



Data Analytics in Life, Disability and Health Insurance

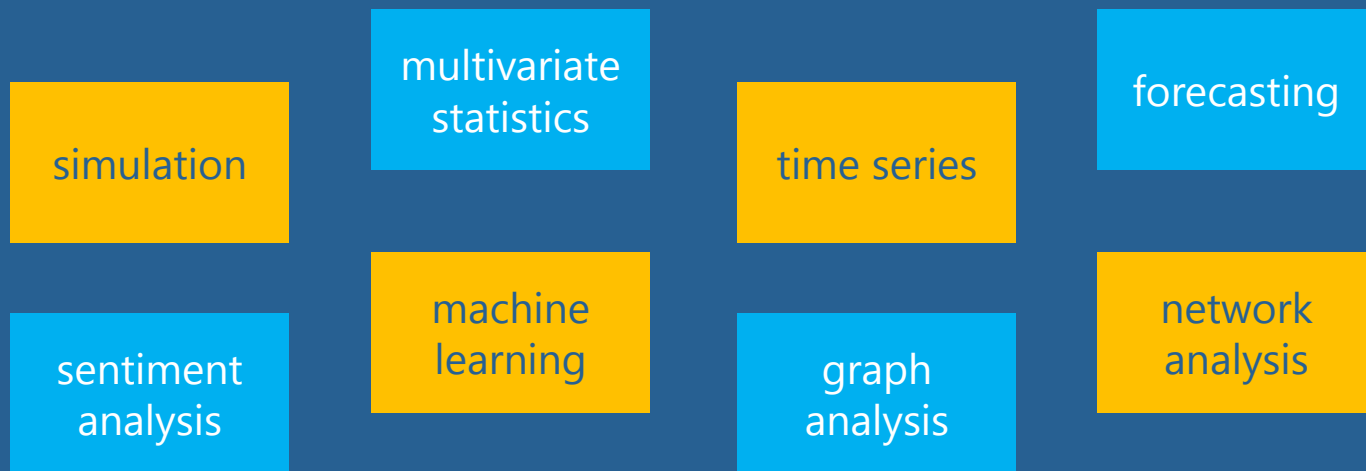
ILAA Convention 19.12.2018, Tel Aviv

Ralph Price

Data analytics

The (semi-)autonomous examination of data using sophisticated techniques and tools, typically beyond traditional analysis, to discover deeper insights, make predictions or recommendations.

(Gartner)



Uses of Data Analytics in Life/Health Insurance

Pricing <ul style="list-style-type: none">• New risk/rating factors• Initiate product differentiation• Other pricing refinements	Lapse Behaviour <ul style="list-style-type: none">• Identify prospects more likely to lapse	Claims Management <ul style="list-style-type: none">• Identify fraud / misrepresentation• Triage claims• Other claims management
Underwriting <ul style="list-style-type: none">• Develop targeted uw• Reduce / simplify uw• Refine uw decisions• Do away with uw	Marketing <ul style="list-style-type: none">• Identify prospects more likely to buy• Target marketing efforts• Distribution management	

© Thinkstock: RamCreativ

Generalized Linear Models (GLM)

- A Generalized Linear Model (GLM) is a generalized form of a Linear Model (LM).
- The purpose of both is to express the relationship between an observed response variable, Y , and a number of explanatory variables or covariates, X .
- GLMs offer considerable flexibility regarding the distribution of Y
- GLMs also can provide information about the certainty of model results

Established actuarial analysis for the German market



Mortality Pool – 2017

- 165 million life years / 680,000 deaths



Disability Pool – 2017

- 53 million life years / 115,000 claims

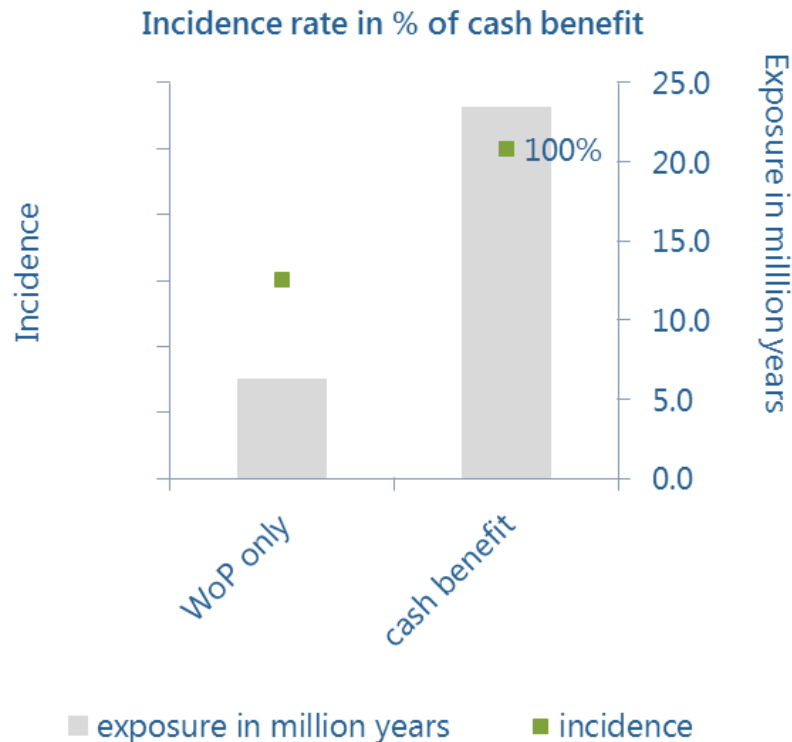


Annuity Pool – 2017

- 100 million life years / 190,000 deaths

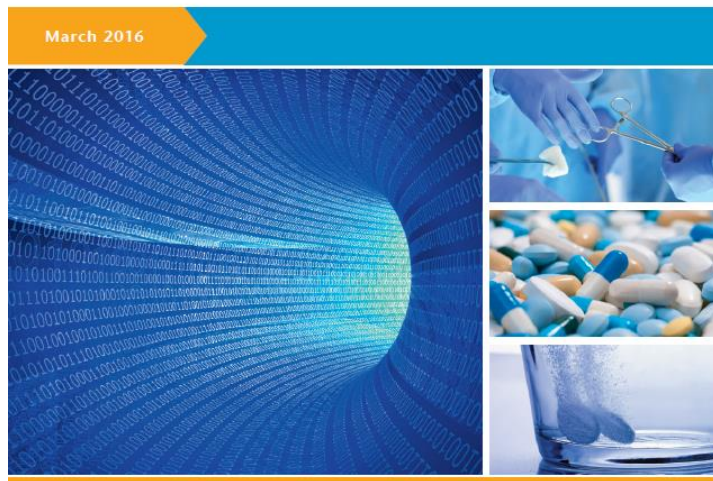
- 30 companies deliver portfolio information on an annual basis
- Annual analysis of data from last 10 to 12 years
- Presentation and reports for the participating companies

GLM analysis – cash benefit versus WoP only



- Classical analysis: only $\frac{1}{3}$ better claims experience for WoP only
- Reason: In the classical analysis there was no differentiation by the occ. class. The WoP only portfolio contains more blue collar.
- The GLM helps us to see the "true" difference between the two benefit types.

Gen Re Asia Dread Disease Surveys



March 2016

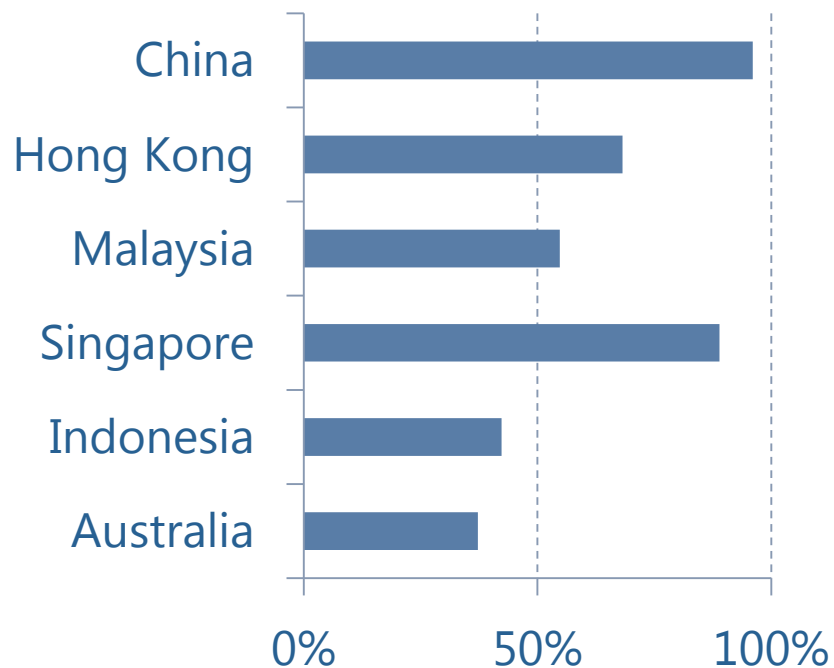
DREAD DISEASE SURVEY 2008-2012

Australia, China, Hong Kong, Indonesia, Malaysia, Singapore and South Korea

		Coun-tries	Com-panies	Claims	Life years
1	'90-'94	3		4,600	
2	'93-'97	3		7,000	3m
3	'96-'00	3	31	16,000	4m
4	'00-'04	6	48	263,000	41m
5	'04-'08	10	95	750,000	> 70m
6	'08-'12	7	82	> 1m	>100m

Investigation Areas in the DD Survey

Market share

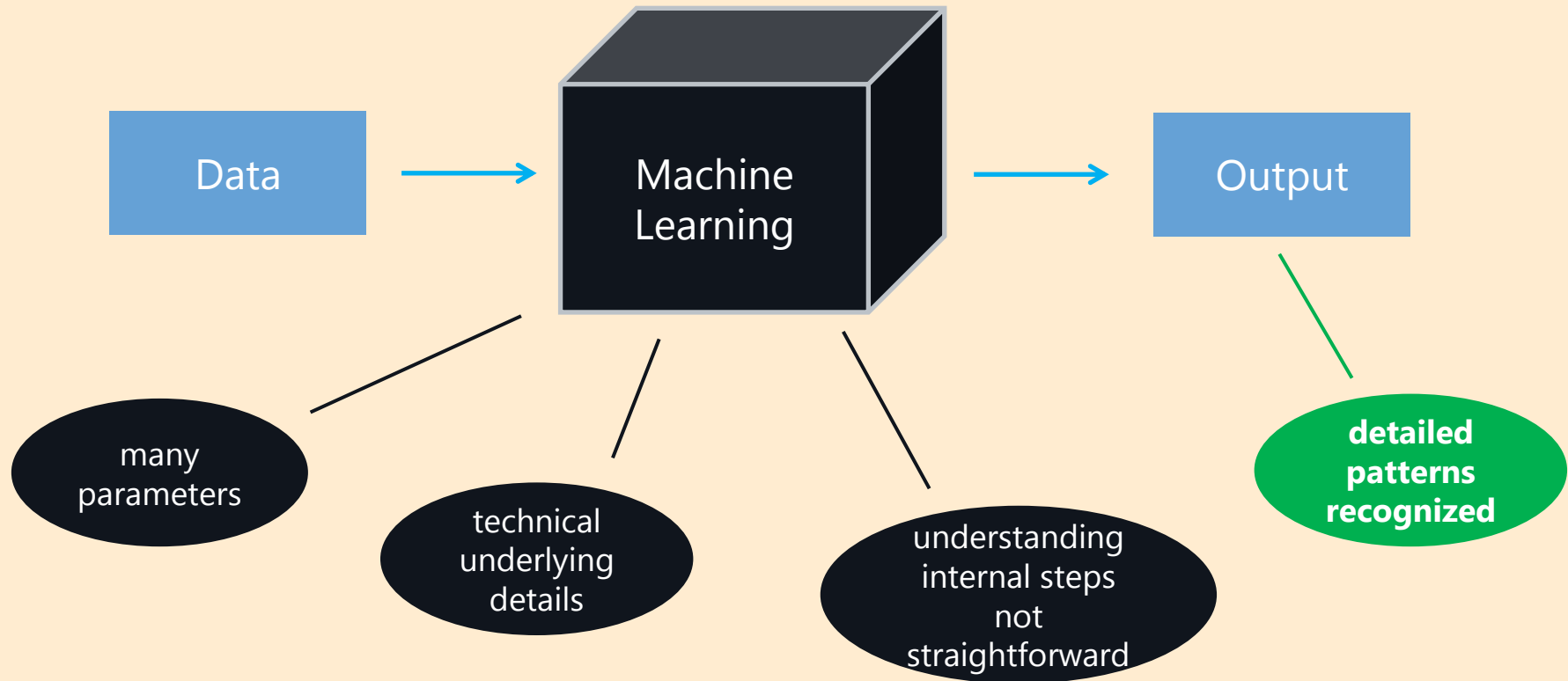


Areas of Interest

- I. Distribution of In-force and New Business
 - Policy Size, Smoking Status, Guarantees, Benefit Type, Lapses
- II. Analysis of Claims
 - Declinature Rates and Reasons, Ex-gratia Claims, Time Lag Between Diagnosis, Reporting and Payment
- III. Claims Experience
 - Comparison by Market
 - Trends
 - Selection Effect
 - Variation by Company, Occupation and Region

What can GLM do better than traditional analysis?

- GLM can detect new relevant pricing factors (e.g. postcode)
- GLM can analyse multiple factors simultaneously
- GLM is able to detect interactions and dependencies between rating factors (e.g. correlations between occupation and smoking status)





Medical insurance cost

Age
Sex
BMI
Children
Smoker
Region in US

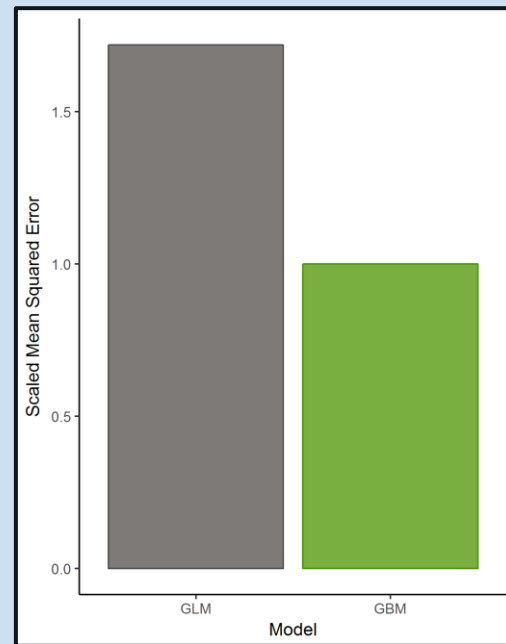
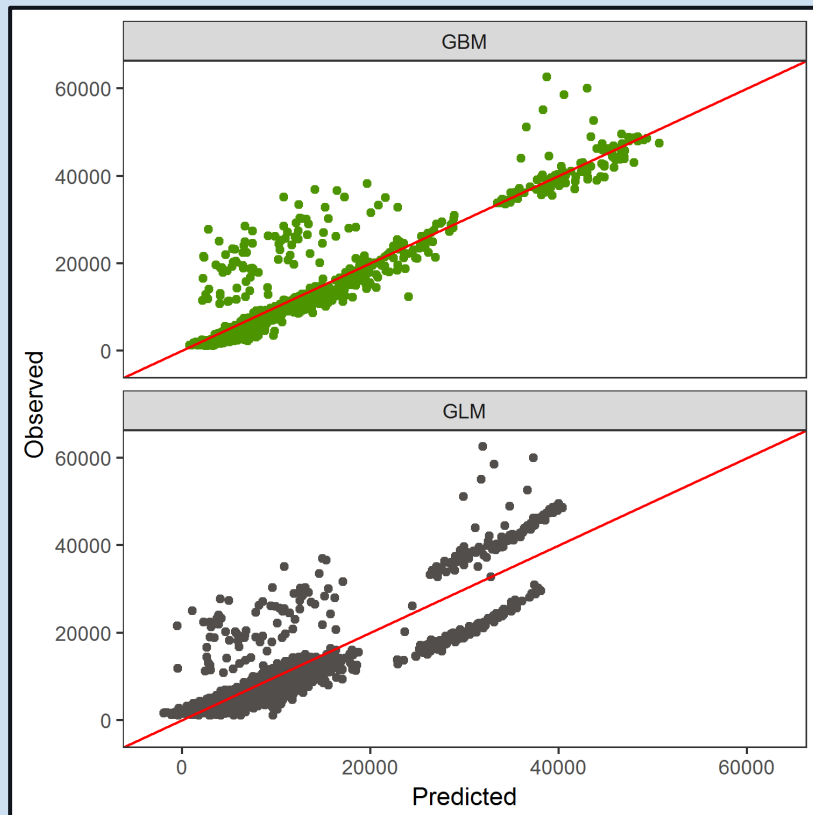


Diabetes (female older than 21)

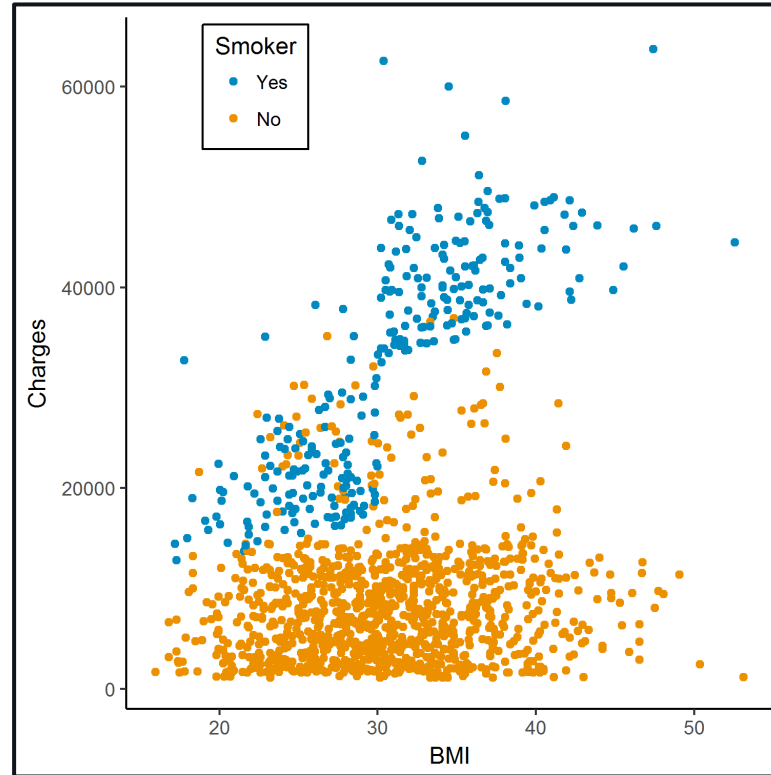
Age
BMI
Glucose
Blood Pressure
Insulin
Skin Thickness
Number of pregnancies
Diabetes Pedigree Function

Generic small data; Models with default parameters; Demonstration purpose only!

Medical insurance cost

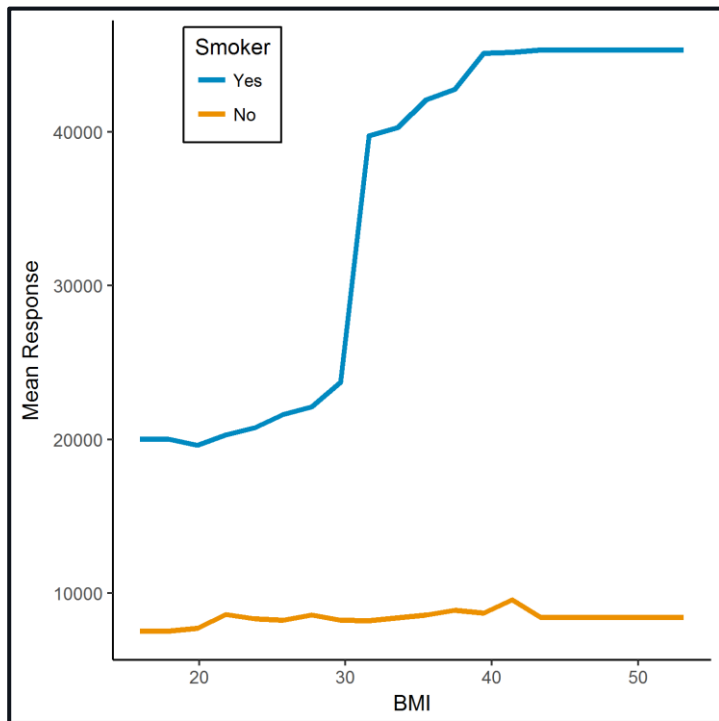


BMI and Smoking Status

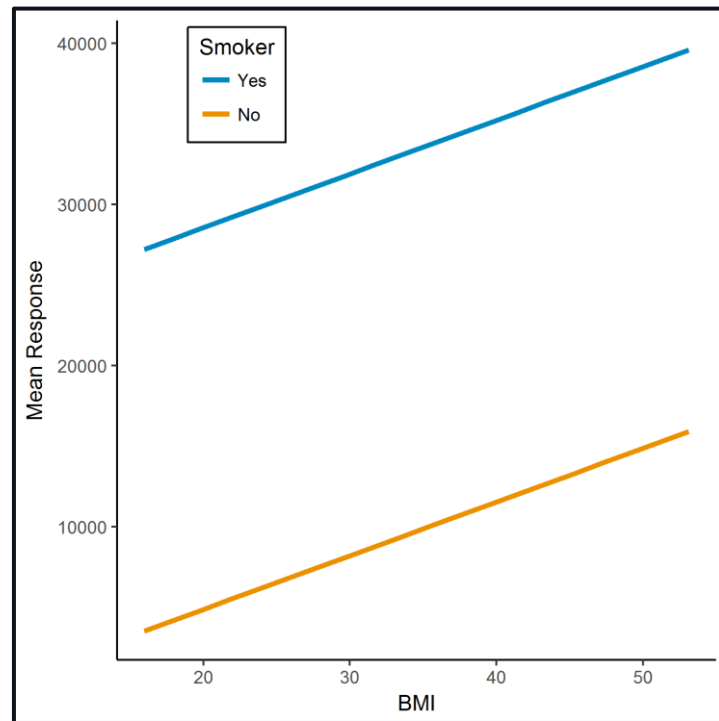


BMI and Smoking Status

GBM



GLM



Claims Triaging

- **Business question**

- Assign claims to the right manager based on complexity of disability income claims

- **Analyse Claims to identify patterns**

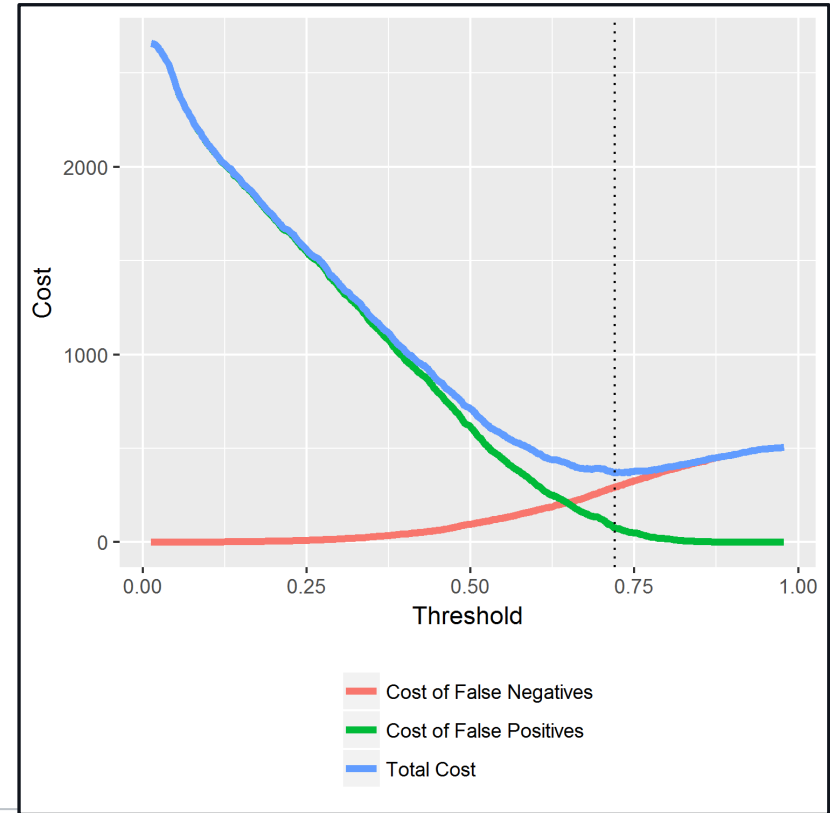
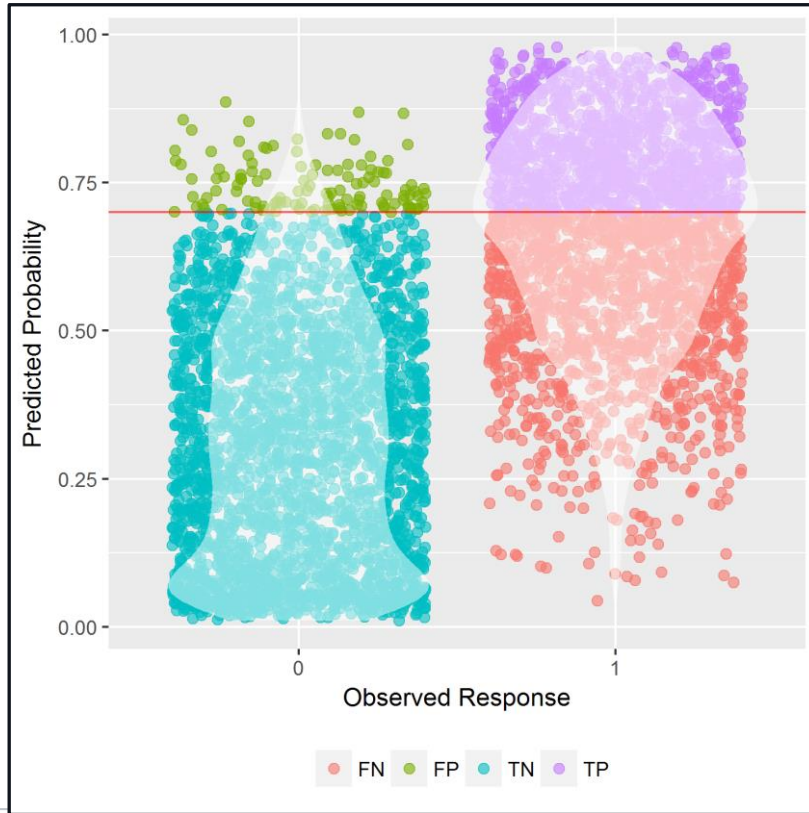
- Which information that are available early at claims stage lead to complexity classification
- Binary model (0 for complex, 1 for simple)
- False Negative = Jr. managers decide on complex cases
- False Positive = Sr. managers waste time on simple cases
- Grey area **business requirements suggest what thresholds are best to minimise cost**

- **Deployment**

- Assign score for new claims

Trade-off between threshold and false positives/negatives

Increasing threshold, lowers number of FP cases, while increases number of FN cases



What can Machine Learning do better than GLM and/or traditional analysis?

- ML can apply more advanced models
- ML is not limited to linear framework, and is better at picking up non additive dependencies
- ML can often better approximate data leading to improved predictions

Challenges

- Data
 - Accesability of Data
 - Quality of Data
 - Data protection aspects
- Lack of expertise
- Interpretability of models / results
- Regulation

Challenges

German financial regulator BaFin publishes study “Big Data meets Artificial Intelligence”

No black box excuses – explainability/traceability of models is necessary and can improve the analysis process | It is the responsibility of supervised firms to guarantee the explainability/traceability⁴⁷⁰ of BDAI-based decisions. In particular, chapter 3.5.2 details how new approaches could provide at least some insight into how models work and the reasons behind decisions, even in the case of highly complex models, thereby preventing models from being categorised purely as black boxes. Supervisory and regulatory authorities will therefore not accept any models presented as an unexplainable black box. In addition, a better understanding of models provides an opportunity to improve the analysis process – allowing, for instance, the responsible units in the supervised firm to identify overfitting and data bias (see Chapter 3.3.1).

Challenges

ILAA 70th Anniversary Convention June 2016:

Models don't make decisions. People do.

Challenges

Data analytics is also revealing what's actually happening. And some of the findings are surprising.

For example, one pet food maker assumed that having its product positioned next to its main competitor wouldn't be good for sales. But it was. Why?

"We are not bothered by explaining why, we're just following the data," says Mr Bar-EI. "It could be about colour, brain psychology, but we don't know. Our conclusions are evidence-based."

Outlook

- Data Analytics will increase in importance
- Gen Re has a dedicated department headed by a data scientist dealing purely with Data Analytics

Outlook

- Let's have some fun.....
- NOW

Outlook

- People will still make the decisions, however Data Analytics will play a more important role in preparing decisions
- In pricing Data Analytics will lead to new costing factors, more accurate models
- Straightforward tasks will be identified by Data Analytics and automated
- Data Analytics will allow better customer segmentation and control of sales channels
- Data Analytics will allow better detection of anomalies (e.g. non-disclosure; claims fraud)