

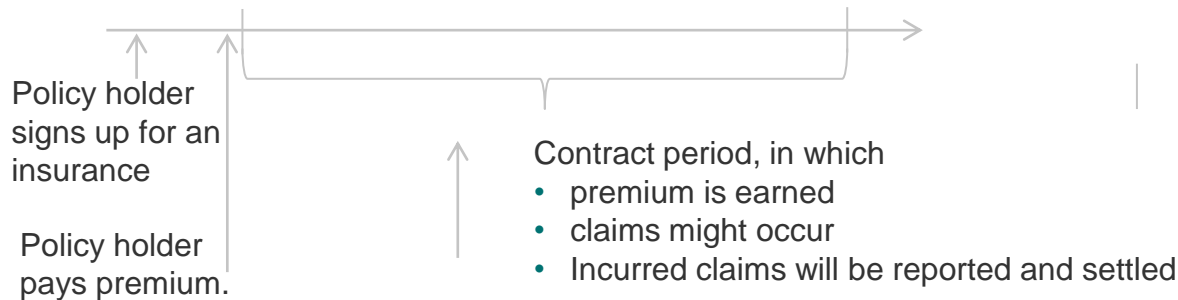
Non-life insurance mathematics

Nils F. Haavardsson, University of Oslo and DNB
Skadeforsikring

Introduction to reserving

Non-life insurance from a financial perspective:

for a premium an insurance company commits itself to pay a sum if an event has occurred



Issues that need to be solved:

- How much premium is earned?
- How much premium is unearned?
- How do we measure the number and size of unknown claims?
- How do we know if the reserves on known claims are sufficient?

Premium reserves

The premium reserve is split in two parts:

- Provision for unearned premiums
- Provisions for unexpired risks

Earned and unearned premium:

- Written premium is earned evenly/uniformly over the cover period
- The share of the premium that has been earned is the past time's proportion of the total period
- If a larger premium has been received the difference is the unearned premium

Example:

An insurance policy starts on September 1 2012 and is valid until August 31 2013.

The premium for the entire period is 2400.

At 31 December we have received two quarterly premiums or 1200.

We have then earned $(4/12) \cdot 2400 = 800$.

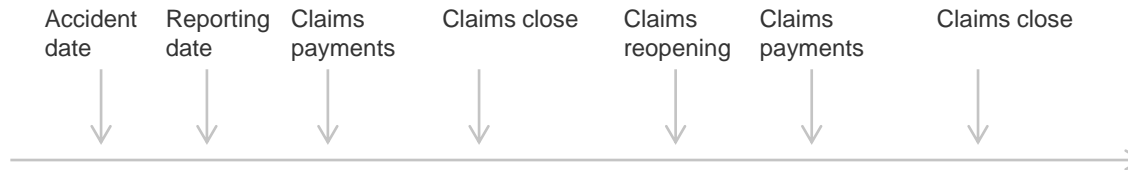
Unearned is $1200 - 800 = 400$



Unexpired risk reserve

- Regard entire period covered by the insurance
- From a point in time, say 31/12-2012, we look forward to all the claims and expenses that could occur after this point. Call them FC_{3112}
- If $FC_{3112} > \text{Future premiums yet not due (FP)} + \text{unearned premium reserve (UP)}$ the difference is accounted as unexpired risk reserve
- In example assume $FC_{3112} = 1800 > \text{FP} + \text{UP} = 1200 + 400 = 1600$, so unexpired risk reserve is 200

Claims reserves



Claims reserving issues:

- How do we measure the number and size of unknown claims? (IBNR reserve, i.e., Incurred But Not Reported)
- How do we know if the reserves on known claims are sufficient? (RBNS reserve, i.e., Reserved But Not Settled)

Claims occurrence year	Claims losses settled
2005	3963
2006	4975
2007	5873
2008	6401
2009	6563
2010	6358
2011	6918
2012	3072

Claim payments plus claims handling expenses

The development of claims losses settled

- Claims losses settled for each claims occurrence year are often not paid on one date but rather over a number of years

Incremental claims loss settlement data presented as a run-off triangle

Incremental claims loss settlements		Development year							
		0	1	2	3	4	5	6	7
Claims occurrence year	2005	1232	946	520	722	316	165	48	14
	2006	1469	1201	708	845	461	235	56	
	2007	1652	1416	959	954	605	287		
	2008	1831	1634	1124	1087	725			
	2009	2074	1919	1330	1240				
	2010	2434	2263	1661					
	2011	2810	4108						
	2012	3072							

Comments:

- The development year* for a claims settlement amount reflects how long after the claims occurrence year the amount was settled.
 - An amount settled during the claims occurrence year was settled in development year 0
- In the example the largest development year for any claims occurrence years is 7
- The data shown represents the incremental claims losses settled in the development year
- For any cell in the table, the value shown represents the incremental claims loss amount that was settled in calendar year
- Each diagonal set of data represents the amounts settled in a single calendar year
- Green cells represent observed data – all red represent time periods in the future for which we wish to estimate the expected claims settlements amounts

Assumptions underlying the CLM

- Patterns of claims loss settlement observed in the past will continue in the future
- The development of claims loss settlement over the development years follows an identical pattern for every claims occurrence year
- But the observed claims loss settlement patterns may change over time:
 - Changes in product design and conditions
 - Changes in the claims reporting, assessment and settlement processes (example: different owners)
 - Change in the legal environment
 - Abnormally large or small claim settlement amounts
 - Changes in portfolio so that the history is not representative for predicting the future (example: strong growth)

CLM in practice

Determining the CLM estimator for the cumulative claims loss settlement factor

Cumulative claims loss settlements		Development year							
		0	1	2	3	4	5	6	7
Claims occurrence year	2005	1232	2178	2698	3420	3736	3901	3949	3963
	2006	1469	2670	3378	4223	4684	4919	4975	
	2007	1652	3068	4027	4981	5586	5873		
	2008	1831	3465	4589	5679	6401			$3736+4684+5586+6401=20407$
	2009	2074	3993	5323	6563				$3736+4684+5586+6401=20407$
	2010	2434	4697	6358					
	2011	2810	6918						
	2012	3072							$20407/18300=1,1151$
CLM estimator for claims loss settlement factor			1,9989	1,3140	1,2422	1,1151	1,0491	1,0118	1,0035

Comments:

- These CLM estimators for the cumulative claims loss settlement factors are used to estimate the cumulative claims loss settlement amount in the future
- For each claims occurrence year the last historical observation is used together with the appropriate CLM estimator for the development factor to estimate the cumulative settlement amount in the next development year
- This value is, in turn, multiplied by the estimator for the development factor for the next development year and so on.

CLM in practice

Determining the estimated cumulative claims loss settlements in future periods

Cumulative claims loss settlements		Development year							
		0	1	2	3	4	5	6	7
Claims occurrence year	2005	1232	2178	2698	3420	3736	3901	3949	3963
	2006	1469	2670	3378	4223	4684	4919	4975	4993
	2007	1652	3068	4027	4981	5586	5873	5942	5963
	2008	1831	3465	4589	5676	6401	6401 *1,0491 = 6715	6715 *1,0118 = 6794	6794 *1,0035 = 6818
	2009	2074	3993	5323	6563	7319	7678	7768	7796
	2010	2434	4697	6358	7898	8807	9239	9348	9381
	2011	2810	4918	6462	8027	8952	9391	9502	9535
	2012	3072	5686	7471	9280	10349	10857	10985	11024
	CLM estimator for claims			1,8508	1,3140	1,2422	1,1151	1,0491	1,0118

Comments:

- The values shown in the red cells are the estimators for future cumulative claims settled
- These estimates are always based on the latest available cumulative claims settlement amounts for the relevant claims occurrence year, i.e., the estimated future cumulative claims settlements are always based on the last green diagonal of data
- It is now simple to derive the estimated incremental claims settlement amounts for the future periods
- An incremental settlement amount is the difference between two consecutive cumulative settlement amounts

CLM in practice

Determining the estimated incremental settlement amounts from the estimated cumulative amounts

Incremental claims loss settlements		Development year							
		0	1	2	3	4	5	6	7
Claims occurrence year	2005	1232	946	520	722	316	165	48	14
	2006	1469	1201	708	845	461	235	56	18
	2007	1652	1416	959	954	605	287	69	21
	2008	1831	1634	1124	1087	725	6715-6401 = 314	6794-6715 = 79	6818-6794 = 24
	2009	2074	1919	1330	1240	756	359	91	28
	2010	2434	2263	1661	1540	909	432	109	33
	2011	2810	4108	1544	1565	924	439	111	34
	2012	3072	2614	1785	1810	1069	508	128	39

CLM in practice

Determining the estimated incremental settlement amounts from the estimated cumulative amounts

Incremental claims loss settlements		Development year							
		0	1	2	3	4	5	6	7
Claims occurrence year	2005								
	2006								18
	2007							69	21
	2008						314	79	24
	2009					758	359	91	28
	2010				1540	909	432	109	32
	2011			1574	1565	924	439	117	34
	2012		2614	1785	1810	1069	500	129	39

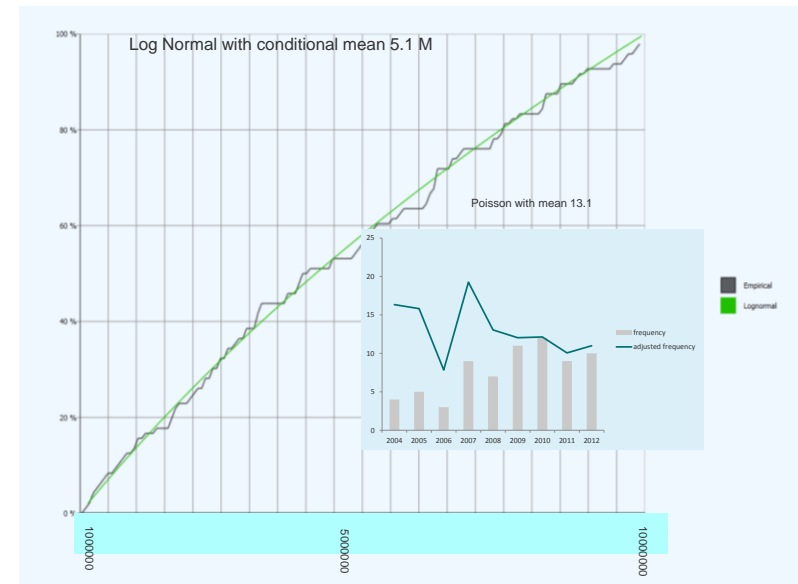
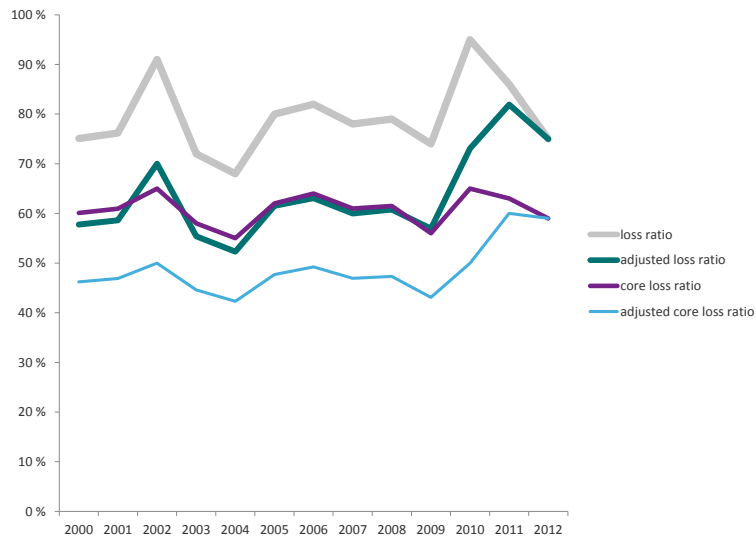
Calendar year	Estimated claims loss settlement amounts
2013	6855
2014	4718
2015	2181
2016	1069+439+109+28=1645
2017	652
2018	162
2019	39

Comments:

- Group the estimated incremental claims loss settlement amounts by the year in which they will be settled
- These cash flows can then be discounted to determine the technical provisions
- Norwegian State Treasury Bonds (Statsobligasjoner in Norwegian) may be used as discount factor
 - Example: a cash flow due in 2017 is discounted with a 4 year old Norwegian State Treasury Bond etc. Why do we hope that the development year does not exceed 10 ??

It can be useful to be able to predict loss ratio

	Frequency	Severity	Loss ratio
Core	1	220	55,0 %
Large	13,1	5,1	16,7 %
			71,7 %



Overview

Important issues	Models treated	Curriculum	Duration (in lectures)
What is driving the result of a non-life insurance company?	insurance economics models	Lecture notes	0,5
How is claim frequency modelled?	Poisson, Compound Poisson and Poisson regression	Section 8.2-4 EB	1,5
How can claims reserving be modelled?	Chain ladder, Bernhuetter Ferguson, Cape Cod,	Note by Patrick Dahl	2
How can claim size be modelled?	Gamma distribution, log-normal distribution	Chapter 9 EB	2
How are insurance policies priced?	Generalized Linear models, estimation, testing and modelling. CRM models.	Chapter 10 EB	2
Credibility theory	Buhlmann Straub	Chapter 10 EB	1
Reinsurance		Chapter 10 EB	1
Solvency		Chapter 10 EB	1
Repetition			1

Overview of this session

The Bornhuetter Ferguson model

An example

Stochastic claims reserving in non-life insurance

The Bornhuetter-Ferguson method

- The Bornhuetter-Ferguson method is more sophisticated than the Naive loss method
- It looks on where in time claims will be reported or paid
- It is very similar to an ordinary budgeting model used by businesses
- You budget for future claims by period
- The sum of these future budgeted claims is the IBNR reserve

The Bornhuetter-Ferguson method

- More formally, the following principles apply:

(BF1) Expected claims are considered known, i.e. we have a

predictor of final claims identical to the mean $C_{im}^N = E[C_{im}]$

(BF2) Unemerged claims are independent of emerged claims,

or C_{ij} is independent of $C_{im} - C_{ij}$

(BF3) The F_{ij} are known, in the meaning that

$$\text{we know } F_{ij} = \frac{E[C_{im}]}{E[C_{ij}]}$$

- F_{ij} in (BF3) is the factor that would develop losses from development period j to the end for accident year i .
- F_{ij} could have been determined by the Chain Ladder technique

The Bornhuetter-Ferguson method

- The unbiased Bornhuetter-Ferguson predictor is given by

$$C_{im}^{BF} = C_{ij} + \left(1 - \frac{1}{F_{ij}}\right) * \hat{C}_{im} \quad (1)$$

- This predictor takes emerged claims into account as it swaps past expected emergence with real emergence (i.e. it is better than the Naive Loss Ratio)
- (1) can be re-written

$$\begin{aligned} C_{im}^{BF} &= \frac{1}{F_{ij}} C_{ij} F_{ij} + \left(1 - \frac{1}{F_{ij}}\right) * \hat{C}_{im} = \\ &= \underbrace{W_{ij} C_{ij} F_{ij}}_{\text{Chain ladder type estimate}} + (1 - W_{ij}) \underbrace{\hat{C}_{im}}_{\text{«Known» expected claims}} \quad \text{with} \quad W_{ij} = \frac{1}{F_{ij}} \end{aligned} \quad (2)$$

- We make a further assumption

$$(BF4) \quad \text{var}[C_{ij} * F_{ij}] = F_{ij} * \text{var}[C_{im}]$$

The Bornhuetter-Ferguson method

(Theorem) The weights implicitly defined in (2) produces the best combination of the two predictors $C_{im}^{CL} = C_{ij} * F_{ij}$ and \hat{C}_{im}^N (in the meaning minimizing quadratic loss)

- Proof: we ultimately want to weight between chain ladder and the naive loss ratio method.
- Introduce the two random variables

$$\varepsilon_{CL} = C_{im} - C_{ij} F_{ij} \quad \text{Error from Chain Ladder estimate}$$

$$\varepsilon_{NL} = C_{im} - \hat{C}_{im}^N \quad \text{Error from naive loss method}$$

- How should ε_{CL} and ε_{NL} be weighted to minimize total error?
- Note that

The Bornhuetter-Ferguson method

Bornhuetter Ferguson

An example

Stochastic claims reserving

$$\begin{aligned}
 \varepsilon_{\text{Total}} &= W_{ij} \varepsilon_{\text{CL}} + (1 - W_{ij}) \varepsilon_{\text{NL}} = \\
 &= W_{ij} (C_{im} - C_{ij} F_{ij}) + (1 - W_{ij}) (C_{im} - \hat{C}_{im}^N) \\
 &= W_{ij} C_{im} - W_{ij} C_{ij} F_{ij} + C_{im} - \hat{C}_{im}^N - W_{ij} C_{im} + W_{ij} \hat{C}_{im}^N \\
 &= C_{im} - \hat{C}_{im}^{BF}
 \end{aligned} \tag{3}$$

- We want to find the best combination of the two predictors (CL + NL) so that the error in (3) is minimized
- Thus, we need to solve the problem

$$\min_{W_{ij}} \{E\{\varepsilon_{\text{Total}}\}^2\} = \min_{W_{ij}} \{E(C_{im} - \hat{C}_{im}^{BF})^2\} \tag{4}$$

The Bornhuetter-Ferguson method

- In general for an estimator W of a parameter θ :

Mean square error of estimator W is $E(W - \theta)^2$:

$$E(W - \theta)^2 = E(W^2 - 2W\theta + \theta^2) = EW^2 - 2EW\theta + \theta^2 \quad (5)$$

$$= EW^2 - (EW)^2 - 2EW\theta + \theta^2 + (EW)^2$$

$$= \text{var } W + (EW - \theta)^2 = \text{var } W \text{ if } EW = \theta$$

- From (4) and (5) we see that we want to minimize the variance of \hat{C}_{im}^{BF}
- The strategy is now to use Lemma 6.2 on the two variables ε_{CL} and ε_{NL}
- We can see that ε_{CL} and ε_{NL} have the same mean, 0.
- We also need to prove that ε_{CL} and ε_{NL} are uncorrelated
- To show the uncorrelatedness of the components we need the auxiliary result,

$$\begin{aligned} \text{Cov}[-C_{ij}F_{ij}, C_{im}] &= -F_{ij}\text{Cov}[C_{ij}, C_{im}] = \\ &= -F_{ij}\text{Cov}[C_{ij}, (C_{im} - C_{ij}) + C_{ij}] = -F_{ij}\text{Cov}[C_{ij}, (C_{im} - C_{ij})] \\ &- F_{ij}\text{Cov}[C_{ij}, C_{ij}] = -F_{ij}0 - F_{ij}\text{var}[C_{ij}] = -F_{ij}\text{var}[C_{ij}] / F_{ij} \\ &= -\text{var}[C_{ij}] = -v^2 \end{aligned}$$

The Bornhuetter-Ferguson method

- Using the calculation rules for covariances and remembering that covariances between a random variable and a constant vanishes, we get that

$$\begin{aligned} \text{Cov}[\varepsilon_{CL}, \varepsilon_{NL}] &= \text{Cov}[C_{im} - C_{ij}F_{ij}, C_{im} - C_{im}^N] = \\ &= \text{Cov}[C_{im}, C_{im}] + \text{Cov}[-C_{ij}F_{ij}, C_{im}] = v^2 - v^2 = 0 \end{aligned} \quad (6)$$

- (6) proves uncorrelatedness.
- We also need to calculate the variance of ε_{CL} and ε_{NL}

$$\begin{aligned} \text{var}[\varepsilon_{NL}] &= \text{var}[C_{im} - C_{im}^N] = v^2, \\ \text{var}[\varepsilon_{CL}] &= \text{var}[C_{im} - C_{ij}F_{ij}] = \\ &= \text{var}[C_{im}] + F_{ij}^2 \text{var}[C_{ij}] + 2\text{Cov}[C_{im}, -C_{ij}F_{ij}] \\ &= v^2 + F_{ij}^2 v^2 / F_{ij} - 2v^2 = v^2(F_{ij} - 1) \end{aligned}$$

- Using Lemma 6.2 we find that the optimal weights are

$$W_{ij} = \frac{1}{\frac{v^2(F_{ij} - 1)}{1} + \frac{1}{v^2}} = \frac{1}{F_{ij}}$$

Bornhuetter-Ferguson example

EksempelChainLadder - Microsoft Excel

Formler Data Se gjennom Visning Tillegg

144079245,64

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Tidspersiden	Tidspersiden	Periode + 0	Periode + 1	Periode + 2	Periode + 3	Periode + 4	Skadereserver (Tidspersiden)	opptjent	periode+4 beregnet	skadereserve beregnet		skaderese chain lac
2	0	2008	7008147,76	25877312,92	31723256,17	32718766,17	33019648,17	0	34689213				0
3	1	2009	30105219,65	65758082,38	76744304,57	79560296,33		724970,20432	99392051	=F3+(1-1/E19)*H21/100*3	=J3-F3		731637,036332
4	2	2010	89181137,64	171787014,81	201380708,65			8532036,1183	178213379	=E4+(1-1/(D19*E19))*H22/100*4	=J4-E4		8993401,65690
5	3	2011	109818684,47	198015727,55				40795736,602	277531254	=D5+(1-1/(C19*D19*E19))*H23/100*5	=J5-D5		45300160,9535
6	4	2012	97250541,11					144079245,64	331404506	=C6+(1-1/(B19*C19*D19*E19))*H24/100*6	=J6-C6		136286647,604
7													191311847
8													
9													
10													
11													
12	methode	Dato + 0	Dato + 1	Dato + 2	Dato + 3	Tidspersiden	Tidspersiden	Skadeprosent					
13	Standard triangel	1,9543090269	1,176241115	1,0351395545	1,0091960069								
14	Ar-vektet triangel	2,1052283894	1,1824566176	1,0344165591	1,0091960069								
15	Sisteverdivektet triangel	1,8031150938	1,1722696787	1,0366931693	1,0091960069								
16	Gjennomsnittsvekt et triangel	2,4015305087	1,1884166511	1,034037124	1,0091960069								
17	K last-weighted Chain Ladder	1,9712203148	1,1884166511	1,034037124	1,0091960069								
18	Chain Ladder uten de ekstreme utviklings	2,0552729253	1,1722696787	1,034037124	1,0091960069								
19	Oppdater	2,0484460431	1,180011732	1,0347267758	1,0091960069								
20						2008	0	95,187078193					

ChainLadder - Ark1 BonhuetterFerguson

15:16 24.05.2013

Bornhuetter-Ferguson example

Time period	Year	Period + 0	Period + 1	Period + 2	Period + 3	Period + 4	Claims reserves (The time period)	earned premium	periode+4 calculated	claims reserve calculated
0	2008	7 008 148	25 877 313	31 723 256	32 718 766	33 019 648	0	34 689 213		
1	2009	30 105 220	65 758 082	76 744 305	79 560 296		724 970	99 392 051	80 285 267	724 970
2	2010	89 181 138	171 787 015	201 380 709	.		8 532 036	178 213 379	209 912 745	8 532 036
3	2011	109 818 684	198 015 728	.	.	.	40 795 737	277 531 254	238 811 464	40 795 737
4	2012	97 250 541	144 079 246	331 404 506	241 329 787	144 079 246

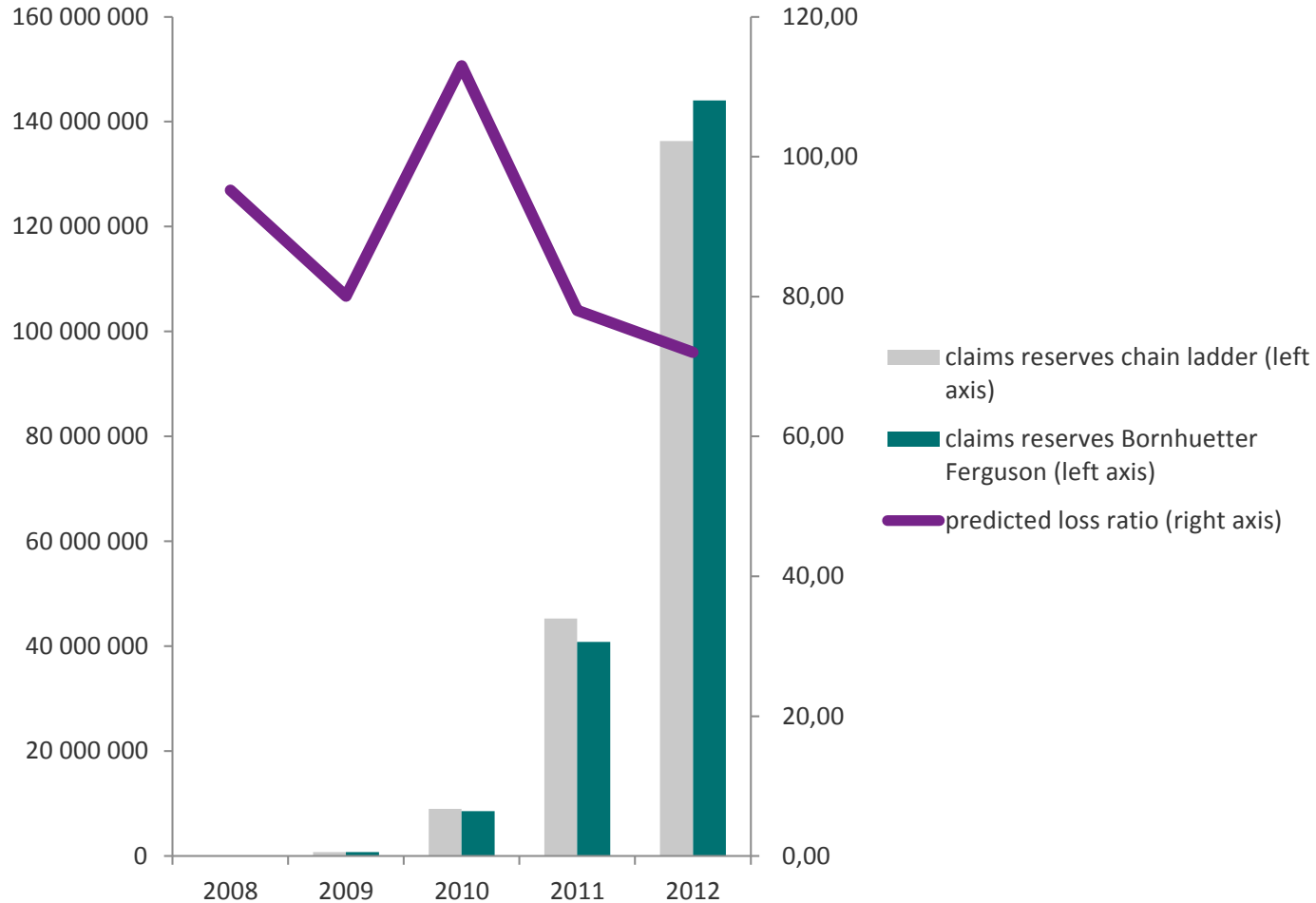
	claims reserves chain ladder (left axis)	claims reserves Bornhuetter Ferguson (left axis)	predicted loss ratio (right axis)
2008	0	0	95,19
2009	731 637	724 970	80,05
2010	8 993 402	8 532 036	113,00
2011	45 300 161	40 795 737	78,00
2012	136 286 648	144 079 246	72,00

Comparison Bornhuetter-Ferguson and chain-ladder

Bornhuetter Ferguson

An example

Stochastic claims reserving



How good is our model?

- Up to now we have (only) given an estimate for the mean/expected ultimate claim
- We would also like to know how good this estimate predicts the outcome of random variables
- How accurate are our reserves estimates?
- Imagine a non-life insurance company with a total claims reserve of 6 047 061 NOK and a profit-loss statement given below
- The earnings statement is slightly positive (+60 000)
- If the claims reserves are reduced by 1% this doubles the income before taxes
- Only a slight change of the claims reserves may have an enormous impact on the earning statement
- Therefore it is very important to know the uncertainties in the estimates

Earning statement at 31 December	
a) Premium earned	4 020 000
b) Claims incurred current accident year	-3 340 000
c) Loss experience prior years	-40 000
d) Underwriting and other expen	-1 090 000
e) Investment income	510 000
Income before taxes	60 000

The mean square error of prediction measures the quality of the estimated claims reserves

- Assume that we have a random variable θ and that W estimates θ
- Then the mean square error of W for θ is given by $E(W - \theta)^2$. We also have that

$$\begin{aligned} E(W - \theta)^2 &= E(W^2 - 2W\theta + \theta^2) = EW^2 - 2EW\theta + \theta^2 \\ &= EW^2 - (EW)^2 - 2EW\theta + \theta^2 + (EW)^2 \\ &= \text{var } W + (EW - \theta)^2 = \text{var } W \text{ if } EW = \theta \end{aligned}$$

- Normally we measure the quality of the estimators and the predictors for the ultimate claims by means of second moments such as the mean square error of prediction defined above
- We really want to derive the whole predictive distribution of stochastic claims reserving.
- Most often it is not feasible to calculate this distribution analytically
- Therefore we have to rely on numerical algorithms such as Bootstrapping methods and Monte Carlo Simulation methods to produce a simulated predictive distribution for the claims reserves

The previous example revisited

Year	Period + 0	Period + 1	Period + 2	Period + 3	Period + 4	Claims reserve (time period)	Claims reserve/ Payments(%)	Standard error	Relative standard error (%)
2008	7 008 148	25 877 313	31 723 256	32 718 766	33 019 648	.	0,00	.	.
2009	30 105 220	65 758 082	76 744 305	79 560 296	80 331 442	771 145	0,97	255 413	33,1
2010	89 181 138	171 787 015	201 380 709	.	210 358 696	8 977 988	4,46	2 201 288	24,5
2011	109 818 684	198 015 728	.	.	246 995 944	48 980 216	24,74	12 094 890	24,7
2012	97 250 541	.	.	.	271 859 031	174 608 490	179,55	55 673 299	31,9

	claims reserves chain ladder	claims reserves Bornhuetter Ferguson	claims reserve stochastic method with Bootstrap
2008	0	0	.
2009	731 637	724 970	771 145
2010	8 993 402	8 532 036	8 977 988
2011	45 300 161	40 795 737	48 980 216
2012	136 286 648	144 079 246	174 608 490

Imagine you want to build a reserve risk model

There are three effects that influence the best estimate and the uncertainty:

- Payment pattern
- RBNS movements
- Reporting pattern

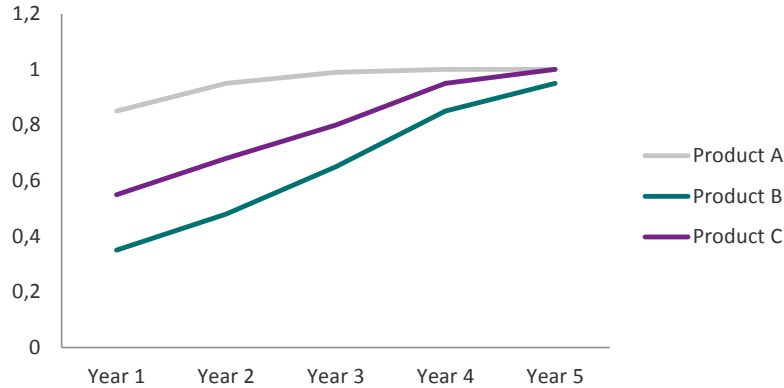
Up to recently the industry has based model on payment triangles:

Year	Period + 0	Period + 1	Period + 2	Period + 3	Period + 4
2008	7 008 148	25 877 313	31 723 256	32 718 766	33 019 648
2009	30 105 220	65 758 082	76 744 305	79 560 296	
2010	89 181 138	171 787 015	201 380 709		?
2011	109 818 684	198 015 728			
2012	97 250 541				

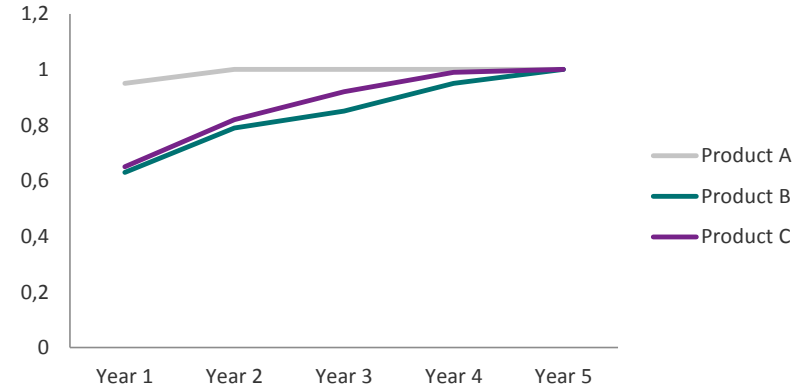
What will the future payments amount to?

Payment pattern, reporting pattern and RBNS movement

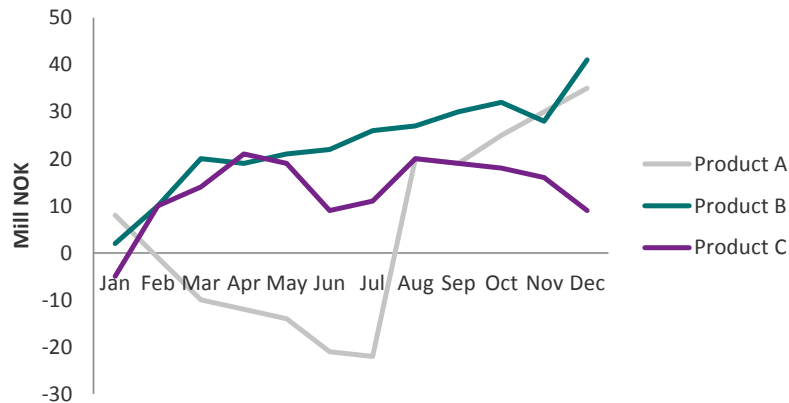
Payment pattern



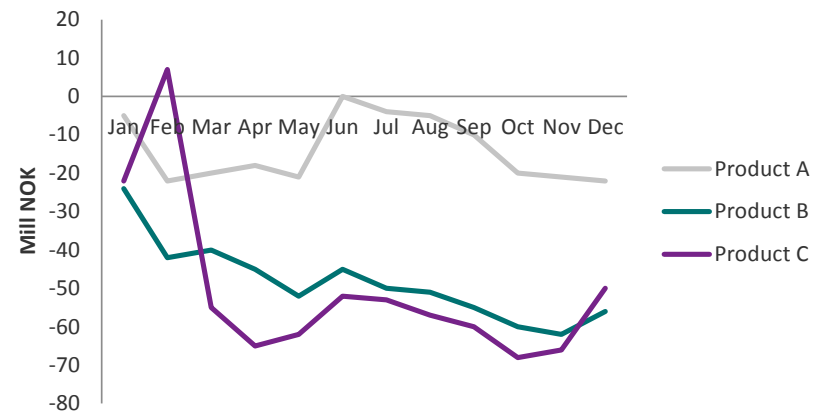
Reporting pattern



RBNS movement year 1



RBNS movement year 2

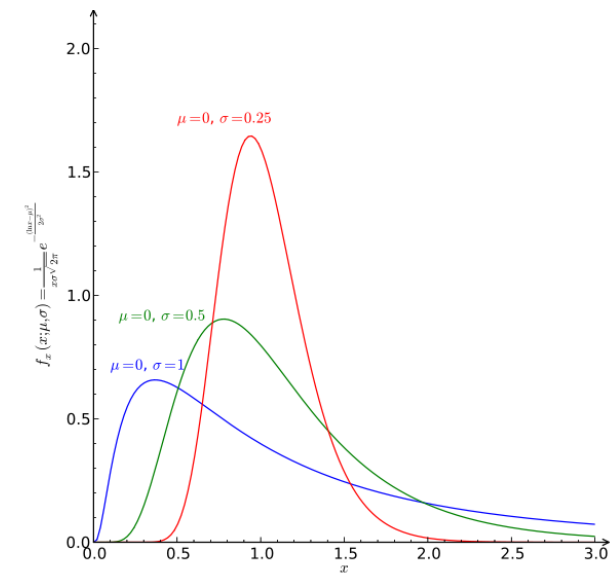


Reserve risk distribution

Year	Period + 0	Period + 1	Period + 2	Period + 3	Period + 4	Relative standard error (%)
2008	7 008 148	25 877 313	31 723 256	32 718 766	33 019 648	.
2009	30 105 220	65 758 082	76 744 305	79 560 296	80 291 933	33,1
2010	89 181 138	171 787 015	201 380 709	208 457 137	210 374 110	24,5
2011	109 818 684	198 015 728	232 914 240	241 098 743	243 315 889	24,7
2012	97 250 541	190 057 610	223 553 576	231 409 149	233 537 189	31,9

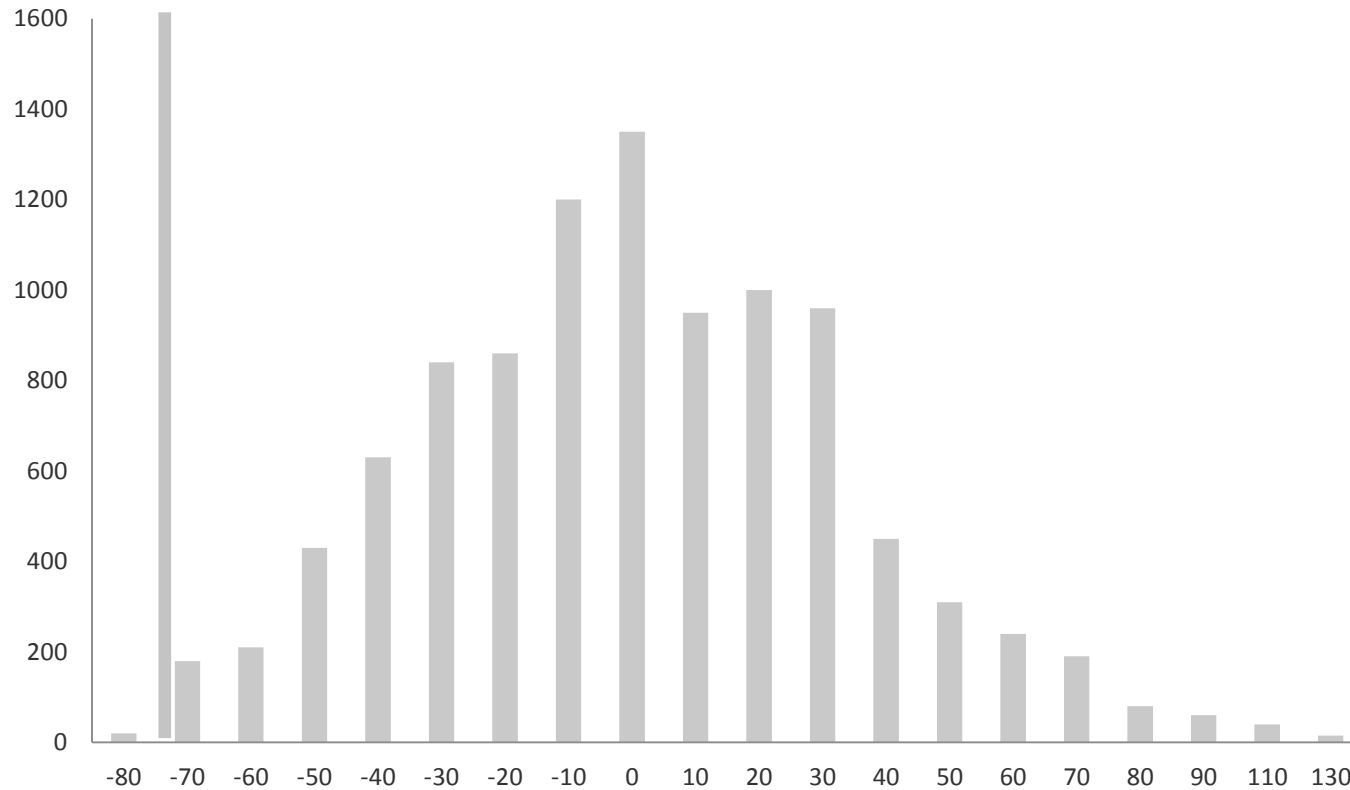
One possible stochastic models for the chain-ladder technique:

- the logarithm of the cumulative claims amounts $Y_{ij} = \log(C_{ij})$ and the log-Normal class of models $Y_{ij} = m_{ij} + \varepsilon_{ij}$ with ε_{ij} as independent normal random errors



Worst case for reserve risk

Development result product A



99.5% = -74 M