Antitrust Notice

- The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

- Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding, expressed or implied, that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

- It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.

Bayesian Loss Development for Real People
David R. Clark, FCAS
Munich Re America – March 2021

Agenda

1. Business Context: Why we are doing this?
2. Basic Model: Combining Triangles
3. Extended Model: What about the tail?
4. Next Steps: Setting the Parameters
Business Context: Why are we doing this?

Goal is to improve estimation of loss development patterns for individual clients. Including benchmark patterns helps stabilize this estimation.

- Avoid two extremes of relying solely on client data (variance) and using benchmark for everyone (bias).

Basic Model

- Use conjugate distributions for simple implementation
  [we are skipping the math for today]
- Related to Chain Ladder method and applies to each age-to-age (ATA) factor

We will start with a blending example to build intuition.

Combining Triangles: Possible Even for Different Sizes

<table>
<thead>
<tr>
<th>Company A</th>
<th>Triangle is complete for old years, but we did not get latest diagonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>22,010 44,059 58,548 68,746 76,157 81,650 85,778</td>
</tr>
<tr>
<td>2012</td>
<td>22,010 44,059 58,548 68,746 76,157 81,650</td>
</tr>
<tr>
<td>2013</td>
<td>22,010 44,059 58,548 68,746 76,157</td>
</tr>
<tr>
<td>2014</td>
<td>22,010 44,059 58,548</td>
</tr>
<tr>
<td>2015</td>
<td>22,010 44,059 58,548</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Company B</th>
<th>Triangle is complete for old years, but not many evaluations on some years</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>81,011 84,250 86,778 89,000</td>
</tr>
<tr>
<td>2012</td>
<td>75,010 81,011 84,250 86,778</td>
</tr>
<tr>
<td>2013</td>
<td>66,973 75,010 81,011 84,250</td>
</tr>
<tr>
<td>2014</td>
<td>52,380 66,973 75,010 81,011</td>
</tr>
<tr>
<td>2015</td>
<td>33,015 52,380 66,973 75,010</td>
</tr>
</tbody>
</table>

ATA

| Col. 1 | 2.002 1.329 1.174 1.108 1.072 1.051 1.040 |
| Col. 2 | 1.587 1.279 1.120 1.080 1.040 1.030 1.026 |

COMBINED:

| Col. 1 | 231,105 377,435 435,111 431,268 395,347 250,150 86,778 |
| Col. 2 | 421,494 493,659 500,014 471,504 416,050 259,334 89,000 |

ATA

| 1.824 | 1.308 | 1.149 | 1.093 | 1.052 | 1.037 | 1.026 |

Numbers for illustration only
An Industry consolidated triangle may be the source of a benchmark pattern. But it does not need to be a full triangle as we have seen: it can be a weighted average from selections for each company.
Basic Model

In the basic model, the actual client data is smoothed by supplementing it with "pseudo data" from the benchmark, which acts as ballast.

This is equivalent to a Bayesian credibility formula using a conjugate prior. Two alternative derivations can be found in the two papers below.


The concept of pseudo data:

"Conjugate priors... have the desirable feature that prior information can be viewed as 'fictitious sample information' in that it is combined with the sample in exactly the same way that additional sample information would be combined. "The only difference is that the prior information is 'observed' in the mind of the researcher, not in the real world." - Bayesian Econometric Methods; Koop, Poirier & Tobias

PS: This is also what is done in ISO state advisory loss cost circulars.

Extended Model

A limitation of the Basic Model:

• Each age-to-age (ATA) factor, or column of the triangle, is treated independently
• This means that we would use the benchmark "tail" even if ATA factors from the client were consistently less than the benchmark.

Shi & Hartman address this by introducing a correlation structure in the model.

An alternative is to first "nudge" the benchmark before applying the Basic Model.
Actuaries select a benchmark development pattern (growth curve) for representative business segments. The selected benchmarks may be based on data from various sources and judgmentally smoothed.

As a simplified method for setting a range around the benchmark, we can start by setting “fast” and “slow” patterns. In this derivation, the benchmark will always be the exact midpoint between the fast and slow patterns.

We will assume that each client has a development pattern that is a weighted average of the “Fast” and “Slow” patterns around the benchmark. If the weight for company $j$ is exactly 50%/50%, then the benchmark pattern is used. To start, we will constrain the weights to be between 0% and 100%. The parameter $p$ is assumed to be a random variable from a beta distribution.

$$G(\text{age}|j) = p_j \cdot G_{\text{Fast}}(\text{age}) + (1 - p_j) \cdot G_{\text{Slow}}(\text{age})$$

$$0 \leq p_j \leq 1$$
The form of the model is a linear combination of two “basis functions” (fast and slow).

- A simple form of Regression Spline

The weight parameter $p$ can be estimated various ways, along with its standard error. Ideally, we damp this parameter close to .500 based on assigning a prior distribution (e.g., a Beta Distribution). If we assume a prior uniformly distributed between 0 and 1, then the variance of hypothetical means = 1/12.

How well does this work? Example here uses Products Liability payment patterns from Schedule P. The fast and slow patterns reasonably bracket the range of patterns across companies.

Next Steps: Selecting the Prior Distribution

How do we set the spread around the benchmark parameter?

Subjective Bayes:
- Business expertise selects the range of possible values
  For example: how much faster or slower than average can a company settle its claims?

"Subjunctive Bayes" (Stephen Senn):
- Set prior parameters to get the credibility-weighted result that makes sense

Empirical Bayes:
- How much actual spread is there among the companies (or states)?
- Data Scientists call this cross validation
Next Steps: Selecting the Prior Distribution

**Good News!**

Even if we cannot estimate the optimal credibility perfectly, we can select a value that produces a blended estimate that is an improvement on either estimator alone.

We are just looking for a sensible weighted average.

---

Thank you!

---

Selected References

[https://www.casact.org/pubs/forum/16sforum/Clark.pdf](https://www.casact.org/pubs/forum/16sforum/Clark.pdf)


Racine, J.S., “A Primer on Regression Splines”. CRAN library
[https://cran.r-project.org/web/packages/splines/spline_primer.pdf](https://cran.r-project.org/web/packages/splines/spline_primer.pdf)

Senn, S. “Two Cheers for $P$-Values?” Journal of Epidemiology and Biostatistics (2001)

Strategies for Working with Loss Development Factors

Uri Korn, FCAS
Ratemaking, Product, and Modeling Seminar
March 16, 2021

Blending LDFs

- LDFs are volatile
- To reduce LDF volatility, leverage 2 related pieces of information
  1. Adjacent LDFs – fit a curve
  2. Related LDFs – blend with credibility

Diagram: Graphs showing data trends and blending of LDFs

The LDF Ninja
Part 1) LDF Curves

Inverse Power Curve
(Sherman 1984)

\[ \log(LDF - 1) = A + B \times \log(\text{age}) \]

- Easy to implement
- But often poor fit to the data

* Using age instead of \( 1/\text{age} \), since the regression equations are equivalent.
  Also, ignoring the \( c \) parameter

IPOC Fit

Trouble making the “turn”

Starts too high

Tail too high
Problem with the IPOC

- Weights (assumed variances) aren’t accurate
  - Tail LDFs are more volatile
  - High volatility at initial ages due to lack of volume (longer tailed lines)

IN MY DAY, WE NEVER USED CURVES

Double IPOC (DIPOC)

- Modify the weights of the Inverse Power Curve
- Weights are a function of age and loss volume
  - Use a weighted Gamma regression instead
- Fit a curve to the variance/weights by age
  1. Fit simultaneously with LDFs
  2. Or calculate directly from triangle beforehand
DIPOC Fit

Improved, but still flawed ...

Smoothed IPOC (SMIPOC)

- Double IPOC with regression splines
  - (Concept borrowed from England & Verall 2001)
- Adds flexibility to the curve
- Can still be done in Excel

How do Splines work?

- Performs a special transformation on a variable (such as age)
- Runs a regular regression on the new variables instead
- Enables a better fit to the data at the cost of additional variables
Inverse Power Distributions

- Model the percent completion distribution instead
  - Idea inspired by Clark 2003
- Use a similar inverse power function to define the CDF (and likelihood)

Smoothed Inverse Power Distribution (SMIPOD)

- Similar to before, use regression splines
- Better fit to data
Part 2) Credibility

Credibility

- Only 3 parameters!

- Best answer to the trade off between:
  - Fewer stable heterogeneous segments
  - Many volatile homogenous segments
Bayesian Credibility

- You find a toothbrush on the subway!
- It looks semi-clean!
- Should you use it?

![Graph showing likelihood, prior, and posterior distributions]

**Likelihood:** The toothbrush looks clean.

**Prior:** Most things on the subway are not clean.

**Posterior:** Taking all information into account, you should probably not use the toothbrush (it's a good thing you're an actuary).

Bayesian Credibility on a Curve or Distribution

- Performs credibility weighting on the parameters simultaneously while fitting the curve/distribution

Implementing Bayesian Credibility in Excel

- Maximum Likelihood Estimation (e.g. via Solver) returns the mode of the distribution
- Same as the mean for the Normal distribution

![Graph showing likelihood, prior, and posterior distributions]
Implementing Bayesian Credibility in Excel

Variance of the data is implied

Variance of the prior needs to be provided though

The theory is fine but...

• [This only happens when credibility weighting multiple parameters]

Fixing the Credibility

• The prior should be calculated on the curve parameters
• What are the parameters?
  • Intercept & Slope
• But what if we...
  • LDF1 & LDF2 → Intercept & Slope → Entire LDF Curve
• Calculate the prior on the two predicted LDFs (even if the inversion wasn’t performed)

Note: In practice, using log(LDF – 1) works better
Calculating the Prior Variance

- How do we calculate this prior variance?
  - (Equivalent to Between Variance and Z)

- Options:
  - Build a Bayesian model
  - Holdout/Cross validation
  - Buhlmann-Straub

Buhlmann-Straub

- Remember: We are using LDF parameters

- Use the Buhlmann-Straub formulas on the actual LDFs as an approximation

- Fit a curve by age to smooth them out
Part 3) Excel Template
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
<td>1.8</td>
<td>1.9</td>
<td>2.0</td>
<td>2.1</td>
<td>2.2</td>
</tr>
<tr>
<td>2</td>
<td>2.2</td>
<td>2.3</td>
<td>2.4</td>
<td>2.5</td>
<td>2.6</td>
<td>2.7</td>
<td>2.8</td>
<td>2.9</td>
<td>3.0</td>
<td>3.1</td>
<td>3.2</td>
</tr>
<tr>
<td>3</td>
<td>3.3</td>
<td>3.4</td>
<td>3.5</td>
<td>3.6</td>
<td>3.7</td>
<td>3.8</td>
<td>3.9</td>
<td>4.0</td>
<td>4.1</td>
<td>4.2</td>
<td>4.3</td>
</tr>
</tbody>
</table>

**Table 1: Description of Data**

- **Column 1**: This column represents the first data series.
- **Column 2**: This column represents the second data series.
- **Column 3**: This column represents the third data series.
- **Column 4**: This column represents the fourth data series.
- **Column 5**: This column represents the fifth data series.
- **Column 6**: This column represents the sixth data series.
- **Column 7**: This column represents the seventh data series.
- **Column 8**: This column represents the eighth data series.
- **Column 9**: This column represents the ninth data series.
- **Column 10**: This column represents the tenth data series.
- **Column 11**: This column represents the eleventh data series.
Thank You!

For more details, refer to:
http://www.variancejournal.org/issues/11-01-02/95.pdf