Estimation of the conditional mean square error of a product of random variables

presentation at the ASTIN Colloquium 2022

Alois Gisler and Gisela Menzel

21 June 2022

Preliminary Remark

- Motivation/Trigger for looking at this problem
  old discussion about estimating the msep in the CL-reserving model
  - 1993 paper by Mack Distribution-free calculation of the standard error of chain ladder reserve estimates
    Mack estimator
  - 2006 paper by BBMW The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited)
    BBMW estimator
    new estimator of the estimation error based on bootstrapping arguments
  - 2006 great discussion
    - three more papers
    - however, no answer to the question, which of the two estimators should be preferred
  - 2020 paper by Gisler Estimation error and bootstrapping in the chain ladder model of Mack
    the two estimators were compared by use of the telescope formula
    conclusion: the BBMW estimator is not a sensible alternative to the Mack estimator
- new discovery
  - problem encountered with CL is more general and inherent with other products of r.v.

Introduction

- simple situation in statistics
  - $X_1, X_2, \ldots, X_{n+1}$ independent r.v. with

$$E[X_i] = \mu, \quad \mathrm{Var}(X_i) = \frac{\sigma^2}{w_i}, \quad \text{where } w_i \text{ are known weights.}$$

  - $\{\mathcal{D} = X_1, X_2, \ldots, X_n\}$ observations;
  - estimators of the unknown parameters

$$\widehat{\mu} = \left(\sum_{i=1}^{n} w_i X_i\right) / \left(\sum_{i=1}^{n} w_i\right) \qquad \text{(minimum variance unbiased estimator)}$$

$$\widehat{\sigma^2} = \frac{1}{n-1} \sum_{i=1}^{n} w_i \left(X_i - \widehat{\mu}\right)^2 \qquad \text{(unbiased estimator)}$$

  - predictor of X_{n+1}

$$\widehat{X_{n+1}} = \widehat{\mu}. \tag{1}$$

  - *conditional* mean square error of prediction

$$\operatorname{msep}_{X_{n+1}|\mathcal{D}}(\widehat{X_{n+1}}) := E\left[\left(X_{n+1} - \widehat{X_{n+1}}\right)^2 \middle| \mathcal{D}\right] = PV + SEE \tag{2}$$

where $PV =$ process variance (random error)

$SEE =$ squared estimation error (parameter risk)

Introduction

- simple situation in statistics
  - **Process Variance**

$$\begin{aligned} PV &:= E\left[(X_{n+1}-\mu)^2 \middle| \mathcal{D} \right] \\ &= \mathrm{Var}\,(X_{n+1}) \\ &= \sigma^2 / w_{n+1} \end{aligned} \tag{3}$$

=> estimator

$$\widehat{PV} = \widehat{\sigma^2} / w_{n+1}$$

  - **Squared Estimation Error**

$$\begin{aligned} SEE &:= \left(\widehat{X_{n+1}} - E\left[\widehat{X_{n+1}} \right] \right)^2 \\ &= (\widehat{\mu} - \mu)^2 . \end{aligned} \tag{4}$$

Note that

$$SEE = E\left[SEE \,|\, \mathcal{D} \right]$$

Conditional on $\mathcal{D}$, $\widehat{\mu}$ is known but SEE is still unknown because μ is not known.

  - **Estimation of the squared estimation error**
    - ★ Replacing μ in (4) by $\widehat{\mu}$ would result in $\widehat{SEE} = 0$; not meaningful!
    - ★ $\nexists$ an estimator of EE which is best for all parameters values μ.

Introduction

- simple situation in statistics
  - **Estimation of the squared estimation error**
    - ★ to find an estimator of SEE we have to study the volatility of $\widehat{\mu}$ around μ

      the observed $\widehat{\mu}$ is *one specific* realisation of the r.v. $\widehat{\mu}$; it could have been different and another realisation of the r.v. $\widehat{\mu}$
    - ★ => *reasonable* volatility measure:

$$E\left[(\widehat{\mu}-\mu)^2\right] = \mathrm{Var}(\widehat{\mu}) = \frac{\sigma^2}{\sum_{i=1}^n w_n}.$$

      Note that

$$\mathrm{Var}(\widehat{\mu}) = E[SEE] \quad \textit{unconditional} \text{ expected value.}$$

    - ★ => estimator

$$\widehat{SEE} = \frac{\widehat{\sigma^2}}{\sum_{i=1}^n w_n}. \tag{5}$$

- **topic of this lecture: products of r.v.**
  - In the following: the r.v. X_i are themselves products of other r.v.
    - ★ Products of r.v. are encountered in many areas such as economics, engineering, insurance
    - ★ for instance:
      total claims load = (# of year risks) · (claims frequency) · (claims average) .

Introduction

- we will derive estimators of the *conditional mean square error of predictions* (msep)
  - in a model with independent r.v.
    - we will show that the estimator can be improved by use of the Telescope formula
  - in a model with uncorrelated r.v.
    - this model occurs in many situations of the insurance practice such as in pricing, solvency and reserving ;
      we will derive and discuss the corresponding estimators
  - in the CL reserving model of Mack
    - by reconsidering the CL-reserving in this broader context we gain further insight and new results

- **Telescope Formula**
  - For any real numbers x_j and y_j, $j = 1, 2, \ldots, J$, it holds that

$$\prod_{j=1}^{J} x_i - \prod_{j=1}^{J} y_i = \sum_{j=1}^{J} \left(\prod_{k=1}^{j-1} x_k \right) (x_j - y_j) \left(\prod_{m=j+1}^{J} y_m \right). \qquad (6)$$

Model with independent r.v.

- **Model Assumptions**

$$X_i = w_i \prod_{j=0}^{J-1} F_{i,j}$$

where $F_{i,j}$, $j = 0, \ldots, J-1$ are independent r.v. with

$$\begin{aligned} E\left[F_{i,j}\right] &= f_j \\ \mathrm{Var}\left(F_{i,j}\right) &= \sigma_j^2 / w_{i,j} \end{aligned}$$

and where w_i and $w_{i,j}$ are known weights.

- $=>$

$$E\left[X_i\right] = w_i \prod_{j=0}^{J-1} f_j.$$

$$\mathrm{Var}\left(X_i\right) = w_i^2 \left\{ \prod_{j=0}^{J-1} \left(f_j^2 + \sigma_j^2 / w_{i,j} \right) - \prod_{j=1}^{J-1} f_j^2 \right\}$$

- Observed data

$$\mathcal{D} = \left\{ F_{i,j}, i = 1, \ldots, n,\ j = 0, \ldots, J-1 \right\}.$$

further notation

$$\mathcal{B}_j = \left\{ F_{i,k} \in \mathcal{D},\ k \leq j \text{ data up to column } j \right\}.$$

Model with independent r.v.

- Estimators of unknown parameters

$$\widehat{f}_j = \frac{\sum_{i=1}^{n} w_{i,j} F_{i,j}}{w_{\bullet,j}} \quad \text{where } w_{\bullet,j} = \sum_{i=1}^{n} w_{i,j}, \qquad \text{(min. variance unbiased linear estimator)}$$

$$\widehat{\sigma_j^2} = \frac{1}{n-1} \sum_{i=0}^{n} w_{i,j} (F_{i,j} - \widehat{f}_j)^2. \qquad \text{(unbiased estimator)}$$

- Predictor of X_{n+1}

$$\widehat{X_{n+1}} = w_{n+1} \prod_{j=0}^{J-1} \widehat{f}_j.$$

- *conditional* mean squared error of prediction

$$\mathrm{msep}_{X_{n+1}|\mathcal{D}}(\widehat{X_{n+1}}) := E\left[\left(X_{n+1} - \widehat{X_{n+1}}\right)^2 \middle| \mathcal{D}\right] = PV + SEE$$

  - **Process Variance**

$$PV := \mathrm{Var}\left(X_{n+1}\right) \tag{7}$$

$$= w_{n+1}^2 \left\{ \prod_{j=0}^{J-1} \left(f_j^2 + \sigma_j^2 / w_{n+1,j}\right) - \prod_{j=0}^{J-1} f_j^2 \right\} \tag{8}$$

=>

$$\widehat{PV} = w_{n+1}^2 \left\{ \prod_{j=0}^{J-1} \left(\widehat{f}_j^{\,2} + \widehat{\sigma_j^2} / w_{n+1,j}\right) - \prod_{j=0}^{J-1} \widehat{f}_j^{\,2} \right\}. \tag{9}$$

Model with independent r.v.

- *conditional* mean squared error of prediction

  - **Squared Estimation Error**

$$SEE := E\left[\left(\widehat{X_{n+1}} - E\left[\widehat{X_{n+1}}\right]\right)^2 \middle| \mathcal{D}\right]$$
$$= w_{n+1}^2 \left(\prod_{j=0}^{J-1} \widehat{f}_j - \prod_{j=0}^{J-1} f_j\right)^2 . \qquad (10)$$

    - ★ Replacing the f_j in (10) by $\widehat{f}_j$ would result in $\widehat{EE} = 0$, which is not meaningful!
    - ★ $\nexists$ an estimator of SEE which is best for all parameters values f_j

  - *unconditional* estimator of SEE (analogous estimator as in Section 2)
    Volatility measure: *unconditional* variance

$$\mathrm{Var}\left(\widehat{X_{n+1}}\right) = w_{n+1}^2 \left\{\prod_{j=0}^{J-1}(f_j^2 + \frac{\sigma_j^2}{w_{\bullet,j}}) - \prod_{j=0}^{J-1} f_j^2\right\}$$

=>

$$\widehat{SEE}^{\mathrm{uncond}} = w_{n+1}^2 \left\{\prod_{j=0}^{J-1}\left(\widehat{f}_j^{\,2} + \widehat{\sigma_j^2}/w_{\bullet,j}\right) - \prod_{j=0}^{J-1} \widehat{f}_j^{\,2}\right\}. \qquad (11)$$

  - But this time we can do it better by use of the telescope formula

Model with independent r.v.

- Telescope representation of SEE (telescope representation)

$$SEE = w_{n+1}^2 \left(\sum_{j=0}^{J-1} A_j \right)^2 = w_{n+1}^2 \left\{ \sum_{j=0}^{J-1} A_j^2 + 2 \sum_{j=0}^{J-1} \sum_{k=j+1}^{J-1} A_j A_k \right\} \tag{12}$$

where

$$A_j = \left(\prod_{k=0}^{j-1} \widehat{f}_k \right) \left(\widehat{f}_j - f_j \right) \left(\prod_{m=j+1}^{J-1} f_m \right) \tag{13}$$

To find an estimator we have to find an estimator for each summand in (12)

- *unconditional* estimator of *SEE*
volatility measure: *unconditional* variance and *unconditional* covariance

$$\begin{aligned} \mathrm{Var}(A_j) &= \prod_{k=0}^{j-1} \left(\widehat{f}_k^{\,2} + \frac{\sigma_k^2}{w_{\bullet,k}} \right) \frac{\sigma_j^2}{w_{\bullet,j}} \prod_{m=j+1}^{J-1} f_m^2, \\ \mathrm{Cov}(A_j A_k) &= 0. \end{aligned}$$

$$\widehat{SEE}^{\text{uncond}} = w_{n+1}^2 \left\{ \sum_{j=0}^{J-1} \widehat{A_j^2}^{\text{uncond}} \right\} \quad \text{where } \widehat{A_j^2}^{\text{uncond}} = \prod_{k=0}^{j-1} \left(\widehat{f}_k^{\,2} + \frac{\widehat{\sigma_k^2}}{w_{\bullet,k}} \right) \frac{\widehat{\sigma_j^2}}{w_{\bullet,j}} \prod_{m=j+1}^{J-1} \widehat{f_m}^2 \tag{14}$$

Model with independent r.v.

- *conditional* estimator of *SEE*

  take into account for each summand in (12) as many data from $\mathcal{D}$ as possible
  => volatility measure: *conditional* variance and *conditional* covariance

$$\begin{aligned} \operatorname{Var}\left(A_j \,|\, \mathcal{B}_{j-1} \right) &= \prod_{k=0}^{j-1} \widehat{f}_k^{\,2} \, \frac{\sigma_j^2}{w_{\bullet,j}} \prod_{m=j+1}^{J-1} f_m^2, \\ \operatorname{Cov}\left(A_j A_k \,|\, \mathcal{B}_{k-1} \right) &= 0. \end{aligned}$$

$$\widehat{SEE}^{\text{cond}} = w_{n+1}^2 \left\{ \sum_{j=0}^{J-1} \widehat{A_j^2}^{\text{cond}} \right\} \quad \text{where } \widehat{A_j^2}^{\text{cond}} = \prod_{k=0}^{j-1} \widehat{f}_k^{\,2} \, \frac{\widehat{\sigma_j^2}}{w_{\bullet,j}} \prod_{m=j+1}^{J-1} \widehat{f_m}^2 \tag{15}$$

- **Theorem**

$$E\left[\widehat{SEE}^{\text{uncond}} - SEE \right] > E\left[\widehat{SEE}^{\text{cond}} - SEE \right] > 0$$

  - Conclusion

    Both estimators are "upward biased", but $\widehat{SEE}^{\text{cond}}$ to a smaller extent than $\widehat{SEE}^{\text{uncond}}$.

    => $\widehat{SEE}^{\text{cond}}$ has better properties and is preferable to $\widehat{SEE}^{\text{uncond}}$.

Model with uncorrelated r.v

- example from insurance
  - data usually listed in a calculation statistics for periods $i = 0, \ldots, n$

$$
\begin{aligned}
JR_i &= \text{number of year risks in year } i \\
N_i &= \textstyle\sum_{\nu=1}^{JR_i} N_i^{(\nu)} = \text{number of claims in year } i \\
F_i &= N_i / JR_i \text{ claim frequency in year } i \\
S_i &= \textstyle\sum_{\nu=1}^{N_i} Y_i^{(\nu)} \text{ total claim amount in year } i \\
Y_i &= S_i / N_i \text{ claims average} \\
X_i &= S_i / JR_i = F_i Y_i \quad \text{average claim per risk (pure risk premium)}
\end{aligned}
$$

  - assumptions (collective risk model) and properties

$$
\begin{aligned}
E\left[F_i\right] &= \lambda, \qquad \mathrm{Var}\left(F_i \mid JR_i\right) = \sigma_N^2 / JR_i, & (16)\\
E\left[Y_i \mid N_i \geqslant 1\right] &= \mu_Y, \qquad \mathrm{Var}\left(Y_i \mid N_i\right) = \sigma_Y^2 / N_i, & (17)\\
E\left[X_i\right] &= \lambda \mu_Y =: \mu_X. & (18)
\end{aligned}
$$

  - remarks
    - ★ collective risk model: claim number and claim severities are assumed to be independent; however F_i and Y_i are not independent but only uncorrelated; reason: the weight in the variance-condition (17) is stochastic.

Model with uncorrelated r.v.

- example from insurance (collective risk model)
  - questions of interest
    - ★ how can we model such situations ?
    - ★ how accurate are the estimators of λ, μ_Y, μ_X given the observed data ?
    - ★ how accurate is the forecast of the claims load of next year ?
- Uncorrelated model

  Model Assumptions (uncorrelated model)

  M1 $\{C_{i,0}, \dots, C_{i,J}\}$ and $\{C_{k,0}, \dots, C_{k,J}\}$ are independent for $i \neq k$.
  M2 $C_{i,0}$ are known weights and $C_{i,1}, \dots, C_{i,J}$ are r.v. with

$$E[C_{i,j+1}|C_{i,0}, \dots, C_{i,j}] = f_j C_{i,j}, \tag{19}$$

$$\mathrm{Var}(C_{i,j+1}|C_{i,0}, \dots, C_{i,j}) = \sigma_j^2 C_{i,j}. \tag{20}$$

- Link Ratios

$$F_{i,j} := \frac{C_{i,j+1}}{C_{i,j}} \text{ for } j = 0, \dots, J-1. \tag{21}$$

- observed data

$$\mathcal{D} \ := \{C_{i,j}, i = 0, 1, \dots, n, \ j = 1, \dots, J\} \quad \text{observed data}$$

  we further define

$$\mathcal{B}_j := \{C_{i,k} \in \mathcal{D}, \ k \leq j\} \quad \text{observed data up to column } j$$

Model with uncorrelated r.v.

- estimators of unknown parameters

$$\widehat{f}_j = \frac{\sum_{i=1}^n C_{i,j+1}}{\sum_{i=1}^n C_{i,j}} = \frac{\sum_{i=1}^n C_{i,j} F_{i,j}}{C_{\bullet,j}}, \quad \text{where } C_{\bullet,j} = \sum_{i=1}^n C_{i,j}$$

$$\widehat{\sigma_j^2} = \frac{1}{n-1} \sum_{i=0}^{n} C_{i,j} (F_{i,j} - \widehat{f}_j)^2.$$

- products of r.v.
  from definition of $F_{i,j}$ and $\widehat{f}_k$ =>

$$C_{i,j} = C_{i,0} \prod_{k=0}^{j-1} F_{i,k} \quad \text{for } j = 0, \ldots, J-1, \tag{22}$$

$$C_{\bullet,j} = C_{\bullet,0} \prod_{k=0}^{j-1} \widehat{f}_k. \tag{23}$$

  However

  $\{F_{i,0}, \ldots, F_{i,J}\}$ and $\left\{\widehat{f}_0, \ldots, \widehat{f}_J\right\}$ are not independent, but only uncorrelated

- Properties

$$E[F_{i,j} | C_{i,0,\ldots,} C_{i,j}] = E[F_{i,j} | C_{i,j}] \quad = f_j,$$

$$\mathrm{Var}(F_{i,j} | C_{i,0,\ldots,} C_{i,j}) = \mathrm{Var}(F_{i,j} | C_{i,j}) = \frac{\sigma_j^2}{C_{i,j}}.$$

$$E\left[\widehat{f}_j \middle| \mathcal{B}_j \right] = f_j,$$

$$\mathrm{Var}\left(\widehat{f}_j \middle| \mathcal{B}_j \right) = \frac{\sigma_j^2}{C_{\bullet,j}} \text{ for } C_{\bullet,j} > 0.$$

Model with uncorrelated r.v

- Accuracy of statistical observations
  - squared estimation errors

$$SEE_j := \left(\widehat{f}_j - f_j\right)^2,$$

$$SEE := \left(C_{\bullet,J} - E\left[C_{\bullet,J}\right]\right)^2 = C_{\bullet,0}^2 \left(\prod_{J=0}^{J-1} \widehat{f}_j - \prod_{J=0}^{J-1} f_j\right)^2$$

  - Telescope Representation

$$SEE = \left(\sum_{j=0}^{J-1} A_j^2 + 2\sum_{j=1}^{J-1}\sum_{k=j+1}^{J-1} A_j A_k\right)$$

where

$$A_j = C_{\bullet,j}\left(\widehat{f}_j - f_j\right)\prod_{m=j+1}^{J-1} f_m$$

  - volatility measure
    conditional variance and *conditional* covariance

$$\mathrm{Var}\left(\widehat{f}_j \middle| \mathcal{B}_j\right) = \frac{\sigma_j^2}{C_{\bullet,j}},$$

$$\mathrm{Var}\left(A_j \middle| \mathcal{B}_j\right) = C_{\bullet,j}\sigma_j^2 \prod_{m=j+1}^{J-1} f_m^2,$$

$$\mathrm{Cov}\left[A_j A_{k,} \middle| \mathcal{B}_k\right] = E\left[A_j A_{k,} \middle| \mathcal{B}_k\right] = 0.$$

Model with uncorrelated r.v

- Accuracy of statistical observations
  - =>

$$\widehat{SEE}_j^{\text{cond}} = \frac{\widehat{\sigma_j^2}}{C_{\bullet,j}}; \quad \widehat{SEE}^{\text{cond}} = C_{\bullet,J}^2 \sum_{j=0}^{J-1} \frac{\widehat{\sigma_j^2}}{\widehat{f}_j^2} \frac{1}{C_{\bullet,j}},$$

$$\frac{\sqrt{\widehat{SEE}_j^{\text{cond}}}}{\widehat{f}_j} = \frac{\widehat{\sigma}_j}{\widehat{f}_j}; \quad \frac{\sqrt{\widehat{SEE}^{\text{cond}}}}{C_{\bullet,J} =} = \sqrt{\sum_{j=0}^{J-1} \frac{\widehat{\sigma_j^2}}{\widehat{f}_j^2} \frac{1}{C_{\bullet,j}}} = \sqrt{\sum_{j=0}^{J-1} \frac{\widehat{SEE}_j^{\text{cond}}}{\widehat{f}_j^2}}.$$

- Alternative derivation of the estimator
  - Given the data, the weights in the variance-condition of $F_{i,j}$ are known.
  - replace the $F_{i,j}$ by independent r.v. $F_{i,j}^*$ with

$$E\left[F_{i,j}^* \middle| \mathcal{B}_j\right] = f_{j,} \quad \mathrm{Var}\left(F_{i,j}^* \middle| \mathcal{B}_j\right) = \frac{\sigma_j^2}{C_{i,j}}.$$

=>

$$A_j^* = C_{\bullet,j}\left(\widehat{f}_j^* - f_j\right) \prod_{m=j+1}^{J-1} f_m$$

=>

$$\mathrm{Var}\left(A_{i,j}^* \middle| \mathcal{B}_j\right) = C_{i,j}\sigma_j^2 \prod_{m=j+1}^{J-1} f_m^2 = \mathrm{Var}\left(A_{i,j} \middle| \mathcal{B}_j\right)$$

=>

$$\widehat{SEE^*}^{\text{cond}} = \widehat{SEE}^{\text{cond}}$$

Model with uncorrelated r.v

- conditional msep of next years' forecast
  - Predictor of $X_{n+1,j}$

$$\widehat{C_{n+1,j}} = C_{n+1,0} \prod_{k=0}^{j-1} \widehat{f}_k. \tag{24}$$

  - *Conditional* mean square error of prediction

$$\begin{aligned} \mathrm{msep}_{C_{n+1,J}|\mathcal{D}}\left(\widehat{C_{n+1,J}}\right) := & \quad E\left[\left(C_{n+1,J} - \widehat{C_{n+1,J}}\right)^2 \middle| \mathcal{D}\right] \\ = & \quad PV + SEE \end{aligned}$$

where

$$\begin{aligned} PV &= C_{n+1,0}^2 E\left[\left(\prod_{j=0}^{J-1} F_{n+1,j} - \prod_{j=0}^{J-1} f_j\right)^2\right], \\ SEE &= C_{n+1,0}^2 \left\{\left(\prod_{j=0}^{J-1} \widehat{f}_j - \prod_{j=0}^{J-1} f_j\right)^2\right\} \end{aligned}$$

  - estimators (obtained by analogous calculations)

$$\begin{aligned} \widehat{PV} &= \widehat{C_{n+1,J}}^2 \left\{\sum_{j=0}^{J-1} \frac{\widehat{\sigma_j^2}}{\widehat{f}_j^2} \frac{1}{\widehat{C_{n+1,j}}}\right\}, \\ \widehat{SEE}^{\text{cond}} &= \widehat{C_{n+1,J}}^2 \left\{\sum_{j=1}^{J-1} \frac{\widehat{\sigma_j^2}}{\widehat{f}_j^2} \frac{1}{C_{\bullet,j}}\right\}. \end{aligned}$$

Model with uncorrelated r.v

- Numerical example

accuracy of statistical observations in a calculation statistics

year i	JR_i $(C_{i,0})$	N_i $((C_{i,1})$	S_i $(C_{i,1})$	F_i $(F_{i,0})$ in %o	$\frac{\widehat{\sigma}_0}{\sqrt{JR_i}} \frac{1}{F_i}$	Y_i $(F_{i,1})$	$\frac{\widehat{\sigma}_1}{\sqrt{N_i}} \frac{1}{Y_i}$	X_i		$\frac{\sqrt{\widehat{SEE}_i^{\text{cond}}}}{X_i}$
1	10'000	2'060	5'792'720	206	± 2.3%	2'812	± 10.9%	579	±	11.2%
2	10'000	1'980	5'104'440	198	± 2.3%	2'578	± 11.2%	510	±	11.4%
3	10'000	1'955	4'306'865	196	± 2.3%	2'203	± 11.2%	431	±	11.5%
4	10'000	2'010	4'641'090	201	± 2.3%	2'309	± 11.1%	464	±	11.3%
Total	40'000	8'005	19'845'115	200	± 1.1%	2'479	± 5.6%	496	±	5.7%

estimation of parameters

$$\widehat{f_0} = 200‰, \quad \widehat{f_1} = 2'479$$
$$\widehat{\sigma_0} = 452‰, \quad \widehat{\sigma_1} = 12'313$$
$$\frac{\widehat{\sigma_0}}{\widehat{f_0}} = 2.3, \quad \frac{\widehat{\sigma_1}}{\widehat{f_1}} = 5.0$$

next years forecast and accuracy

$$\widehat{S_5} = 10'000\, \widehat{f_0}\, \widehat{f_1} = 4'961'279; \qquad \widehat{X_5} = 496$$

$$\frac{\sqrt{\widehat{PV}}}{\widehat{S_5}} = 11.3\% \qquad \frac{\sqrt{\widehat{SEE}^{\text{cond}}}}{\widehat{S_5}} = 5.7\% \qquad \frac{\sqrt{\widehat{msep}}}{\widehat{S_5}} = 12.7\%$$

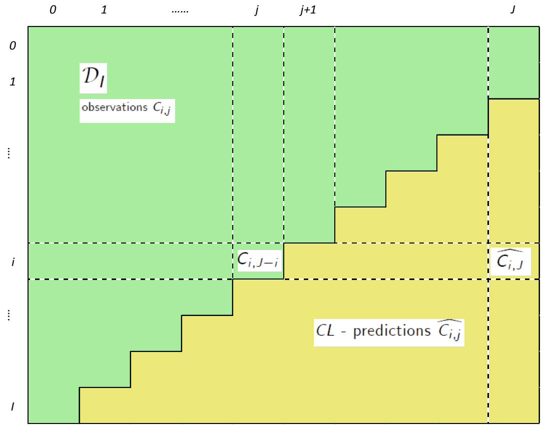
Chain Ladder Model of Mack

observed data $\mathcal{D} := \{C_{i,j} : i = 0, \dots, I,\ i + j \le I\}$ we further define

$\mathcal{B}_j := \{C_{i,k} \in \mathcal{D},\ k \le j\}$ observed data up to development year j

$\widehat{C_{i,J}}$ CL-prediction of ultimate claim of accident year i

$R_i = \widehat{C_{i,J}} - C_{i,J-i}$ claims reserves for accident year i; $\sum_{i=I-J+1}^{I} R_i$ total claims reserves



Chain Ladder Model of Mack

- Model Assumptions
  M1 $\{C_{i,0}, \ldots, C_{i,J}\}$ and $\{C_{k,0}, \ldots, C_{k,J}\}$ are independent for $i \neq k$.
  M2 $C_{i,j}$ are r.v. with

$$E[C_{i,j+1}|C_{i,0}, \ldots, C_{i,j}] = f_j C_{i,j}, \tag{25}$$

$$\mathrm{Var}(C_{i,j+1}|C_{i,0}, \ldots, C_{i,j}) = \sigma_j^2 C_{i,j}. \tag{26}$$

  The same model assumptions as in Section 4
  But $\mathcal{D}$ is now a triangle and not a rectangle and we have to forecast $C_{i,J}$ for all accident years i

- Link Ratios

$$F_{i,j} := \frac{C_{i,j+1}}{C_{i,j}} \text{ for } j = 0, \ldots, J-1.$$

- estimators of unknown parameters

$$\widehat{f_j} = \frac{\sum\limits_{i=0}^{I-j-1} C_{i,j+1}}{\sum\limits_{i=0}^{I-j-1} C_{i,j}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{l=0}^{I-j-1} C_{l,j}} F_{i,j}, \quad j = 0, \ldots, J-1,$$

$$\widehat{\sigma_j^2} = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(F_{i,j} - \widehat{f_j}\right)^2.$$

Chain Ladder Model of Mack

- CL-forecasts

$$\widehat{C_{i,j}} = \begin{cases} C_{i,I-i} \prod_{k=I-i}^{j-1} \widehat{f}_k, & \text{for } C_{i,j} \in \mathcal{D}^c \\ C_{i,j} & \text{for } C_{i,j} \in \mathcal{D} \end{cases} .$$

from definition of $\widehat{f}_j$ =>

$$\widehat{C_{\bullet,j}} = C_{\bullet,0} \prod_{k=0}^{j-1} \widehat{f}_k$$

- *conditional* mean squared error of prediction (msep)
  - msep of reserves is identical to msep of the CL-prediction of the ultimate claim

$$\mathrm{msep}_i := E\left[\left(C_{i,J} - \widehat{C}_{i,J}\right)^2 \middle| \mathcal{D}\right] = PV_i + SEE_i \qquad \text{single accident years,}$$

$$\mathrm{msep}_{tot} := E\left[\left(C_{tot,J} - \widehat{C}_{tot,J}\right)^2 \middle| \mathcal{D}\right] = PV_{tot} + SEE_{to.} \quad \text{total over all accident years,}$$

$$PV_i = C_{i,I-i}^2 \sum_{j=I-i}^{J-1} E\left[C_{n+1,j}\right] \sigma_j^2 \prod_{k=j+1}^{J-1} f_k^2; \quad PV_{tot} = \sum_{i=0}^{I} PV_i. \tag{27}$$

$$SEE_i = \left(\sum_{j=I-i}^{J-1} A_j\right)^2 \text{ where } A_j = C_{i,I-i} \prod_{k=I-i}^{j-1} \widehat{f}_k \left(\widehat{f}_j - f_j\right) \prod_{m=j+1}^{J-1} f_m, \tag{28}$$

$$SEE_{tot} = \left(\sum_{j=0}^{J-1} A_j\right)^2 \text{ where } B_j = \left(\sum_{i=I-j}^{I} \widehat{C_{i,j}}\right) \left(\widehat{f}_j - f_j\right) \prod_{m=j+1}^{J-1} f_m \tag{29}$$

Chain Ladder Model of Mack

- **Mack Estimators** (1993)

$$\widehat{PV_i}^{\text{Mack}} = \widehat{C_{i,J}}^2 \left\{ \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j} \frac{1}{\widehat{C_{i,j}}} \right\}, \qquad \widehat{PV_{tot}}^{\text{Mack}} = \sum_{i=0}^{I} \widehat{PV_i}^{\text{Mack}}, \tag{30}$$

$$\widehat{SEE_i}^{\text{Mack}} = \widehat{C_{i,J}}^2 \left\{ \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j} \frac{1}{\sum_{i=0}^{I-j-1} C_{i,j}} \right\}, \qquad \widehat{SEE_{tot}}^{\text{Mack}} = \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{\left(\sum_{i=I-j}^{I} \widehat{C_{i,J}}\right)^2}{\sum_{i=0}^{I-j-1} C_{i,j}}. \tag{31}$$

the original Mack estimators for SEE_{tot} was different, but the above estimator is equivalent

- **BBMW Estimators** (2006)
  the estimators of the process variance were identical to the ones of Mack;

$$\widehat{SEE_i}^{\text{BBMW}} = C_{i,I-i}^2 \left(\prod_{j=I-i}^{J-1} \left(\widehat{f}_j^2 + \frac{\widehat{\sigma_j^2}}{\widehat{C_{i,j}}} \right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \right) \tag{32}$$

$$\overset{\text{telescope}}{=} C_{i,I-i}^2 \sum_{j=I-i}^{J-1} \left(\prod_{k=I-i}^{j-1} \left(\widehat{f}_k^2 + \frac{\widehat{\sigma_k^2}}{\sum_{l=0}^{I-k-1} C_{l,j}} \right) \frac{\widehat{\sigma_j^2}}{\sum_{l=0}^{I-j-1} C_{l,j}} \prod_{m=j+1}^{J-1} \widehat{f_m}^2 \right) \tag{33}$$

$$\widehat{SEE_{tot}}^{\text{BBMW}} = \sum_{j=0}^{J-1} \left(\sum_{i=I-j+1}^{I} \widehat{C_{i,j}} \right)^2 \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \left(\prod_{k=j+1}^{J-1} \left(\widehat{f}_k^2 + \frac{\widehat{\sigma_k^2}}{\sum_{l=0}^{I-k-1} C_{l,j}} \right) \right) \tag{34}$$

(32) was the original BBMW estimator of SEE_i. The original BBMW estimator of SEE_{tot} is different, but (34) is equivalent

Chain Ladder Model of Mack

- Discussion in 2006
  - great discussion in 2006 about the two estimators (4 papers)
    however no answer to the question, which of the two estimators should be preferred
  - Focus of the discussion in 2006
    - ★ independence assumption
      BBMW replaced $\widehat{f}_j$ by resampled $\widehat{f}_j^*$ which were independent
    - ★ Mack, Quarg and Braun: $\widehat{f}_j^2$ and $\widehat{f_{j+1}}^2$ are conditionally on $\mathcal{B}_j$ negatively correlated
      => independence assumption not in line with CL-assumptions and therefore
      BBMW-estimator > Mack-estimator and BBMW overestimate the *SEE*
  - However discussion focused on the wrong item; this is not the relevant point
    - ★ Results

$$\widehat{SEE_i}^{\text{Mack}} = \widehat{SEE_i}^{\text{cond}}, \qquad \widehat{SEE_{tot}}^{\text{Mack}} = \widehat{SEE_{tot}}^{\text{cond}}$$
$$\widehat{SEE_i}^{\text{BBMW}} = \widehat{SEE_i^*}^{\text{uncond}}, \qquad \widehat{SEE_{tot}}^{\text{BBMW}} = \widehat{SEE_{tot}^*}^{\text{uncond}}$$

    - ★

$$\widehat{SEE_i^*}^{\text{cond}} = \widehat{SEE_i}^{\text{cond}} = \widehat{SEE_i}^{\text{Mack}}$$
$$\widehat{SEE_{tot}^*}^{\text{cond}} = \widehat{SEE_{tot}}^{\text{cond}} = \widehat{SEE_{tot}}^{\text{Mack}}$$

Chain Ladder Model of Mack

- Mack and BBMW estimator
  - => The essential point is not the independence assumption made by BBMW, but that they have taken the *unconditional* instead of the *conditional* estimator
    otherwise they would have received the same estimator as Mack
  - implication and conclusion

$$E\left[\widehat{SEE_i}^{\text{BBMW}}\right] > E\left[\widehat{SEE_i}^{\text{Mack}}\right] > E\left[SEE_i\right]$$

$$E\left[\widehat{SEE_{tot}}^{\text{BBMW}}\right] > E\left[\widehat{SEE_{tot}}^{\text{Mack}}\right] > E\left[SEE_{tot}\right]$$

=> both estimators are positively biased, but the Mack estimator to a smaller extent
the Mack estimator has better properties and the BBMW estimator is not a sensible alternative to the Mack-estimator.

some literature

- Buchwalder M., Bühlmann, H., Merz M., Wüthrich M. (2006) The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited). ASTIN Bulletin 36/2:521-542.
- Buchwalder M., Bühlmann, H., Merz M., Wüthrich M. (2006) The Mean Square Error of Prediction in the Chain Ladder Reserving Method-Final Remark. ASTIN Bulletin 36/2:553.
- Gisler, A. (2006) The estimation error in the chain-ladder reserving method: a Bayesian approach. ASTIN Bulletin 36/2:554-565.
- Gisler, A. (2019) The Reserve Uncertainties in the Chain-Ladder model of Mack revisited. ASTIN Bulletin 49/3:787-821.
- Gisler, A. (2021). Estimation error and bootstrapping in the chain ladder model of Mack. *European Actuarial Journal, 169-283*.
- Mack, T. (1993) Distribution-free calculation of the standard error of chain ladder reserve estimates. ASTIN Bulletin 23/2:213–225.
- Mack, T., Quarg B., Braun C. (2006) The Mean Square Error of Prediction in the Chain Ladder Reserving Method- A Comment. ASTIN Bulletin 36/2:523-552.