GENERAL INSURANCE PRICING SEMINAR

13 JUNE 2008, LONDON

Application of predictive modelling in commercial lines
Ryan Warren
Watson Wyatt Limited

Agenda

- The rating process
- Generalised linear models
- A predictive modelling case study
- Investigating uncertainty
The rating process

Integrated business processes

- Rate monitoring
- Allocating cost of capital
  - Line of business
  - Policy
- Communication
  - Reserving
  - Underwriting
  - Capital modelling
Processes and controls

- Managing operational risk is increasingly seen as an important aspect of the business
- Rigorous controls are the norm for claims
- Pricing has had less attention but arguably more important to the profitability of the business
- How do you protect your business against:
  - Loss of key staff
  - Accusations of unfairness in pricing
  - Errors in key calculations

Processes and controls

- Key is to have:
  - Clear and persistent records of analysis
  - Documentation of decisions
  - Standard methods to allow task sharing
- To be effective these should be:
  - Automatic and embedded within systems
  - Universal: Actuaries, Underwriters, Claims Managers
  - Regularly reviewed to check compliance
  - Not too onerous
Agenda

- The rating process
- Generalised linear models
- A predictive modelling case study
- Investigating uncertainty

Modelling in non-traditional areas

- Predictive modelling widely used:
  - Excel, @Risk, ...
- GLMs:
  - Commercial property
  - Commercial motor and Fleet
  - Marine
  - Mortality/Morbidity
  - Accident and Health
  - Aviation
  - D&O
- Used for:
  - Underwriting
  - Fraud detection
  - Marketing
Modelling in non-traditional areas

- Types of models used:
  - GLM
  - Clustering (eg CHAID)
  - Simulation and Bayesian models (MCMC)
- Complex models can be blended with simpler models where appropriate

Generalised linear models

\[ E[Y] = \mu = g^{-1}(X.\beta + \xi) \]
\[ \text{Var}[Y] = \phi.V(\mu) / \omega \]

- Consider all factors simultaneously
- Allow for nature of random process
- Robust and transparent
- EU industry standard for personal lines
Applying GLMs in commercial lines

- Market databases vs own claims experience
- Standard rates for risk considered
- Adjust using typical experience rating methods
- Combine with / consider alongside other "traditional" methods
Agenda

- The rating process
- Generalised linear models
- A predictive modelling case study
- Investigating uncertainty

Marine liability example - objective

Model

- Age
- Flag
- Vessel
- Tonnage
- Excess
- NCD

Expected cost of claims
Modelling the cost of claims

Car: \( \text{Freq} \times \text{Amt} = \text{Cost 1} \)
Col: \( \text{Freq} \times \text{Amt} = \text{Cost 2} \)
Pax: \( \text{Freq} \times \text{Amt} = \text{Cost 3} \)
Pol: \( \text{Freq} \times \text{Amt} = \text{Cost 4} \)
Oth: \( \text{Freq} \times \text{Amt} = \text{Cost 5} \)

Some marine examples

Marine Cargo numbers model

Vessel age

P \( \text{value} = 0.0\% \)
Rank 43

Exposure (years)

The Actuarial Profession
making financial sense of the future
Some marine examples

Marine
Cargo numbers model

-5 1%
5%
-10%
-16%
28% 25%
10%
-19%
0%
-12%
7% 6%
20% 41% 28%

Flag state

Log of multiplier

0 2000 4000 6000 8000 10000 12000 14000 16000 18000

Japan Sweden Greece England Norway Cuba Denmark Germany Others China Liberia Panama Cyprus Bahamas Korea

Approx 95% confidence interval Unsmoothed estimate Smoothed estimate P value = 0.0% Rank 2/2

The Actuarial Profession
making financial sense of the future

Some marine examples

Marine
Pollution numbers model

-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 2

Vessel type

Bulker Chemical (clean) Chemical (dirty) Comballiners Container carriers Container carriag Carriage General cargo Passenger Reefer Supply Support Supply (clean) Trawler (dirty)

P value < 0.1%
Rank 2/2

The Actuarial Profession
making financial sense of the future
Some marine examples

Dealing with large claims
Predictive power of models

Validation on 20% subset - frequency analysis - cargo

Total over factor levels - Where Random number >=0.8

Expected frequency

Exposure

Actual frequency

Agenda

- The rating process
- Generalised linear models
- A predictive modelling case study
- Investigating uncertainty
Investigating uncertainty

- Simulation from GLM possible
- Monte Carlo type simulation using current or desired portfolio
- Allows for stochastic features
- Combine other methods or models

Generalised linear models

<table>
<thead>
<tr>
<th>Linear Models</th>
<th>Generalised Linear Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[Y_i] = \mu_i = \Sigma X_{ij} \beta_j$</td>
<td>$E[Y_i] = \mu_i = g^{-1}(\Sigma X_{ij} \beta_j + \xi_i)$</td>
</tr>
<tr>
<td>$\text{Var}[Y_i] = \sigma^2$</td>
<td>$\text{Var}[Y_i] = \sigma^2 = \phi V(\mu_i)/\omega_i$</td>
</tr>
<tr>
<td>$Y_i \sim N(\mu_i, \sigma^2)$</td>
<td>$Y$ from a distribution from the exponential family</td>
</tr>
</tbody>
</table>
Typical GLM model forms

<table>
<thead>
<tr>
<th></th>
<th>Claim frequency</th>
<th>Claim number</th>
<th>Average claim amount</th>
<th>Probability (eg lapses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>( g(x) )</td>
<td>( \ln(x) )</td>
<td>( \ln(x) )</td>
<td>( \ln(x/(1-x)) )</td>
</tr>
<tr>
<td>Error</td>
<td>Poisson</td>
<td>Poisson</td>
<td>Gamma</td>
<td>Binomial</td>
</tr>
<tr>
<td>( \phi )</td>
<td>1</td>
<td>1</td>
<td>estimate</td>
<td>1</td>
</tr>
<tr>
<td>( V(x) )</td>
<td>( x )</td>
<td>( x )</td>
<td>( x^2 )</td>
<td>( x(1-x) )</td>
</tr>
<tr>
<td>( \omega )</td>
<td>exposure</td>
<td>1</td>
<td># claims</td>
<td>1</td>
</tr>
<tr>
<td>( \xi )</td>
<td>0</td>
<td>( \ln(\text{exposure}) )</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Interpreting the GLM

[Graph showing expected mean, variation about mean, and quite certain, less certain areas with P-value and rank information]
Simulating from the GLM

- To determine a sample number of claims, for each record first:
  - Calculate $\eta_i$ as $\sum X_i \beta_j$
  - Calculate $\sigma_i^2$ as $\sum \sum X_i \sigma_{jk} X_{ik}$
- Where
  - $\beta_j$ is the vector of parameter estimates
  - $\sigma_{jk}$ is the covariance matrix

Simulating from the GLM

- Simulate a value for the linear predictor $\eta_i^s$ from $N(\eta_i, \sigma_i^2)$
- Convert into Poisson mean $\lambda_i = \exp(\eta_i^s)$
- Sample number of claims from Poisson distribution
Simulating from the GLM

- To determine the sample total cost of claims, for each record:
  - For each record, calculate $\eta_i$ and $\sigma_i^2$ from the amounts model
  - For each sampled claim, simulate a value for the linear predictor $\eta^s_i$ from $N(\eta_i, \sigma_i^2)$
  - Convert into Gamma mean $\lambda_i = \exp(\eta^s_i)$
  - Sample each claim from Gamma distribution
  - Add all the sampled claims together

Some marine examples

![Projected total large claims cost for cargo claims](chart.png)
Conclusions

- GLMs perform well in non-traditional areas
- Results appear to be very predictive of future experience
- Fits with ideal general pricing process
- Can be combined with other methods
- Robust framework for assessing uncertainty

Contact details

Ryan Warren
Watson Wyatt Limited
Tel: 01737 284898
Email: ryan.warren@watsonwyatt.com
GENERAL INSURANCE PRICING SEMINAR

13 JUNE 2008, LONDON

Application of predictive modelling in commercial lines
Ryan Warren
Watson Wyatt Limited