RP-3 Predictive Modeling

- A Regulator's Perspective

- Daniel Davis, FCAS, MAAA
- Alabama Department of Insurance

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.

This slide show is comprised of a mixture of:

- complaints I have about GLM filings
- How a different approach from insurers could help
- Comments regulators share amongst themselves
- Better Questions for regulators to use to "peel that elusive onion"

Experience Reviewing Models

➢ GLMs

➢ Tiering/scoring Models

Regulator to Regulator discussions: Finding Questions that get an answer

OR: approaches to questioning that actually get an answer you can use

What is the problem?

- Some filers just want us to "go away"
- Sometimes filing unit personnel only have a very rudimentary knowledge of the GLM subject
- Some companies don't believe we want answers but rather we want the appearance of having asked.

NAIC Survey regarding Predictive Analytics

- The NAIC in 2015 conducted a survey of Regulators and had come up with a Comprehensive list of representative Questions

- 4 different sections of GLM/Predictive analytics questions

Why Multiple Questions on the same topic?

- Came across this idea by accident
- NAIC survey questions (4 different sections)
- Overview, assumptions
- Dataset / variables
- GLM model and math involved
- Goodness of fit tests

Tactics I have had to use to "pull the facts out"

- The OVERVIEW section and the 3 others (dataset, Math model and goodness of fit tests) have much overlap
- This is good for the regulator for 3 reasons (but, isn't this inefficient?) .. Yes, but

Things regulators say to one another

- Answers to the NAIC GLM Survey questions are not always answered in a consistent manner
- Regulators can point this out
- Force the (real) modelers to come forward, really address the questions
- Ties back to the advantage of many, many (some might argue too many) questions

Refer to earlier filing (Regulator to Regulator suggestions)

- In the last filing you said _____ and now __

Sometimes it really doesn't have to be a perfect question/response.

Sometimes it doesn't have to be a perfect question/response

- "Your company sent me an 844 page document with no table of contents or index or Section Header pages. Please fix."
- Funny, no elaboration was actually necessary.

NEW SECTION:

Here are some questions that have been improved – in order to combat some of the "avoiding the question" techniques

The original question

"an intuitive argument for why an increase in each variable should increase or decrease frequency, severity, loss cost or expenses incurred"

The original answer

GLMs find the right coefficients without regard to intuition or "reasonable" explanation

Implication: that is a stupid question.

A 2nd implication: we never thought about that (which is simply false).

Tell the filer what you do know (skip right to phase 2)

"Because both company employees and agents have a need to understand & explain these concerns, the filers of GLM models have consistently provided the AL DOI with reasonable/plausible answers to this question:"

"an intuitive argument for why an increase in each variable should increase or decrease frequency, severity, loss cost or expenses incurred"

Please supply me with your explanations/rationales

The original question

An explanation of why you believe this model is better than the one it is replacing.

The second question

An explanation of why you believe this model is better than the one it is replacing. How did you form that conclusion? What metrics did you rely on?

The 3rd question

An explanation of why you believe this model is better than the one it is replacing. How did you form that conclusion? What metrics did you rely on? (Model gain is not a goodness of fit measure. It is useful but not totally adequate by itself). Did you measure and compare model stability, A) versus B)? Did you consider penalized measures of fit such as the Akaike Information Criterion? When reviewing model lifts, did you also consider both predictive accuracy and monotonicity?

"A Final Chapter":

- Assorted questions (and possible approaches to answering them)

Taking a step back: Compliments can be better than complaints

- An overview of the model including:
- a.   a technical memorandum describing the steps taken in the construction of the model;
- Please reference the GLM Explanatory Memo.
- Sometimes these are much better because they read in an almost historical order – we thought about A), then we considered B) but rejected it and so on and so forth

Problematic Question: source of any non-insurance data (customer provided or other) + who owns it, how customers can obtain and correct errors

- not well answered
- can the insured appeal to insurer?
- CARFAX is a good example

Many different models – a useful trend

- With all the different models impacting rates, it is sometimes very useful when an insurer sends in a schematic that displays them all and how they interrelate. Some insurers are very proficient at this.

On the other hand -- Many different models

- Some insurers are not helpful in this regard and this necessitates much back and forth correspondence
- They also have trouble with our request for information regarding prior filings

Here is what we have now: (suggestions?)

- Predictive Analytics or GLM Questionnaire
- Description
- If this filing has rating factors (variable differentials) that are supported by an analysis that contained Generalized Linear Models or other Predictive Analytic Tools, please review and answer the questions in the attachment found below.

- If the Predictive Analytics work or the attached questionnaire has been submitted in a previous filing, please respond with "please see prior filing, SERFF # _____."

  If this filing has no rating factors (variable differentials) supported by Predictive Analytic Tools, please respond with "Not Applicable".

  In the event of methodology changes, please indicate in detail what was changed and explain how and why the changes were made.

More Compliments: how overfitting was addressed

- We appreciated the "big picture" responses we got
- Unlike some of the other standard responses
- Are perhaps 3 possible solutions or alternatives to handle the problem
- Hold-out sample / correlation / AIC
- Ultimately the thoroughness of this answer increased the speed to market

Specifics about the GLM model(s) including link function: (comments please)

- "All of the models used log link functions so that the indicated rating factors directly support our multiplicative rating structure".

- **Is this backwards??**