Motivation

- Sophisticated insurers can have dozens of models...
- Aggregators are increasing our reliance on technical models
- A slick monitoring process delivers significant competitive advantage and improves risk management
- Pricing doesn’t get the recognition that it deserves!
- (Validation and monitoring are areas where Actuaries can work together)

Workshop agenda

- Introduction
- Tools of the trade
- Validation – (very briefly)
- Monitoring
  - Conversion
  - Claims
- Example model monitoring dashboards
- Discussion
Questions for delegates

- Are you comfortable with your current validation process?
- Are you comfortable with your current monitoring process?

Questions we discussed

- How can we monitor the performance of claims models and other models?
- Where can we add value?
  - quickly detect departures between modelled and actual results
  - migrate away from calendar based model refresh/rebuild cycles
  - make suggestions for "technical model MI"
  - risk management

Tools of the trade - Gains Curve

[Diagram showing a Gains Curve with percentages labeled: 20%, 55%, 100% of all sales and quotes (ranked highest to lowest).]
Validation

- Covered well in statistical literature
  - we’ll bring out points relating to the use of GLMs in personal lines
- Companies generally have processes to do this (unlike monitoring)?

Example 1 – AD frequency model
Example 2 k-fold cross validation to test for overfitting

Propensity/Elasticity Models

Background
- Zurich UK aggregator and telesales conversion data (2008)
- Production strength conversion/price elasticity model(s)
- We’re aiming to track “global”, as opposed to “segmental” model performance

Metrics

Desirable properties of model monitoring metrics
- Invariant to changes in volume
- Invariant to changes in average response rate
- Invariant to seasonality
- Suggests financial significance of departures?
Gains Curves

- % of quotes (sorted in order of descending predicted conversion)

- % of actual sales

Gains Curves - GINI Coefficient

- AUC

- GINI = 2 × AUC

AUC

GINI

% of all quotes

(-ranked highest to lowest)

% of all sales

100%
GINI coefficient over time

Evolution of GINI statistic over time for aggregator conversion model

First three months used to create model

Month (first three months used to create model)

Bootstrap – GINI with confidence intervals

Evolution of GINI statistic over time for aggregator conversion model - with bootstrapped error bars

First three months used to create model

Month (first three months used to create model)

Shortfall of the GINI Coefficient

% of all sales 100%

% of all quotes (ranked highest to lowest)
Telesales example

Evolution of GINI and U statistics over time for telesales conversion model - with bootstrapped error bars

Summary

Propensity/elasticity models

- Tracking the Gini and/or U statistics gives quick way of assessing whether the model is broken
- Other statistics could be used
- Bootstrap confidence intervals could define thresholds for more detailed review
- Investigations teach you about shelf life of models

Claims model monitoring

Why are claims far more complex?

- Seasonality
- Frequency and severity
- Constantly evolving mix of underlying events
- Getting both the model structure and the inflation rate correct
- Development
- Accident period vs. underwriting period
Claims model monitoring

**Developed data**
- BAU technical model monitoring part 1
- Data can be used to refit “production” model

**Undeveloped data**
- BAU technical model monitoring part 2
- Provides early warnings
- Needs version of model adapted to historical undeveloped data

---

Claims models example

**Background**
- RBSI accidental damage UK motor claims data
- Frequency and severity models built by working party to be of an average standard for UK motor market
- Again, we’re aiming to track “global”, as opposed to “segmental” model performance

---

Gains curves for claims data – ??

**Gains Curve for AD Frequency Model**

- Cumulative %
- % of Observations
- 2005 Q1, 2005 Q2, 2005 Q3, 2005 Q4, 2006 Q1, 2006 Q2, 2006 Q3, 2006 Q4
GINI Coefficient – not so useful here

Evolution of Gini Coefficient over time for AD Frequency Model

<table>
<thead>
<tr>
<th>Year</th>
<th>Gini Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-Q1</td>
<td>0.288</td>
</tr>
<tr>
<td>2005-Q2</td>
<td>0.305</td>
</tr>
<tr>
<td>2005-Q3</td>
<td>0.296</td>
</tr>
<tr>
<td>2005-Q4</td>
<td>0.251</td>
</tr>
<tr>
<td>2006-Q1</td>
<td>0.261</td>
</tr>
<tr>
<td>2006-Q2</td>
<td>0.287</td>
</tr>
<tr>
<td>2006-Q3</td>
<td>0.296</td>
</tr>
<tr>
<td>2006-Q4</td>
<td>0.285</td>
</tr>
<tr>
<td>2007-Q1</td>
<td>0.254</td>
</tr>
<tr>
<td>2007-Q2</td>
<td>0.249</td>
</tr>
<tr>
<td>2007-Q3</td>
<td>0.271</td>
</tr>
<tr>
<td>2007-Q4</td>
<td>0.272</td>
</tr>
<tr>
<td>2008-Q1</td>
<td>0.280</td>
</tr>
<tr>
<td>2008-Q2</td>
<td>0.291</td>
</tr>
<tr>
<td>2008-Q3</td>
<td>0.273</td>
</tr>
<tr>
<td>2008-Q4</td>
<td>0.273</td>
</tr>
</tbody>
</table>

Evolution of Deviance Statistic over Time

<table>
<thead>
<tr>
<th>Year</th>
<th>Deviance Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-Q1</td>
<td>0.000</td>
</tr>
<tr>
<td>2005-Q2</td>
<td>0.020</td>
</tr>
<tr>
<td>2005-Q3</td>
<td>0.040</td>
</tr>
<tr>
<td>2005-Q4</td>
<td>0.060</td>
</tr>
<tr>
<td>2006-Q1</td>
<td>0.080</td>
</tr>
<tr>
<td>2006-Q2</td>
<td>0.100</td>
</tr>
<tr>
<td>2006-Q3</td>
<td>0.120</td>
</tr>
<tr>
<td>2006-Q4</td>
<td>0.140</td>
</tr>
<tr>
<td>2007-Q1</td>
<td>0.160</td>
</tr>
<tr>
<td>2007-Q2</td>
<td>0.180</td>
</tr>
<tr>
<td>2007-Q3</td>
<td>0.200</td>
</tr>
<tr>
<td>2007-Q4</td>
<td>0.220</td>
</tr>
<tr>
<td>2008-Q1</td>
<td>0.240</td>
</tr>
<tr>
<td>2008-Q2</td>
<td>0.260</td>
</tr>
<tr>
<td>2008-Q3</td>
<td>0.280</td>
</tr>
<tr>
<td>2008-Q4</td>
<td>0.300</td>
</tr>
</tbody>
</table>

An alternative model monitoring metric

We'll refer to it as “the deviance based metric”

\[ 1 - \frac{\text{deviance}_{\text{model}}}{\text{deviance}_{\text{null}}} \]

where \( \text{deviance}_{\text{model}} \) is the deviance of the selected model and \( \text{deviance}_{\text{null}} \) is the deviance of a model containing just a mean parameter.

Deviance based metric

Evolution of Deviance Statistic over Time

First two years used to create model
Summary

- No single metric tells the full story
- Need to understand:
  - changes in cover (e.g., drift in excess seen throughout UK market)
  - changes in mix of events leading to claims
  - changes in business mix
  - …..and indicators for these should be included on regular model monitoring dashboards

Dashboards

- Dashboards which follow are prototypes based on (disguised) real data
- They are the output of a semi-automated process
- Production does not require skilled actuary/statistician
- Circulation
  - Management with overall responsibility for signing off technical models
  - Users of the models
  - Builders of the models
  - Circulated alongside with “master dashboard” listing performance of all models and some measure of “aggregate” lift
Conclusions

- Possible to use a simple metric to monitor models of propensity
- Situation more complex for claims models
- Dashboards can be developed that:
  - detect departures between modelled and actual results
  - assist in the migration away from calendar based model refresh/rebuild cycles
  - aid risk management
- Approach can be extended to undeveloped claims data

Members of the Working Party

- John Berry, EMB
- Gary Hemming, Zurich
- Georgy Matov, RBSI
- Owen Morris, AVIVA
Report… is on the web