

4

Simulação e Análise dos Resultados

Cada um dos estimadores apresentados foi implementado em C++. Utilizando um banco de dados foi possível simular todos os modelos e assim compará-los. O banco de dados contém informações sobre seguro de automóvel durante um período de dois anos. Mas só há informação sobre os sinistros ocorridos no primeiro ano. Já os avisos estendem-se para o segundo ano. Como se deseja observar o comportamento do IBNR ao longo de um período, a simulação será feita para diversas datas presentes consecutivas.

Além dos modelos também foi simulado o *IBNR-Observado* na data presente. Este é definido como a soma dos valores de todos os sinistros ocorridos antes da data presente e avisados depois desta. É claro que o seu valor não é exatamente o valor real do IBNR, uma vez que devem existir sinistros com ocorrência anterior à data presente e aviso posterior à última data do banco de dados. Mesmo assim é um bom valor para comparações.

A simulação começa com a data presente 01/fev e termina em 01/dez, ambas do primeiro ano. Ela não foi iniciada em janeiro pois como esta é a data inicial do banco de dados os estimadores não teriam dados para o seu cálculo e por isso forneceriam valor zero. Além disso a última data foi dezembro do primeiro ano para que o valor do IBNR-Observado fosse representativo, uma vez que ainda existe um ano de observação para que os sinistros ocorridos sejam avisados.

Também é importante destacar que os valores fornecidos pelos estimadores foram multiplicados por uma constante, com o objetivo de não revelar sua ordem de grandeza. Apesar disso a comparação e análise dos resultados não foi comprometida, uma vez que todos foram multiplicados pelo mesmo número.

A tabela 4.1 fornece os valores do resultado da simulação.

Data Presente	$IB\hat{N}R_1$	$IB\hat{N}R_2$	$IB\hat{N}R_3$	$IB\hat{N}R_{novo}$	IBNR-Observado
01/fev	0,085	0,085	0,085	2,848	0,582
01/mar	0,860	0,878	0,828	2,786	0,897
01/abr	1,787	1,834	1,889	2,626	1,256
01/mai	5,941	6,036	6,077	2,462	1,455
01/jun	1,340	1,680	2,437	2,292	1,152
01/jul	1,510	1,954	2,327	2,230	1,343
01/ago	2,364	2,990	3,378	2,181	1,485
01/set	2,352	2,982	3,528	2,107	1,487
01/out	2,094	2,566	3,058	2,000	1,589
01/nov	2,398	2,974	3,392	1,877	1,654
01/dez	1,368	1,590	1,955	1,749	1,768

Tabela 4.1: Resultado da simulação.

Com os valores da tabela foi possível construir o gráfico apresentado na figura 4.1, que facilita a comparação e análise dos estimadores.



Figura 4.1: Gráfico gerado pelo resultado da simulação.

4.1

Comportamento do Novo Estimador

Nessa seção serão vistos os detalhes da simulação relativos ao novo estimador, analisando cada parâmetro estimado. Assim será possível entender melhor o seu comportamento.

Toda vez que o $IB\hat{N}R_{novo}$ foi calculado, estimativas para α e λ também foram determinadas, através das equações 3-4 e 3-5. A tabela 4.2 apresenta essas estimativas.

Data Presente	$\bar{\alpha}$	$\bar{\lambda}$
01/fev	0,302264	63,830700
01/mar	0,299380	63,879200
01/abr	0,297687	63,911800
01/mai	0,295113	63,956500
01/jun	0,291809	64,031600
01/jul	0,298647	63,711300
01/ago	0,305744	63,483700
01/set	0,310600	63,299700
01/out	0,313210	63,230800
01/nov	0,313513	63,297500
01/dez	0,313531	63,307300

Tabela 4.2: Estimativas para α e λ fornecidas pela simulação.

Lembre-se que o parâmetro α representa a média da razão entre o valor total assegurado e o valor pago por um certo sinistro, e o parâmetro λ representa a média do intervalo de tempo entre a ocorrência e aviso de um sinistro. Ambos foram estimados pela média amostral.

Observando a tabela 4.2 conclui-se que não existe grande variação nas estimativas de α e λ . Além disso, para datas presentes consecutivas, a variação desses valores é muito pequena. Essa afirmação pode ser verificada na tabela 4.3, que mostra a razão entre os $\bar{\alpha}$'s e $\bar{\lambda}$'s consecutivos.

α_k/α_{k+1}	λ_k/λ_{k+1}
1,010	0,999
1,006	0,999
1,009	0,999
1,011	0,999
0,977	1,005
0,977	1,004
0,984	1,003
0,992	1,001
0,999	0,999
1,000	1,000

Tabela 4.3: Comparação entre $\bar{\alpha}$ e $\bar{\lambda}$ consecutivos.

Isso mostra que ao longo do tempo nem $\bar{\alpha}$ nem $\bar{\lambda}$ influenciam para o decréscimo no valor do $IBNR_{novo}$. Além disso, sua estabilidade contribui para a estabilidade do estimador.

Além de α e λ , o parâmetro β também foi estimado. Este representa a média do intervalo de tempo entre o início da exposição e acontecimento de um sinistro, como foi explicado na seção 3.3. A equação 3-9 nos mostra

que a estimativa é o limite de uma seqüência. Como computacionalmente isso pode virar um problema muito grande, foi feita uma pequena aproximação. O valor de β fornecido pela simulação é o primeiro termo da seqüência β_k tal que $\beta_{k+1} - \beta_k \leq 0.5$. Os valores de $\hat{\beta}$ para cada data presente podem ser vistos na tabela 4.4.

Data Presente	$\hat{\beta}$
01/fev	812,60
01/mar	1311,82
01/abr	1765,80
01/mai	2136,60
01/jun	2480,67
01/jul	2764,38
01/ago	3027,66
01/set	3301,46
01/out	3611,83
01/nov	3947,43
01/dez	4305,56

Tabela 4.4: Estimativas para β fornecidas pela simulação.

Analisando agora os valores apresentados na tabela 4.4, pode-se verificar um crescimento significativo de $\hat{\beta}$. Isso indica que o parâmetro β não está sendo bem estimado, o que pode ser justificado pelo pequeno período que se encontra o banco de dados ou por falhas nas suposições do modelo.

Para compreender melhor o motivo pelo qual β está crescendo, foram observados os valores de β_0 em cada data presente. Estes valores são resultados da equação 3-7 e representam a média amostral do tempo entre o início da exposição e a ocorrência de um sinistro. Tais valores estão apresentados na tabela 4.5.

Data Presente	β_0
01/fev	127,86
01/mar	129,87
01/abr	130,11
01/mai	129,68
01/jun	128,20
01/jul	125,95
01/ago	123,00
01/set	120,08
01/out	118,16
01/nov	116,46
01/dez	115,57

Tabela 4.5: Valores de β_0 fornecidas pela simulação.

A tabela mostra que os valores de β_0 não estão crescendo, logo não é este o motivo do crescimento de $\hat{\beta}$. Analisando o cálculo da estimativa $\hat{\beta}$ acredita-se que o seu crescimento é causado pelo aumento do período de exposição devido a mudança da data presente, como pode ser visto em 3-1. Com períodos de exposição maiores, é esperado que $\hat{\beta}$ seja maior.

Se os dados abrangessem um período relativamente maior, a mudança na data presente não geraria um crescimento tão significativo no valor de $\hat{\beta}$, uma vez que o aumento nos períodos de exposição seria proporcionalmente pequeno em relação à soma de todos os períodos de exposição. Dessa forma seria possível fazer uma análise mais detalhada do valor de $\hat{\beta}$ e observar se o seu crescimento continua indefinidamente ou se estabiliza ao longo do tempo, o que é o mais esperado.

É interessante destacar que o crescimento de $\hat{\beta}$ é responsável pelo decréscimo do valor do $IBNR_{novo}$. Essa relação pode ser verificada pela equação 3-11 : conforme $\hat{\beta}$ cresce, $IBNR_{novo}$ decresce. Como o valor estimado para β está crescendo bastante seria bom estudar o caso onde β é fixo, pois se este parâmetro fosse conhecido o estimador do IBNR deveria ser melhor comportado. Fez-se então uma nova simulação para determinar o valor do $IBNR_{novo}$ escolhendo $\hat{\beta} = 4305,56$, uma vez que essa foi a estimativa para β determinada com o maior número de dados. O resultado está apresentado na tabela 4.6.

Substituindo os valores do $IBNR_{novo}$ da tabela 4.1 pelos da tabela 4.6, foi gerado o gráfico 4.2. Com ele é possível uma comparação melhor entre os estimadores.

Data Presente	$IBNR_{novo}$	IBNR-Observado
01/fev	0,546	0,582
01/mar	0,862	0,897
01/abr	1,092	1,256
01/mai	1,237	1,455
01/jun	1,335	1,152
01/jul	1,445	1,343
01/ago	1,546	1,485
01/set	1,625	1,487
01/out	1,685	1,589
01/nov	1,725	1,654
01/dez	1,749	1,768

Tabela 4.6: Valores do $IBNR_{novo}$ supondo $\hat{\beta}$ fixo.Figura 4.2: Gráfico gerado pelo resultado da simulação supondo β fixo.

O comportamento do novo estimador para essa última simulação foi excelente, mas nada pode ser afirmado sobre o valor real do parâmetro β . Porém é notória a necessidade de um banco de dados maior, pois se nessas condições $\hat{\beta}$ for estável, o novo estimador pode gerar valores muito bons.

4.2

Comparação do Novo Estimador com os Demais

É importante destacar que no primeiro mês a simulação não retrata muito bem a realidade dos dados. Isso ocorre devido a pequena quantidade de informação existente em um mês de observação. Com a data presente sendo fevereiro, os estimadores 1, 2 e 3 possuem apenas as informações do mês de janeiro do primeiro ano. Por outro lado, para o mesmo mês, o novo estimador utiliza em sua estimativa todos as apólices existentes, além dos

sinistros ocorridos e avisados em janeiro.

Um fato que chama muita atenção no gráfico é o pico existente nos estimadores 1, 2 e 3 para o mês de maio. Justificar isso para os estimadores 1 e 2 é bastante simples, uma vez que estes são baseados nos fatores de desenvolvimento de cada coluna. Já para o estimador 3, a justificativa é mais complicada devido à complexidade da equação 2-17.

Primeiro será feita uma análise nos fatores de desenvolvimento determinados pelo método *Chain Ladder*, na tentativa de justificar o pico gerado por este estimador. A justificativa do pico gerado pelo segundo estimador é análoga, e não será apresentada neste trabalho. Em seguida o pico no terceiro estimador será analisado.

Para cada data presente pode-se gerar o seu triângulo de desenvolvimento ou acumulado. Dessa forma é possível observar os fatores em cada caso. Na tabela 4.7 encontra-se o triângulo acumulado para data presente igual a 01/dez do primeiro ano.

$$\begin{bmatrix} 0,013 & 0,151 & 0,366 & 0,485 & 0,542 & 0,558 & 0,564 & 0,567 & 0,572 & 0,574 & 0,574 & 0,574 \\ 0,008 & 0,138 & 0,317 & 0,417 & 0,447 & 0,458 & 0,467 & 0,471 & 0,474 & 0,475 & 0,475 & \\ 0,021 & 0,239 & 0,547 & 0,632 & 0,668 & 0,687 & 0,692 & 0,694 & 0,696 & 0,696 & & \\ 0,031 & 0,415 & 0,589 & 0,672 & 0,717 & 0,729 & 0,739 & 0,745 & 0,748 & & & \\ 0,109 & 0,295 & 0,510 & 0,613 & 0,650 & 0,665 & 0,670 & 0,673 & & & & \\ 0,029 & 0,227 & 0,484 & 0,595 & 0,641 & 0,650 & 0,659 & & & & & \\ 0,052 & 0,352 & 0,594 & 0,704 & 0,745 & 0,757 & & & & & & \\ 0,087 & 0,354 & 0,622 & 0,722 & 0,746 & & & & & & & \\ 0,096 & 0,427 & 0,668 & 0,785 & & & & & & & & \\ 0,075 & 0,409 & 0,690 & & & & & & & & & \\ 0,109 & 0,402 & & & & & & & & & & \\ 0,035 & & & & & & & & & & & \end{bmatrix}$$

Tabela 4.7: Triângulo de desenvolvimento para o mês de dezembro.

Os demais triângulos podem ser determinados através deste. Por exemplo, o triângulo acumulado para a data presente 01/abr do primeiro ano seria o triângulo formado pelas quatro primeiras linhas e pelas quatro primeiras colunas, desconsiderando os elementos abaixo da diagonal não principal, como pode ser visto a seguir.

$$\begin{bmatrix} 0,013 & 0,151 & 0,366 & 0,485 \\ 0,008 & 0,138 & 0,317 & \\ 0,021 & 0,239 & & \\ 0,031 & & & \end{bmatrix}$$

Tabela 4.8: Triângulo de desenvolvimento para o mês de abril.

Com os dados da tabela 4.7 foram determinados, para o primeiro ano de observação, os triângulos acumulados para as datas presentes 01/abr, 01/mai e 01/jun. Em cada caso foram calculados os fatores. Os resultados podem ser vistos na tabela 4.9:

	01/abr	01/mai	01/jun
f_1	12,22	12,63	6,75
f_2	2,37	2,33	1,93
f_3	1,32	1,32	1,25
f_4	-	1,12	1,10
f_5	-	-	1,03

Tabela 4.9: Fatores para os meses de abril, maio e junho do primeiro ano de simulação.

Os valores de f_1 para abril e maio são muito maiores que os demais fatores, isso pode ser facilmente compreendido. Observe no triângulo acumulado, tabela 4.7, que os valores das quatro primeiras linhas da coluna 1 são números com ordem de grandeza 10^{-2} ou 10^{-3} , enquanto os quatro primeiros números da segunda coluna variam com ordem de grandeza 10^{-1} . Isso mostra um grande crescimento entre a primeira e segunda coluna para as quatro primeiras linhas. Por causa disso f_1 é muito grande para as datas presentes que usam apenas as quatro primeiras linhas no seu cálculo: 01/fev, 01/mar, 01/abr e 01/mai do primeiro ano.

Já a quinta linha da primeira coluna possui um valor muito grande em comparação aos anteriores, na ordem de 10^{-1} . Dessa forma, no triângulo acumulado para 01/mai do primeiro ano este elemento foi multiplicado pelo elevado valor de f_1 . O que gerou um valor muito alto e influenciou no IBNR final daquele mês.

Observe que quando a quinta linha da primeira coluna passa a ser levada em consideração no cálculo de f_1 , este diminui seu valor significativamente, como pode ser visto na tabela 4.10. Por causa disso, apesar da décima primeira linha do triângulo acumulado possuir um valor muito alto na primeira coluna, não ocorreu pico na estimativa do IBNR.

Data Presente	f_1
01/fev	11, 18
01/mar	13, 19
01/abr	12, 22
01/mai	12, 63
01/jun	6, 75
01/jul	6, 91
01/ago	6, 89
01/set	6, 19
01/out	5, 81
01/nov	5, 75
01/dez	5, 40

Tabela 4.10: Primeiro fator para cada data presente.

Para analisar o pico no estimador 3 é preciso entender porque a equação 2-17 fornece uma saída grande quando a data presente é 01/mai. Por isso foram observados os valores de ${}_0F_1(.;.)$ em cada uma das datas presentes. Esse resultado está apresentado na tabela 4.11.

Data Presente	${}_0F_1(.;.)$
01/fev	1,000
01/mar	1,022
01/abr	1,010
01/mai	1,006
01/jun	1,137
01/jul	1,121
01/ago	1,098
01/set	1,095
01/out	1,085
01/nov	1,099
01/dez	1,100

Tabela 4.11: Valor da função ${}_0F_1(.;.)$ para cada data presente.

Como ${}_0F_1(.;.)$ apresenta um comportamento bastante estável, é razoável pensar que este não é o motivo do pico do estimador 3 para o mês de maio. Provavelmente a estimativa \hat{v} para este mês é tal que o somatória da equação 2-17 fornece um valor grande, tornando assim o valor do $IBNR_3$ grande também.

Enquanto os estimadores 1, 2 e 3 mudam bruscamente a sua estimativa por causa de uma entrada fora do padrão, o Novo Estimador se manteve estável. Isso mostra o seu comportamento suave em comparação aos demais. Esta é uma característica muito favorável, uma vez que as seguradoras não

desejam grandes variações nas estimativas do IBNR.

Também é interessante observar a semelhança entre $IB\hat{N}R_1$ e $IB\hat{N}R_2$. Isso é mais que esperado uma vez que ambos são baseados nos fatores de desenvolvimento, diferindo apenas na forma em que eles são calculados.

Infelizmente não foi possível obter o valor para a provisão do IBNR usado por uma seguradora. Dessa forma poderia ser feita uma análise mais profunda sobre o comportamento dos métodos, viabilizando sugestões para melhorar o seu desempenho.