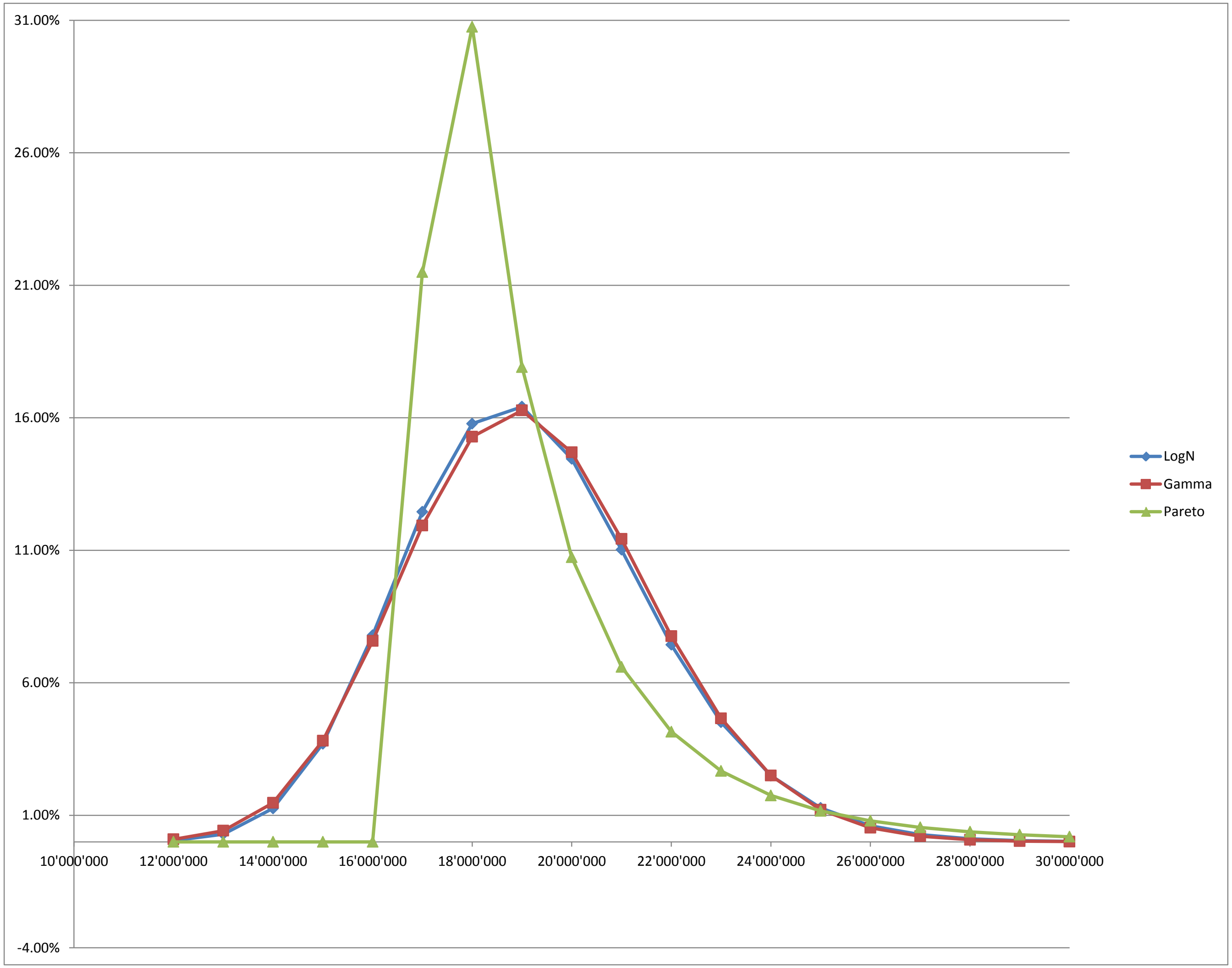
Application of skewness to non-life reserving

ASTIN colloquium 2012

SCOR

Reserve risk distribution – Why an interest on skewness ?

- 3 distributions
  - Same mean
  - Same CoV
- Skewness gives an indicator of the asymmetry of the risk distribution.



Reserve risk distribution – What do we know in a chain-ladder framework ?

From Mack (1993):

UWY	Dvpt 1	2	3	4	5	Ultimate
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$	$\hat{C}_{1,I}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$		$\hat{C}_{2,I}$
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$			$\hat{C}_{3,I}$
4	$C_{4,1}$	$C_{4,2}$				$\hat{C}_{4,I}$
5	$C_{5,1}$					$\hat{C}_{5,I}$
						$\sum_{j=1}^{I} \hat{C}_{j,I}$

Reserves

$$\hat{R}_j = \hat{C}_{j,I} - C_{j,I-j+1}$$

Best Estimate

$$\sum_{j=1}^{I} \hat{C}_{j,I} = \sum_{j=1}^{I} C_{j,I-j+1} \prod_{k=I-j+1}^{I} \hat{f}_k$$

$$\hat{f}_k = \frac{\sum_{j=1}^{I-k} C_{j,k+1}}{\sum_{j=1}^{I-k} C_{j,k}}, 1 \leq k \leq I\text{-}1$$

$$\hat{\sigma}_k^2 = \frac{1}{I-k-1} \sum_{i=1}^{I-k} C_{i,k} \left(\frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k \right)^2 \quad for\ 1 \leq k \leq I\text{-}2$$

$$mse\left(\sum_{i=2}^{I} \hat{R}_i\right) = \sum_{i=2}^{I} \left(mse(\hat{R}_i) + \hat{C}_{i,I} \left(\sum_{j=i+1}^{I} \hat{C}_{j,I} \right) \sum_{k=I+1-i}^{I-1} \frac{2\hat{\sigma}_k^2 / \hat{f}_k^2}{\sum_{n=1}^{I-k} C_{nk}} \right) \quad (1)$$

with

$$mse(\hat{R}_i) = \hat{C}_{i,I}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{i,k}} + \frac{1}{\sum_{n=1}^{I-k} C_{nk}} \right)$$

Reserve risk distribution – What do we know in other frameworks ?

- Bornhuetter – Ferguson
  - Best estimate known
  - Estimate of the standard deviation known (see Mack 2008)
- Mix Bornhuetter – Ferguson / Chain-ladder
  - Best estimate known
  - Hybrid chain-ladder method provides an estimate of the standard deviation (see Arbenz 2010)
- GLM based on incremental triangles
  - Best estimate known
  - Different estimates of the standard deviation given (See Merz-Wüthrich 2008)
- Cape-Cod
  - Best estimate known
  - No estimate of the standard deviation

Reserve risk distribution – How do we get the distribution today ?

- Assume a lognormal distribution with mean given by Best Estimate and standard deviation given by Mack standard equation (see equation (1) on slide 3).
- Bootstrapping techniques based on Pearson residuals (see England and Verall 2006)
- Generalized Linear Models based on incremental triangles (see Merz and Wüthrich 2008)
  - Usual assumption: The distribution of the random element of the incremental claim $X_{i,j}$ belongs to the Exponential Dispersion Family (e.g. Poisson, Gamma …)
- Model of Salzman, Wüthrich, Merz on higher moments of the Claim Development Result in General Insurance (ASTIN Bulletin 2012)
  - Two models assumed for the distribution of the individual claims development factor $F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$: Gamma and Lognormal models

- All of the above models use some distributional assumptions.
- However, in Madrid, last year, a presentation indicated that the attempts to backtest the hypothesis of a lognormal distribution for the reserving risk was not successful (See Meyers 2011).

Reserve risk distribution – Why is it important to know more about it ?

- The reserving risk distribution is used in many instances:
  - Calculation of the Solvency Capital in Internal models
  - Use in Solvency 2 standard model

- Example
  - Best Estimate = 18 680 856, Standard deviation = 2 447 095

Lognormal		Gamma	
μ	16.735	k	58.276
σ	0.130	θ	320'556.691
VaR 99%	25'089'172		24'840'064

  - 1% difference on the VaR 99% depending only on the choice of the distribution

Reserving risk distribution – Skewness: A proposal for a new approach

UWY	Dvpt					Ultimate
	1	2	3	4	5	
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$	$\hat{C}_{1,I}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$		$\hat{C}_{2,I}$
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$			$\hat{C}_{3,I}$
4	$C_{4,1}$	$C_{4,2}$				$\hat{C}_{4,I}$
5	$C_{5,1}$					$\hat{C}_{5,I}$
						$\sum_{j=1}^{I}\hat{C}_{j,I}$

Assumption on skewness: $\dfrac{SK\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right)}{\left[Var\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right)\right]^{3/2}}$ depends on k but not on i

where: $SK\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right) = E\left[\left(C_{i,k+1} - E\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right)\right)^3 \mid C_{i,1},\ldots,C_{i,k}\right]$

For one development year within a triangle, the skewness is the same for any UWY.

Context : Reserving portfolio which risks are similar for every UWY.

Skewness : Use of the new approach

- With the above assumption, under a Mack model for the volatility, we have:

$$\exists \gamma_k \text{ such that } \gamma_{\mathrm{k}} = \frac{SK\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right)}{\left[Var\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right)\right]^{3/2}} = \frac{SK\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right)}{\left[\sigma_k^2 C_{i,k}\right]^{3/2}}$$

$$\Rightarrow SK\left(C_{i,k+1} \mid C_{i,1},\ldots,C_{i,k}\right) = \gamma_{\mathrm{k}}\left[\sigma_k^2 C_{i,k}\right]^{3/2} = Sk_k^3 C_{i,k}^{3/2}$$

- Then, it is possible to show that the estimator below is unbiased.:

$$\hat{S}k_k^3 = \frac{1}{\left(I-k-\frac{\left(\sum_{i=1}^{I-k} C_{i,k}^{3/2}\right)^2}{\left(\sum_{i=1}^{I-k} C_{i,k}\right)^3}\right)} \sum_{i=1}^{I-k} C_{i,k}^{3/2}\left(\frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k\right)^3 \quad \text{for } 1 \le k \le I\text{-}3$$

Comments:

1 – The formula for $\hat{S}k_k^3$ has a similar "shape" as the formula for $\hat{\sigma}_k^2$

2 – The formula for $\hat{S}k_k^3$ uses the usual weighted average (power 1.5) of cubic differences.

3 – As for the formula of $\hat{\sigma}_k^2$, outliers can play a major role in the estimation of $\hat{S}k_k^3$

4 – The homogeneity formulas in terms of power of $C_{i,k}$ is kept in the above formulas.

Skewness : Hint to the proof of unbiasedness

- We want to show that $E\left(\hat{S}k_k^3\right) = Sk_k^3$
- We have:

$$(Eq) = E\left[\left(I-k-\frac{\left(\sum_{i=1}^{I-k} C_{i,k}^{3/2}\right)^2}{\left(\sum_{i=1}^{I-k} C_{i,k}\right)^3}\right)\hat{S}k_k^3 \mid B_k\right] = E\left[\sum_{i=1}^{I-k} C_{i,k}^{3/2}\left(\frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k\right)^3 \mid B_k\right] \text{ for } 1 \le k \le I\text{-}3$$

with $B_k = \left\{C_{i,j} \mid i+j \le I+1, j \le k\right\}, 1 \le k \le I\text{-}1$

$$(Eq) = \sum_{i=1}^{I-k}\left[\frac{1}{C_{i,k}^{3/2}} E\left(C_{i,k+1}^3 \mid B_k\right)\right] - 3\sum_{i=1}^{I-k}\left[\frac{1}{\sqrt{C_{i,k}}} E\left(\hat{f}_k C_{i,k+1}^2 \mid B_k\right)\right] + 3\sum_{i=1}^{I-k}\left[\sqrt{C_{i,k}}\, E\left(\hat{f}_k^2 C_{i,k+1} \mid B_k\right)\right]$$
$$- \sum_{i=1}^{I-k} C_{i,k}^{3/2} E\left(\hat{f}_k^3 \mid B_k\right)$$

- For each element of (Eq), it is possible to find a formula.
- Overall, simplifications are possible.

Skewness : How to use the new parameter Sk^3_k ?

- Law of total skewness:

$$SK(A) = E[SK(A|B)] + SK[E(A|B)] + 3\operatorname{cov}[E(A|B), \operatorname{var}(A|B)]$$

- In the current context, we have the following scheme (without any mathematical formalism):

$$\begin{aligned} SK(C_{i,I}) = {} & E[SK(C_{i,I} | C_{i,1},\dots,C_{i,I-1})] + SK[E(C_{i,I} | C_{i,1},\dots,C_{i,I-1})] \\ & + 3\operatorname{cov}[E(C_{i,I} | C_{i,1},\dots,C_{i,I-1}), \operatorname{var}(C_{i,I} | C_{i,1},\dots,C_{i,I-1})] \end{aligned}$$

$$SK(C_{i,I}) = E\left[Sk^3_{I-1} C_{i,I-1}^{3/2}\right] + SK\left[f_{I-1} C_{i,I-1}\right] + 3\operatorname{cov}\left[f_{I-1} C_{i,I-1}, \sigma^2_{I-1} C_{i,I-1}\right]$$

$$SK(C_{i,I}) = Sk^3_{I-1} E\left[C_{i,I-1}^{3/2}\right] + f^3_{I-1} SK\left[C_{i,I-1}\right] + 3 f_{I-1} \sigma^2_{I-1} \operatorname{var}\left[C_{i,I-1}\right]$$

No easy formula for such an equation

$\Rightarrow$ Simulation is necessary

Skewness : Simulation to the ultimate

UWY	Dvpt					Ultimate
	1	2	3	4	5	
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$	$\hat{C}_{1,I}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$		$\hat{C}_{2,I}$
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$			$\hat{C}_{3,I}$
4	$C_{4,1}$	$C_{4,2}$				$\hat{C}_{4,I}$
5	$C_{5,1}$					$\hat{C}_{5,I}$
						$\sum_{j=1}^{I}\hat{C}_{j,I}$

$$\hat{f}_k = \frac{\sum_{j=1}^{I-k} C_{j,k+1}}{\sum_{j=1}^{I-k} C_{j,k}}, 1 \le k \le I\text{-}1$$

$$\hat{\sigma}_k^2 = \frac{1}{I-k-1}\sum_{i=1}^{I-k} C_{i,k}\left(\frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k\right)^2 \quad for\ 1 \le k \le I\text{-}2$$

$$\hat{Sk}_k^3 = \frac{1}{\left(I-k-\frac{\left(\sum_{i=1}^{I-k} C_{i,k}^{3/2}\right)^2}{\left(\sum_{i=1}^{I-k} C_{i,k}\right)^3}\right)}\sum_{i=1}^{I-k} C_{i,k}^{3/2}\left(\frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k\right)^3 \quad \text{for } 1 \le k \le I\text{-}3$$

Supporting distribution to estimate the distributional properties of $C_{i,I}$.

$\Rightarrow$ Recursive calculation (see Mack 1999)

Assumption: The chosen supporting distribution should not be influencing the overall simulated skewness.

Skewness : Simulation to the ultimate – Generalized Pareto Distribution

i	C_{i1}	C_{i2}	C_{i3}	C_{i4}	C_{i5}	C_{i6}	C_{i7}	C_{i8}	C_{i9}	C_{i10}
2										
3										
4										
5										
6										
7										
8										
9										
10										
Asymmetry		ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}

$$\hat{C}_{i,k} \propto GPD(\mu_{i,k}; s_{i,k}; \zeta_k)$$

or

$$\hat{C}_{i,k} \propto GEV(\tilde{\mu}_{i,k}; \tilde{s}_{i,k}; \tilde{\zeta}_k)$$

Parameters of all distributions at step k are known

Estimation of parameters of all distributions at step k+1

$$\hat{C}_{i,k} = \mu_{i,k} + \frac{s_{i,k}\left(U_{i,k}^{-\zeta_k} - 1\right)}{\zeta_k} \propto GPD(\mu_{i,k}; s_{i,k}; \zeta_k)$$

$$\left\{\begin{aligned}
&\xi_{k+1} = \frac{\hat{Sk}^3_{k+1}}{\hat{\sigma}^3_{k+1}} = \frac{2\left(1+\zeta_{k+1}\right)\sqrt{1-2\zeta_{k+1}}}{1-3\zeta_{k+1}} \\
&\hat{Variance}_{i,k+1} = \hat{C}_{i,k}\hat{\sigma}_k^2\left(1+\frac{\hat{C}_{i,k}}{\sum_{j=1}^{I-k} C_{j,k}}\right) = \frac{s^2_{i,k+1}}{\left(1-\zeta_{k+1}\right)^2\left(1-2\zeta_{k+1}\right)} \\
&\hat{C}_{i,k+1} = \hat{f}_k \hat{C}_{i,k} = \mu_{i,k+1} + \frac{s_{i,k+1}}{1-\zeta_{k+1}}
\end{aligned}\right.$$

Skewness : Simulation to the ultimate – Generalized Extreme Value Distribution

i	C_{i1}	C_{i2}	C_{i3}	C_{i4}	C_{i5}	C_{i6}	C_{i7}	C_{i8}	C_{i9}	C_{i10}
2										
3										
4										
5										
6										
7										
8										
9										
10										
Asymmetry		ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}

$$\hat{C}_{i,k} \propto GPD(\mu_{i,k}; s_{i,k}; \zeta_k)$$

or

$$\hat{C}_{i,k} \propto GEV(\tilde{\mu}_{i,k}; \tilde{s}_{i,k}; \tilde{\zeta}_k)$$

Parameters of all distributions at step k are known → Estimation of parameters of all distributions at step k+1

$$\hat{C}_{i,k} = \tilde{\mu}_{i,k} + \frac{\tilde{s}_{i,k}\left(\left(-\ln\left(U_{i,k}\right)\right)^{-\tilde{\zeta}_k} - 1\right)}{\tilde{\zeta}_k} \propto GEV(\tilde{\mu}_{i,k}; \tilde{s}_{i,k}; \tilde{\zeta}_k)$$

$$\xi_{k+1} = \frac{\hat{Sk}^3_{k+1}}{\hat{\sigma}^3_{k+1}} = \frac{\Gamma\left(1-3\tilde{\zeta}_{k+1}\right) - 3\Gamma\left(1-\tilde{\zeta}_{k+1}\right)\Gamma\left(1-2\tilde{\zeta}_{k+1}\right) + 2\Gamma\left(1-\tilde{\zeta}_{k+1}\right)^3}{\left(\Gamma\left(1-2\tilde{\zeta}_{k+1}\right) - \Gamma\left(1-\tilde{\zeta}_{k+1}\right)^2\right)^{3/2}}$$

$$\hat{Variance}_{i,k+1} = \hat{C}_{i,k}\hat{\sigma}_k^2\left(1 + \frac{\hat{C}_{i,k}}{\sum_{j=1}^{I-k} C_{j,k}}\right) = \tilde{s}^2_{i,k+1}\frac{\Gamma\left(1-2\tilde{\zeta}_{k+1}\right) - \Gamma\left(1-\tilde{\zeta}_{k+1}\right)^2}{\tilde{\zeta}_{k+1}^2}$$

$$\hat{C}_{i,k+1} = \hat{f}_k\hat{C}_{i,k} = \tilde{\mu}_{i,k+1} + \tilde{s}_{i,k+1}\frac{\Gamma\left(1+\tilde{\zeta}_{k+1}\right) - 1}{\tilde{\zeta}_{k+1}}$$

Skewness : Application on one triangle – Chain-ladder framework

i	C_{i1}	C_{i2}	C_{i3}	C_{i4}	C_{i5}	C_{i6}	C_{i7}	C_{i8}	C_{i9}	C_{i10}
1	397	793	868	889	964	964	964	964	964	964
2	523	1'098	1'475	1'489	1'489	1'489	1'489	1'489	1'489	
3	1'786	2'951	3'370	3'029	3'211	3'289	3'325	3'325		
4	241	465	536	596	652	652	652			
5	327	622	577	583	583	583				
6	275	520	529	529	541					
7	89	327	378	382						
8	295	301	396							
9	151	406								
10	315									

k	1	2	3	4	5	6	7	8	9
$\hat{f}_k$	1.832	1.149	0.969	1.046	1.011	1.006	1.000	1.000	1.000
$\hat{\sigma}_k^2$	90	13	6	2	0	0	0	0	0

Skewness : Application on one triangle – 10 000 simulations

k	1	2	3	4	5	6	7	8	9
$\hat{f}_k$	1.832	1.149	0.969	1.046	1.011	1.006	1.000	1.000	1.000
$\hat{\sigma}_k^2$	90	13	6	2	0	0	0	0	0
$\hat{Sk}_k^3$	401	12	-3	-1	0	0	0	0	NA
ξ_κ	0.472	0.256	-0.249	-0.285	0.378	0.165	0.773	0.153	NA
GPD ζ_κ	-0.565	-0.738	-1.325	-1.377	-0.635	-0.824	-0.378	-0.836	-1.000
GEV ζ_κ	-0.141	-0.199	-0.361	-0.374	-0.165	-0.226	-0.070	-0.230	-0.278

$$\xi_k = \frac{\hat{Sk}_k^3}{\hat{\sigma}_k^3}$$

Comments:

1 – Skewness ξ_k of development years 1 and 2 are the most significant.

2 – Negative skewness is also present on certain development years for which the study of underlying data could be necessary.

Skewness : Application on one triangle - Results

	Actual data		GPD model				GEV model			
i	Mean	Mack mse	Mean	Difference to actual	Mse	Difference to actual	Mean	Difference to actual	Mse	Difference to actual
2	0	0	0	0.1%	0	0.0%	0	-0.1%	0	-0.2%
3	0	0	0	0.1%	0	1.5%	0	-1.0%	0	-0.3%
4	0	0	0	0.1%	0	-1.0%	0	0.5%	0	-0.4%
5	3	6	3	-1.4%	6	-0.6%	3	0.0%	6	-0.9%
6	9	14	9	-1.1%	14	0.4%	9	-0.8%	14	-0.2%
7	24	28	24	2.9%	28	0.8%	25	-1.5%	28	0.3%
8	12	59	12	0.6%	59	0.4%	12	0.5%	60	-0.2%
9	75	101	76	-1.1%	101	0.2%	75	-0.3%	103	-2.3%
10	369	240	370	-0.3%	238	0.8%	369	-0.1%	239	0.5%
Total	494	274	495	-0.2%	270	1.8%	494	-0.2%	271	1.2%

Comment: 10 000 simulations give consistent results with Mack formulaic estimates.

Skewness : Application on one triangle - Results

	Actual data		GPD model	GEV model
i	Mean	Mack mse	Skewness	Skewness
2	0	0	0.0056	-0.0115
3	0	0	0.0154	0.0466
4	0	0	0.0200	0.0168
5	3	6	0.7871	0.7916
6	9	14	0.2016	0.2102
7	24	28	0.3442	0.2953
8	12	59	-0.0768	-0.0383
9	75	101	0.0133	0.0246
10	369	240	0.3774	0.3694
Total	494	274	0.2702	0.2314

k	ξ_k
9	NA
8	0.153
7	0.773
6	0.165
5	0.378
4	-0.285
3	-0.249
2	0.256
1	0.472

Comments:
1 – Overall, we have a positive skewness as expected.
2 – Consistency of overall skewness between GPD and GEV model: The underlying “skewness support” distribution does not seem to play a significant role.

Skewness : Application on one triangle - Results

Lognormal		Gamma		Simulation GPD	Simulation GEV
μ	6.067	k	3.234		
σ	0.519	θ	152.638		
Skewness	1.840	Skewness	1.112	0.2702	0.2314
VaR 99%	1'443		1'343	1'152	1'182

Lognormal and Gamma distribution parameters are chosen on the basis of the mean (494) and of the mse (274) corresponding to the overall reserve risk distribution.

Comments:
VaR 99% estimated with a fitted Lognormal or Gamma distribution is higher than the VaR 99% coming from the simulations.

Reserve risk distribution – Conclusions and next steps

- This is only a first step to try and increase the knowledge of the reserve risk distribution characteristics.
- Bad news: The LogNormal assumption used to model the reserve risk may not be appropriate
- Good News: The LogNormal assumption used to model the reserve risk is conservative
- Next steps:
  - Study of Kurtosis
  - Use of co-skewness to estimate accident year skewness instead of simulating results ?
  - Review of general skewness/kurtosis properties of aggregated accident year reserves
- See you in Den Haag, Holland next year

References

- MACK Thomas, 1993 : “Distribution-free calculation of the standard error of chain-ladder reserve estimates”, ASTIN Bulletin Vol 23, No2, 1993
- MACK Thomas, 1999: “The standard error of chain-ladder reserve estimates, Recursive calculation and inclusion of a tail factor”, ASTIN Bulletin Vol 29, No2, 1999
- MACK Thomas, 2008 : “The Prediction Error of Bornhuetter/Ferguson”, ASTIN Bulletin Vol 38, No1, 2008
- WÜTHRICH, M. V., MERZ, M., 2008: “Stochastic Claims Reserving Methods in Insurance“, Wiley, Chichester
- SALZMANN R., WÜTHRICH, M. V., MERZ, M., 2012: “Higher Moments of the Claim Development Result in General Insurance“, ASTIN Bulletin Vol 42, No1, 2012
- MEYERS G., SHI P., 2011: “The Retrospective Testing of Stochastic Loss Reserve Models“, ASTIN Colloquium 2011 Madrid
- ARBENZ P., SALZMANN R., 2010: “A robust distribution-free loss reserving method with weighted data- and expert-reliance”, http://www.risklab.ch/hclmethod