The Actuarial Profession making financial sense of the future

GIRO Conference and Exhibition 2012

Juggling uncertainty the actuary's part to play



© 2012 The Actuarial Profession • www.actuaries.org.uk



GIRO Conference and Exhibition 2012

Analytics: A transatlantic comparison

James Rakow, James Guszcza & Matthew Wilson

Deloitte

Analytics: A Model Definition



Data

- Historically, use of data driven by legislation and customer attitudes
 - Regulatory requirements
 - Route to market
 - What is acceptable to the customer
- More recently, an explosion in global data availability has led to insurers looking elsewhere
 Delotte.
 - 3rd party data
 - Non-traditional data
 - Open data



Models: GLM and Beyond

- Two levels of Discussion:
 - Tools and Methodology
 - Application
- At the level of methodology:
 - Regression and its Relations
 - Statistics vs machine learning
 - Traditional vs "modern" approaches to data analysis
 - Hallmarks of "Data Science"
 - "Actuarial Science as Data Science"

Deloitte.	R
Actuarial Science as I Actuarial Modeling in	Data Science R
Revolution Analytics Webinar	Jim Guszcza, FCAS, MAAA
March 28, 2012	Deloitte Consulting LLP University of Wisconsin-Madisor

Data Science: An Implicit Definition



Image borrowed from Drew Conway's blog http://www.dataists.com/2010/09/the-data-science-venn-diagram 5

Putting GLM in Statistical Perspective

(Or: Why Programming with Data is Important)

- Linear Model Assumptions:
 - Courtesy of Andrew Gelman and Jennifer Hill.
- 1. Valid model design and variable selection
- 2. Additivity and Linearity
- 3. Independence of errors
- 4. Equal variance of errors (homoscedasticity)
- 5. Normality of errors

 $Y = \beta \cdot X + \varepsilon$ ε ~_{iid} N(0,σ²)

(Paraphrase from Data Analysis Using Regression and Multilevel/Hierarchical Models by Andrew Gelman and Jennifer Hill)

GLM and Beyond: Tools and Techniques

- Supervised Learning:
 - Regression \rightarrow GLM
 - Multilevel / Hierarchical and Bayesian Models
 - Generalized Additive Models
 - CART, MARS, Random Forests
 - Neural networks, Support Vector Machines
- Unsupervised Learning
 - Clustering
 - Factor / Principal components Analysis
 - Association Rules

- ...

Beyond GLM: Methodology and Mindset

- Predictive variables are the "fuel" that "power" scoring engines.
- This means that the most important part of the modeling process occurs before the modeling begins
 - EDA: Exploratory Data Analysis
 - KDD: Knowledge Discovery in Databases
- Focus on:
 - Model design
 - Programming with data
 - Variable creation and selection
 - Remember Gelman's #1 assumption
 - Early example: credit scoring has been a major innovation for US actuarial work
 - Out-of-sample model validation

Lift Curves and Out-of-Sample Validation

- The availability of a multitude of variables and powerful machine learning tools heightens the risk of overfitting.
- Out-of-sample validation became a "gold standard" in US actuarial practice around the same time the "data science" approach took hold.
- Key tool: the lift curve.



Profitability Model Lift Curve

Training

What do actuaries (and others) learn about Analytics?



Training

When in your education did you meet these distributions?



4.10 -4.18

Analytics – the Full Picture?



Analytics – the Full Picture



It All Starts with a Strategy...

Moneyball – A great book/film about

- Business analytics
- Innovation through data science
- Behavioral economics
- Human psychology
- Culture change
- (... and, oh yes, baseball)

... And ends with decisions

Decision support – Where Analytics adds/destroys value

- Divergent outcomes
 - Human decision-makers remaining central, and models supplementing and anchoring human judgment; not replacing it.
 - Use of analytical techniques resulting in people being removed from the decision-making process

Closing Remarks and discussion



Questions or comments?

Expressions of individual views by members of The Actuarial Profession and its staff are encouraged.

The views expressed in this presentation are those of the presenters.

