The Case of the Credulous Actuary: Rediscovering the Importance of Judgment

Outline

- Introduction  
  Mark Graham
- Analyses and results  
  Alex Glencross
- Recommendations and issues  
  Mark Graham
- Discussion

Introduction

- Why?
- How?
Prior hypothesis

Long held beliefs that:

- [Actuaries] consistently give too much credibility to data
- Necessary consequence is tendency to over-fit and, hence, underestimation of uncertainty

Methodology (1)

- Pick a “true” distribution, e.g. LogNormal ($\mu, \sigma$)
- Sample a sequence of 100 data points from that distribution
- Calculate the sample mean, sample standard deviation and corresponding 99.5th %ile as each data point is added to the sequence, e.g. after 10 data points, after 100 data points
- Repeat through 10,000 simulations

Methodology (2)

- End up with 10,000 simulations of 100 sequential data points from the same distribution
- Determine distribution of sample means, standard deviations and 99.5th %iles at each data point in the sequence
- Provides real insight into how much information is really contained in the data
Analyses and results

- Equity returns
- Large loss experience
  - Frequency
  - Severity
  - Aggregate claims

Recent example (or something like it)

- Eleven years of equity returns
- Mean: 5.3%
- StDev: 15.0%
- Would you define a distribution with these parameters?

<table>
<thead>
<tr>
<th>Year</th>
<th>Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>24.7%</td>
</tr>
<tr>
<td>1998</td>
<td>14.6%</td>
</tr>
<tr>
<td>1999</td>
<td>17.8%</td>
</tr>
<tr>
<td>2000</td>
<td>10.2%</td>
</tr>
<tr>
<td>2001</td>
<td>-14.3%</td>
</tr>
<tr>
<td>2002</td>
<td>-24.9%</td>
</tr>
<tr>
<td>2003</td>
<td>13.6%</td>
</tr>
<tr>
<td>2004</td>
<td>7.5%</td>
</tr>
<tr>
<td>2005</td>
<td>16.1%</td>
</tr>
<tr>
<td>2006</td>
<td>10.7%</td>
</tr>
<tr>
<td>2007</td>
<td>3.8%</td>
</tr>
</tbody>
</table>

Basic Model: methodology

- First define our ‘true’ distribution
  - Equity-type asset returns
  - Mean return: 10%, Standard Deviation: 16%
  - 0.5th Percentile: -49%
- Repeatedly sample ten years of observations
- Look at errors between ‘true’ values and those implied by each approximation
  - Mean, Standard Deviation, 0.5th Percentile
Basic Model: Points to note

- Distribution of standard deviation is skewed
  - P( Error < 0 ) = 0.62
  - So there’s a decent chance of systematic under-estimation of standard deviations
- However, errors are not significantly greater than for estimating the mean
- Skewed distribution of 0.5th percentile
  - P( Understate extreme value) = 0.61

Equity Returns Example
Distribution of Parameter Errors
Equity Returns Example
Distribution of Parameter Errors

100 Years Data

-25.0% -20.0% -15.0% -10.0% -5.0% 0.0% 5.0% 10.0% 15.0% 20.0% 25.0%

Mean StDev 0.5th Percentile

Frequency / Severity Example

- Another simple model
  - Poisson Frequency / LogNormal Severity
  - Mean Frequency 5
  - Mean Severity 100
  - StDev Severity 100
- Sample 10 years of data and estimate parameters and percentile points

Frequency / Severity Example
Distribution of Parameter Errors (Frequency)

10 Years Data

-4 -3 -2 -1 0 1 2 3 4

Mean 25.5th Percentile
Average Frequency = 5
Frequency / Severity Example Distribution of Parameter Errors (Frequency)

- No need for StDev
  - Square root of mean
  - Estimate of mean is unbiased and unskewed
- Not so for estimate of 99.5\textsuperscript{th} percentile
  - Biased: average error is -0.4
  - Not particularly skewed

Frequency / Severity Example Distribution of Parameter Errors (Severity)

- No need for StDev
  - Square root of mean
  - Estimate of mean is unbiased and unskewed
- Not so for estimate of 99.5\textsuperscript{th} percentile
  - Biased: average error is -0.4
  - Not particularly skewed
Frequency / Severity Example Distribution of Parameter Errors (Severity)

- Mean is still unbiased and unskewed
- StDev is unbiased and skewed
  - Skewness = 0.6
- 99.5th percentile is biased and skewed
  - 'True' value = 603
  - Average error = -33.4
  - Skewness = 2.1

Average Severity = 100
Technical Problems

- Equity Return example used a single distribution
  - Invert distribution to find percentile points
- Aggregate claims is more difficult
  - No closed form that we can invert
- Simulation approach used instead

Technical Problems (2)

- So for each of the 10,000 simulations…
  - do another 10,000 simulations
  - estimate the 99.5th percentile
- Quickly run out of processing power
  - Managed it on 4,000 simulations before it crashed…

Frequency / Severity Example
Distribution of Parameter Errors (Aggregate Claims)

10 Years Data

- Mean
- StDev
- 99.5th Percentile

Average Aggregate Claims = 500
Technical Problems (3)

- To avoid simulation error…
  - Approximate aggregate distribution by a shifted Gamma distribution
  - Shift + Gamma( alpha, beta )
  - Can solve shift, alpha and beta to match the first three cumulants of the aggregate claims distribution
- Provides very good fit in the tail
  - Removes need for extra dimension of simulations

Frequency / Severity Example
Distribution of Parameter Errors
(Aggregate Claims)

10 Years Data

Mean StDev 99.5th Percentile
Average Aggregate Claims = 500

Frequency / Severity Example
Distribution of Parameter Errors
(Aggregate Claims)

- Mean unbiased, but slightly skewed
  - Skewness = 0.4
- StDev marginally biased, and more skewed
  - Average error = -7
  - Skewness = 1.7
- 99.5th percentile is biased and skewed
  - Average error = -25
  - Skewness = 3.6
Frequency / Severity Example

- True mean of aggregate distribution is 500
  - Estimates of the mean carry a range of errors of approximately 500
  - Estimates of 99.5th percentile carry a range of approximately 1500
- Error in each estimate will be correlated
  - Underestimate capital and overestimate profit

Frequency / Severity Example
Return on Capital

- Suppose premium received = 600
- ‘True’ statistics
  - mean aggregate claims = 500
  - 99.5th percentile aggregate claims = 1575
  - mean profit = 100
  - Capital = 1075
- ‘True’ mean Return on Capital = 9.3%

Frequency / Severity Example
Distribution of Parameter Errors
Return on Capital

10 Years Data

- True Mean Return on Capital = 9.3%
**Frequency / Severity Example**

**Distribution of Parameter Errors**

**Return on Capital**

- Average error = 6.0%
- Skewness = 93

---

**Issues**

- Understates problem
- Why over rely on data vs. judgment?
- Poorly understood
- How to respond?

---

**Understates problem**

- Investigation assumed known distribution
  - Actual distribution unknown and unknowable
  - Much modelling based on best fit distribution, rather than best fit parameters for a given distribution
- Investigation assumed stationary distribution
  - Certainly non-stationary, in practice
Why over rely on data vs. judgment?

- Actuarial nature
- Lack of understanding
  - Of statistical limitations of data
  - Of underlying exposures
- External influences
  - Fear of being sued
  - Fair value accounting
  - Financial mathematics

Poorly understood

- Little or no actuarial literature
- FRC Discussion Paper “Promoting Actuarial Quality” makes no mention of understanding the credibility of data as a driver of actuarial quality
- 10 sigma events
- EC/CEIOPS expectations of data in Solvency II

How to respond? (1)

- Invest time in understanding underlying exposures
- Invest time in understanding the limitations of your data
- Use judgment, informed by data, to parameterise
- Use emerging data to test parameters
- Test over varying time periods, due to non-stationarity
- Do not expect every parameter to pass every test (1:20 fail at 95% confidence interval)
How to respond? (2)

- Your parameters will be wrong but…
- … judgment can ensure consistency between related exposures (different asset classes, different business types) …
- … and parameters can be adjusted to reflect changes in the underlying exposure much earlier than would be the case if you wait for the data to emerge