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
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


---

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.

### The Claims Run-off Triangle

- **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.

It is the simplest nonparametric density estimator.

That we can learn...

2. The multiplicative structure

The outstanding numbers are predicted assuming the simple mean structure:

\[ \mathbb{E}[N_{ij}] = \alpha_i \beta_j \]
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. 
    \[ f(x, y) = f_1(x)f_2(y) \]
  - 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)

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

Illustration: results for small claims

Estimated time effects

Predictions (future calendar years)

<table>
<thead>
<tr>
<th>Future</th>
<th>CLM</th>
<th>LL</th>
<th>MBC</th>
<th>Sieve-CLM</th>
<th>GAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.312</td>
<td>1,105</td>
<td>1,005</td>
<td>80,418</td>
<td>1,303</td>
</tr>
<tr>
<td>2</td>
<td>98</td>
<td>95</td>
<td>92</td>
<td>65,584</td>
<td>46</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>40</td>
<td>37</td>
<td>52,615</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>49</td>
<td>18</td>
<td>16</td>
<td>40,921</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>30,262</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>20,056</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>13,167</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6,717</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1,433</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-2,583</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-5,575</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-8,009</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>-9,745</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>-10,524</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-10,443</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-9,920</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-8,933</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>1,501</td>
<td>1,284</td>
<td>1,235</td>
<td>228,145</td>
<td>1,435</td>
</tr>
</tbody>
</table>
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
   \[ err_1^c = \frac{1}{\# \{ (i, j) \in J_c \}} \sum_{(i, j) \in J_c} (\hat{N}_{ij} - N_{ij})^2 \]
2. Predictions by calendar years
   \[ err_2^c = \frac{1}{c} \sum_{k=1}^{c} (\hat{D}_{k,c} - D_{k,c})^2 \]
3. The prediction of the overall total
   \[ err_3^c = |\hat{R}_c - R_c| \]
Illustration: testing results against experience

<table>
<thead>
<tr>
<th>Large claims</th>
<th>Small claims</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective</strong></td>
<td><strong>c</strong></td>
</tr>
<tr>
<td><strong>Cells</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.11</td>
</tr>
<tr>
<td>2</td>
<td>1.02</td>
</tr>
<tr>
<td>3</td>
<td>1.11</td>
</tr>
<tr>
<td>4</td>
<td>1.13</td>
</tr>
<tr>
<td>5</td>
<td>1.03</td>
</tr>
<tr>
<td><strong>Calendar</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.09</td>
</tr>
<tr>
<td>2</td>
<td>1.05</td>
</tr>
<tr>
<td>3</td>
<td>1.12</td>
</tr>
<tr>
<td>4</td>
<td>1.42</td>
</tr>
<tr>
<td>5</td>
<td>1.32</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.09</td>
</tr>
<tr>
<td>2</td>
<td>1.05</td>
</tr>
<tr>
<td>3</td>
<td>1.13</td>
</tr>
<tr>
<td>4</td>
<td>1.60</td>
</tr>
<tr>
<td>5</td>
<td>1.32</td>
</tr>
</tbody>
</table>

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
   \[ \text{CCL} + \text{DCL} = \text{CDCL} \]

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:** $\alpha_i$
- **Reporting delay:** $\beta_{ji}$
- **Settlement delay:** $\pi_l$
- **Development delay:** $\bar{\beta}_j$
- **Ultimate payment numbers:** $\bar{\alpha}_i$

**Severity:**
- ✓ **underwriting inflation:** $\gamma_i$
- ✓ **delay mean dependencies:** $\tilde{\mu}_{ji}$

---

**The Double Chain Ladder in practice**

- **Data**
- **The kernel:** calibrating the model
- **Expert knowledge**
- **Full cash-flow (RBNS/IBNR)**
- **Best estimate (RBNS/IBNR)**
DCL a R-Package implementing Double Chain Ladder

- dcl.estimation
- bdcl.estimation
- idcl.estimation
- Plot.dcl.par
- clm
- Plot.clm.par

The kernel: calibrating the model

- dcl.boot
- dcl.boot.prior
- Plot.cashflow
- dcl.predict
- dcl.predict.prior
- validating.incurred

Data

8 run-off triangles
Plot.triangle
Aggregate, get.incremental,
get.cumulative

Expert knowledge

extract.prior

Full cash-flow (RBNS/IBNR)

Best estimate (RBNS/IBNR)

Validation

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


Thank you!!!

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