Asymptotic Consistency and Inconsistency of the Chain Ladder

(O banálnom aktuárskom probléme, ktorý ignorujeme, a jeho nebanálnom, ale triviálnom riešení)

Michal Pešta

(joint work with Šárka Hudecová)

Charles University in Prague

Faculty of Mathematics and Physics

Actuarial Seminar, Prague

5 October 2012

Overview

- Claims reserving in non-life insurance
- Chain ladder model
- Open problem: consistency of the development factors
- Simulations and real data example

- Based on:

  Pešta and Hudecová (2012). Asymptotic consistency and inconsistency of the chain ladder. *Insurance: Mathematics and Economics*, 51(2): 472–479.

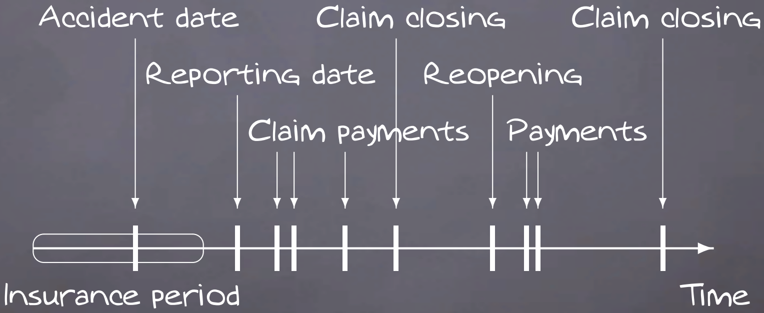
- Support:

  Czech Science Foundation project "DYME Dynamic Models in Economics" No. P402/12/G097

Non-life insurance

- Operates on the lines of business (LoB):
  - motor/car insurance (motor third party liability, motor hull)
  - property insurance (private and commercial insurance against fire, water, flooding, business interruption, . . .)
  - liability insurance
  - accident insurance
  - health insurance
  - marine insurance (including transportation)
  - other (aviation, travel insurance, legal protection, credit insurance, epidemic insurance, . . .)
- Life insurance products are rather different, e.g., terms of contracts, type of claims, risk drivers

Timeline of a claim



Settlement of a claim

- Reporting delay (between occurrence and reporting) – can take several years (liability insurance: asbestos or environmental pollution claims)
- After being reported to the insurer – several years may elapse before the claim is finally settled (fast in property insurance, liability or bodily injury claims: long time before the total circumstances are clear and known)
- Reopening – (unexpected) new developments, or if a relapse occurs

Reserving

- Claims reserves represent the money which should be held by the insurer so as to be able to meet all future claims arising from policies currently in force and policies written in the past
- Most non-life insurance contracts are written for a period of one year
- Only one payment of premium at the start of the contract in exchange for coverage over the year
- Reserves are calculated by forecasting future losses from past losses

Terminology

- $X_{i,j} \ldots$ claim amounts in development year j with accident year i
- $X_{i,j}$ stands for the incremental claims in accident year i made in accounting year $i+j$
- $n \ldots$ current year – corresponds to the most recent accident year and development period
- Our data history consists of right-angled isosceles triangles $X_{i,j}$, where $i = 1, \ldots, n$ and $j = 1, \ldots, n+1-i$

Notation

- $C_{i,j}$ ... cumulative payments in origin year i after j development periods

$$C_{i,j} = \sum_{k=1}^{j} X_{i,k}$$

- $C_{i,j}$ ... a random variable of which we have an observation if $i + j \leq n + 1$
- Aim is to estimate the ultimate claims amount $C_{i,n}$ and the outstanding claims reserve

$$R_i = C_{i,n} - C_{i,n+1-i}, \quad i = 2, \ldots, n$$

Run-off triangle

Accident year i	Development year j				
	1	2	$\cdots$	$n-1$	n
1	$C_{1,1}$	$C_{1,2}$	$\cdots$	$C_{1,n-1}$	$C_{1,n}$
2	$C_{2,1}$	$C_{2,2}$	$\cdots$	$C_{2,n-1}$	
			$\ddots$		
$\vdots$	$\vdots$	$\vdots$	$C_{i,n+1-i}$		
$n-1$	$C_{n-1,1}$	$C_{n-1,2}$			
n	$C_{n,1}$				

Reasonable Estimate for Reserves

- nonsense estimate, e.g., $\widehat{R}_i = 10^6$ or $\widehat{R}_i = -1$
  - why bad? the most precise estimate (in terms of variability) $\mathsf{Var}\widehat{R}_i = 0$
- unbiased estimate, i.e., $\mathsf{E}\widehat{R}_i = \mathsf{E}R_i$, or conditionally unbiased
  - firstly, introduce a model with assumptions, then construct an estimate
  - have you ever heard of a reserving model with unbiased estimates of reserves?
- consistent estimate (but where's n ?)

$$\widehat{R}_i \xrightarrow[n\to\infty]{(stochastically)} R_i$$

Chain ladder

Mack (1993)

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

[2] Var$[C_{i,j+1}|C_{i,1},\ldots,C_{i,j}] = \sigma_j^2 C_{i,j}$

[3] Accident years $[C_{i,1},\ldots,C_{i,n}]$ are independent vectors

Development factors f_j

$$\widehat{f}_j^{(n)} = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}}, \quad 1 \leq j \leq n-1$$

$\widehat{f}_n^{(n)} \equiv 1$ (assuming no tail)

Properties

- Ultimate claims amounts $C_{i,n}$ are estimated by

$$\widehat{C}_{i,n} = C_{i,n+1-i} \times \widehat{f}^{(n)}_{n+1-i} \times \cdots \times \widehat{f}^{(n)}_{n-1}$$

- Under the assumptions [1], [3], and
  [4] $\sum_{i=1}^{n-j} C_{i,j} > 0$
  $\widehat{f}^{(n)}_j$ are <u>unbiased and mutually uncorrelated</u>
- Assumption [2] is essential for the standard error of $\widehat{C}_{i,n}$

Unbiasedness

- unbiasedness of development factors $\nRightarrow$ unbiasedness of reserves' estimate !
- Given data $D_j = \{C_{i,k} : k \leq j,\, i \in \mathbb{N}\}$ or $\{C_{i,1}, \ldots, C_{i,j}\}$, where $j = n + 1 - i$

$$
\begin{aligned}
&\mathsf{E}\left[R_i | D_{n+1-i}\right] = \mathsf{E}\left[C_{i,n} - C_{i,n+1-i} | D_{n+1-i}\right] \\
&= \mathsf{E}\left[C_{i,n} | D_{n+1-i}\right] - C_{i,n+1-i} \\
&= \mathsf{E}\left[\mathsf{E}\left\{C_{i,n} | C_{i,1}, \ldots, C_{i,n-1}\right\} | D_{n+1-i}\right] - C_{i,n+1-i} \\
&= \mathsf{E}\left[f_{n-1} C_{i,n-1} | D_{n+1-i}\right] - C_{i,n+1-i} = \ldots \\
&= f_{n-1} \ldots f_{n+2-i} \mathsf{E}\left[f_{n+1-i} C_{i,n+1-i} | D_{n+1-i}\right] - C_{i,n+1-i} \\
&= C_{i,n+1-i}\left(f_{n+1-i} \times \ldots \times f_{n-1} - 1\right), \quad a.s.
\end{aligned}
$$

(Un)Biasedness

- $\widehat{R}_i = \widehat{C}_{i,n} - C_{i,n+1-i} = C_{i,n+1-i}\left(\widehat{f}^{(n)}_{n+1-i} \times \cdots \times \widehat{f}^{(n)}_{n-1} - 1\right)$

$$\mathsf{E}\left[\widehat{R}_i | D_{n+1-i}\right] = C_{i,n+1-i}\left(\mathsf{E}\left[\widehat{f}^{(n)}_{n+1-i} \times \cdots \times \widehat{f}^{(n)}_{n-1} | D_{n+1-i}\right] - 1\right)$$
$$= \ldots = \mathsf{E}\left[R_i | D_{n+1-i}\right], \quad a.s.$$

- how much is unbiasedness important? and can we always achieve it? do we need to?

(In)consistency

- disadvantage:
  - asymptotic property (data history sufficiently long)
  - only a qualitative property
- advantages:
  - easier to verify (?)
  - characterize the accuracy (and, hence, meaningfulness) of an estimate
  - can be quantified (e.g., rate of consistency)
  - is retained by algebraic operations

Open problem

- From some point of view, consistency is more important than unbiasedness
- E.g., $\widehat{f}^{(n)}_{n+1-i} \times \cdots \times \widehat{f}^{(n)}_{n-1}$ or Bornhuetter-Ferguson method uses

$$\widehat{\beta}_j^{(n)} = \prod_{k=j}^{n-1} \frac{1}{\widehat{f}_k^{(n)}}$$

- The unbiasedness of $\widehat{f}_j^{(n)}$ does not "transfer" to $\widehat{\beta}_j^{(n)}$ in any sense
- Ex: $Y_1, \ldots, Y_n$ iid with finite EY
- $T_1(Y_1, \ldots, Y_n) = Y_1$ vs $T_2(Y_1, \ldots, Y_n) = \frac{1}{n}\sum_{i=1}^n Y_i + \frac{1}{n}$

Stochastic Convergence

- deterministic: one convergence (of real numbers)

$$X_n \xrightarrow[n\to\infty]{(stochastically)} X$$

- almost sure

$$\mathsf{P}\left[\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right] = 1$$

- in probability

$$\forall \varepsilon > 0 : \lim_{n\to\infty} \mathsf{P}\left[|X_n - X| \geq \varepsilon\right] = 0$$

- L_p, $p \geq 1$

$$\lim_{n\to\infty} \mathsf{E}|X_n - X|^p = 0$$

- $\xrightarrow[n\to\infty]{a.s.} \quad \Rightarrow \quad \xrightarrow[n\to\infty]{\mathsf{P}} \quad \Leftarrow \quad \xrightarrow[n\to\infty]{\mathsf{L}_p}; \qquad \xrightarrow[n\to\infty]{\mathsf{P}} \quad \Rightarrow \quad \xrightarrow[n\to\infty]{\mathsf{D}}$

Conditional convergence

$\xi_n \xrightarrow[n\to\infty]{[\mathsf{P}_\zeta]\text{-}a.s.} \chi$, $[\mathsf{P}]$-$a.s.$ means

$$\mathsf{P}\left[\mathsf{P}_\zeta\left\{\lim_{n\to\infty}\xi_n = \chi\right\} = 1\right] = 1$$

$\xi_n \xrightarrow[n\to\infty]{\mathsf{P}_{\zeta_n}} \chi$, $[\mathsf{P}]$-$a.s.$ means

$$\forall\varepsilon > 0 : \mathsf{P}\left[\lim_{n\to\infty}\mathsf{P}_{\zeta_n}\left\{|\xi_n - \chi| \geq \varepsilon\right\} = 0\right] = 1$$

$\xi_n \xrightarrow[n\to\infty]{\mathsf{L}_p(\mathsf{P}_{\zeta_n})} \chi$, $[\mathsf{P}]$-$a.s. \, (p \geq 1)$ means

$$\mathsf{P}\left[\lim_{n\to\infty}\mathsf{E}_{\zeta_n}|\xi_n - \chi|^p = 0\right] = 1$$

Conditioning

- Conditional convergence in probability and in L_p along some sequence of random variables $\{\zeta_n\}_{n=1}^{\infty}$ can be defined, because the concept of these two types of convergence comes from a topology
- Despite of that, the almost sure convergence does not correspond to a convergence with respect to any topology and, hence, it is not metrizable
- Thereafter, the conditional convergence almost surely cannot be defined along a sequence of random variables, but only given one random variable ζ

Consistency

Denote $D_j^{(n)} = \{C_{i,k} : k \le j, i \le n-j+1\}$ and $D_j = \{C_{i,k} : k \le j, i \in \mathbb{N}\}$. Then (i)–(iv) are <u>equivalent</u>:

(i)

$$\widehat{f}_j^{(n)} \xrightarrow[n\to\infty]{[\mathsf{P}_{D_j}]\text{-}a.s.} f_j, \quad [\mathsf{P}]\text{-}a.s.;$$

(ii)

$$\widehat{f}_j^{(n)} \xrightarrow[n\to\infty]{\mathsf{P}_{D_j^{(n)}}} f_j, \quad [\mathsf{P}]\text{-}a.s.;$$

(iii)

$$\widehat{f}_j^{(n)} \xrightarrow[n\to\infty]{\mathsf{L}_2\left(\mathsf{P}_{D_j^{(n)}}\right)} f_j, \quad [\mathsf{P}]\text{-}a.s.;$$

(iv)

$$\sum_{i=1}^{n-j} C_{i,j} \xrightarrow[n\to\infty]{} \infty, \quad [\mathsf{P}]\text{-}a.s.$$

Remark 1

Due to the independence of the different accident years (assumption [3]), the statements (ii) and (iii) can be equivalently replaced by

$$\widehat{f}_j^{(n)} \xrightarrow[n\to\infty]{\mathsf{P}_{D_j}} f_j,\ [\mathsf{P}]\text{-}a.s.$$

and

$$\widehat{f}_j^{(n)} \xrightarrow[n\to\infty]{\mathsf{L}_2(\mathsf{P}_{D_j})} f_j,\ [\mathsf{P}]\text{-}a.s.,$$

respectively.

Remark II

- Unconditional consistency in case of the L_2 convergence
- $\widehat{f}_j^{(n)} \to f_j$ in L_2 (unconditionally) as $n \to \infty$ iff

$$\mathrm{E}\left[\frac{1}{\sum_{i=1}^{n-j} C_{i,j}}\right] \to 0, \quad n \to \infty$$

- This condition is obviously more complicated than the condition (iv), and it is practically unverifiable
- Thus, the conditional convergence is not only more natural one in this case, but even more convenient one

Rate of convergence

- Consistency of an estimator is a very important but only qualitative property
- Measure consistency – quantitative way
- Denote the conditional mean square error of the estimate of development factor f_j as

$$MSE\left(\widehat{f}_j^{(n)}\right) := \mathsf{E}\left\{\left[\widehat{f}_j^{(n)} - \mathsf{E}\left(\widehat{f}_j^{(n)}\right)\right]^2 \middle| D_j^{(n)}\right\}.$$

Then, with probability one holds

$$MSE\left(\widehat{f}_j^{(n)}\right) = \mathcal{O}\left(\left[\sum_{i=1}^{n-j} C_{i,j}\right]^{-1}\right), \quad n \to \infty.$$

On the rate of convergence

- Complete characterization of the conditional convergence of development factors' estimate
- The slower (faster) divergence of

$$\sum_{i=1}^{\infty} C_{i,j}$$

implies the slower (faster) realization of consistency of the development factors' estimates

On the necessary and sufficient condition

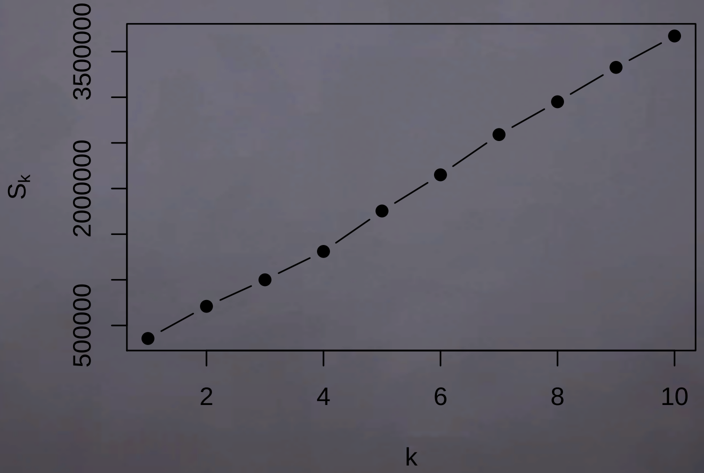
Let $j \in \mathbb{N}$ be fixed. Then the following conditions are equivalent.

1. The condition (iv) holds.
2. $$\sum_{i=1}^{\infty} \mathsf{E} C_{i,1} = \infty. \tag{1}$$
3. The condition (iv) holds for $j_0 \in \mathbb{N}$, $j \neq j_0$.

Practical aspects

- Either $\widehat{f}_j^{(n)}$ is consistent for f_j <u>for all</u> $j \in \mathbb{N}$, or <u>none</u> of them is consistent
- The consistency of $\widehat{f}_j^{(n)}$ is <u>equivalent</u> to the condition $\sum_{i=1}^{n} C_{i,1} \to \infty$, $[\mathrm{P}]$-$a.s.$ as $n \to \infty$
- Denote $S_k = \sum_{i=1}^{k} C_{i,1}$, $k = 1, \ldots, n$ the cumulative sums of the cumulative claims $C_{i,1}$ in the <u>first development year</u>
- For instance, the ratios $C_{k+1,1}/C_{k,1}$ or the sequence $\sqrt[k]{C_{k,1}}$ can be studied
- Artificial data set Taylor and Ashe (1983)

Real data example



Inconsistency

- What kinds of business behavior corresponds to the violation of consistency?
- For instance, condition (iv) can be violated if one observes a decreasing trend (decreasing fast enough) in payments across the accident years
- So to speak, the corresponding line of business is worsening maybe due to new insurance companies entering the market or changing (decreasing) prices of such insurance product
- Furthermore, splitting one existing line of business into several others can also cause inconsistency in the estimation of development factors

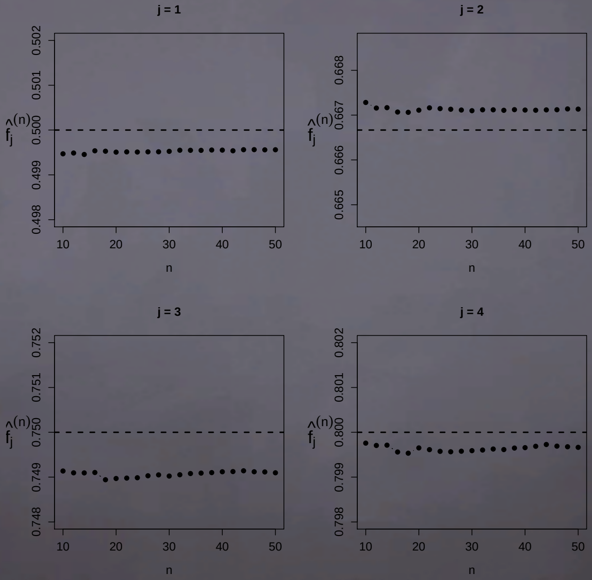
Simulations

- $C_{i,1}$ was generated such that $C_{i,1} \in \mathsf{L}_2$ and $C_{i,j} \geq 0$
- f_j was set to $f_j = j/(j+1)$
- $C_{i,2}, \ldots C_{i,N}$ were generated successively such that $C_{i,j}$ satisfies [1] and [2]
- Hence, for each j, $C_{i,j}$ is drawn from a distribution with mean $f_{j-1}C_{i,j-1}$ and variance $\sigma_j^2 C_{i,j-1}$ for some $\sigma_j^2 \in (0, \infty)$
- Since the accident years are assumed to be independent (assumption [3]), the rows of the data sets were generated separately, using the same approach
- $C_{i,1}$ were drawn from the exponential distribution, and the $C_{i,j}$ was generated from the Poisson distribution with the parameter $f_{j-1}C_{i,j-1}$ for $j = 2, \ldots, N$

Decreasing Business

- consider a fast decreasing business, i.e., the situation where condition (I) does not hold
- $C_{i,1}$ was generated from the exponential distribution with the mean $i^{-2} \times 10^6$
- $\widehat{f}_j^{(n)}$ do not converge to the true value f_j
- Their values are close to f_j in this setting, but the estimates are indeed not consistent
- The same simulations were run also for $C_{i,1}$ with the uniform distribution: the differences between values of estimates $\widehat{f}_j^{(n)}$ and the true values f_j are more noticeable

Example 1

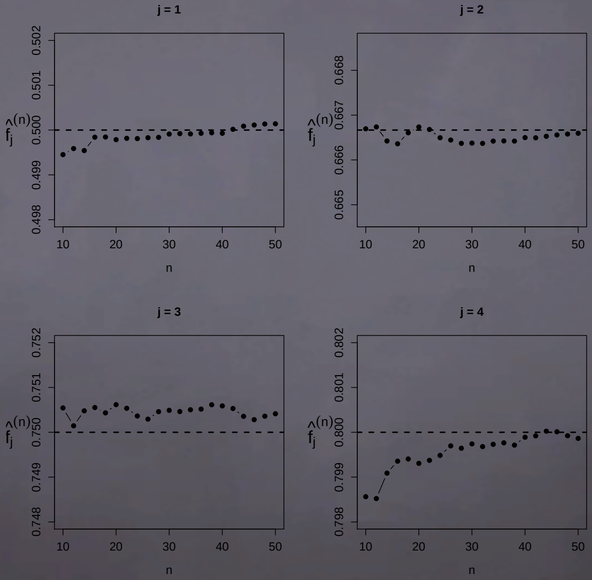


j = 1
j = 2
j = 3
j = 4
$\hat{f}_j^{(n)}$
n
0.498
0.499
0.500
0.501
0.502
0.665
0.666
0.667
0.668
0.748
0.749
0.750
0.751
0.752
0.798
0.799
0.800
0.801
0.802
10
20
30
40
50

Slowly decreasing business

- situation where $\mathrm{E}C_{i,1}$ decreases with increasing i, but in a slow manner such that condition (I) holds
- $\mathrm{E}C_{i,1} = i^{-1/2} \times 10^6$ in the exponential distribution
- clear convergence pattern can be observed, confirming that the estimates $\widehat{f}_j^{(n)}$ are consistent in this case

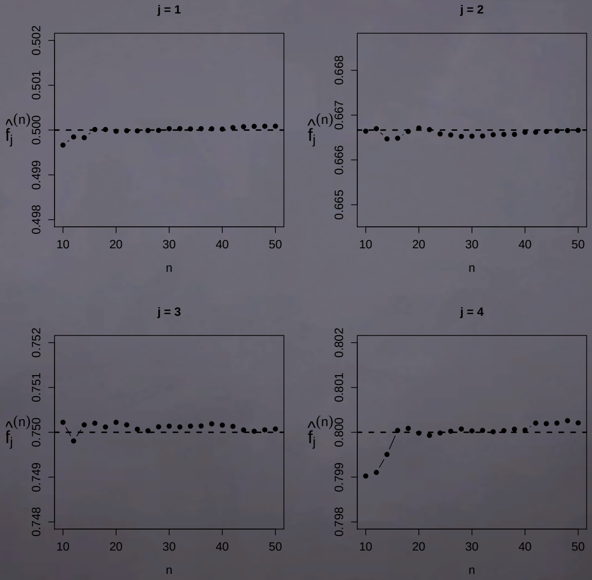
Example II


j = 1
j = 2
j = 3
j = 4
$\hat{f}_j^{(n)}$
n
0.498
0.499
0.500
0.501
0.502
0.665
0.666
0.667
0.668
0.748
0.749
0.750
0.751
0.752
0.798
0.799
0.800
0.801
0.802
10
20
30
40
50

Growing Business

- $\mathrm{E}C_{i,1}$ increases with increasing i
- The parameters of the exponential distribution were set such that $\mathrm{E}C_{i,1} = \sqrt{i} \times 10^6$
- The figure obviously confirms that the estimates are consistent
- Moreover, $\widehat{f}_j^{(n)}$ converges to the true values f_j much faster than in previous case

Example III



One Year Prospective

- in the classical claims reserving, one usually studies the total uncertainty in the claims development until the total ultimate claim is finally settled
- Solvency II purposes
- run-off = opening balance − expenses − closing balance
- we predict the total ultimate claim at time n (with the available information up to time n), and one period later at time $n+1$ we predict the same total ultimate claim with the updated information available at time $n+1$
- difference between these two successive predictions is the <u>claims development result</u> (CDR) for accounting year $(n, n+1]$

CDR

- a direct impact on the P&L statement and on the financial strength of the insurance company
- CDR for accident year i and accounting year $(n, n+1]$

$$\begin{aligned} CDR_i(n+1) &= \mathsf{E}[C_{i,n}|D^{(n)}] - \mathsf{E}[C_{i,n}|D^{n+1}] \\ &= \mathsf{E}[R_i^{(n)}|D^{(n)}] - (X_{i,n+2-i} + \mathsf{E}[R_i^{n+1}|D^{n+1}]) \end{aligned}$$

where

$$\begin{aligned} D^{(n)} &= \{C_{i,j} : i + j \leq n + 1\} \\ D^{n+1} &= \{C_{i,j} : i + j \leq n + 2 \,\&\, i \leq n + 1\} \\ &= D^{(n)} \cup \{C_{i,n+2-i} : i \leq n + 1\} \\ R_i^{(n)} &= C_{i,n} - C_{i,n+1-i} \\ R_i^{n+1} &= C_{i,n} - C_{i,n+2-i} \end{aligned}$$

Merz-Wüthrich

- implied Chain Ladder assumptions (time series):

(1i)
$$C_{i,j} = f_{j-1} C_{i,j-1} + \sigma_{j-1} \sqrt{C_{i,j-1}} \varepsilon_{i,j},$$

where $\varepsilon_{i,j}$ are iid with $\mathrm{E}\varepsilon_{i,j} = 0$ and $\mathrm{Var}\varepsilon_{i,j} = 1$

(2i) Accident years $[C_{i,1}, \ldots, C_{i,n}]$ are independent vectors

Reasonable results?

- using the martingale property

$$\mathsf{E}[CDR_i(n+1)|D^{(n)}] = 0$$

- prediction uncertainty in the budget value 0 for the observable claims development result at the end of the accounting period

$$MSE_{\widehat{CDR}_i(n+1)|D^{(n)}}(0) = \mathsf{E}[(\widehat{CDR}_i(n+1) - 0)^2|D^{(n)}]$$

- in the solvency margin, we need to hold risk capital for possible negative deviations of $CDR_i(n+1)$ from 0

? are the "results" by MW reasonable (e.g., consistent estimate of $MSE_{\widehat{CDR}_i(n+1)|D^{(n)}}(0)$)?

Conclusions

- conditional consistency and inconsistency of the development factors' estimate in the distribution-free chain ladder is investigated
- necessary and sufficient condition is derived
- weak, strong consistency, and consistency in the mean square are equivalent
- convergence rate is provided
- practical recommendations, how to check this necessary and sufficient condition, are discussed
- real data example and numerical simulations to illustrate the performance of the estimates
- possible violation of the condition with the consequences is demonstrated

References

- Mack, T. (1993)
  Distribution-free calculation of the standard error of chain ladder reserve estimates.
  Astin Bulletin, 23(2), 213–225.
- Merz, M. and Wuthrich, M. V. (2008)
  Modeling the claims development results for Solvency purposes.
  CAS E-Forum, Fall 2008, 542–568.
- Pesta, M. and Hudecova, S. (2012)
  Asymptotic consistency and inconsistency of the chain ladder.
  Insurance: Mathematics and Economics, 51(2), 472–479.

Thank you !

pesta@karlin.mff.cuni.cz