Institute and Faculty of Actuaries

GIRO40

8 – 11 October, Edinburgh

Institute and Faculty of Actuaries

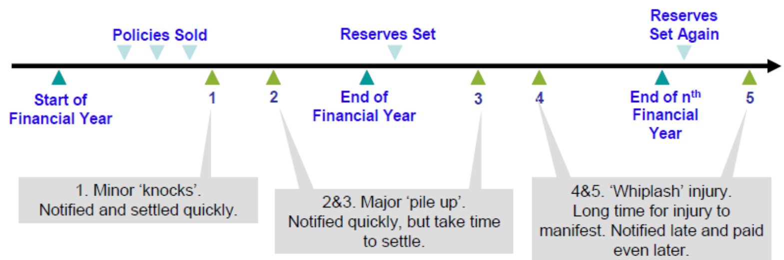
Don't throw the baby out with the bathwater

Going granular in reserving and respecting the conventional chain ladder approach

María Dolores Martínez-Miranda
Jens Perch Nielsen
Cass Business School, City University London

17 September 2013

The Claims Reserving exercise



- Claims are first notified and then (at a later date) settled - **reporting delays** and **settlement delays** exist.
- The amount and timing of future claims is not known and this creates an **uncertainty over the amount of reserves** that needs to be held.
- Companies have an **outstanding liability** for claims events that have already happened and for claims that have not yet been fully settled.

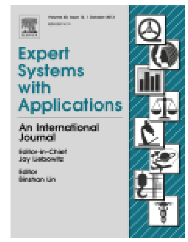
An exercise which amounts to about 5% of GNP

- **Insurance amounts about to 5% of the GNP in western countries**. In the UK the greatest number work in the banking industry (454,200), followed by insurance (345,600). [Source TheCityUK 2012]
- The output from the reserving exercise is probably the **most important number on a non-life insurance balance sheet**.
- The apparent **profitability** of a business as well as its **solvency** is highly dependent upon the value of the reserves and the reserving philosophy.

Our proposal: reformulating the problem

Martínez-Miranda M.D., Nielsen, J.P., Sperlich, S., Verrall, R. (2013).

Continuous Chain Ladder: Reformulating and generalizing a classical insurance problem. *Experts Systems with Applications*, 40(14), 5588–5603.



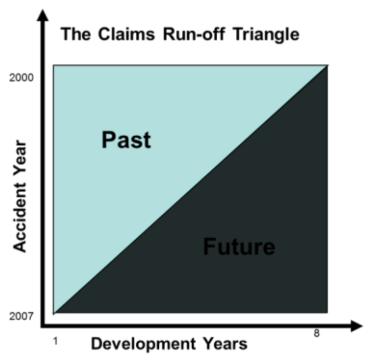
It is time to modernising claims reserving methodology

- Classical reserving methods rely on aggregate run-off triangles since only recently has micro-level information been available at companies.
- Now **the challenge is to use micro-level information in an efficient way**.

- There is a growing awareness among non-life actuaries that modern statistical expert models should be used when analysing this type of data.

An important issue: the available data

- The available information matters: **look at your data…**
- **Aggregated run-off triangles** lead to classical collective methods such as the popular Chain Ladder.



Accident (underwriting) year: year in which the claim arose or was underwritten

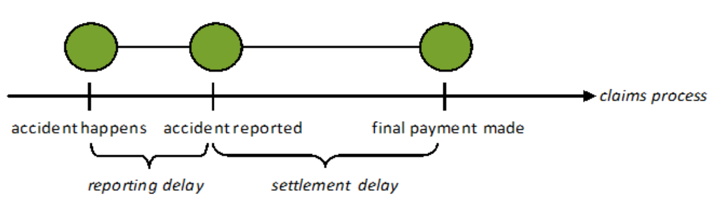
Development year: difference between the payment (or other action) year and the accident year

Periods: years, quarters …

Data: payments, number of claims …

When you have "more data": going granular

- **Micro-level data** leading to **individual claim loss models** (among others Taylor et al. 2008, Zhao and Zhou 2010, Antonio and Platz 2013)
- These approaches aim is to **understand** and **model** the **individual claim process** in the general claims process

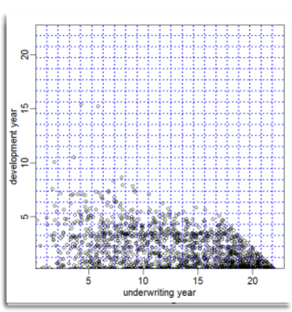


Going granular in reserving…
…but respecting the chain ladder approach

- We suggest to **reformulating classical chain ladder** into a modern statistical framework.
- Then, a natural way to improve it will come: Continuous Chain Ladder.

- Some good reasons to proceed in such way:
  1. Actuaries have **tacit knowledge worth millions.**
  2. When you build a system from many small systems you get **bias**. Keep the chain ladder mean as a benchmark.
  3. Simpler models are preferred for forecasting.

Reformulating claims reserving as a density problem

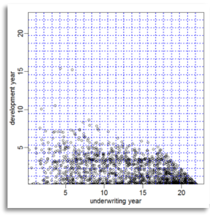


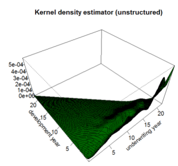
Specifications:
1. Use data on a individual claim base: granular data.
2. The data are arranged in a two-dimensional space: **still a triangle.**

- Outstanding liabilities can be derived by **integrating a two-dimensional density.**
- Thus, the aim is to **estimate/forecast a density which is only observed into a triangle**.

Solving the problem in two steps

1. **Density estimation with a triangular support**



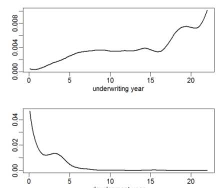


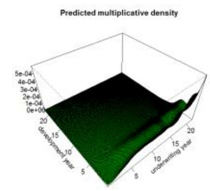
Look for **the best density estimator in the market!**

2. **Forecasting problem: the density in the whole square**

Start with a simple model: **multiplicative structure**

$$f(x, y) = f_1(x) f_2(y)$$

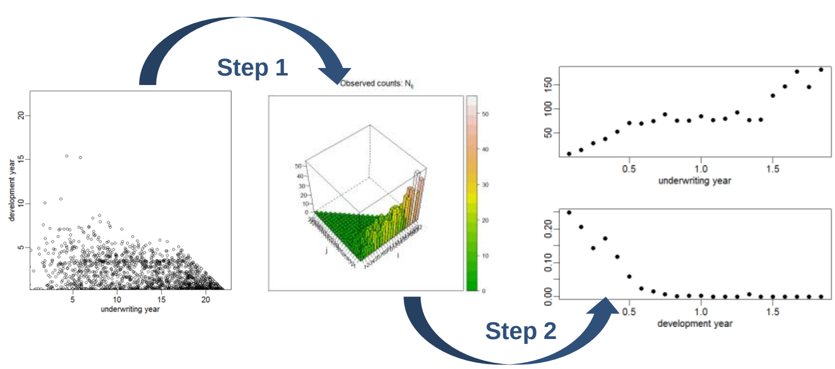




Predict the **outstanding numbers** by integrating the estimated density

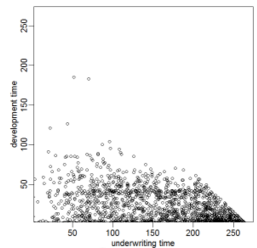
Reformulating classical chain ladder in this framework

Chain ladder starts from a **histogram** of the granular data. Then, this histogram is projected on a **multiplicative structure** for forecasting the future



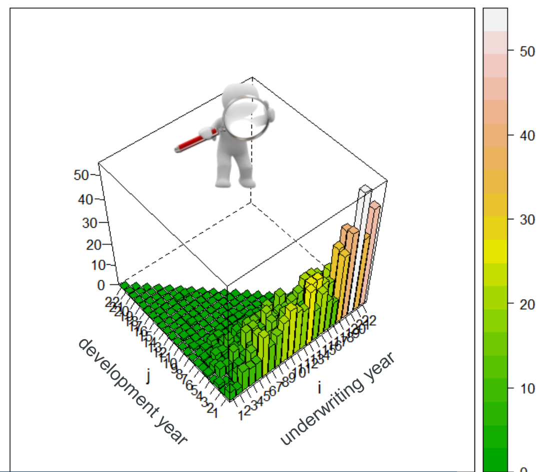
That we can learn…

1. Chain ladder is indeed granular!



A **histogram** is a common graphical tool to discover and show the underlying distribution of **continuous data.**

Observed counts: N_{ij}

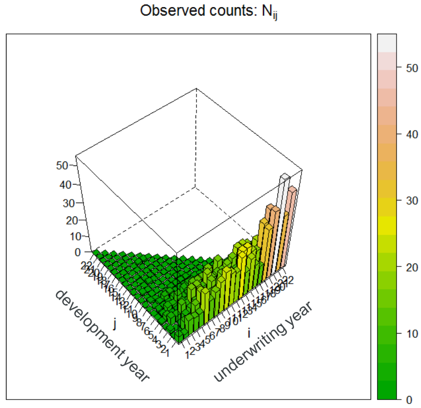


It is the **simplest nonparametric density estimator.**

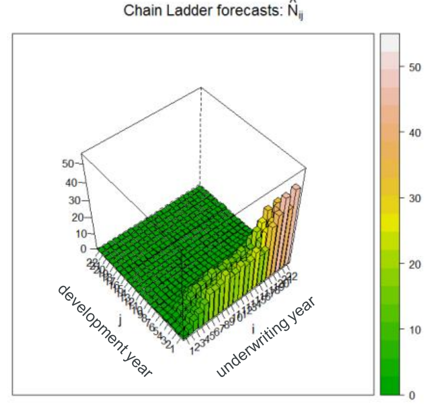
That we can learn…

2. The multiplicative structure

Observed counts: N_{ij}



Chain Ladder forecasts: $\hat{N}_{ij}$



The outstanding numbers are predicted assuming the simple **mean structure**: $\mathrm{E}\left[N_{ij}\right] = \alpha_i \beta_j$

Summary:

- Classical chain ladder estimates the density in the triangle using a histogram.
- Assumptions for forecasting the target density in the future:
  - A **multiplicative structure** for the 2-dimensional density.

$$f(x, y) = f_1(x) f_2(x)$$

  - The densities in the underwriting and development directions are **piece-wise constant.**

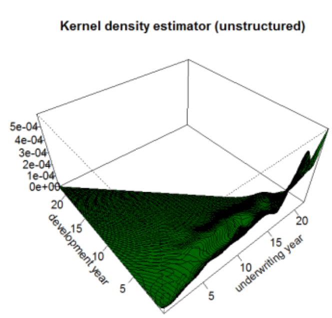
Advantages of this approach: simplicity, the problem can be treated as a parametric problem with maximum likelihood solutions.

Drawbacks:

- The **histogram is an inefficient estimator of the density.**
- It leads to **discrete time effects.**

Continuous Chain Ladder: the natural improvement

1. Replace the histogram by a **kernel estimator** of the density: the natural way to improve on histograms
2. Assume a **multiplicative structure** but with **non-parametric time effects** (continuous densities)



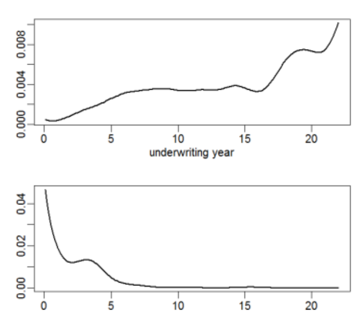


Illustration. Prediction of the outstanding number of claims

We consider two data sets provided by a major insurer on a monthly base. The data are the number of reported claims, and it has been arranged in a triangle where the **development period** corresponds with the **reporting period**.

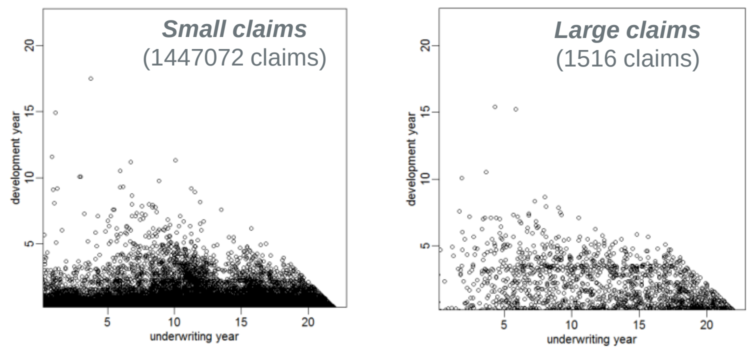


Illustration: comparing four methods to solve the problem

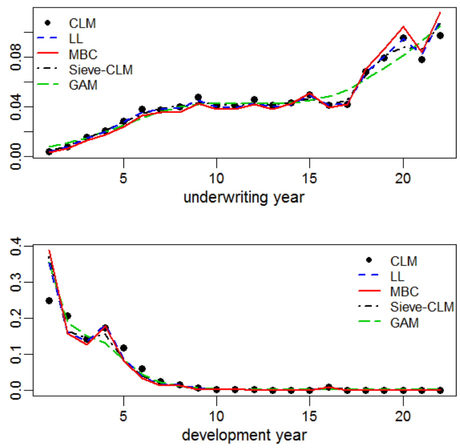
- ✓ Classical **Chain Ladder** from a yearly run-off triangle.
- ✓ Two versions of **Continuous Chain Ladder** with two kernel unstructured density estimators: local linear (LL) estimator and multiplicative bias corrected (MBC) estimator.
- ✓ **GAM method of England and Verrall (2001)**: starting from the histogram the time effects are estimated using smoothing splines

$$\log(N_{ij}) = s_{\theta_i}(i) + s_{\theta_j}(j) + \varepsilon_{ij}$$

- ✓ A **sieve method on monthly chain ladder parameters**: providing smoothed chain ladder time effects using local regression.

Illustration: results for large claims

Estimated time effects

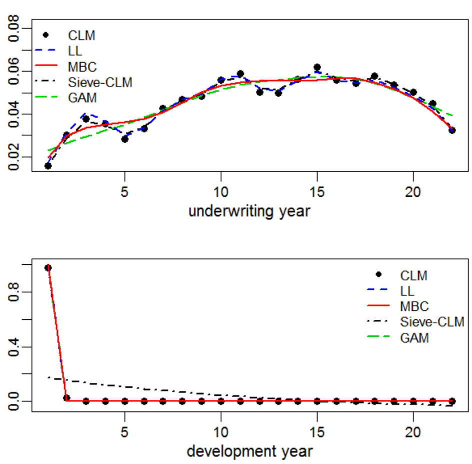


Predictions (future calendar years)

Future	CLM	LL	MBC	Sieve-CLM	GAM
1	118	132	104	133	127
2	86	93	71	93	88
3	65	72	54	70	64
4	38	43	32	43	40
5	19	21	15	21	22
6	9	10	7	10	11
7	5	6	4	6	6
8	2	3	2	3	3
9	1	1	1	1	2
10	1	1	1	1	1
11	1	1	1	1	1
12	1	1	1	1	1
13	1	1	1	1	0
14	1	1	1	1	0
15	1	1	1	1	0
16	0	1	0	1	0
17	0	0	0	0	0
Total	350	390	295	388	367

Illustration: results for small claims

Estimated time effects



Predictions (future calendar years)

Future	CLM	LL	MBC	Sieve-CLM	GAM
1	1,312	1,105	1,065	80,418	1,303
2	98	95	92	65,584	46
3	42	40	37	52,615	33
4	19	18	16	40,921	22
5	9	9	8	30,262	13
6	6	6	5	20,956	7
7	4	4	3	13,167	4
8	2	2	2	6,717	3
9	2	2	1	1,433	2
10	1	1	1	-2,583	1
11	1	1	1	-5,575	1
12	1	1	1	-8,009	0
13	1	0	1	-9,745	0
14	1	0	1	-10,524	0
15	0	0	0	-10,443	0
16	0	0	0	-9,920	0
17	0	0	0	-8,933	0
Total	1,501	1,284	1,235	228,145	1,435

Illustration: testing results against experience

The validation strategy:

1. Cut c=1,2,…,5 diagonals (years) from the observed triangle.
2. Apply the four estimation methods.
3. Compare forecasts and actual values.

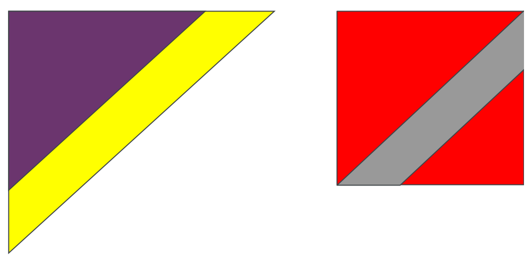
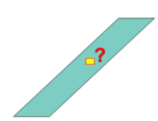


Illustration: testing results against experience

Three possible objectives:

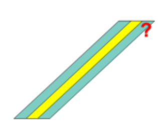
1. Predictions of the individual cells

$$err_1^c = \frac{1}{\#\{(i,j) \in \widetilde{\mathcal{J}_c}\}} \sum_{(i,j)\in\widetilde{\mathcal{J}}_c} (\widehat{N}_{ij} - N_{ij})^2$$



2. Predictions by calendar years

$$err_2^c = \frac{1}{c}\sum_{k=1}^{c}(\widehat{D}_{k;c} - D_{k;c})^2$$



3. The prediction of the overall total

$$err_3^c = |\widehat{R}_c - R_c|$$

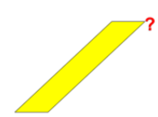


Illustration: testing results against experience

Large claims

Objective	c	LL	MBC	Sieve-CLM	GAM
Cells	1	1.11	0.82	1.23	1.06
	2	1.02	0.59	1.22	1.03
	3	1.11	0.84	1.15	0.76
	4	1.13	0.94	1.33	0.86
	5	1.03	0.99	0.88	0.79
Calendar	1	1.09	0.34	1.73	1.72
	2	1.05	0.56	1.45	1.02
	3	1.12	0.66	1.76	0.93
	4	1.42	0.82	2.57	0.96
	5	1.32	0.85	0.89	1.47
Total	1	1.09	0.34	1.73	1.72
	2	1.05	0.54	1.46	1.03
	3	1.13	0.30	1.95	1.00
	4	1.60	0.26	3.07	1.04
	5	1.32	0.21	0.89	1.46

Small claims

Objective	c	LL	MBC	GAM
Cells	1	0.77	0.43	0.81
	2	0.83	0.54	0.77
	3	0.67	0.50	0.71
	4	0.75	0.49	0.67
	5	0.70	0.96	0.53
Calendar	1	0.78	0.46	0.69
	2	0.85	0.58	0.69
	3	0.65	0.46	0.53
	4	0.76	0.52	0.61
	5	0.70	0.91	0.40
Total	1	0.78	0.46	0.69
	2	0.86	0.60	0.62
	3	0.69	0.52	0.53
	4	0.77	0.52	0.52
	5	0.76	0.27	0.15

Relative errors with respect to the classical chain ladder method (values lower than 1 indicate an improvement on chain ladder)

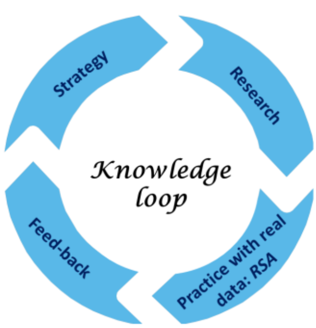
Summary

- We have established a link between classical chain ladder and modern mathematical statistics.
- **The interpretation of classical chain ladder as a structured histogram estimator** has a number of immediate implications for further developments.
- **"Continuous Chain Ladder"** is **the natural kernel smoother improving the histogram of classical chain ladder**.

Conclusion

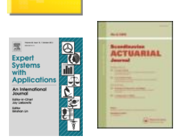
Remember your (continuous) Chain Ladder when going granular

Where to go from here?



2010 Including Count Data in Claims Reserving
2011 Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers
2012 Double Chain Ladder

2012 Statistical modelling and forecasting in Non-life insurance
2013 Double Chain Ladder and Bornhuetter-Ferguson



2013 Double Chain Ladder, Claims Development Inflation and Zero Claims
2013 Continuous Chain Ladder

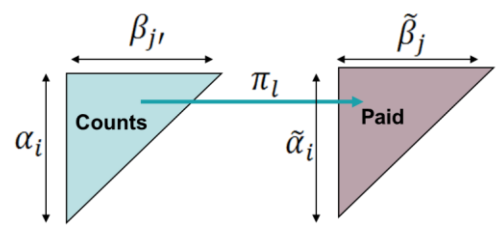
Continuous versions doing mathematical statistical theory on optimizing reserving type of structured models

Granular data for a better description of the distribution

1. Just Continuous Chain Ladder as well as classical chain ladder could be used to provide the full cash-flow: Poisson approximation.
2. But with payments the Poisson assumption is not suitable: a description of the underlying dependencies is required...
3. Our proposal: **Continuous Double Chain Ladder**

CCL + DCL = CDCL

Double Chain Ladder model (Martínez-Miranda, Nielsen and Verrall 2012)



The Double Chain Ladder Model

What is Double Chain Ladder?

A firm **statistical model** which **breaks down the chain ladder estimates into individual components.**

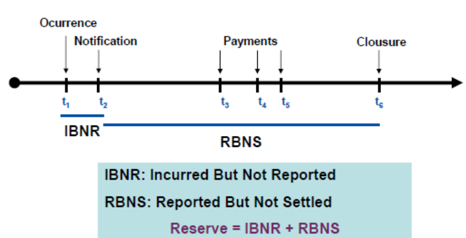
Why?

✓**Connection with classical reserving (tacit knowledge)**

✓Intrinsic **tail** estimation

✓**RBNS** and **IBNR** claims

✓The distribution: **full cash-flow**



What is required? It works on **run-off triangles** (**adding expert knowledge** if available).

Describing the model

Parameters involved in the model:

Ultimate claim numbers: α_i

Reporting delay: $\beta_{j'}$

Settlement delay: π_l

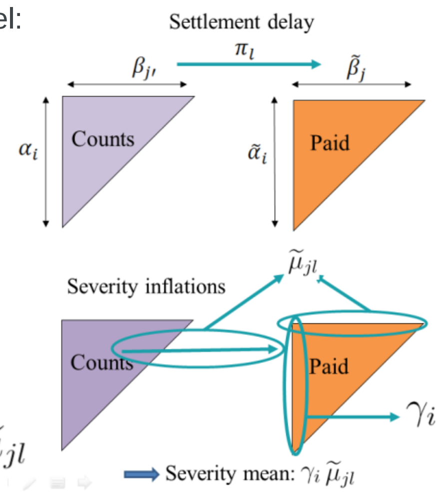
Development delay: $\tilde{\beta}_j$

Ultimate payment numbers: $\tilde{\alpha}_i$

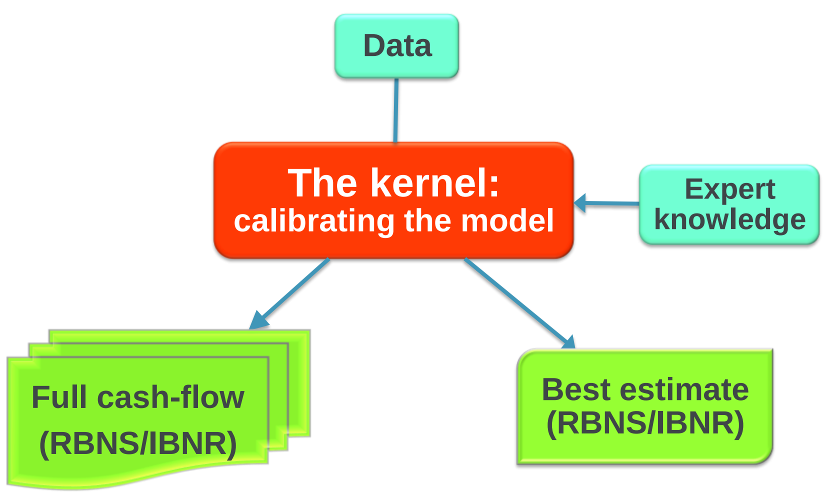
Severity:

✓ underwriting inflation: γ_i

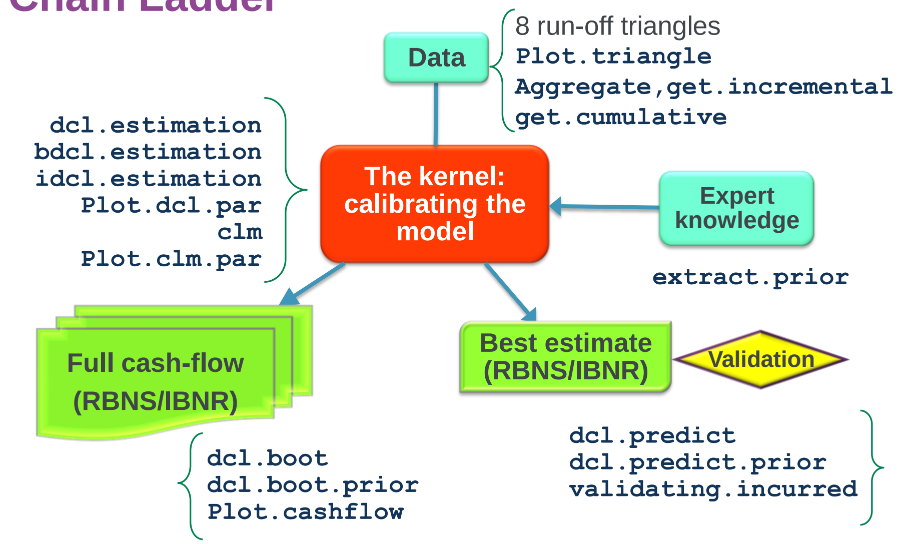
✓ delay mean dependencies: $\tilde{\mu}_{jl}$



The Double Chain Ladder in practice



DCL a R-Package implementing Double Chain Ladder



It is free open-source software, please try it!

- We look for a **wide audience** (academics, practitioners, students).
- Your feedback is very valuable…
- Reference papers+package+documentation+examples are available at:

  http://www.cassknowledge.com/research/article/double-chain-ladder-cass-knowledge

- Variations and extensions are expected to come soon from the knowledge loop.

Other references in the slides

- Antonio, K. and Plat, H.J. (2013) Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*. DOI: 10.1080/03461238.2012.755938
- England, P.D. and Verrall, R.J. (2001) A flexible framework for stochastic claims reserving. *Proceedings of the Casualty Actuarial Society* LXXXVIII, 1-38.
- Taylor, G. McGuire, G. and Sullivan, J. (2008) Individual claim loss reserving conditioned by case estimates. *Annals of Actuarial Science*, 3, 215-256.
- Zhao, X. and Zhou, X. (2010) Applying copula models to individual claim loss reserving methods. *Insurance: Mathematics and Economics*, 46, 290-299.

Cass Business School
CITY UNIVERSITY LONDON

Email: Maria.Miranda.1@city.ac.uk
Homepage: www.cass.city.ac.uk/experts/M.Miranda