Data Variance & Model Consistency

1 April 2014

Christopher Cooksey, FCAS, MAAA
EagleEye Analytics

EagleEye Analytics

Data Variance & Model Consistency

A reasonable thought…

For each customer find the 5-year loss ratio. Sort them from highest to lowest.

If we were to compare this to actual loss ratios for business sorted by our model predictions, we should see similar results, yes?

This idea relies on there being enough consistent losses across our data.

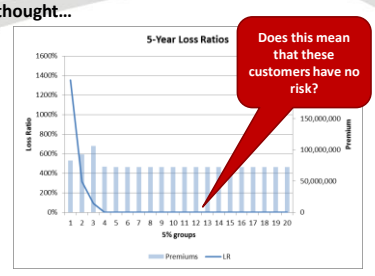
2

Data Variance & Model Consistency

A reasonable thought…

Oops…

This illustrates the low frequency of insurance data.



What we have here is essentially a Gini of one (or rather, a maximized Gini).

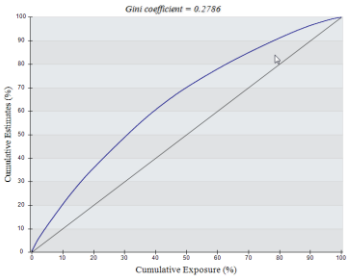
3

Data Variance & Model Consistency

And speaking of the Gini coefficient…

What do typical Gini values tell us about the extreme variance of insurance data?

About the significant difference between modeled and observed?

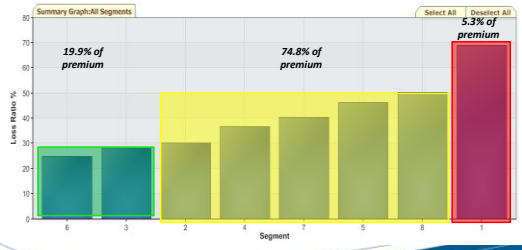


4

Data Variance & Model Consistency

And what does this say about effective models?

This model was used to categorize agents – those who wrote the most green business and those who wrote the mode red.

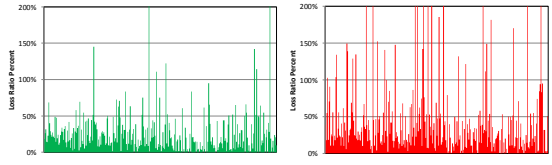


5

Data Variance & Model Consistency

And what does this say about effective models?

These are pictures of the difference between a 30% loss ratio and a 70% loss ratio.



6

Questions?

Contact Info
Christopher Cooksey, FCAS, MAAA
EagleEye Analytics

ccooksey@eeanalytics.com

www.eeanalytics.com