Mackの公式 : 支払備金の区間推定

斎藤, 新悟
九州大学大学院数理学研究院

出版情報：COE Lecture Note. 13, pp.11-21, 2009-02-06. 九州大学大学院数理学研究院

Mack の公式：支払備金の区間推定

斎藤新悟
九州大学大学院数理学研究院

1　はじめに

この文書の目的は，Mack [1] の内容を概説し，そこで与えられている Mack の公式の一般化について紹介することである．執筆の際は Wüthrich and Merz [2] を適宜参照した．

チェインラダー法 (chain-ladder method) は支払備金の統計的見積法のうち最も有名なものであり，実務でも広く利用されている．チェインラダー法は従来は単なる計算方法に過ぎないと考えられていたが，実際には多くの確率モデルによって導出されることが分かってきた．Mack はチェインラダー法を導く新たな確率モデルを与えたが，これは**分布によらない** (distribution-free) という点で画期的であった．

さらに，Mack はこの確率モデルに基づいて支払備金を区間推定する公式を提示した．これはシミュレーションや乱数を必要とせず，計算が簡単にできる有用な公式である．

この文書では，第 2 節で Mack モデルとそれによる支払備金の点推定について述べ，第 3 節で支払備金を区間推定する Mack の公式とその一般化を扱う．

2　Mack モデルによる点推定

2.1　設定

事故年度 i の経過年数 j までの累計保険金を $S_{i,j}$ $(i, j = 1, \ldots, n)$ とする．$S_{i,j}$ は正の実数値をとる確率変数である．$i + j \leqq n + 1$ なる i, j に対しては $S_{i,j}$ が既知であるとして，$i + j \geqq n + 2$ に対する $S_{i,j}$ を推定する．

Mack モデルでは次の仮定をおく：

- $\mathbb{R}^n$ 値確率変数 $(S_{i,1}, \ldots, S_{i,n})$ $(i = 1, \ldots, n)$ は独立である（事故年度に関する独立性）．
- 各 $j = 1, \ldots, n-1$ に対してある定数 f_j が存在し，
$$E[S_{i,j+1}|S_{i,1}, \ldots, S_{i,j}] = S_{i,j} f_j$$
がすべての $i = 1, \ldots, n$ に対して成立する．
- 各 $j = 1, \ldots, n-1$ に対してある定数 σ_j が存在し，
$$V(S_{i,j+1}|S_{i,1}, \ldots, S_{i,j}) = S_{i,j} \sigma_j^2$$
がすべての $i = 1, \ldots, n$ に対して成立する．

これらの仮定において $S_{i,j}$ の具体的な分布形は定められていないので，「分布によらない」と表現される．

各 $i,j=1,\dots,n$ に対して $\mathcal{F}_{i,j}=\{S_{i,1},\dots,S_{i,j}\}$ とおくと，2 つめ，3 つめの仮定内の式はそれぞれ $E[S_{i,j+1}|\mathcal{F}_{i,j}]=S_{i,j}f_j$, $V(S_{i,j+1}|\mathcal{F}_{i,j})=S_{i,j}\sigma_j^2$ と書ける．既知の $S_{i,j}$ 全体の集合を $\mathcal{F}=\{S_{i,j}\mid i+j\leqq n+1\}$ とおく．

$i+j\geqq n+2$ のときの $S_{i,j}$ を予測するのが目標であるが，$S_{i,j}$ は確率変数なので点推定としては $E[S_{i,j}|\mathcal{F}]$ を推定するのが妥当である．したがって，$S_{i,j}$ と推定量 $\hat{S}_{i,j}$ の差は，$S_{i,j}$ が確率変数であることから生じる $S_{i,j}$ と $E[S_{i,j}|\mathcal{F}]$ の差と，推定の誤差から生じる $E[S_{i,j}|\mathcal{F}]$ と $\hat{S}_{i,j}$ の差の 2 つに分解できる．このことは 3.1 節でより詳しく考察する．

また，各 $j=1,\dots,n$ に対して $\mathcal{G}_j=\{S_{i,k}\mid i+k\leqq n+1,\ 1\leqq k\leqq j\}=\mathcal{F}\cap\bigcup_{i=1}^n\mathcal{F}_{i,j}$ とおく．$\mathcal{G}_j$ は後に様々な計算をするときに重要な役割を果たす．

2.2 f_j の推定

推定 2.1 $j=1,\dots,n-1$ に対して，f_j を次で推定する：

$$\hat{f}_j=\frac{S_{1,j+1}+\cdots+S_{n-j,j+1}}{S_{1,j}+\cdots+S_{n-j,j}}.$$

注意 2.2 $\hat{f}_j$ は $\mathcal{G}_{j+1}$ の元の式で書ける．

以下で示すように，$\hat{f}_j$ は f_j の不偏推定量であり，さらにある意味で最良推定量となっている．この小節を通じて $j=1,\dots,n-1$ を 1 つとって固定しておく．

命題 2.3 $\alpha_1,\dots,\alpha_{n-j}$ を和が 1 であるような非負確率変数で $\mathcal{G}_j$ の元の式で書けるものとし，

$$\hat{f}'_j=\alpha_1\frac{S_{1,j+1}}{S_{1,j}}+\cdots+\alpha_{n-j}\frac{S_{n-j,j+1}}{S_{n-j,j}}$$

とおくと，$E[\hat{f}'_j|\mathcal{G}_j]=f_j$ が成立する．特に $E[\hat{f}'_j]=f_j$，すなわち $\hat{f}'_j$ は f_j の不偏推定量である．

証明 α_i, $S_{i,j}$ はすべて $\mathcal{G}_j$ の元の式で書けることと，事故年度に関する独立性を用いると，

$$E[\hat{f}'_j|\mathcal{G}_j]=E\left[\sum_{i=1}^{n-j}\alpha_i\frac{S_{i,j+1}}{S_{i,j}}\middle|\mathcal{G}_j\right]=\sum_{i=1}^{n-j}\alpha_i\frac{E[S_{i,j+1}|\mathcal{G}_j]}{S_{i,j}}=\sum_{i=1}^{n-j}\alpha_i\frac{E[S_{i,j+1}|\mathcal{F}_{i,j}]}{S_{i,j}}=\sum_{i=1}^{n-j}\alpha_i f_j=f_j$$

を得る． ∎

例 2.4 (1) $\alpha_i=S_{i,j}/(S_{1,j}+\cdots+S_{n-j,j})$ とおくと $\hat{f}'_j=\hat{f}_j$ となる．

(2) $\alpha_1,\dots,\alpha_{n-j}$ を和が 1 であるような非負定数とした場合，明らかにこれらは $\mathcal{G}_j$ の元の式で書ける．特にこれらがすべて $1/(n-j)$ に等しいとき

$$\hat{f}'_j=\frac{1}{n-j}\left(\frac{S_{1,j+1}}{S_{1,j}}+\cdots+\frac{S_{n-j,j+1}}{S_{n-j,j}}\right)$$

となる．

命題 2.3，例 2.4 (1) より，特に次が成立する：

系 2.5 $E[\hat{f}_j|\mathcal{G}_j] = f_j$ が成立する．特に $E[\hat{f}_j] = f_j$，すなわち $\hat{f}_j$ は f_j の不偏推定量である．

さらに，次の意味で $\hat{f}_j$ は $\hat{f}'_j$ の中で最良である：

命題 2.6 命題 2.3 の $\hat{f}'_j$ の中で，$\hat{f}'_j = \hat{f}_j$ は $V(\hat{f}'_j|\mathcal{G}_j)$ および $V(\hat{f}'_j)$ を最小とする．

証明 まず $V(\hat{f}'_j|\mathcal{G}_j)$ について考える．事故年度に関する独立性より

$$\begin{aligned} V(\hat{f}'_j|\mathcal{G}_j) &= V\left(\sum_{i=1}^{n-j} \alpha_i \frac{S_{i,j+1}}{S_{i,j}}\middle|\mathcal{G}_j\right) = \sum_{i=1}^{n-j} V\left(\alpha_i \frac{S_{i,j+1}}{S_{i,j}}\middle|\mathcal{G}_j\right) = \sum_{i=1}^{n-j} \frac{\alpha_i^2}{S_{i,j}^2} V(S_{i,j+1}|\mathcal{G}_j) \\ &= \sum_{i=1}^{n-j} \frac{\alpha_i^2}{S_{i,j}^2} V(S_{i,j+1}|\mathcal{F}_{i,j}) = \sigma_j^2 \sum_{i=1}^{n-j} \frac{\alpha_i^2}{S_{i,j}} \end{aligned}$$

なので，Cauchy-Schwarz の不等式より

$$V(\hat{f}'_j|\mathcal{G}_j) \sum_{i=1}^{n-j} S_{i,j} = \sigma_j^2 \left(\sum_{i=1}^{n-j} \frac{\alpha_i^2}{S_{i,j}}\right)\left(\sum_{i=1}^{n-j} S_{i,j}\right) \geqq \sigma_j^2 \left(\sum_{i=1}^{n-j} \alpha_i\right)^2 = \sigma_j^2$$

となる．$\alpha_i = S_{i,j}/(S_{1,j} + \cdots + S_{n-j,j})$ のとき，すなわち $\hat{f}'_j = \hat{f}_j$ のとき，等号が成立するので $V(\hat{f}'_j|\mathcal{G}_j)$ は最小値 $\sigma_j^2/\sum_{i=1}^{n-j} S_{i,j}$ をとる．

次に $V(\hat{f}'_j)$ について考える．命題 2.3 より

$$V(\hat{f}'_j) = V(E[\hat{f}'_j|\mathcal{G}_j]) + E[V(\hat{f}'_j|\mathcal{G}_j)] = V(f_j) + E[V(\hat{f}'_j|\mathcal{G}_j)] = E[V(\hat{f}'_j|\mathcal{G}_j)]$$

なので，上に示した不等式より

$$V(\hat{f}'_j) \geqq E\left[\frac{\sigma_j^2}{\sum_{i=1}^{n-j} S_{i,j}}\right]$$

となる．$\alpha_i = S_{i,j}/(S_{1,j}+\cdots+S_{n-j,j})$ のとき，すなわち $\hat{f}'_j = \hat{f}_j$ のとき，等号が成立するので $V(\hat{f}'_j)$ は最小値 $E[\sigma_j^2/\sum_{i=1}^{n-j} S_{i,j}]$ をとる． ∎

次の補題は第 3 節で必要となる：

補題 2.7 $j = 1, \ldots, n-1$ に対して，次が成立する：

$$E[(f_j - \hat{f}_j)^2|\mathcal{G}_j] = \frac{\sigma_j^2}{\sum_{i=1}^{n-j} S_{i,j}}.$$

証明 系 2.5 より，

$$\begin{aligned} E[(f_j - \hat{f}_j)^2|\mathcal{G}_j] &= V(f_j - \hat{f}_j|\mathcal{G}_j) + E[f_j - \hat{f}_j|\mathcal{G}_j]^2 = V(\hat{f}_j|\mathcal{G}_j) = \frac{\sum_{i=1}^{n-j} V(S_{i,j+1}|\mathcal{G}_j)}{\left(\sum_{i=1}^{n-j} S_{i,j}\right)^2} \\ &= \frac{\sum_{i=1}^{n-j} V(S_{i,j+1}|\mathcal{F}_{i,j})}{\left(\sum_{i=1}^{n-j} S_{i,j}\right)^2} = \frac{\sum_{i=1}^{n-j} S_{i,j}\sigma_j^2}{\left(\sum_{i=1}^{n-j} S_{i,j}\right)^2} = \frac{\sigma_j^2}{\sum_{i=1}^{n-j} S_{i,j}} \end{aligned}$$

となる． ∎

2.3 $S_{i,j}$ の推定

命題 2.8 $i+j \geqq n+1$ のとき次が成立する：

$$E[S_{i,j}|\mathcal{F}] = S_{i,n+1-i} f_{n+1-i} \cdots f_{j-1}.$$

証明 $i=1,\ldots,n$ を固定し，帰納的に $j=n+1-i,\ldots,n$ に対して主張を証明する．事故年度に関する独立性より，左辺は $E[S_{i,j}|\mathcal{F}_{i,n+1-i}]$ に等しい．$j=n+1-i$ のときは $S_{i,j}=S_{i,n+1-i} \in \mathcal{F}_{i,n+1-i}$ なので主張は明らかである．j で正しいと仮定すると，

$$\begin{aligned} E[S_{i,j+1}|\mathcal{F}_{i,n+1-i}] &= E\big[E[S_{i,j+1}|\mathcal{F}_{i,j}]\big|\mathcal{F}_{i,n+1-i}\big] = E[S_{i,j} f_j|\mathcal{F}_{i,n+1-i}] \\ &= E[S_{i,j}|\mathcal{F}_{i,n+1-i}] f_j = S_{i,n+1-i} f_{n+1-i} \cdots f_{j-1} \cdot f_j \qquad \text{（帰納法の仮定）} \end{aligned}$$

となり，$j+1$ でも正しいことが分かる． ∎

この命題から，次のように推定することが妥当であることが分かる：

推定 2.9 $i+j \geqq n+1$ のとき $S_{i,j}$ を次で推定する：

$$\hat{S}_{i,j} = S_{i,n+1-i} \hat{f}_{n+1-i} \cdots \hat{f}_{j-1}.$$

注意 2.10 $i+j=n+1$ のときは $S_{i,j}=S_{i,n+1-i}$ は既知で推定の必要はないが，このときも $\hat{S}_{i,j}$ を定義しておくと後に記法が簡便になる．

命題 2.11 $i+j \geqq n+1$ のとき $E[\hat{S}_{i,j}] = E[S_{i,j}]$ が成立する．

証明 より強く $E[\hat{S}_{i,j}|\mathcal{G}_{n+1-i}] = E[S_{i,j}|\mathcal{G}_{n+1-i}]$ であることを証明する．右辺は命題 2.8 より

$$E[S_{i,j}|\mathcal{G}_{n+1-i}] = E[S_{i,j}|\mathcal{F}_{i,n+1-i}] = E[S_{i,j}|\mathcal{F}] = S_{i,n+1-i} f_{n+1-i} \cdots f_{j-1}$$

となるので，$\hat{S}_{i,j}$ の定義と $S_{i,n+1-i} \in \mathcal{G}_{n+1-i}$ であることより $E[\hat{f}_{n+1-i} \cdots \hat{f}_{j-1}|\mathcal{G}_{n+1-i}] = f_{n+1-i} \cdots f_{j-1}$ であることを証明すればよい．$j=n+1-i$ のときは明らかである．j のときに等式が成立すると仮定すると，注意 2.2，系 2.5 より

$$\begin{aligned} E[\hat{f}_{n+1-i} \cdots \hat{f}_j|\mathcal{G}_{n+1-i}] &= E\big[E[\hat{f}_{n+1-i} \cdots \hat{f}_j|\mathcal{G}_j]\big|\mathcal{G}_{n+1-i}\big] = E[\hat{f}_{n+1-i} \cdots \hat{f}_{j-1} f_j|\mathcal{G}_{n+1-i}] \\ &= f_{n+1-i} \cdots f_{j-1} \cdot f_j \qquad \text{（帰納法の仮定）} \end{aligned}$$

となり，$j+1$ でも正しいことが分かる． ∎

注意 2.12 $E[\hat{S}_{i,j}|\mathcal{F}] = E[S_{i,j}|\mathcal{F}]$ が成立するわけではない．実際，$\hat{S}_{i,j}$ は $\mathcal{F}$ の元の式で書けるので左辺は $\hat{S}_{i,j} = S_{i,n+1-i} \hat{f}_{n+1-i} \cdots \hat{f}_{j-1}$ に等しく，右辺は命題 2.8 より $S_{i,n+1-i} f_{n+1-i} \cdots f_{j-1}$ に等しい．

2.4　σ_j の推定

定義 2.13 $j=1,\dots,n-1$ に対して σ_j^2 を次で推定する：

$$\hat\sigma_j^2 = \begin{cases} \frac{1}{n-j-1}\sum_{i=1}^{n-j} S_{i,j}(S_{i,j+1}/S_{i,j}-\hat f_j)^2 & (j=1,\dots,n-2), \\ \min\{(\hat\sigma_{n-2}^2)^2/\hat\sigma_{n-3}^2, \hat\sigma_{n-2}^2, \hat\sigma_{n-3}^2\} & (j=n-1). \end{cases}$$

命題 2.14 $j=1,\dots,n-2$ に対して $E[\hat\sigma_j^2]=\sigma_j^2$，すなわち $\hat\sigma_j^2$ は σ_j^2 の不偏推定量である．

証明 より強く，$E[\hat\sigma_j^2|\mathcal{G}_j]=\sigma_j^2$ を証明する．簡単のため $T=\sum_{i=1}^{n-j}S_{i,j+1}$, $U=\sum_{i=1}^{n-j}S_{i,j}$ とおく．$\hat f_j = T/U$ であり，U は $\mathcal{G}_j$ の元の式で書け，$E[T|\mathcal{G}_j]=Uf_j$ である．よって

$$\begin{aligned}(n-j-1)\hat\sigma_j^2 &= \sum_{i=1}^{n-j} S_{i,j}\left(\frac{S_{i,j+1}}{S_{i,j}}-\hat f_j\right)^2 = \sum_{i=1}^{n-j}\frac{S_{i,j+1}^2}{S_{i,j}} - 2\hat f_j\sum_{i=1}^{n-j}S_{i,j+1} + \hat f_j^2\sum_{i=1}^{n-j}S_{i,j} \\ &= \sum_{i=1}^{n-j}\frac{S_{i,j+1}^2}{S_{i,j}} - 2\cdot\frac{T}{U}\cdot T + \left(\frac{T}{U}\right)^2 U = \sum_{i=1}^{n-j}\frac{S_{i,j+1}^2}{S_{i,j}} - \frac{T^2}{U}\end{aligned}$$

となるので，

$$(n-j-1)E[\hat\sigma_j^2|\mathcal{G}_j] = \sum_{i=1}^{n-j}\frac{E[S_{i,j+1}^2|\mathcal{G}_j]}{S_{i,j}} - \frac{E[T^2|\mathcal{G}_j]}{U}$$

を得る．ここで，$i=1,\dots,n-j$ に対して

$$E[S_{i,j+1}^2|\mathcal{G}_j] = V(S_{i,j+1}|\mathcal{G}_j) + E[S_{i,j+1}|\mathcal{G}_j]^2 = V(S_{i,j+1}|\mathcal{F}_{i,j}) + E[S_{i,j+1}|\mathcal{F}_{i,j}]^2 = S_{i,j}\sigma_j^2 + S_{i,j}^2 f_j^2$$

であり，

$$\begin{aligned}E[T^2|\mathcal{G}_j] &= V(T|\mathcal{G}_j) + E[T|\mathcal{G}_j]^2 = \sum_{i=1}^{n-j}V(S_{i,j+1}|\mathcal{G}_j) + \left(\sum_{i=1}^{n-j}E[S_{i,j+1}|\mathcal{G}_j]\right)^2 \\ &= \sum_{i=1}^{n-j}V(S_{i,j+1}|\mathcal{F}_{i,j}) + \left(\sum_{i=1}^{n-j}E[S_{i,j+1}|\mathcal{F}_{i,j}]\right)^2 = \sum_{i=1}^{n-j}S_{i,j}\sigma_j^2 + \left(\sum_{i=1}^{n-j}S_{i,j}f_j\right)^2 \\ &= U\sigma_j^2 + U^2 f_j^2\end{aligned}$$

なので，

$$(n-j-1)E[\hat\sigma_j^2|\mathcal{G}_j] = \sum_{i=1}^{n-j}(\sigma_j^2 + S_{i,j}f_j^2) - (\sigma_j^2 + Uf_j^2) = (n-j-1)\sigma_j^2$$

となり，命題が従う． ∎

3　Mack モデルによる区間推定

3.1　平均 2 乗誤差と信頼区間

今考察している問題を抽象的に表現すると，未知の確率変数 X（例えば $S_{n,n}$）を $\mathcal{F}$ の元の式で書ける確率変数 $\hat X$（例えば $\hat S_{n,n} = S_{n,1}\hat f_1\cdots\hat f_{n-1}$）で推定するということになる．この推定がど

れくらい良いかを測る1つの指標が**平均2乗誤差** (mean squared error) であり，

$$\operatorname{mse}\hat{X} = E[(\hat{X}-X)^2|\mathcal{F}]$$

で定義される．$\hat{X}$ が $\mathcal{F}$ の元の式で書けることに注意して変形すると

$$\operatorname{mse}\hat{X} = V(\hat{X}-X|\mathcal{F}) + E[\hat{X}-X|\mathcal{F}]^2 = V(X|\mathcal{F}) + (E[X|\mathcal{F}]-\hat{X})^2$$

となり，平均2乗誤差は分散と推定誤差の和であることが分かる．前者は X と $E[X|\mathcal{F}]$ の差を表す指標である．

平均2乗誤差が分かると，Chebyshev の不等式

$$P\big(\{|\hat{X}-X| \geqq r\}\big|\mathcal{F}\big) \leqq \frac{E[(\hat{X}-X)^2|\mathcal{F}]}{r^2} = \frac{\operatorname{mse}\hat{X}}{r^2}$$

を用いて信頼区間を求めることができる．実際，$\alpha > 0$ (例えば $\alpha = 0.05$) として $r = \sqrt{\alpha^{-1}\operatorname{mse}\hat{X}}$ とおくと，$P\big(\{|\hat{X}-X| \geqq r\}\big|\mathcal{F}\big) \leqq \alpha$ すなわち $P\big(\{|\hat{X}-X| < r\}\big|\mathcal{F}\big) \geqq 1-\alpha$ なので，開区間 $(\hat{X}-r, \hat{X}+r)$ は X の信頼水準 $1-\alpha$ 以上の信頼区間となる．

3.2　平均2乗誤差の推定

この小節では，いくつかの重要な確率変数に対する平均2乗誤差の推定を与える公式を述べる．根拠は次小節で与えられる．

推定 3.1 $i+j \geqq n+2$ のとき，$\operatorname{mse}\hat{S}_{i,j}$ を次で推定する：

$$(\operatorname{mse}\hat{S}_{i,j})\widehat{\ } = \hat{S}_{i,j}^2 \sum_{l=n+1-i}^{j-1} \frac{\hat{\sigma}_l^2}{\hat{f}_l^2}\left(\frac{1}{\hat{S}_{i,l}} + \frac{1}{\sum_{m=1}^{n-l} S_{m,l}}\right).$$

次の推定は [1] の主結果であり，[2] で Mack の公式と呼ばれている：

推定 3.2 支払備金 $T = \sum_{i=2}^n (S_{i,n} - S_{i,n+1-i})$ について，その推定量 $\hat{T} = \sum_{i=2}^n (\hat{S}_{i,n} - S_{i,n+1-i})$ の平均2乗誤差 $\operatorname{mse}\hat{T}$ を次で推定する：

$$\begin{aligned}(\operatorname{mse}\hat{T})\widehat{\ } = \sum_{i=2}^n &\left(\hat{S}_{i,n}^2 \sum_{l=n+1-i}^{n-1} \frac{\hat{\sigma}_l^2}{\hat{f}_l^2}\left(\frac{1}{\hat{S}_{i,l}} + \frac{1}{\sum_{m=1}^{n-l} S_{m,l}}\right)\right) \\ &+ 2\sum_{i=2}^{n-1} \hat{S}_{i,n}\left(\sum_{i'=i+1}^{n} \hat{S}_{i',n}\right)\left(\sum_{l=n+1-i}^{n-1} \frac{\hat{\sigma}_l^2}{\hat{f}_l^2 \sum_{m=1}^{n-l} S_{m,l}}\right).\end{aligned}$$

注意 3.3 上式に現れる $S_{m,l}$ は，$m+l \leqq (n-l)+l = n$ なので $\mathcal{F}$ の元である．

推定 3.4 次年度の支払保険金 $T = \sum_{i=2}^n (S_{i,n+2-i} - S_{i,n+1-i})$ について，その推定量 $\hat{T} = \sum_{i=2}^n (\hat{S}_{i,n+2-i} - S_{i,n+1-i})$ の平均2乗誤差 $\operatorname{mse}\hat{T}$ を次で推定する：

$$(\operatorname{mse}\hat{T})\widehat{\ } = \sum_{i=2}^n \hat{S}_{i,n+2-i}^2 \frac{\hat{\sigma}_{n+1-i}^2}{\hat{f}_{n+1-i}^2}\left(\frac{1}{\hat{S}_{i,n+1-i}} + \frac{1}{\sum_{m=1}^{i-1} S_{m,n+1-i}}\right).$$

3.3　平均 2 乗誤差の推定の根拠

3.3.1　一般化

前小節で述べた平均 2 乗誤差の推定は，すべて次の推定の特別な場合である：

推定 3.5 各 $i = 1, \dots, n$ に対して $n+1-i \leqq j_i \leqq k_i \leqq n$ なる j_i, k_i が定まっているとする．$T = \sum_{i=1}^{n}(S_{i,k_i} - S_{i,j_i})$ について，その推定量 $\hat{T} = \sum_{i=1}^{n}(\hat{S}_{i,k_i} - \hat{S}_{i,j_i})$ の平均 2 乗誤差 $\mathrm{mse}\,\hat{T}$ を次で推定する：

$$(\mathrm{mse}\,\hat{T})\widehat{\ } = \sum_{i,l=1}^{n} \frac{\hat{\varphi}_{i,l}^2 \hat{\sigma}_l^2}{\hat{S}_{i,l}\hat{f}_l^2} + \sum_{i,i',l=1}^{n} \frac{\hat{\varphi}_{i,l}\hat{\varphi}_{i',l}\hat{\sigma}_l^2}{\hat{f}_l^2 \sum_{m=1}^{n-l} S_{m,l}}.$$

ただし，$i, l = 1, \dots, n$ に対して $\hat{\varphi}_{i,l}$ を次で定義する：

$$\hat{\varphi}_{i,l} = \begin{cases} \hat{S}_{i,k_i} - \hat{S}_{i,j_i} & (n+1-i \leqq l < j_i), \\ \hat{S}_{i,k_i} & (j_i \leqq l < k_i), \\ 0 & (1 \leqq l < n+1-i,\ k_i \leqq l \leqq n). \end{cases}$$

注意 3.6 $(\mathrm{mse}\,\hat{T})\widehat{\ }$ の式の 1 つめの $\sum$ の中に $\hat{S}_{i,l}$ が現れており，これは $i + l \leqq n$ のときは定義されていないが，このときは分子の $\hat{\varphi}_{i,l}$ が 0 なので問題ない．また，$l = n$ に対する $\hat{f}_l$, $\hat{\sigma}_l$ も定義されていないが，同様の理由で問題ない．

例 3.7　(1) $j_i = n+1-i$, $k_i = n$ とおくと $T = \sum_{i=2}^{n}(S_{i,n} - S_{i,n+1-i})$ となり，推定 3.2 を得る．

(2) $j_i = n+1-i$, $k_1 = n$, $k_i = n+2-i$ $(i \geqq 2)$ とおくと $T = \sum_{i=2}^{n}(S_{i,n+2-i} - S_{i,n+1-i})$ となり，推定 3.4 を得る．

(3) $p + q \geqq n+2$ なる p, q を固定し，$j_i = k_i = n+1-i$ $(i \neq p)$, $j_p = n+1-p$, $k_p = q$ とおくと $T = S_{p,q} - S_{p,n+1-p}$ となる．このとき $\hat{T} = \hat{S}_{p,q} - S_{p,n+1-p}$ より $\hat{T} - T = \hat{S}_{p,q} - S_{p,q}$ なので $\mathrm{mse}\,\hat{T} = \mathrm{mse}\,\hat{S}_{p,q}$ となり，推定 3.1 を得る．

以下，推定 3.5 の根拠を述べる．j_i, k_i, T, $\hat{\varphi}_{i,l}$ は推定 3.5 のように定義されているとする．$\mathrm{mse}\,\hat{T}$ を分散と推定誤差に分解すると

$$\begin{aligned} \mathrm{mse}\,\hat{T} &= V(T|\mathcal{F}) + (E[T|\mathcal{F}] - \hat{T})^2 \\ &= V\left(\sum_{i=1}^{n}(S_{i,k_i} - S_{i,j_i}) \middle| \mathcal{F} \right) + \left(\sum_{i=1}^{n}\bigl(E[S_{i,k_i} - S_{i,j_i}|\mathcal{F}] - (\hat{S}_{i,k_i} - \hat{S}_{i,j_i})\bigr) \right)^2 \end{aligned}$$

となることに注意し，これらを別々に推定する．記法の簡便のため，$i + j \geqq n+1$ のとき $g_j = f_{n+1-i} \cdots f_{j-1}$, $\hat{g}_j = \hat{f}_{n+1-i} \cdots \hat{f}_{j-1}$ と書く．定義より $g_{j+1} = g_j f_j$, $\hat{g}_{j+1} = \hat{g}_j \hat{f}_j$ が成立する．注意 2.2，系 2.5 より $\hat{g}_j$ は $\mathcal{G}_j$ の元の式で書け，$E[\hat{g}_{j+1}|\mathcal{G}_j] = f_j \hat{g}_j$ が成立する．また，命題 2.8 より $E[S_{i,j}|\mathcal{F}] = S_{i,n+1-i} g_j$ であり，推定 2.9 より $\hat{S}_{i,j} = S_{i,n+1-i}\hat{g}_j$ である．

3.3.2 分散の推定

補題 3.8 $i+j \geqq n+1$ のとき次が成立する：

$$V(S_{i,j}|\mathcal{F}) = \sum_{l=n+1-i}^{j-1} \frac{S_{i,n+1-i}g_j^2\sigma_l^2}{g_l f_l^2}.$$

証明 $l = n+1-i, \ldots, j-1$ に対して

$$\begin{aligned} V(S_{i,l+1}|\mathcal{F}) &= V(S_{i,l+1}|\mathcal{F}_{i,n+1-i}) = E[V(S_{i,l+1}|\mathcal{F}_{i,l})|\mathcal{F}_{i,n+1-i}] + V(E[S_{i,l+1}|\mathcal{F}_{i,l}]|\mathcal{F}_{i,n+1-i}) \\ &= E[S_{i,l}\sigma_l^2|\mathcal{F}_{i,n+1-i}] + V(S_{i,l}f_l|\mathcal{F}_{i,n+1-i}) = S_{i,n+1-i}g_l\sigma_l^2 + V(S_{i,l}|\mathcal{F})f_l^2 \end{aligned}$$

なので，両辺を g_{l+1}^2 で割ると

$$\frac{V(S_{i,l+1}|\mathcal{F})}{g_{l+1}^2} = \frac{S_{i,n+1-i}g_l\sigma_l^2}{g_{l+1}^2} + \frac{V(S_{i,l}|\mathcal{F})f_l^2}{g_{l+1}^2} = \frac{S_{i,n+1-i}\sigma_l^2}{g_l f_l^2} + \frac{V(S_{i,l}|\mathcal{F})}{g_l^2}$$

となり，l に関する和をとると

$$\frac{V(S_{i,j}|\mathcal{F})}{g_j^2} = \sum_{l=n+1-i}^{j-1} \frac{S_{i,n+1-i}\sigma_l^2}{g_l f_l^2} + V(S_{i,n+1-i}|\mathcal{F})$$

を得る．$S_{i,n+1-i} \in \mathcal{F}$ より $V(S_{i,n+1-i}|\mathcal{F}) = 0$ なので，補題が成立することが分かる． ∎

ここで，$i, l = 1, \ldots, n$ に対して $\varphi_{i,l}$ を次で定義する：

$$\varphi_{i,l} = \begin{cases} E[S_{i,k_i}|\mathcal{F}] - E[S_{i,j_i}|\mathcal{F}] & (n+1-i \leqq l < j_i), \\ E[S_{i,k_i}|\mathcal{F}] & (j_i \leqq l < k_i), \\ 0 & (1 \leqq l < n+1-i,\ k_i \leqq l \leqq n). \end{cases}$$

補題 3.9 次が成立する：

$$V\left(\sum_{i=1}^{n}(S_{i,k_i} - S_{i,j_i})\middle|\mathcal{F}\right) = \sum_{i,l=1}^{n} \frac{\varphi_{i,l}^2\sigma_l^2}{E[S_{i,l}|\mathcal{F}]f_l^2}.$$

証明 事故年度に関する独立性より左辺は $\sum_{i=1}^n V(S_{i,k_i} - S_{i,j_i}|\mathcal{F})$ に等しいので，任意の i に対して

$$V(S_{i,k_i} - S_{i,j_i}|\mathcal{F}) = \sum_{l=n+1-i}^{k_i-1} \frac{\varphi_{i,l}^2\sigma_l^2}{E[S_{i,l}|\mathcal{F}]f_l^2}$$

であることを証明すればよい（$n+1-i \leqq l \leqq k_i - 1$ でないような l については $\varphi_{i,l} = 0$ であることに注意）．$i = 1, \ldots, n$ を任意にとって固定し，$j_i = j,\ k_i = k$ と書く．

$l=j,\dots,k-1$ に対して

$$\begin{aligned}
V\left(\frac{S_{i,l+1}g_k}{g_{l+1}}-S_{i,j}\middle|\mathcal{F}\right) &= V\left(\frac{S_{i,l+1}g_k}{g_{l+1}}-S_{i,j}\middle|\mathcal{F}_{i,n+1-i}\right)\\
&= E\left[V\left(\frac{S_{i,l+1}g_k}{g_{l+1}}-S_{i,j}\middle|\mathcal{F}_{i,l}\right)\middle|\mathcal{F}_{i,n+1-i}\right]+V\left(E\left[\frac{S_{i,l+1}g_k}{g_{l+1}}-S_{i,j}\middle|\mathcal{F}_{i,l}\right]\middle|\mathcal{F}_{i,n+1-i}\right)\\
&= E\left[\frac{S_{i,l}\sigma_l^2 g_k^2}{g_{l+1}^2}\middle|\mathcal{F}_{i,n+1-i}\right]+V\left(\frac{S_{i,l}f_l g_k}{g_{l+1}}-S_{i,j}\middle|\mathcal{F}_{i,n+1-i}\right)\\
&= \frac{S_{i,n+1-i}g_l\sigma_l^2 g_k^2}{g_{l+1}^2}+V\left(\frac{S_{i,l}f_l g_k}{g_{l+1}}-S_{i,j}\middle|\mathcal{F}\right)\\
&= \frac{S_{i,n+1-i}\sigma_l^2 g_k^2}{g_l f_l^2}+V\left(\frac{S_{i,l}g_k}{g_l}-S_{i,j}\middle|\mathcal{F}\right)
\end{aligned}$$

であり，このような l に関する和をとると

$$V(S_{i,k}-S_{i,j}|\mathcal{F})=\sum_{l=j}^{k-1}\frac{S_{i,n+1-i}\sigma_l^2 g_k^2}{g_l f_l^2}+V\left(\frac{S_{i,j}g_k}{g_j}-S_{i,j}\middle|\mathcal{F}\right)$$

となる．ここで，

$$\sum_{l=j}^{k-1}\frac{S_{i,n+1-i}\sigma_l^2 g_k^2}{g_l f_l^2}=\sum_{l=j}^{k-1}\frac{(S_{i,n+1-i}g_k)^2\sigma_l^2}{S_{i,n+1-i}g_l f_l^2}=\sum_{l=j}^{k-1}\frac{E[S_{i,k}|\mathcal{F}]^2\sigma_l^2}{E[S_{i,l}|\mathcal{F}]f_l^2}=\sum_{l=j}^{k-1}\frac{\varphi_{i,l}^2\sigma_l^2}{E[S_{i,l}|\mathcal{F}]f_l^2}$$

であり，補題 3.8 より

$$\begin{aligned}
V\left(\frac{S_{i,j}g_k}{g_j}-S_{i,j}\middle|\mathcal{F}\right) &= \left(\frac{g_k}{g_j}-1\right)^2 V(S_{i,j}|\mathcal{F})=\left(\frac{g_k}{g_j}-1\right)^2\sum_{l=n+1-i}^{j-1}\frac{S_{i,n+1-i}g_j^2\sigma_l^2}{g_l f_l^2}\\
&= \sum_{l=n+1-i}^{j-1}\frac{(S_{i,n+1-i}g_k-S_{i,n+1-i}g_j)^2\sigma_l^2}{S_{i,n+1-i}g_l f_l^2}\\
&= \sum_{l=n+1-i}^{j-1}\frac{(E[S_{i,k}|\mathcal{F}]-E[S_{i,j}|\mathcal{F}])^2\sigma_l^2}{E[S_{i,l}|\mathcal{F}]f_l^2}=\sum_{l=n+1-i}^{j-1}\frac{\varphi_{i,l}^2\sigma_l^2}{E[S_{i,l}|\mathcal{F}]f_l^2}
\end{aligned}$$

なので求める等式が示された． ∎

この補題より $V\big(\sum_{i=1}^n(S_{i,k_i}-S_{i,j_i})\big|\mathcal{F}\big)$ の推定量

$$\widehat{V\left(\sum_{i=1}^n(S_{i,k_i}-S_{i,j_i})\middle|\mathcal{F}\right)}=\sum_{i,l=1}^n\frac{\hat{\varphi}_{i,l}^2\hat{\sigma}_l^2}{\hat{S}_{i,l}\hat{f}_l^2}$$

が得られる．

3.3.3 推定誤差の推定

$i=1,\dots,n$ に対して $\Phi_i=E[S_{i,k_i}-S_{i,j_i}|\mathcal{F}]-(\hat{S}_{i,k_i}-\hat{S}_{i,j_i})$ とおくと，推定すべき推定誤差は

$$\left(\sum_{i=1}^n\Phi_i\right)^2=\sum_{i,i'=1}^n\Phi_i\Phi_{i'}$$

なので，各 $i, i' = 1, \ldots, n$ に対して $\Phi_i\Phi_{i'}$ を

$$(\Phi_i\Phi_{i'})\widehat{\ } = \sum_{l=1}^{n} \frac{\hat{\varphi}_{i,l}\hat{\varphi}_{i',l}\hat{\sigma}_l^2}{\hat{f}_l^2 \sum_{m=1}^{n-l} S_{m,l}}$$

で推定する根拠を述べればよい．$i, i' = 1, \ldots, n$ を任意にとって固定し，$j_i = j$, $k_i = k$, $j_{i'} = j'$, $k_{i'} = k'$ と書く．

まず，

$$\begin{aligned}
\Phi_i &= E[S_{i,k} - S_{i,j}|\mathcal{F}] - (\hat{S}_{i,k} - \hat{S}_{i,j}) = S_{i,n+1-i}(g_k - g_j - \hat{g}_k + \hat{g}_j) \\
&= S_{i,n+1-i}\left(1 - \frac{\hat{g}_j}{g_j}\right)(g_k - g_j) + S_{i,n+1-i}g_k\left(\frac{\hat{g}_j}{g_j} - \frac{\hat{g}_k}{g_k}\right) \\
&= S_{i,n+1-i}\sum_{l=n+1-i}^{j-1}\left(\frac{\hat{g}_l}{g_l} - \frac{\hat{g}_{l+1}}{g_{l+1}}\right)(g_k - g_j) + S_{i,n+1-i}g_k\sum_{l=j}^{k-1}\left(\frac{\hat{g}_l}{g_l} - \frac{\hat{g}_{l+1}}{g_{l+1}}\right)
\end{aligned}$$

であることに注意して

$$\psi_l = \begin{cases} S_{i,n+1-i}(\hat{g}_l/g_l - \hat{g}_{l+1}/g_{l+1})(g_k - g_j) & (n+1-i \leqq l < j), \\ S_{i,n+1-i}g_k(\hat{g}_l/g_l - \hat{g}_{l+1}/g_{l+1}) & (j \leqq l < k), \\ 0 & (1 \leqq l < n+1-i,\ k \leqq l \leqq n) \end{cases}$$

とおくと $\Phi_i = \sum_{l=1}^n \psi_l$ が成立する．同様に ψ'_l を定めると $\Phi_{i'} = \sum_{l=1}^n \psi'_l$ が成立する．これより $\Phi_i\Phi_{i'} = \sum_{l,l'=1}^n \psi_l\psi'_{l'}$ となる．

補題 3.10 $l, l' = 1, \ldots, n$ に対して，次が成立する：

$$E[\psi_l\psi'_{l'}|\mathcal{G}_{\max\{l,l'\}}] = \begin{cases} \frac{\varphi_{i,l}\varphi_{i',l}\hat{g}_l^2\sigma_l^2}{g_{l+1}^2\sum_{m=1}^{n-l}S_{m,l}} & (l = l'), \\ 0 & (l \neq l'). \end{cases}$$

証明 $\hat{g}_l$ が $\mathcal{G}_l$ の元の式で書けることから ψ_l, ψ'_l は $\mathcal{G}_{l+1}$ の元の式で書ける．また，$E[\hat{g}_{l+1}|\mathcal{G}_l] = f_l\hat{g}_l$ より

$$E\left[\frac{\hat{g}_l}{g_l} - \frac{\hat{g}_{l+1}}{g_{l+1}}\middle|\mathcal{G}_l\right] = \frac{\hat{g}_l}{g_l} - \frac{f_l\hat{g}_l}{g_{l+1}} = 0$$

なので，$E[\psi_l|\mathcal{G}_l] = 0$, $E[\psi'_l|\mathcal{G}_l] = 0$ が成立する．これらを用いて補題の等式を示す．

まず $l \neq l'$ のときを考える．$l < l'$ としても一般性を失わず，このとき

$$E[\psi_l\psi'_{l'}|\mathcal{G}_{\max\{l,l'\}}] = E[\psi_l\psi'_{l'}|\mathcal{G}_{l'}] = \psi_l E[\psi'_{l'}|\mathcal{G}_{l'}] = 0$$

となり確かに補題が成立する．

次に $l = l'$ のときを考える．$l = n$ のときは $\psi_l = \psi'_l = 0$, $\varphi_{i,l} = \varphi_{i',l} = 0$ となり補題は成立するので，以下 $l = 1, \ldots, n-1$ とする．補題 2.7 より

$$E\left[\left(\frac{\hat{g}_l}{g_l} - \frac{\hat{g}_{l+1}}{g_{l+1}}\right)^2\middle|\mathcal{G}_l\right] = \frac{\hat{g}_l^2}{g_{l+1}^2}E[(f_l - \hat{f}_l)^2|\mathcal{G}_l] = \frac{\hat{g}_l^2\sigma_l^2}{g_{l+1}^2\sum_{m=1}^{n-l}S_{m,l}}$$

なので，$n+1-i \leqq l < j$ かつ $n+1-i' \leqq l < j'$ のときは

$$\begin{aligned}
E[\psi_l \psi_l' | \mathcal{G}_l] &= S_{i,n+1-i}(g_k - g_j) \cdot S_{i',n+1-i'}(g_{k'} - g_{j'}) \cdot \frac{\hat{g}_l^2 \sigma_l^2}{g_{l+1}^2 \sum_{m=1}^{n-l} S_{m,l}} \\
&= \frac{(E[S_{i,k}|\mathcal{F}] - E[S_{i,j}|\mathcal{F}])(E[S_{i',k'}|\mathcal{F}] - E[S_{i',j'}|\mathcal{F}])\hat{g}_l^2 \sigma_l^2}{g_{l+1}^2 \sum_{m=1}^{n-l} S_{m,l}} \\
&= \frac{\varphi_{i,l}\varphi_{i',l}\hat{g}_l^2 \sigma_l^2}{g_{l+1}^2 \sum_{m=1}^{n-l} S_{m,l}}
\end{aligned}$$

となる．他の範囲にある l に対しても同様である． ∎

この補題より，$\Phi_i \Phi_{i'} = \sum_{l,l'=1}^{n} \psi_l \psi_{l'}'$ を

$$\sum_{l,l'=1}^{n} E[\psi_l \psi_{l'}' | \mathcal{G}_{\max\{l,l'\}}] = \sum_{l=1}^{n} \frac{\varphi_{i,l}\varphi_{i',l}\hat{g}_l^2 \sigma_l^2}{g_{l+1}^2 \sum_{m=1}^{n-l} S_{m,l}}$$

で近似し，さらに右辺を推定することで

$$(\Phi_i \Phi_{i'})\widehat{} = \sum_{l=1}^{n} \frac{\hat{\varphi}_{i,l}\hat{\varphi}_{i',l}\hat{g}_l^2 \hat{\sigma}_l^2}{\hat{g}_{l+1}^2 \sum_{m=1}^{n-l} S_{m,l}} = \sum_{l=1}^{n} \frac{\hat{\varphi}_{i,l}\hat{\varphi}_{i',l}\hat{\sigma}_l^2}{\hat{f}_l^2 \sum_{m=1}^{n-l} S_{m,l}}$$

を得る．

参考文献

[1] Thomas Mack, *Distribution-free calculation of the standard error of chain ladder reserve estimates*, ASTIN Bulletin, Vol. 23, No. 2, 1993, pp. 213–225.

[2] Mario V. Wüthrich and Michael Merz, *Stochastic claims reserving methods in insurance*, Wiley, 2008.