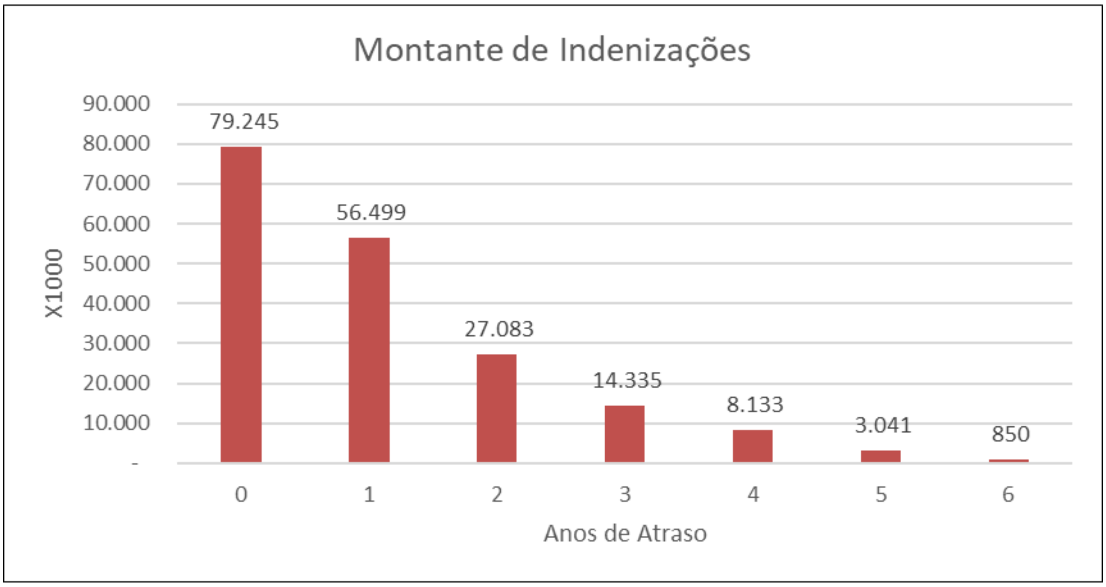

Figura 8 - Sinistralidade (Seguros Gerais)

4.2. O Modelo de *Chain Ladder*

Para a realização das estimativas das indenizações finais e da provisão de sinistros ocorridos e não avisado (IBNR) pelo modelo de *Chain Ladder* seguiremos os passos descritos por Friedland (2009) mencionados no capítulo 3.

Primeiramente compilaremos os dados incrementais de indenizações avisadas em triângulos de desenvolvimentos, isto é, as linhas serão os anos de ocorrência de 2012 a 2018 e as colunas serão o tempo entre a ocorrência e o aviso do sinistro, no nosso caso de 0 a 6 anos.

A partir do triângulo de indenizações incrementais, construiremos o triângulo de indenizações acumuladas, ou seja, os valores serão acumulados coluna por coluna de modo que cada diagonal representa os sinistros avisados até um determinado momento de análise.

Uma vez, elaborado o triângulo de desenvolvimento acumulado, precisamos calcular os fatores de desenvolvimento de um período de atraso para o outro. Lembrando que neste trabalho iremos utilizar a média volume conforme descrita na equação (2) do capítulo 3. Posteriormente, acumularemos os fatores de desenvolvimento. Assim

podemos calcular a projeção final de sinistros e a provisão de IBNR. Maiores detalhes serão apresentados no capítulo 5 de resultados.

4.3. Modelo de Regressão Baseado no Logaritmo das Indenizações Incrementais

Para implementar o modelo de regressão exemplificado no capítulo 3, foi utilizado o software estatístico R com as bibliotecas "*lattice*" e "*ChainLadder*". A modelagem realizada observou o seguinte macroprocesso:

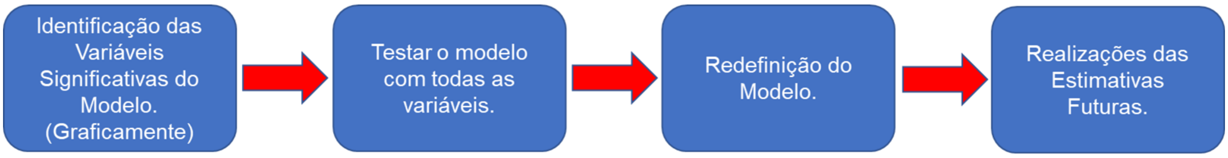

Figura 9 - Macroprocesso - Implementação do Modelo de Regressão

Primeiramente deve-se analisar graficamente o montante de indenizações na base logarítmica para cada período de desenvolvimento. Desta maneira indicativos de possíveis períodos onde não há contribuição significativa ao modelo já seria apontado.

Em um primeiro momento testaríamos o modelo de regressão com todas as informações, isto é, a matriz design teria todos os elementos do modelo de *Chain Ladder*.

Esse primeiro teste demostrará quais variáveis são significativas ao modelo, ou seja, quais anos de ocorrência contribuem para as estimativas finais. Uma vez conhecido os períodos mais significativos, redefinimos o modelo, através de uma nova matriz design, contendo somente as variáveis mais significativas. Realizamos um novo ajuste aos dados.

Por fim, um novo modelo com menos variáveis é definido produzindo estimativas de indenizações futuras, IBNR além das estimativas de erro padrões para cada projeção realizada.

5. Resultados

Neste capítulo iremos utilizar informações de uma carteira hipotética de seguros gerais referente ao grupo Patrimonial referente ao período entre 2012 e 2018 para demonstrar os principais resultados encontrados tanto pela aplicação do modelo de *Chain Ladder* quanto pelo modelo de regressão baseado no logaritmo das indenizações incrementais.

Na sequência, são apresentados comparativamente os resultados obtidos por cada modelo tanto para a projeção final dos sinistros futuros quanto para o cálculo da provisão de sinistros ocorridos e não avisados (IBNR).

Apenas para ressaltar que todos os valores presentes nas análises desse capítulo estão divididos por mil.

5.1. Aplicação do Modelo de *Chain Ladder*

Para aplicação do modelo de *Chain Ladder* aos dados iremos seguir, conforme proposto por Friedland (2009), os seguintes passos:

- Compilação das indenizações em triângulos de desenvolvimentos;
- Cálculo dos fatores de desenvolvimento dos sinistros;
- Análise de diversas médias dos fatores de desenvolvimento;
- Seleção do melhor padrão de desenvolvimento das indenizações; e
- Projeção futura das indenizações

Desta maneira conseguiremos estimar o desenvolvimento das indenizações futuras e a provisão de sinistros ocorridos e não avisados (IBNR).

O primeiro procedimento será alocar as indenizações avisadas segundo a sua data de ocorrência e período de desenvolvimento no triângulo de desenvolvimento.

Abaixo apresentamos o triângulo de desenvolvimento das indenizações incrementais.

Ano de Ocorrência	Período de Desenvolvimento						
	0	1	2	3	4	5	6
2012	8.778	8.038	5.665	4.280	2.648	1.468	850
2013	10.003	9.255	5.695	2.950	2.390	1.573	
2014	10.888	9.830	4.865	3.805	3.095		
2015	10.738	8.638	5.058	3.300			
2016	10.375	9.368	5.800				
2017	12.755	11.370					
2018	15.708						

Figura 10 - Triângulo Incremental

Apenas para exemplificar a metodologia descrita no Capítulo 3, a ideia contida nesse triângulo de desenvolvimento, é que as indenizações são alocadas de acordo com o atraso em relação ao seu ano de ocorrência. Por exemplo, o montante de R$ 12.755 refere-se ao volume de sinistros ocorridos e avisados no ano de 2017. Já o montante de R$ 3.300 refere-se ao volume de sinistros ocorridos no ano de 2015 e avisado 3 anos após o ando de sua ocorrência.

A ideia central do modelo de *Chain Ladder* é preencher a área do triângulo pintada com as informações de indenizações futuras e assim estimar a provisão de sinistros ocorridos e não avisados (IBNR).

O Procedimento seguinte é elaborar o triângulo de desenvolvimento das indenizações avisadas acumuladas.

Ano de Ocorrência	Período de Desenvolvimento						
	0	1	2	3	4	5	6
2012	8.778	16.816	22.481	26.761	29.409	30.877	31.727
2013	10.003	19.258	24.953	27.903	30.293	31.866	
2014	10.888	20.718	25.583	29.388	32.483		
2015	10.738	19.376	24.434	27.734			
2016	10.375	19.743	25.543				
2017	12.755	24.125					
2018	15.708						

Figura 11 - Triângulo Acumulado

Este triângulo de desenvolvimento demonstra para cada período de ocorrência o montante de indenizações avisadas até um determinado período. Para exemplificar,

o montante de R$ 19.743 corresponde a soma das indenizações do período de desenvolvimento “0” e “1” do ano de ocorrência 2016, isto é, voltando ao triângulo incremental, devemos somar os valores R$ 10.375 e R$ 9.368.

Cada valor contido na diagonal principal desse triângulo fornece o montante total de indenizações de cada ano de ocorrência.

Um dos pilares fundamentais da metodologia padrão de *Chain Ladder* é que os sinistros irão se desenvolver da mesma forma que se desenvolverão no passado, posto isto, o próximo passo será analisar como as indenizações são avisadas ano após ano de desenvolvimento. Para fins desse trabalho, somente é a média volume total será considerada para estimar o montante de indenizações futuras, conforme apresentado na figura abaixo.

	0/1	01/02	02/03	03/04	04/05	05/06	06/07
Média - Volume Total	**1,8892**	**1,2824**	**1,1471**	**1,0968**	**1,0509**	**1,0275**	**1,0000**
Fatores Acumulados	**3,2914**	**1,7422**	**1,3586**	**1,1844**	**1,0799**	**1,0275**	**1,0000**

Figura 12 - Fatores de Desenvolvimento

O fator 0/1 demonstra o quanto o volume de sinistros do período “1” se desenvolveu em relação ao período “0”. Isto é:

$$f_{0/1} = \frac{\sum_{l=1}^{n-k}(\, S_{l,k+1})}{\sum_{l=1}^{n-k}(\, S_{l,k})} = \frac{(\, 16.816 + 19.258 + \cdots + 24.125)}{(8.778 + 10.003 + \cdots + 15.708)} \qquad (20)$$

Como o modelo de *Chain Ladder* não prevê nenhum desenvolvimento de indenização após o período “6”, o modelo atribui para o fator 6/7 valor igual a um.

Uma vez calculado os fatores de desenvolvimento para todos os períodos de desenvolvimento, pode-se calcular o fator acumulado que nada mais é que o acúmulo dos fatores de desenvolvimento.

Com esses fatores calculados, podemos realizar a projeção das indenizações futuras e do cálculo da provisão dos sinistros ocorridos e não avisados (IBNR).

Ano de Ocorrência	Período de Desenvolvimento							Sinistro Incorrido	Sinistro Final Projetado	IBNR
	0	1	2	3	4	5	6			
2012	8.778	16.816	22.481	26.761	29.409	30.877	31.727	31.727	31.727	-
2013	10.003	19.258	24.953	27.903	30.293	31.866	**32.743**	31.866	32.743	877
2014	10.888	20.718	25.583	29.388	32.483	**34.138**	**35.077**	32.483	35.077	2.594
2015	10.738	19.376	24.434	27.734	**30.418**	**31.967**	**32.847**	27.734	32.847	5.113
2016	10.375	19.743	25.543	**29.300**	**32.136**	**33.772**	**34.702**	25.543	34.702	9.159
2017	12.755	24.125	**30.937**	**35.488**	**38.922**	**40.905**	**42.031**	24.125	42.031	17.906
2018	15.708	**29.676**	**38.056**	**43.654**	**47.878**	**50.317**	**51.702**	15.708	51.702	35.994
										71.643

Figura 13 - Sinistro Final Projetado e IBNR

A figura acima, demonstra a evolução das indenizações futuras acumuladas segundo o padrão de desenvolvimento escolhido. Em vermelho estão todo o volume de sinistros acumulados estimados para cada período de desenvolvimento presente no triângulo de desenvolvimento. Por exemplo, o montante de R$ 29.676 é obtido multiplicando o valor total das indenizações do ano de ocorrência de 2018, R$ 15.708, pelo fator de desenvolvimento $f_{0/1}$. Já o valor de R$ 29.300 referente ao ano de ocorrência de 2016 é obtido multiplicando o total das indenizações desse avisadas nesse ano, R$ 25.543 pelo fator de desenvolvimento $f_{02/03}$. Esse procedimento é realizado para cada período futuro a ser estimado.

Os valores presentes na coluna Sinistro Incorrido se referem aos valores da diagonal principal, ou seja, montante total das indenizações avisadas em cada período de ocorrência.

Já na coluna Sinistro Final Projetado, estão o montante estimado de sinistros para cada período de ocorrência, segundo o padrão de desenvolvimento escolhido, neste caso a média volume total.

Na coluna IBNR, estão os valores estimados para a provisão de sinistros ocorridos e não avisados. Conforme vimos no capítulo 3, esses valores são obtidos para cada ano de ocorrência pela diferença entre o Sinistro Final Projetado e o Sinistro Incorrido. No exemplo da figura 13, o valor de IBNR para período de ocorrência 2018 é R$ 35.994, conforme estabelecido pelo modelo de *Chain Ladder*. Este valor estimado é obtido pela diferença entre o sinistro final projetado para 2018, R$ 51.702 e o total de sinistros do ano de 2018, R$ 15.708.

Uma outra forma de interpretar o valor de R$ 35.994 que o modelo de *Chain Ladder* estimou para o ano de 2018 é em relação a maturidade do ano, visto que o volume de IBNR para 2018 é o maior quando comparamos com todos os outros períodos de ocorrência. Ao analisar a figura 13, observa-se uma diminuição gradual do IBNR à medida que caminhamos para períodos de ocorrências mais antigos, isto porque, períodos de ocorrências antigos tendem a ter uma menor necessidade de IBNR uma vez que não se espera muito mais indenizações em atraso para períodos muito antigos.

A provisão total é fruto do IBNR estimado para cada ano de ocorrência. Assim temos que:

$$IBNR_{total} = \sum_{i=0}^{m} IBNR_i = (0 + 877 + 2.594 + \cdots + 35.994) = 71.643 \tag{21}$$

Como pode-se observar este valor é uma estimativa pontual para os sinistros finais projetados e para o IBNR, não possuímos qualquer informação quanto ao erro de estimativa do modelo que estamos incorrendo ou se estamos utilizando mais parâmetros que o necessário. O modelo de *Chain Ladder* está muito suscetível a mudanças bruscas em suas estimativas casos valores de indenizações *outliers*, ou seja, indenizações com alto valores e com atrasos muito grandes apareçam.

Por exemplo, se tivéssemos uma única indenização avisada no valor de 20 milhões de reais referente ao ano de 2013 sendo avisada três anos após a sua ocorrência, a estimativa final de IBNR seria R$ 89.418 um aumento de 24,8%. Isso porque valores elevados em períodos antigos tendem a distorcer a estimativa por este método.

Na próxima seção mostraremos os resultados com um modelo estatístico mais robusto.

5.2. Aplicação do Modelo de Regressão Baseado no Logaritmo das Indenizações Incrementais

O modelo de regressão apresentado no capítulo 3 utiliza como informações de entrada as mesmas informações do modelo de *Chain Ladder*. Os fatores de desenvolvimento de um período para o outro são interpretados como os coeficientes dessa regressão linear múltipla.

Conforme apresentado em Cristofides (1997) a primeira etapa desse método é um exame visual do desenvolvimento do logaritmo das indenizações em cada período de desenvolvimento para tentar inicialmente verificar algum tipo de linearização ou tendências nos dados.

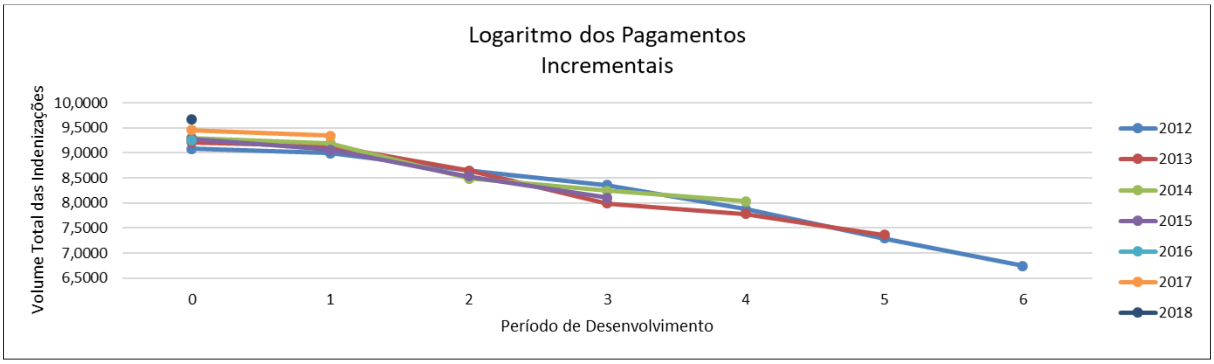

Figura 14 - Logaritmo das Indenizações Incrementais

Analisando a figura 14, observa-se a *priori* no período de desenvolvimento zero, que os dois mais recentes anos de ocorrência (2017 e 2018) estão em um nível superior que os demais. Os outros anos de ocorrência estão mais agrupados, estando todos no mesmo nível, assim um indicativo é que se deve testar se os anos de ocorrência de 2017 e 2018 são significativos para o modelo. A partir do período de desenvolvimento um, a linhas referentes a cada ano de ocorrência parecem decair de forma linear e com a mesma inclinação. Essa constatação pode sugerir que o desenvolvimento das indenizações incrementais a partir do período um está decaindo de forma exponencial uma vez que temos um decaimento linear na base logarítmica. Uma relação linear a partir do período de desenvolvimento um deve ser checada.

O primeiro modelo a ser utilizado, terá todas as variáveis do modelo padrão de *Chain Ladder*. O Objetivo aqui é testar se os anos de 2017 e 2018 são significativos para o modelo.

Conforme descrita na equação (10) do capítulo 3 o modelo de regressão linear múltipla proposto possui a seguinte formulação:

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \varepsilon \tag{22}$$

Como iremos a priori utilizar todos os dados do modelo de *Chain Ladder*, o nosso vetor de dados $,\boldsymbol{Y},$ conterá todas as 28 informações dos 7 anos de ocorrências divididas em 7 períodos de desenvolvimento.

Assim teremos a seguinte formulação a equação a equação do modelo de regressão linear proposto e apresentado acima:

$$\begin{pmatrix} \ln(8.778) \\ \ln(8.038) \\ \ln(5665) \\ .. \\ ln(12.755) \\ \ln(11.370) \\ \ln(15.708) \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \beta + \varepsilon \tag{23}$$

Neste caso temos que as dimensões de $\boldsymbol{Y}$ será (28x1), as dimensões da matriz design é (28x9).

Aplicando as relações matriciais para efetuar o cálculo das estimativas do vetor de parâmetros $\boldsymbol{\beta}$ estabelecida no item (12) do capítulo 3 chegamos nos seguintes resultados para o primeiro ajuste do modelo.

```
Call:

Residuals:
     Min       1Q   Median       3Q      Max 
-0.22142 -0.03974  0.01119  0.03287  0.19617 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  9.4891295  0.0756975 125.356  < 2e-16 ***
originf2013  0.0009373  0.0639407   0.015 0.988458    
originf2014  0.0919769  0.0686818   1.339 0.196316    
originf2015 -0.0187223  0.0752681  -0.249 0.806230    
originf2016  0.0637967  0.0843101   0.757 0.458516    
originf2017  0.2726221  0.0982545   2.775 0.012071 *  
originf2018  0.4689542  0.1316055   3.563 0.002074 ** 
d1          -0.2961583  0.0699093  -4.236 0.000447 ***
d27         -0.4349327  0.0184903 -23.522 1.64e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.113 on 19 degrees of freedom
Multiple R-squared:  0.9832,	Adjusted R-squared:  0.9761 
F-statistic: 139.1 on 8 and 19 DF,  p-value: 3.296e-15
```

Os *outputs* do modelo confirmam a suposição feita a *priori* quando se analisou o gráfico do logaritmo das indenizações incrementais onde a suposição inicial era que os períodos de ocorrência referentes aos anos de 2017 e 2018 não eram significativos ao modelo. Pelo ajuste produzido pelo modelo de regressão linear aos dados de indenizações avisadas, verificou-se que os anos de 2017 e 2018 possuem significância de 1% e 5% respectivamente. Assim, o teste de hipótese realizado pelo modelo onde se testa se o valor dos coeficientes desses dois anos são iguais a zero não é rejeitado.

Uma vez definido os períodos de ocorrência que são significativos para o modelo, o procedimento seguinte e realizar um novo ajuste aos dados com um modelo reduzido, isto é, não considerando a influência dos períodos de ocorrência que não agregam informação ao modelo proposto. Dessa maneira, da mesma forma que fizemos em (23), realizando esse novo ajuste, os seguintes resultados são encontrados:

```
Call:

Residuals:
Min       1Q   Median       3Q      Max
-0.19412 -0.05943  0.01641  0.05114  0.28395

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.47392    0.05399 175.470  < 2e-16 ***
a6           0.28216    0.08186   3.447 0.002300 **
a7           0.47768    0.11524   4.145 0.000424 ***
d1          -0.28968    0.06378  -4.542 0.000161 ***
d27         -0.43008    0.01626 -26.447  < 2e-16 ***
p34          0.06027    0.02917   2.066 0.050817 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1047 on 22 degrees of freedom
Multiple R-squared:  0.9836,    Adjusted R-squared:  0.9799
F-statistic: 263.7 on 5 and 22 DF, p-value: < 2.2e-16
```

A redução da quantidade de parâmetros refletiu diretamente no erro padrão do modelo passando de 11,3% para 10,4% além disso os seus coeficientes não foram rejeitados pelo teste de hipótese do modelo, onde se considera que os coeficientes são diferentes de zero. Assim o vetor de parâmetros $\boldsymbol{\beta}$ é definido como:

$$\boldsymbol{\beta} = \begin{bmatrix} \mathbf{9,47} \\ \mathbf{0,28} \\ \mathbf{0,47} \\ \mathbf{-0,28} \\ \mathbf{-0,43} \\ \mathbf{0,06} \end{bmatrix} \tag{24}$$

Existe uma componente aleatória o ao modelo de regressão estabelecido que é o erro associado ao modelo. Para esta componente fizemos a suposição que seria uma variável aleatória independente e identicamente distribuída segundo uma distribuição normal.

Analisando o QQ plot dos erros residuais do modelo presente na figura abaixo parece que os dados seguem uma distribuição normal. Para embasar essa análise

foi realizado o teste de Shapiro – Wilk para testar a normalidade dos erros residuais. Esse teste teve como estatística de teste um valor W = 0.97 com p-valor = 0.46 o que demonstra a normalidade dos dados.

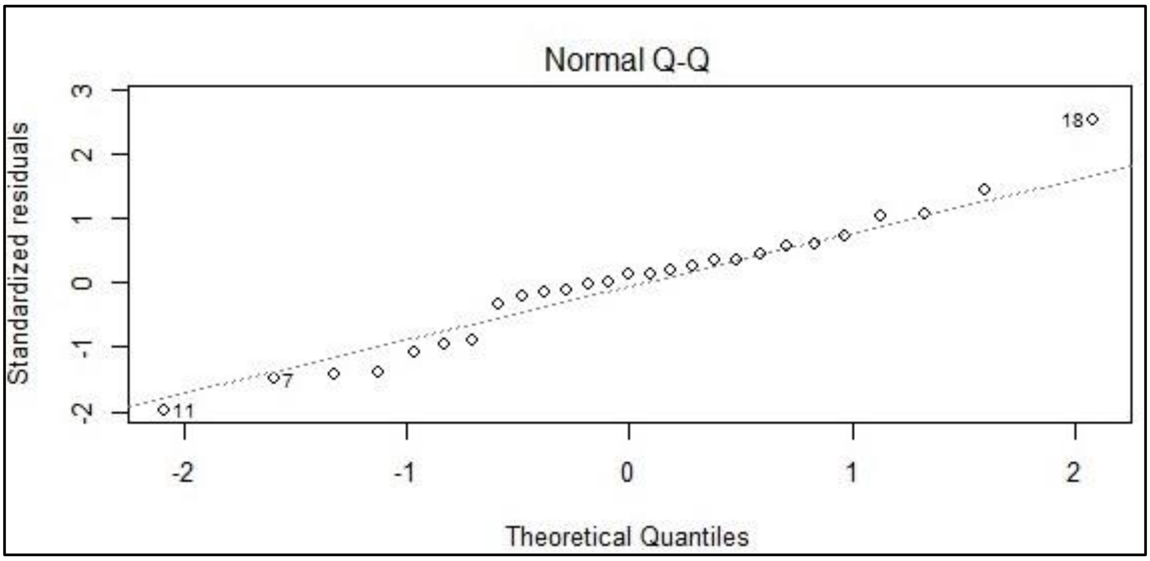

Figura 15 - QQ plot dos Erros do Modelo

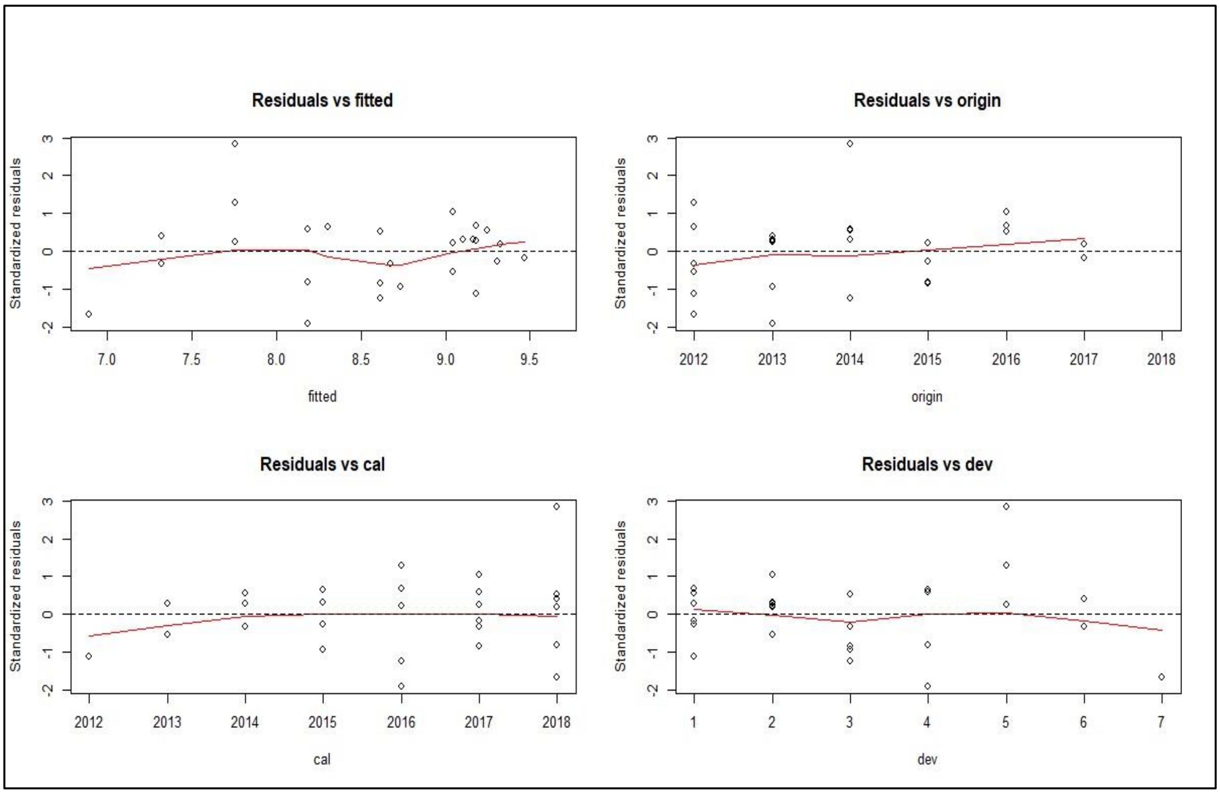

Figura 16 - Resíduos do Modelo

Um outro ponto interessante presente na figura acima é quando analisamos os resíduos do modelo principalmente os que ser referem aos períodos de ocorrência (Residual vs Cal) verificamos um bom ajuste do modelo.

Uma vez testada a adequação do modelo, a normalidade dos erros, o próximo passo é estimar a projeção dos sinistros finais e a reserva de sinistros ocorridos e não avisados (IBNR).

Para realizar esta projeção utilizaremos da formulação apresentada no item (17) utilizando definindo apenas a matriz design futura, $\mathbf{X}_{futuro}$. Para definir essa nova matriz devemos considerar quantos passos a frente devemos realizar a nossa previsão. Com o resultado do produto matricial definido pelo item (17) chegamos no seguinte triângulo de desenvolvimento de sinistros incrementais e de sinistros projetados respectivamente:

Ano de Ocorrência	Período de Desenvolvimento						
	0	1	2	3	4	5	6
2012	8.778	8.038	5.665	4.280	2.648	1.468	850
2013	10.003	9.255	5.695	2.950	2.390	1.573	1.612
2014	10.888	9.830	4.865	3.805	3.095	1.525	993
2015	10.738	8.638	5.058	3.300	2.344	1.525	993
2016	10.375	9.368	5.800	3.603	2.344	1.525	993
2017	12.755	11.370	5.540	3.603	2.344	1.525	993
2018	15.708	8.520	5.540	3.603	2.344	1.525	993

Figura 17 – Triângulo das Indenizações Incrementais – Modelo de Regressão

Ano de Ocorrência	Período de Desenvolvimento							Sinistro Incorrido	Sinistro Final Projetado
	0	1	2	3	4	5	6		
2012	8.778	16.816	22.481	26.761	29.409	30.877	31.727	31.727	31.727
2013	10.003	19.258	24.953	27.903	30.293	31.866	**33.478**	31.866	33.478
2014	10.888	20.718	25.583	29.388	32.483	**34.008**	**35.001**	32.483	35.001
2015	10.738	19.376	24.434	27.734	**30.078**	**31.603**	**32.596**	27.734	32.596
2016	10.375	19.743	25.543	**29.146**	**31.490**	**33.015**	**34.008**	25.543	34.008
2017	12.755	24.125	**29.665**	**33.268**	**35.612**	**37.137**	**38.130**	24.125	38.130
2018	15.708	**24.228**	**29.768**	**33.371**	**35.715**	**37.240**	**38.233**	15.708	38.233

Figura 18 – Projeção dos Sinistros Finais – Modelo de Regressão

No apêndice A, encontra-se de forma mais completa os valores de indenizações estimadas para cada período de ocorrência bem como os seus respectivos erros padrões e coeficientes de variação.

Com base nos valores de sinistros finais projetados pelo modelo de regressão linear múltipla, a seguinte estimativa para o IBNR é calculada:

Tabela 5 - Estimativa de IBNR segundo o Modelo de Regressão

IBNR Total (x1000)	Erro Padrão (x1000)	Coeficiente de Variação
53.987	2.101	0,038

Pelo modelo de regressão o valor de IBNR estimado é de R$ 53.987, representando uma redução de 24,6% em relação ao estimado pelo modelo de *Chain Ladder* além disso agora possuímos uma estimativa de variação da reserva fazendo com que atuários possam criar benchmarks, fatores comparativos entre as diversas linhas de negócios além de medidas comparativas para estudar a estabilidade dos triângulos de desenvolvimentos das linhas de negócios que a companhia opera.

5.3. Análise Comparativa

Neste capítulo, serão analisados os resultados apresentados por cada um dos dois modelos demonstrados nos itens (5.1) e (5.2) acima com o objetivo de identificar qual modelo performou melhor para essa carteira de seguros patrimoniais.

A análise comparativa que será utilizada para esse trabalho terá como base as informações da previsão de desenvolvimento das indenizações futuras geradas pelos dois modelos, os sinistros finais projetados e o IBNR alocado em cada ano de ocorrência e por final um comparativo do IBNR atual com volume de sinistros avisados em atraso, nos moldes do teste de consistência exigido pelo regulador.

Ao analisarmos, no gráfico abaixo, o percentual de evolução das indenizações futuras geradas pelos dois modelos, concluímos que o modelo baseado na

regressão do logaritmo das indenizações possui um padrão, onde o desenvolvimento dos avisos das indenizações futuras é mais rápido que o projetado pelo modelo de *Chain Ladder.* Como pode ser observado o modelo de regressão considera que 40% das indenizações já ocorrem entre os instantes de desenvolvimento 0 e 1 já o modelo de *Chain Ladder* considera que para esse mesmo período apenas 30% das indenizações serão avisadas. Existe uma diferença clara nos dois modelos quando comparamos os dois períodos mais recentes onde a incerteza do modelo é maior tendo em vista que possuímos poucos desenvolvimentos de sinistros ainda registrados. Mais adiante mostraremos outros elementos que motivam o modelo de *Chain Ladder* ter uma projeção mais lenta.

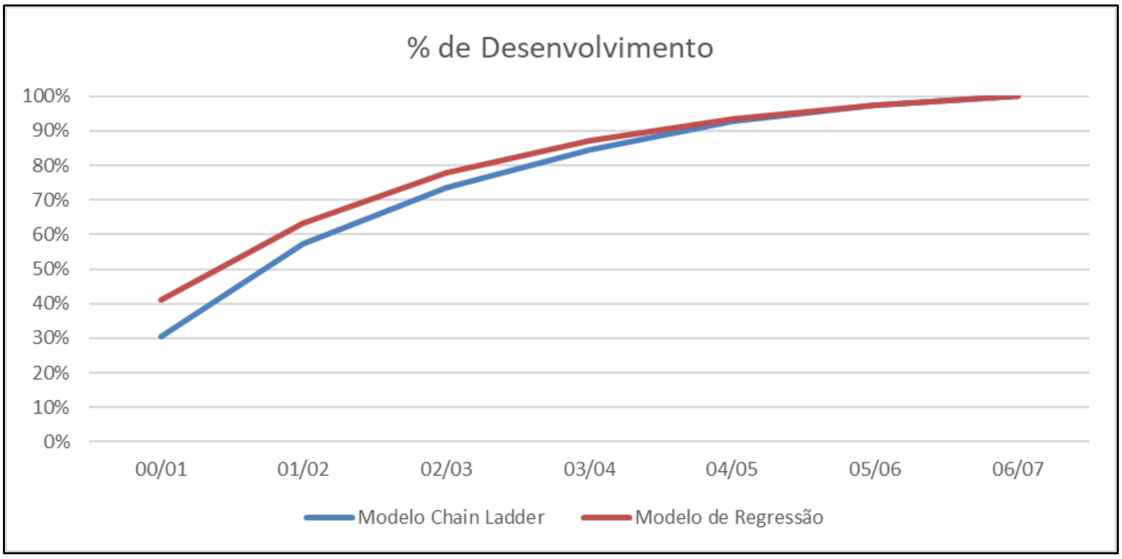

Figura 19 – Percentual de Desenvolvimento Futuro das Indenizações

Analisando os dois modelos em termos de estimativa dos sinistros finais projetados e volume de IBNR para cada ano de ocorrência, o modelo de *Chain Ladder* projeta um volume muito superior para o ano mais recente da análise, 2018, como pode ser observado abaixo.

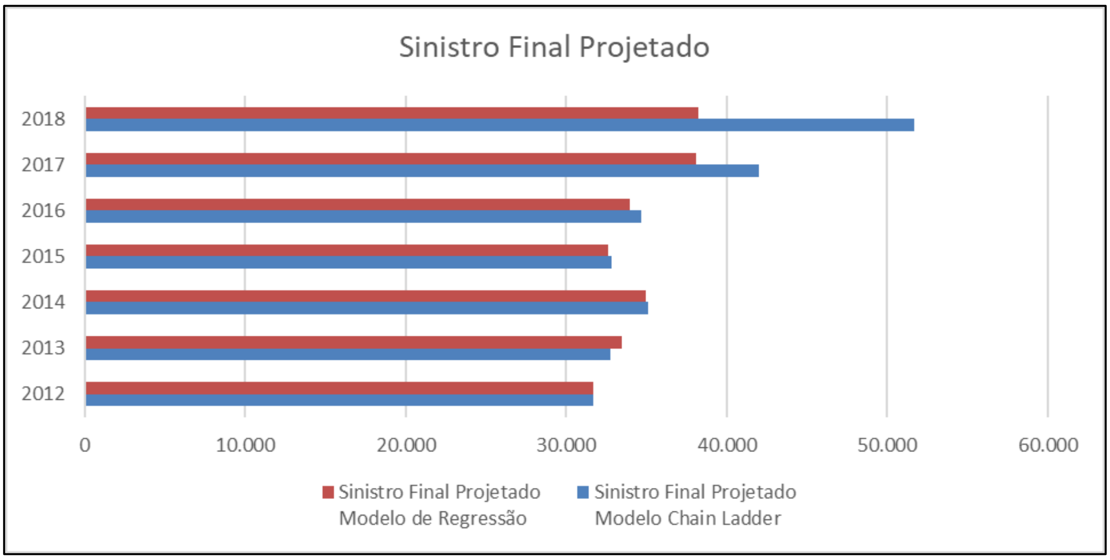

Figura 20 – Sinistro Final Projetado

Tabela 6 – Comparativo entre os Sinistros Finais Projetados

Ano de Ocorrência	Sinistro Final Projetado Modelo Chain Ladder	Sinistro Final Projetado Modelo de Regressão
2012	31.727	31.727
2013	32.743	33.478
2014	35.077	35.001
2015	32.847	32.596
2016	34.702	34.008
2017	42.031	38.130
2018	51.702	38.233

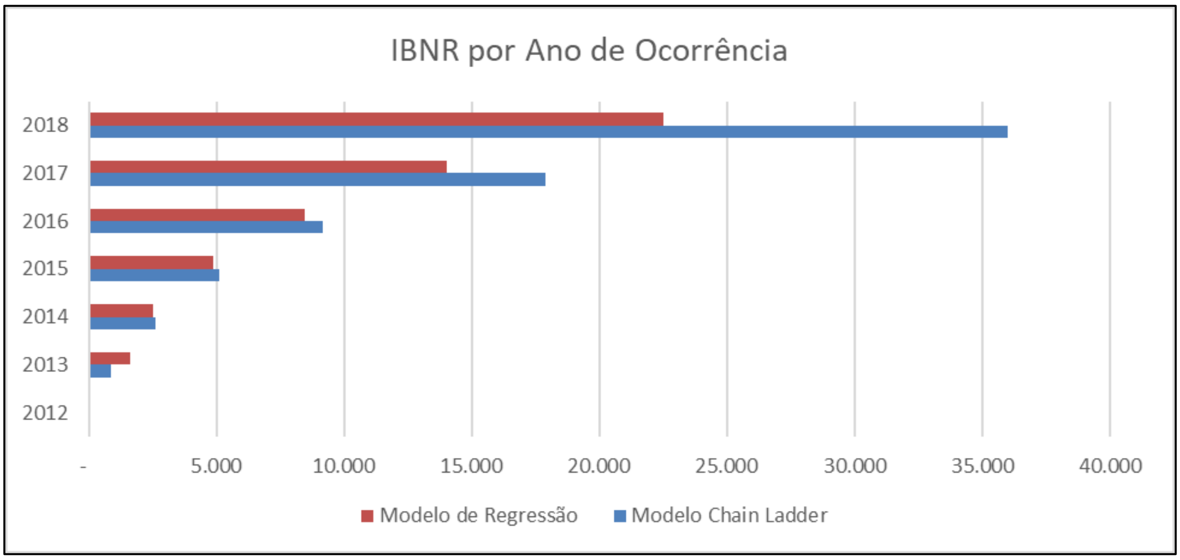

Figura 21 – IBNR por Ano de Ocorrência

Para os períodos de 2017 e 2018 o modelo de *Chain Ladder* projeta um volume de indenizações estimadas que diferem do padrão dos anos anteriores, consequentemente a estimativa de IBNR também fica mal avaliada. Isso pode ser explicado pelo fado do modelo não conseguir eliminar o efeito de um volume de sinistros de R$ 11.370 referente ao ano de ocorrência de 2017 com um ano de atraso. Esse volume destoa do histórico desse ano de desenvolvimento que possui uma média de R$ 9.026.

Quando o modelo de *Chain Ladder* considera um volume atípico de indenizações em suas estimativas há um contágio dos fatores de desenvolvimento fazendo com que estimativas futuras de sinistros sejam distorcidas. Além disso, nesse caso específico, ao considerar esse volume atípico de indenização o modelo projeta uma expectativa alta de desenvolvimento de indenizações isso faz com que a tenha um percentual de desenvolvimento mais lento quando comparado por exemplo com o modelo de regressão.

Ao analisar as estimativas obtidas pelo modelo de regressão, observa-se um volume de sinistros finais projetados em linha com todo o histórico analisado. Um outro ponto importante foi que o modelo conseguiu eliminar o efeito desse volume atípico de sinistros de suas estimativas.

Para os anos de 2012 a 2016 os dois modelos apresentaram estimativas para o volume de sinistros muito parecidas entre si, isso porque, a credibilidade desses períodos é maior, o volume de sinistros avisados em atraso é menor, deixando a estimativa mais estável.

A definição pelo atuário do modelo a ser utilizado para a estimativa do IBNR é muito complicada e depende do tipo de carteira a ser analisada. Em muitos casos, o atuário necessita definir mais de um modelo para serem combinados e utilizados para a definição do IBNR para cada ano de ocorrência.

O objetivo desse trabalho não é provar que de forma geral o modelo de *Chain Ladder* é pior ou melhor que o modelo baseado na regressão do logaritmo das indenizações incrementais e sim mostrar a aplicação dos dois modelos avaliando qual modelo realiza as estimativas de forma mais acurada para essa carteira específica.

Assim, para esta carteira específica o modelo de regressão apresentou de forma mais acurada estimativas para os sinistros finais mostrando ser mais consistente com todo o histórico da carteira.

6. Conclusões

O objetivo do presente trabalho foi implementar o modelo de regressão baseado no logaritmo das indenizações incrementais como forma alternativa ao modelo de *Chain Ladder*. Ainda que a base da abordagem alternativa sejam as informações do modelo de *Chain Ladder*, importantes conclusões com base nas estatísticas e nas premissas do modelo podem ser consideradas.

O trabalho apresentou a estimativa para a provisão de sinistros ocorridos e não avisados para ambos os modelos. O modelo de Chain Ladder apontou um nível de provisionamento superior ao modelo de regressão linear múltipla. Não é o objetivo aqui, dizer se um modelo ou outro é o mais correto até porque isto depende de que tipo de produto está sendo modelado.

O fato é que o modelo de *Chain ladder* mostra-se mais suscetível a valores extremos que são avisados com certo atraso a companhia seguradora, sendo assim menos robusto que o modelo de regressão linear múltipla. Um outro ponto importante é que comportamento atípicos, como um aumento frequência de sinistros avisados em atraso podem prejudicar as conclusões feitas por este modelo.

Atualmente, cada vez mais as companhias seguradoras estão trabalhando com diversos tipos de produtos de seguros estando cada vez mais expostas a perdas significativas. Conhecer como as indenizações se desenvolvem em um portfólio de seguros e consequentemente saber estimar de forma mais acurada o nível de provisão para cobrir sinistros ocorridos e não avisados é um ponto importante no que tange a solvência da companhia.

Garantir a solvência de uma companhia de seguros não está relacionado tão somente ao cumprimento de aspectos regulatórios ou de auditoria externa, mas sim de evitar perdas para acionista e principalmente perdas para a imagem da companhia. Em um mercado onde se vende a promessa de se cumprir uma obrigação em um determinado momento, ter uma imagem junto aos seus clientes de uma companhia que não tem condições de honrar seus compromissos preestabelecidos pode ser o ponto fundamental que determina a falência de uma companhia seguradora.

Assim este trabalho apresentou uma metodologia capaz de trazer informações do erro associado ao modelo e da variabilidade da reserva estimada pelo atuário. Isso passa ser uma importante ferramenta gerencial para comparar portfolios de seguros e comportamento de desenvolvimento de indenizações ao longo dos períodos de ocorrência.

Uma vez que os dados foram ajustados a uma regressão linear múltipla, o método permite extrapolar as estimativas de desenvolvimento de indenizações futuras além dos períodos do triângulo de desenvolvimento. Essa é uma importante limitação do modelo de *Chain Ladder*.

Existem algumas companhias no mercado segurador brasileiro que são focadas em realizar seguros de grandes riscos, como plataforma de petróleo, aviões, barragens etc. Essas companhias estão sujeitas a sinistros de valores muito extremos, outliers. O comportamento desses tipos de sinistros no triângulo de desenvolvimento distorce consideravelmente as estimativas produzidas pelo modelo de *Chain Ladder*. Como desenvolvimento futuro para este trabalho seria utilizar da análise residual do modelo de regressão para encontrar esses comportamentos atípicos de sinistros e tratá-los antes de realizar as estimativas.

Uma outra aplicação para trabalhos posteriores seria comparar as estimativas de outros modelos estocásticos abordados por Carvalho, A com as produzidas pelo modelo de regressão linear múltipla, para entender se existem diferenças significativas entre os valores de sinistros finais projetados e IBNR para cada ano de ocorrência.

Do ponto de vista de negócio, trabalhar com modelos estocásticos que produzem uma distribuição de probabilidade para o cálculo da reserva de IBNR permite avaliar além da reserva média, os percentis dessa distribuição de modo que se possa, por exemplo, analisar a variação da solvência da companhia a medida que aumentamos ou reduzimos os percentis de provisionamento. Os modelos estocásticos constituem uma excelente ferramenta gerencial para se testar, com base nos percentis da distribuição da reserva de IBNR, o nível de provisionamento do atual modelo usado pela companhia ou até mesmo que linha de negócio necessita de um maior nível de provisionamento consumindo assim maior capital da companhia.

Comparar o modelo proposto de regressão linear múltipla com os modelos estocásticos apresentados por Wüthrich, M. V (2008) permite, por exemplo, entender se as estimativas de reserva para cada ano de ocorrência estão próximas de um percentil de 50% ou de 90% da distribuição de probabilidade. Caso estejam próximas de um percentil alto, questionamentos podem aparecer ao atuário que está realizando essas estimativas. Questionamentos como: a existência de sinistros atípicos, aumento na frequência de avisos de sinistros em um determinado ano, mudanças no padrão de aviso dos sinistros, entre outros. Esses questionamentos levam o atuário a estar inserido dentro do contexto de negócio do portfolio de seguros que está sendo analisado.

REFERÊNCIAS

Alves, Ana Margarida Coelho; Provisões para Sinistros: Estudo do Mercado Segurador Português. Dissertação de Mestrado. Universidade Técnica de Lisboa.

Barnett, G. & Zehnwirth, B. (1999); Best estimates for reserves, Proceedings of the CAS, 1999.

Bornhuetter, R. L.; Ferguson, R. R. "The Actuary and IBNR". CAS, 1972.

Carvalho, A; Modelos Estocásticos em Provisões para Sinistros. Dissertação de mestrado. Universidade de Técnica Lisboa.

Christofides, S. (1997), Regression models based on log-incremental payments, Claims Reserving Manual 2, 1995.

Da Rosa Pinto, A., Métodos de Previsão de Sinistros, 2013.

England, P; Verrall, R; Analytic and Bootstrap estimates of prediction erros in claims reserving. Insurance Mathematics and Economics, 1999.

Friedland, J.F., “Estimating Unpaid Claims using Basic Techniques,” CAS, 2009.

Mano, Cristina; Pereira, Paulo; Aspectos Contábeis e Atuariais das Provisões Técnicas. Funenseg, 2009.

Montgomery, Douglas C; Peck, Elizabeth A; Vining, G. Geoffrey; Introduction to Linear Regression Analysis. Wiley Series in Probability and Statistics, 2012.

Murphy, D. M., Unbiased Loss Development Factors. Proceedings of the CAS, 1994.

Taylor, G. C. Loss Reserving – An Actuarial Perspective. Boston: Kluwer Academic Publishers, 2000.

Verrall, R. J., An Investigation into Stochastic Claims Reserving Models and Chain Ladder Technique. Insurance Mathematics and Economics, 2000.

Wüthrich, M. V.; Merz, M. Stochastic Claims Reserving Methods in Insurance. Wiley-Business & Economics, 2008.

Zehnwirth, B., The Chain Ladder Technique – A stochastic Model, Claims Reserving, 1999.

APÊNDICES

APÊNDICE A – Resultados para as Estimativas Futuras - Modelo de Regressão Linear Múltipla

Os resultados apresentados são os *outputs* do modelo confeccionado no software estatístico R.

Os valores de **P** são referentes as indenizações incrementais referentes a cada períodos de ocorrência e desenvolvimento ("dev") futuro. Essas mesmas observações podem ser encontradas na figura 18, no triângulo de desenvolvimento incremental.

Os valores da coluna "**seP**" são referentes ao erro padrão de cada estimativa de **P**. Já a coluna "**CV**" é referente aos coeficientes de variação das estimativas.

origin	dev	Y	VarY	P	seP	CV
2018	2	9,043841	0,0126	8.520	959,3777	0,112606
2017	3	8,61376	0,011851	5.540	604,8422	0,109184
2018	3	8,61376	0,011851	5.540	604,8422	0,109184
2016	4	8,183679	0,01163	3.603	389,6757	0,108156
2017	4	8,183679	0,01163	3.603	389,6757	0,108156
2018	4	8,183679	0,01163	3.603	389,6757	0,108156
2015	5	7,753599	0,011938	2.344	256,8616	0,109588
2016	5	7,753599	0,011938	2.344	256,8616	0,109588
2017	5	7,753599	0,011938	2.344	256,8616	0,109588
2018	5	7,753599	0,011938	2.344	256,8616	0,109588
2014	6	7,323518	0,012775	1.525	172,9437	0,113388
2015	6	7,323518	0,012775	1.525	172,9437	0,113388
2016	6	7,323518	0,012775	1.525	172,9437	0,113388
2017	6	7,323518	0,012775	1.525	172,9437	0,113388
2018	6	7,323518	0,012775	1.525	172,9437	0,113388
2013	7	7,371117	0,027694	1.612	270,0598	0,167576
2014	7	6,893437	0,014141	993	118,475	0,119337
2015	7	6,893437	0,014141	993	118,475	0,119337
2016	7	6,893437	0,014141	993	118,475	0,119337
2017	7	6,893437	0,014141	993	118,475	0,119337
2018	7	6,893437	0,014141	993	118,475	0,119337

Figura 22 – Estimativas Futuras obtidas pelo Modelo de Regressão Linear Múltipla