Loss Cost Modeling vs. Frequency and Severity Modeling

CAS 2011 Ratemaking and Product Management Seminar

March 21, 2011
Marriott Hotel
New Orleans, LA

Chad Davis, FCAS
Safeco Insurance, Liberty Mutual Group

1

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.**

2

Purpose

- Outline an approach to compare accuracy and stability of Loss Cost Modeling vs. Frequency/Severity Modeling
- Execute this approach on a particular insurance loss dataset that includes large losses

3

Procedure

- Attach a random digit (0-9) to the data
- Use the random digit to split the data into Training (0-3), Validation (4-7), and Holdout (8-9) datasets
  - Training Inclusion Criterion – Indication directionally consistent on each random digit
  - Validation Requirement – Remove indications not directionally consistent with Training
  - Holdout Dataset – Not used in developing structure, only for evaluating model results
- Fit main effects only, no interactions

4

Evaluation of Results

- Accuracy
  - Offset Training indications on the Holdout dataset using a Tweedie Error Structure
    - Log-likelihood measure
  - Score the records in the Holdout dataset using Training indications
  - Compare the scored records to the actual Holdout loss experience
    - Lift Curve

5

Evaluation of Results

- Stability
  - Use the Training Model structure (not the indications) to produce indications based on the Holdout dataset
  - Score the records in the Holdout dataset using the Holdout indications
    - Make a histogram of the percentage change in indication between Training indications and Holdout indications

6

Models

- Loss Cost = Frequency * (Capped Severity + Propensity * Excess Severity)

- Loss Cost (1 Model)
- Frequency/Severity (4 Component Models)
  - Frequency = (Claim Count) / Exposure
  - Capped Severity = (Capped Loss) / (Claim Count)
  - Propensity = (Claim Count Exceeding Cap) / (Claim Count)
  - Excess Severity = (Total Loss – Capped Loss) / (Claim Count Exceeding Cap)

7

Error Structures

- Loss Cost
  - Tweedie, p=1.6
- Frequency/Severity
  - Frequency – Overdispersed Poisson
  - Capped Severity – Gamma
  - Propensity – Binomial
  - Excess Severity – Gamma

8

Extracting Indications

- Loss Cost
  - Relativities produced automatically
- Pure Frequency/Severity (2 models)
  - Multiply together relativities produced by each model
- Frequency/Severity/Propensity/Excess (4 models)
  - Create expected loss costs for each record based on indications, then fit a model to that data to produce loss cost indications by variable
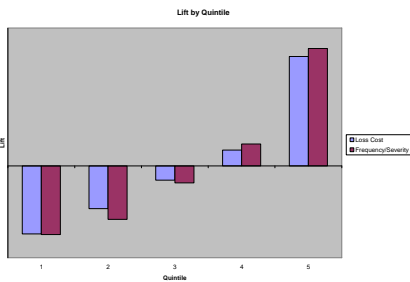
9

Included Variables

10

Accuracy Results

- The Loss Cost model shows a better log-likelihood than the Frequency/Severity model

11

Lift Curve

12

Accuracy Results

- The Frequency/Severity model shows greater lift than the Loss Cost model

13

Model Stability Comparison

Model Stability - Holdout Indications vs. Training Indications

Exposure

□ Loss Cost
■ Frequency/Severity

% Difference in Indicated Loss Cost - Holdout vs. Training

14

Stability Results

- The Frequency/Severity model shows greater stability than the Loss Cost model

15

Loss Cost Pros

- On this particular dataset, the Loss Cost model resulted in a better log-likelihood measure of accuracy
- Requires only one model to build and maintain – may require less resource
- Simpler method to implement offsets
  - A modeled loss cost dataset can be produced for the Frequency/Severity modeling procedure to facilitate offsets, but this may require more work
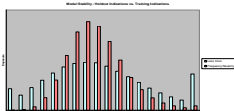
16

Frequency/Severity Pros

- On this particular dataset, the Frequency/Severity model showed greater lift and stability
- Potential for more focused treatment of large losses
- Potential for greater understanding of business

17

Frequency/Severity Pros

- Potential reason for improved stability
  - More Flexibility
    - The Frequency/Severity approach gives the option of including some significant components (e.g. Frequency) of a certain variable in the model while excluding other insignificant components (e.g. Excess Severity) for that variable
    - The Loss Cost approach allows only a binary choice for the inclusion of a variable

18