28 November – 1 December 2004

Xth Accident Compensation Seminar

The Mack-Method and Analysis of Variability

Erasmus Gerigk

Institute of Actuaries of Australia

Contents/outline

- **Introduction**
- **Review of two reserving recipes**:
  Incremental Loss-Ratio Method
  Chain-Ladder Method
- **Mack’s model assumptions and estimating variability**
- **Estimating Risk Margins**
- **Model testing**
- **Example**
- **Further research:** extensions and refinements

Preliminary remarks

- Depends on the type of available data
- There are plenty of methods around
- Many deal with projecting triangulated data
- In this presentation we focus entirely "Chain-Ladder types" of methods
- Manifold terminology in the literature

Incremental Loss-Ratio Method

Look at *loss ratio increments:*

	Exposure	1	2	3	4	5	6
1999	2,721	11%	21%	13%	7%	14%	16%
2000	2,314	22%	57%	19%	13%	18%	?
2001	5,161	6%	7%	4%	5%	?	?
2002	5,789	2%	13%	15%	?	?	?
2003	13,784	3%	7%	?	?	?	?
2004	7,445	9%	?	?	?	?	?

… and complete the rectangle

Method is old and appears under many names

Incremental Loss-Ratio Method

Usually done through weighted averages:

	Exposure	1	2	3	4	5	6
1999	2,721	11%	21%	13%	7%	14%	16%
2000	2,314	22%	57%	19%	13%	18%	**16.2%**
2001	5,161	6%	7%	4%	5%	**16.0%**	**16.2%**
2002	5,789	2%	13%	15%	**7.4%**	**16.0%**	**16.2%**
2003	13,784	3%	7%	**11.6%**	**7.4%**	**16.0%**	**16.2%**
2004	7,445	9%	**13.2%**	**11.6%**	**7.4%**	**16.0%**	**16.2%**

$\boxed{m_2 = 13.2\%}$ $\boxed{m_3 = 11.6\%}$

$$\hat{m}_k = \sum_{i=1}^{n+1-k} \left(\frac{v_i}{\sum_{i=1}^{n+1-k} v_i} \hat{m}_{ik} \right) = \frac{\sum_{i=1}^{n+1-k} S_{ik}}{\sum_{i=1}^{n+1-k} v_i}$$

Summary: Chain-Ladder Method

Look at *individual development factors:*

	1	2	3	4	5	6
1999		= 2.99	1.40	1.15	1.27	1.25
2000		3.60	1.24	1.13	1.17	**1.248**
2001		2.21	1.30	= 1.30	**1.202**	**1.248**
2002		6.47	1.97	**1.172**	**1.202**	**1.248**
2003		2.95	**1.432**	**1.172**	**1.202**	**1.248**
2004		**3.296**	**1.432**	**1.172**	**1.202**	**1.248**

f_3=1.432

1999	290 ÷	868	1,219	1,406	1,784	2,226
2000	511	1,840	2,273	2,568	2,994	**3,736**
2001	316	697	909 ÷	1,184	**1,424**	**1,776**
2002	137	887	1,743	**2,043**	**2,456**	**3,065**
2003	462	1,364	**1,953**	**2,288**	**2,751**	**3,433**
2004	644 →	**2,123**	**3,039**	**3,561**	**4,282**	**5,343**

calculate weighted averages …

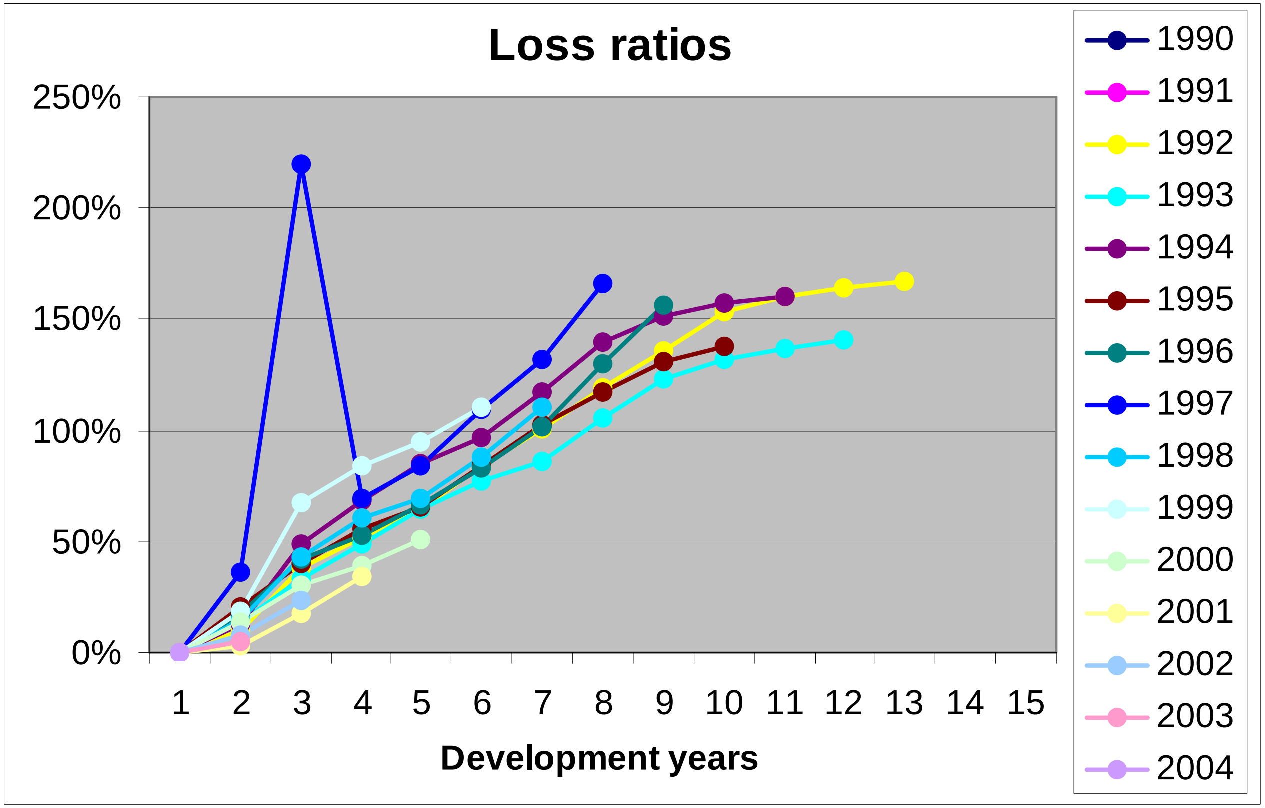
$$\hat{f}_k = \sum_{i=1}^{n-k} \left(\frac{C_{i,k}}{\sum_{i=1}^{n-k} C_{i,k}} \cdot F_{i,k} \right) = \frac{\sum_{i=1}^{n-k} C_{i,k+1}}{\sum_{i=1}^{n-k} C_{i,k}}$$

… and complete the rectangle

$$\hat{C}_{i,k+1} = C_{i,k} \ \hat{f}_k$$

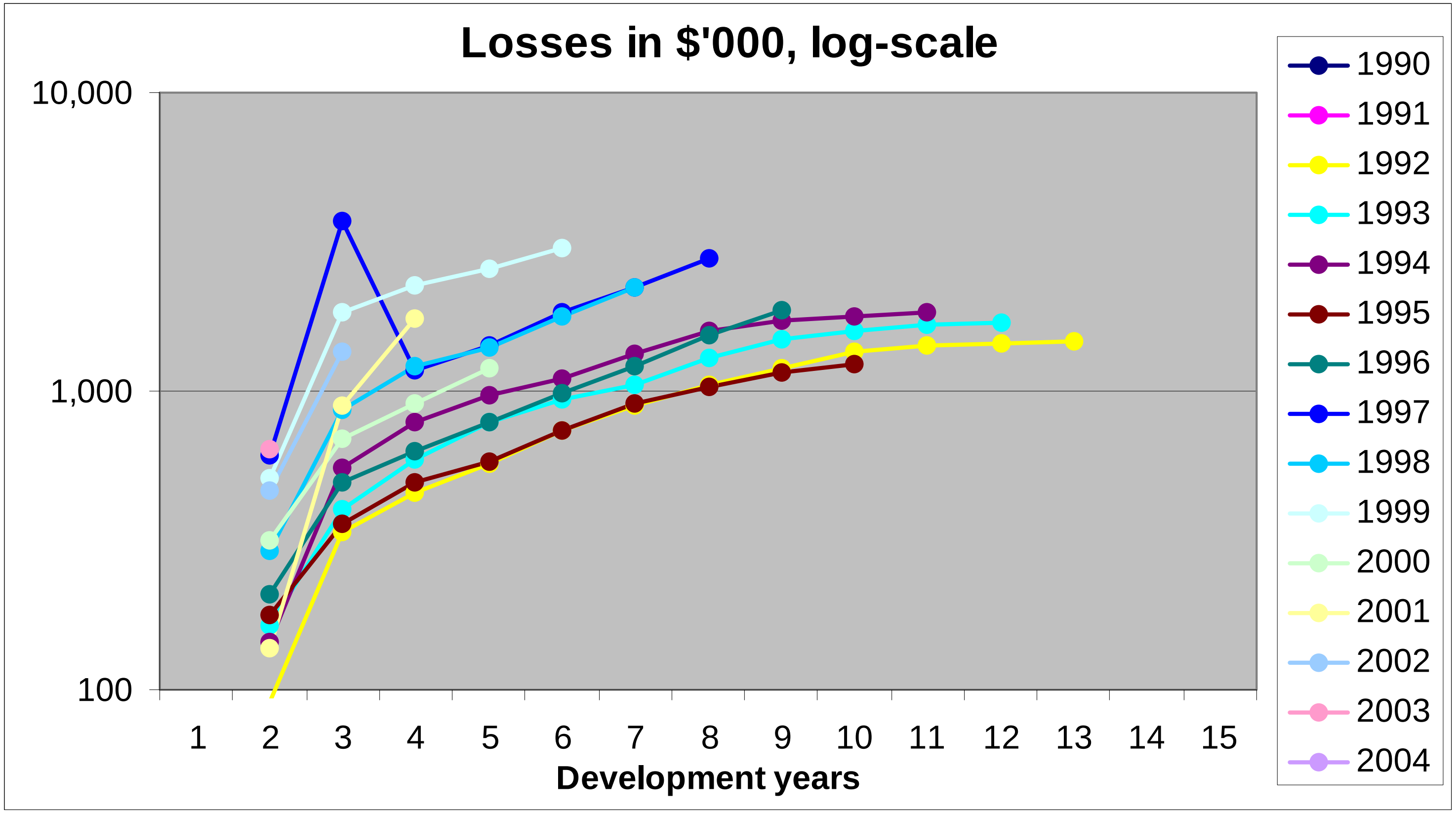
Graphical interpretation

additive projection: $$\frac{C_{i,k+1}}{v_i} = \frac{C_{i,k}}{v_i} + m_k$$



Graphical interpretation

multiplicative projection: $\log C_{i,k+1} = \log C_{i,k} + \log f_k$



Practical issues

1) Can any of the methods be expected to lead to reasonable results?
2) How to decide between the two models?
3) How to select the slope in the graphs, i.e. the development factor? Is there a “best way”?
4) Both recipes provide a point estimate. How to deal with the requirement of risk margins?

The Mack-Method

- Is a textbook example of a proper statistical model with precise model assumptions and estimators.
- Through model-assumptions we will “re-invent” both methods
- In the Mack-Method, both procedures are extended to include variability estimates
- Recall the graphical interpretations. Many formal expressions correspond to “visible” phenomena

Additive model assumptions

The Mack-method makes three model assumptions about the payments S_{ik} in a particular underwriting / development year:

(AM 1) $$E\left(\frac{S_{ik}}{v_i}\right) = m_k \qquad 1 \le i \le n; \quad 1 \le k \le n$$

(AM 2) *The payments S_{ik} are independent for all i,k*

(AM 3) $$Var\left(\frac{S_{ik}}{v_i}\right) = \frac{s_k^2}{v_i} \qquad 1 \le i \le n; \quad 1 \le k \le n$$

(AM 3) may remind you of the “Individual Model”

Estimating the model parameters

The key results are

$$\hat{m}_k = \frac{\sum S_{ik}}{\sum v_i}$$ is *Best Linear Unbiased Estimator* of m_k

$$\hat{s}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} v_i \left(\frac{S_{ik}}{v_i} - \hat{m}_k \right)^2$$ is *Unbiased Estimator* of $\hat{s}_k^2$

All further results follow from straightforward algebraic manipulations

Further consequences

- There are straightforward estimators for the standard errors of the increments $\hat{m}_k$
- There are closed-formula expressions for the standard error of the ultimate loss
- For working with spreadsheets, there are neat recursions-formulas
- It is possible to estimate the random error and the estimation error separately

Chain-Ladder model assumptions

Model assumptions look a bit more difficult, because of conditional expectation

(CL 1) $$E\left(\frac{C_{i,k+1}}{C_{ik}}\Big| C_{i1},\ldots,C_{ik}\right)= f_k \qquad 1\le i\le n;\quad 1\le k\le n-1$$

(CL 2) *The underwriting years* { C_{i1} , …, C_{in}} *are globally independent, i.e. the sets* { C_{i1} , …, C_{in}} *are independent for* $i\neq j$

(CL 3) $$Var\left(\frac{C_{i,k+1}}{C_{i,k}}\Big| C_{i1},\ldots,C_{ik}\right)=\frac{\sigma_k^2}{C_{i,k}} \qquad 1\le i\le n;\quad 1\le k\le n-1$$

Exposure measure is here the loss in the preceding period

Chain-Ladder estimators

The results are very similar to the additive case. However, the proofs are more sophisticated.

Key results:

$$\hat{f}_k = \frac{\sum C_{i,k+1}}{\sum C_{i,k}}$$ is *Best Linear Unbiased Estimator* of f_k

$$\hat{\sigma}_k^2 = \frac{1}{n-k-1} \sum_{i=1}^{n-k} C_{i,k} \left(\frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k \right)^2$$ is *Unbiased* for σ_k^2

And we have nice recursions for the standard error of the ultimate loss estimates (hard to prove!)

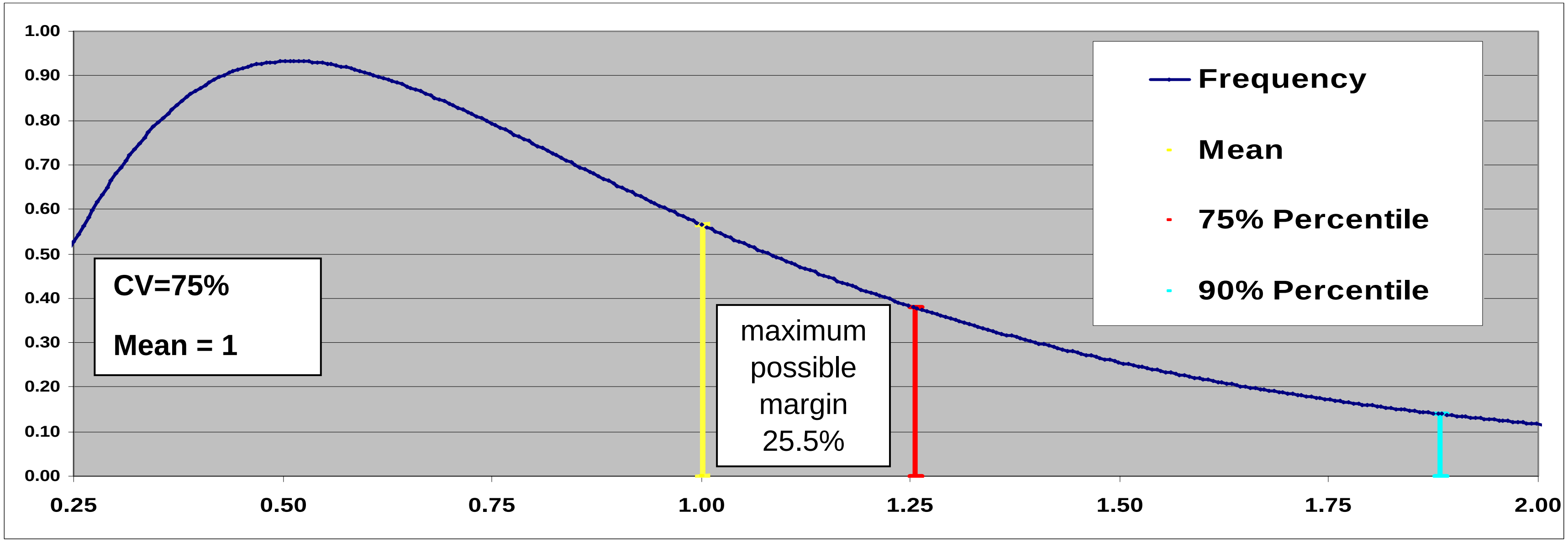
Risk Margins

The total standard error enables us to calculate a stand-alone risk margin, e.g. at a 75% sufficiency level

- ☺ The standard error and thus the risk margin depends only on the original triangle
- ☹ The underlying “sigmas” are often quite volatile and should not be used mechanically

From s.e. to Risk Margins

- Be aware of the properties of your distribution
- Popular choice: log-normal fit



NB: upper bound by one-sided Chebyshev

$$P\left[X - E(X) \geq \alpha \cdot s.e.(X)\right] \leq 1/(1+\alpha^2)$$

Alas, an inequality of such generality cannot lead to sharp results.

Intermediate summary

- Mack method provides two self-contained ways of obtaining central estimates as well as variation estimates from triangulated data
- It is probably the simplest method available for deriving variability estimates from triangles

Caveats

- The reliability of the estimates depends on how well the data is described by the model assumptions
- Which model, if any, shall be preferred?
- Not yet clear how outliers can be dealt with within the model, in particular the calculation of the standard errors

⇨ more work to be done by the actuary: model check and dealing with outliers

Testing the model assumptions

- A lot of information can be extracted from the appropriate graphs, i.e. loss ratios or log-scaled dollars
- Quickly checked: parallel behaviour of graphs and obvious outliers
- More accurate: residual analysis from regression theory
- Regression approach works for both the additive and the multiplicative model

Regression analysis (additive)

- For a fixed development period *k*, look again at

$$\text{(AM 1)} \qquad E(S_{ik}) = v_i m_k$$

$$\text{(AM 2)} \qquad \textit{Independence}$$

$$\text{(AM 3)} \qquad Var(S_{ik}) = v_i s_k^2$$

- A statistician is someone who calls that a heteroskedastic regression without intercept

Model check

For each development year *k*

- Plot $\left(v_i; S_{ik}\right)$. Does it look linear?

- Plot the standardised residuals $\dfrac{S_{i,k} - v_i \hat{m}_k}{\sqrt{v_i}}$

  against the exposure. There should be no pattern!

More on testing: calendar-year effects

Calendar year-effects have many causes, e.g.

- Inflation
- Change in claims handling
- Change in legislation
- ...

They are acting on the diagonal !

Test for calendar-year effects

Plot the standardised residuals

$$\frac{S_{i,k} - v_i \hat{m}_k}{\hat{s}_k \sqrt{v_i}} \qquad 2 \le i + k \le n$$

for each calendar year. If all residuals have the same sign, this could indicate a calendar-year effect.

Ideally, there should be no pattern at all!

What to do with outliers and cy-effects?

Sum only over selected parts of the triangle. All estimators can be adjusted in a straightforward way.

In theory:

$$\hat{m}_k = \frac{\sum (w_{i,k} \cdot S_{ik})}{\sum (w_{i,k} \cdot v_i)} \qquad \hat{s}_k^2 = \frac{1}{|I|-1} \sum w_{i,k} \cdot v_i \left(\frac{S_{ik}}{v_i} - \hat{m}_k \right)^2$$

- Similar for the other formulas
- Same for Chain-Ladder

A brief example

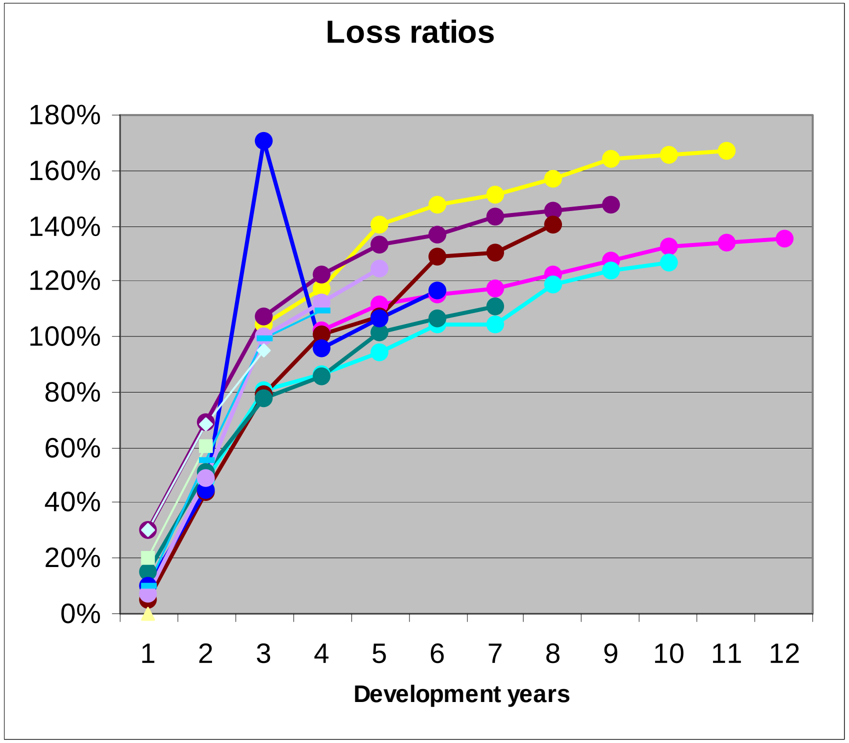
Cumulative payments per development year

	Exposure	1	2	3	4	5	6	7	8	9	10	11	12
1993	1,214				1,240	1,353	1,396	1,421	1,488	1,545	1,608	1,623	1644
1994	1,139			1,185	1,337	1,597	1,677	1,719	1,788	1,869	1,887	1,903	
1995	887		430	717	768	837	923	924	1,055	1,097	1,125		
1996	1,189	357	818	1,276	1,452	1,584	1,627	1,701	1,727	1,753			
1997	1,682	84	740	1,338	1,691	1,808	2,169	2,198	2,366				
1998	2,020	303	1,036	1,566	1,724	2,054	2,150	2,238					
1999	2,721	272	1,211	4,642	2,600	2,891	3,166						
2000	2,314	162	1,135	2,310	2,592	2,890							
2001	5,161	516	2,849	5,132	5,652								
2002	5,789	1,737	3,958	5,485									
2003	7,836	1,567	4,732										
2004	6,936	-											

Cumulative loss ratios per development year

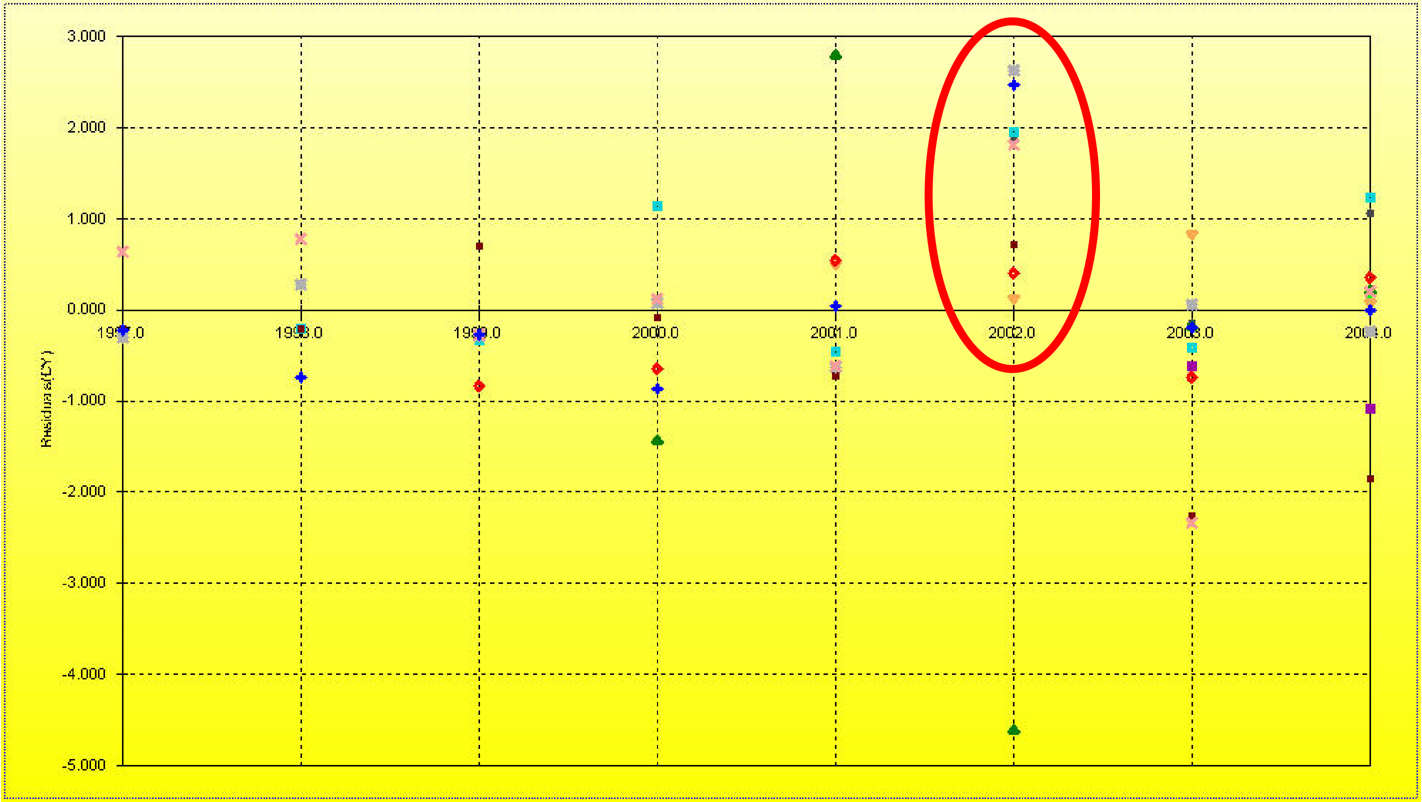
	Exposure	1	2	3	4	5	6	7	8	9	10	11	12
1993	1,214				102%	111%	115%	117%	123%	127%	132%	134%	135%
1994	1,139			104%	117%	140%	147%	151%	157%	164%	166%	167%	
1995	887		48%	81%	87%	94%	104%	104%	119%	124%	127%		
1996	1,189	30%	69%	107%	122%	133%	137%	143%	145%	147%			
1997	1,682	5%	44%	80%	101%	107%	129%	131%	141%				
1998	2,020	15%	51%	78%	85%	102%	106%	111%					
1999	2,721	10%	44%	171%	96%	106%	116%						
2000	2,314	7%	49%	100%	112%	125%							
2001	5,161	10%	55%	99%	110%								
2002	5,789	30%	68%	95%									
2003	7,836	20%	60%										
2004	6,936	0%											

Graphical run-off



Residual analysis

Plot of the additive residuals per calendar year



This could be a CY-effect. In fact, more information is required.

Residual analysis

Closer look at the incremental loss-ratios:

Loss ratio increments per development year

	1	2	3	4	5	6	7	8	9	10	11	12
1993				102%	9%	4%	2%	6%	5%	5%	1%	2%
1994			104%	13%	23%	7%	4%	6%	7%	2%	1%	
1995		48%	32%	6%	8%	10%	0%	15%	5%	3%		
1996	30%	39%	39%	15%	11%	4%	6%	2%	2%			
1997	5%	39%	36%	21%	7%	21%	2%	10%				
1998	15%	36%	26%	8%	16%	5%	4%					
1999	10%	34%	126%	-75%	11%	10%						
2000	7%	42%	51%	12%	13%							
2001	10%	45%	44%	10%								
2002	30%	38%	26%									
2003	20%	40%										
2004	0%											

Do not sum over these values

Conclusion

- Proper application of the Mack-Method is not mechanical
- Judgemental adjustments can be incorporated into calculation of standard errors
- It might provide useful information for establishing loss reserves
- It is all very simple

Something to read

For the *Chain-Ladder Method*:

- Mack, T [1993] *Distribution–free calculation of the standard error of chain–ladder reserve estimates* In: ASTIN Bull. 23, 213–225
- Mack, T [1994] *Measuring the variability of chain–ladder reserve estimates* In: CAS Forum Spring 1994, pp. 101–182

(and many more publications …)

For the *Incremental Loss-Ratio Method*:

- Mack, T [1997, 2002] *Schadenversicherungsmathematik* In: Karlsruhe Verlag Versicherungswirtschaft

Further refinements

1. Weighted regression of development factors and sigmas for smoothing and extending the run-off
2. Munich Chain-Ladder
3. Correlation between triangles