3
Chain ladder estendido

3.1.
Descrição do método chain ladder tradicional

Antes de maior detalhamento do método, serão estabelecidas algumas notações e definições básicas:

1) Seja X um vetor aleatório $n \times 1$ em algum $(\Omega, \Im, P)$ e ℓ uma classe não-vazia de subconjuntos de $\Re^n$. Define-se a *imagem inversa* de ℓ sob X como sendo a classe de eventos
$$X^{-1}(\ell) \equiv \{(X \in A) : A \in \ell\}.$$
2) Denote por Y_i o vetor aleatório formado pelas observações da i-ésima linha do triângulo de runoff (independentemente da ordenação), $i = 1,2,\ldots,J$. A dimensão de Y_i é denotada por n_i e a σ-álgebra gerada por Y_i é denotada por $\Im_i$ (ou seja: $\Im_i \equiv \sigma(Y_i)$). Defina também $\Im_\Delta \equiv \sigma(Y_1, Y_2, \ldots, Y_J)$ e $\Im_{i,j} \equiv \sigma(Y_i, Y_j)$, para todo $i \neq j$.
3) A *σ-álgebra de Borel no* $\Re^n$ (ou seja, aquela gera pelos conjuntos abertos do $\Re^n$) é denotada por $\underline{B}^n$. O *produto cartesiano* – que não deve ser confundido com a *σ-álgebra produto* - das σ-álgebras de Borel no $\Re^n$ e no $\Re^m$ é denotado por $\underline{B}^n \times \underline{B}^m$.
4) Uma classe de subconjuntos de ℓ sobre um conjunto não-vazio qualquer é dita ser um *π-sistema* se for fechada por interseção finita.
5) Sejam X e Y duas variáveis aleatórias quadraticamente integráveis e definidas em um mesmo espaço de probabilidade, e $\wp$ uma sub-σ-álgebra, à qual Y é mensurável. Define-se o *erro médio quadrático teórico condicional de Y (na estimação/predição de X) dada* $\wp$ como sendo
$$EQM(Y \mid \wp) \equiv E\left[(X - Y)^2 \mid \wp\right].$$

Seguindo com a construção da teoria sobre o chain ladder, serão apresentadas as hipóteses enunciadas por Mack (1983) – que doravante serão chamadas *hipóteses de Mack* –, as quais são consideradas para o cálculo do chain ladder (utilizando notação do triângulo acumulado, conforme Figura 7):

1) $E\left(D_{i,k+1} \mid D_{i,1},\ldots,D_{i,k}\right) = D_{i,k} f_k$, *P-q.c.*;

2) Independência entre os anos de acidentes (linhas); e

3) $Var\left(D_{i,k+1} \mid D_{i,1},\ldots,D_{i,k}\right) = D_{i,k} \alpha_k^{(cl)2}$ *P-q.c.*, na qual

$$\alpha_k^{(cl)2} = \frac{1}{J-k-1} \sum_{i=1}^{J-k} D_{i,k} \left(\frac{D_{i,k+1}}{D_{i,k}} - f_k \right)^2, \qquad 1 \le k \le J-2. \tag{1}$$

O método do chain ladder é bastante simples de ser implementado e as expressões para os estimadores de reserva IBNR estão apresentadas abaixo:

$$\hat{D}_{i,k+1} = D_{i,k} f_k, \qquad 2 \le i \le J \quad e \quad 1 \le k \le J-1, \tag{2}$$

sendo que

$$f_k = \frac{\sum_{i=1}^{J-k} D_{i,k+1}}{\sum_{i=1}^{I-k} D_{i,k}}, \qquad 2 \le i \le J \quad e \quad 1 \le k \le J-1. \tag{3}$$

Serão perseguidas, durante toda a Dissertação, expressões para o cálculo dos EQMs teóricos provenientes de cada uma das metodologias a serem estudadas nesta Dissertação e para cada tipo de reserva (vide seção 2.5). Para o chain ladder, quando usado para a obtenção das reservas total e por linha, serão utilizados resultados derivados por Mack (1983), que são expressões para EQMs teóricos condicionais dada $\Im_\Delta$:

$$E\hat{Q}M\left(D^{(i)} \mid \Im_\Delta\right) = \hat{D}_{i,J}^2 \sum_{k=J+1-i}^{J-1} \frac{\alpha_k^{(cl)2}}{\hat{f}_k} \left(\frac{1}{\hat{D}_{i,k}} + \frac{1}{\sum_{l=1}^{J-k} D_{l,k}} \right) \quad P-q.c. \qquad , (4)$$

$$E\hat{Q}M\left(D^{(T)} \mid \Im_{\Delta}\right)=\sum_{i=2}^{J}\left\{E\hat{Q}M\left(D^{(i)}\right)+\hat{D}_{i,J}\left(\sum_{j=i+1}^{J}\hat{D}_{i,J}\right)\sum_{k=J+1-i}^{J-1}\frac{\left(\dfrac{2\alpha_k^{(cl)^2}}{\hat{f}_k^2}\right)}{\sum_{n=1}^{J-k}D_{n,k}}\right\} \quad P-q.c. \;, (5)$$

nos quais, $D^{(i)}$ e $D^{(T)}$ representam as reservas por linha, $i = 2,\ldots,J$, e reserva total respectivamente.

3.2. Extensão do chain ladder para a reserva por ano de calendário

Ainda com base nas hipóteses de Mack, é possível deduzir um estimador, que "imita" o chain ladder tradicional, para a reserva por ano de calendário, assim como um respectivo EQM teórico condicional do mesmo. O ponto de partida é o seguinte lema, o qual depende apenas de uma das hipóteses de Mack.

Lema 1*: Sejam duas linhas diferentes i e k do triângulo. Então, sob a hipótese 2 de Mack, segue-se, com probabilidade 1, que*

$$Cov\left(C_{i,J-i+1}, C_{k,J-k+1} \mid \Im_{i,k}\right)=0\,. \tag{6}$$

Do resultado acima (cuja prova é relegada à subseção 3.3.1), obtém-se o Teorema 1 abaixo (prova na subseção 3.3.2), do qual obtêm-se as quantidades relacionadas à estimação da reserva por ano de calendário. É interessante notar analogia com o chain ladder tradicional em (2), até mesmo por estes dependerem de comum "matéria prima", quais sejam, os fatores que aparecem em (1) e (3).

Teorema 1: *Utilizando a notação de duplo índice do triângulo (cf.* **Erro! Vínculo não válido.** *e* **Erro! Vínculo não válido.***), defina* $Y^d \equiv \sum_{k=2}^{J} C_{k,J-k+1}$. *Sob as hipóteses 1, 2 e 3 de Mack, são válidos os seguintes resultados:*

i) $E\left(Y^d \mid \Im_{\Delta}\right)=\sum_{k=2}^{J}\left(D_{k,J-k}f_{J-k}-D_{k,J-k}\right)\equiv\hat{Y}^d$ *P-q.c.*

ii) $EQM\left(\hat{Y}^d \mid \Im_{\Delta}\right)=\sum_{k=2}^{J}D_{k,J-k}\alpha_{J-k}^{(cl)^2}$. *P-q.c.* (7)

As fórmulas (i) e (ii) em (7) têm um atrativo adicional em relação às expressões do chain ladder tradicional (para as reservas total e por linhas): a fórmula (i) é uma esperança condicional dada $\Im_{\Delta}$, o que, além de prover um estimador não-tendencioso da reserva por ano de calendário, implica – caso as hipóteses de Mack sejam de fato verificadas na prática – que a esperança incondicional da fórmula (ii) tem o EQM teórico *incondicional* mínimo associado a qualquer estimador, da reserva por ano de calendário, que possa ser escrito como função Borel mensurável dos dados dos triângulo.

3.3. Apêndice: Provas

3.3.1. Prova do Lema 1

Seja $A \times B \in \underline{B}^{n_i} \times \underline{B}^{n_k}$. Então, $((Y_i, Y_k) \in A \times B) = (Y_i \in A, Y_k \in B) = (Y_i \in A) \cap (Y_k \in B)$.

Segue-se, da definição de covariância condicional e da identidade acima, que

$$
\begin{aligned}
& E\left(Cov\left(C_{i,J-i+1}, C_{k,J-k+1} \mid \Im_{i,k}\right) I_{((Y_i,Y_k) \in A \times B)}\right) \\
&= E\left\{E\left[\left(C_{i,J-i+1} - E\left(C_{i,J-i+1} \mid \Im_{i,k}\right)\right)\left(C_{k,J-k+1} - E\left(C_{k,J-k+1} \mid \Im_{i,k}\right)\right) \mid \Im_{i,k}\right] I_{(Y_i \in A) \cap (Y_k \in B)}\right\} \\
&= E\left[\left(C_{i,J-i+1} - E\left(C_{i,J-i+1} \mid \Im_i\right)\right)\left(C_{k,J-k+1} - E\left(C_{k,J-k+1} \mid \Im_k\right)\right) I_{(Y_i \in A)} I_{(Y_k \in B)}\right] \\
&= E\left[\left(C_{i,J-i+1} - E\left(C_{i,J-i+1} \mid \Im_i\right)\right) I_{(Y_i \in A)}\right] E\left[\left(C_{k,J-k+1} - E\left(C_{k,J-k+1} \mid \Im_k\right)\right) I_{(Y_k \in B)}\right] = 0. \quad (8)
\end{aligned}
$$

A segunda igualdade em (8) decorre da definição de esperança condicional dada a σ-álgebra $\Im_{i,k}$ e de Shao (2003), Proposição 1.11, (que é aplicável, pois as linhas do triângulo são independentes – cf. hipótese 2 de Mack), e a terceira igualdade decorre de independência entre linhas (observando que $E\left(C_{i,J-i+1} \mid \Im_i\right)$ e $I_{(Y_i \in A)}$ são $\Im_i$-mensuráveis, assim como $E\left(C_{k,J-k+1} \mid \Im_k\right)$ e $I_{(Y_k \in B)}$ são $\Im_k$-mensuráveis).

Por outro lado, $\underline{B}^{n_i} \times \underline{B}^{n_k}$ é π-sistema sobre $\Re^{n_i+n_k}$, o que é suficiente para que $(Y_i, Y_k)^{-1}\left(\underline{B}^{n_i} \times \underline{B}^{n_k}\right)$ seja π-sistema sobre o espaço amostral primitivo Ω. Além disso,

$$
\sigma(Y_i, Y_k) = (Y_i, Y_k)^{-1}\left(\underline{B}^{n_i+n_k}\right) = (Y_i, Y_k)^{-1}\left(\sigma\left(\underline{B}^{n_i} \times \underline{B}^{n_k}\right)\right) = \sigma\left((Y_i, Y_k)^{-1}\left(\underline{B}^{n_i} \times \underline{B}^{n_k}\right)\right),
$$

sendo que a primeira igualdade acima é garantida por Billingsley (1995), Teorema 20.1, a segunda por Billingsley (1995), Exemplo 18.1 e a terceira por Shao (2003), Capítulo 1, Exercício 13. Combinando estas duas últimas constatações com a identidade provada em (8), que é válida para qualquer conjunto arbitrário em $\underline{B}^{n_i} \times \underline{B}^{n_k}$ (ou, equivalentemente, para qualquer evento em $(Y_i, Y_k)^{-1}\left(\underline{B}^{n_i} \times \underline{B}^{n_k}\right)$, tem-se, por Billingsley (1995), Teorema 16.10, que a identidade (6) é válida P-quase certamente. *Q.E.D.*

3.3.2.
Prova do Teorema 1

Pela definição de Y^d, tem-se que

$$E\left(Y^d \mid \Im_\Delta\right) = \sum_{k=2}^{J} E\left(C_{k,J-k+1} \mid \Im_\Delta\right) = \sum_{k=2}^{J} E\left(C_{k,J-k+1} \mid \Im_k\right) =$$
$$\sum_{k=2}^{J} E\left(D_{k,J-k+1} - D_{k,J-k} \mid \Im_k\right) = \sum_{k=2}^{J} \left(D_{k,J-k} f_{J-k} - D_{k,J-k}\right),$$

sendo que a segunda igualdade decorre de Shao (2003), Proposição 1.11 (as linhas são independentes – cf. hipótese 2 de Mack) e a última igualdade vem da hipótese 1 de Mack e de $D_{k,J-k}$ ser $\Im_k$-mensurável. Isto prova a fórmula (i). Quanto à (ii), segue-se, da definição de EQM condicional descrita na seção 3.1, que

$$EQM\left(\hat{Y}^d \mid \Im_\Delta\right) = Var\left(Y^d \mid \Im_\Delta\right) + \left(E\left(Y^d \mid \Im_\Delta\right) - \hat{Y}^d\right)^2$$
$$= \sum_{k=2}^{J} Var\left(C_{k,J-k+1} \mid \Im_\Delta\right) + 2\sum_{i<k} Cov\left(C_{i,J-i+1}, C_{k,J-k+1} \mid \Im_\Delta\right)$$
$$= \sum_{k=2}^{J} Var\left(C_{k,J-k+1} \mid \Im_k\right) + 2\sum_{i<k} Cov\left(C_{i,J-i+1}, C_{k,J-k+1} \mid \Im_{i,k}\right)$$
$$= \sum_{k=2}^{J} Var\left(D_{k,J-k+1} - D_{k,J-k} \mid \Im_k\right) = \sum_{k=2}^{J} D_{k,J-k} \alpha_{J-k}^{(cl)\,2},$$

sendo que a terceira igualdade decorre de Shao (2003), Proposição 1.11, a quarta igualdade é garantida pelo Lema 1 e, finalmente, a última igualdade vem da hipótese 3 de Mack e da $\Im_k$-mensurabilidade de $D_{k,J-k}$. *Q.E.D.*