Insurance pricing analytics: a research perspective

Katrien Antonio
LRisk - KU Leuven and UvA

May 7, 2021

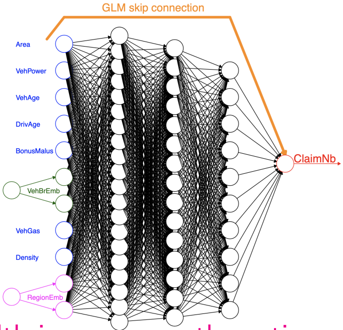
LRISK

UNIVERSITEIT VAN AMSTERDAM

Four principles

1. **Inverted** production cycle.
2. The contributions of the many to cover the **misfortunes of the few**.
3. Segmentation or **risk classification** or differential pricing.
4. **Insurance regulation** and **ethics**.

Inverted production cycle



Data-driven, extensive use of predictive modeling.

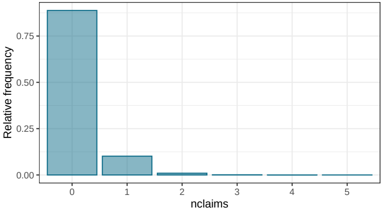
EIOPA (May, 2019) on Big data analytics in motor and health insurance: a thematic review.

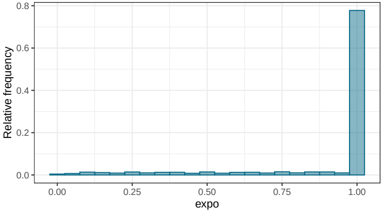
Massive changes:

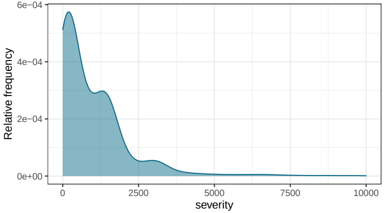
- in techniques: from regression (with GLMs) to neural networks.
- in granularity: from aggregate to fine-grained data (in **pricing**, but also reserving and other tasks).

Picture from Schelldorfer & Wüthrich (2019) on Nesting classical actuarial models into neural networks.

Contributions of the many to cover the misfortunes of the few







With claim **frequency** data (lhs):

- number of (insured) events occurred
- low frequency, very often zero!

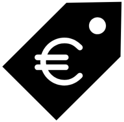
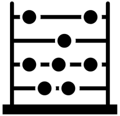
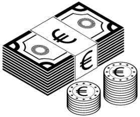
With claim **severity** data (rhs):

- only contracts with frequency > 0
- potentially high impact, (heavy) right tail.

Henckaerts, Antonio et al. (2018, Scandinavian Actuarial Journal). A data driven binning strategy for the construction of insurance tariff classes.

From frequency and severity to pure premium

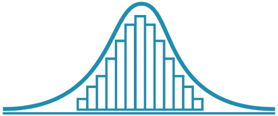
Calculate a **technical, break-even or pure premium** $\rightarrow$ e.g. with frequency-severity approach.

	$=$		$\times$	
pure premium		expected number of claims		expected claim size

Or, in math notation, $\pi = E[F] \times E[S]$ with F for frequency and S for severity.

Technicalities depending on line of business.

From pure $\rightarrow$ risk $\rightarrow$ market premium



Next steps:

- get risk premium $\pi^\star$ by adding **risk or safety loading**, e.g. via actuarial premium principle

$$\pi^\star = \pi + \rho \cdot \pi$$

- **add extra loadings** for capital (k), expenses (e), acquisition (a), taxes (t), etc.

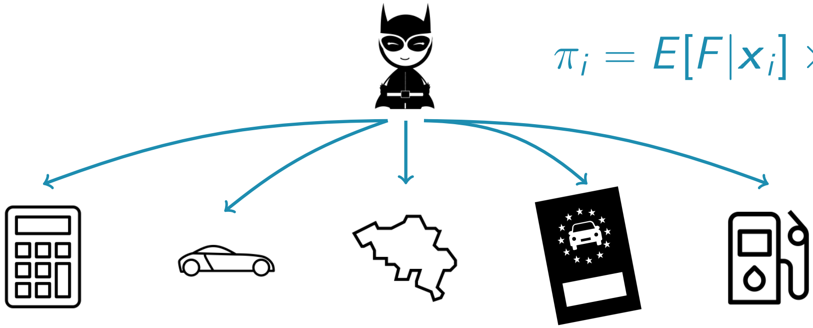
$$\pi^{\text{g}} = \frac{1}{1 - (k + e + a + t)} \pi^\star,$$

- commercial constraints $\rightarrow$ **market premium**.

Deprez, Antonio & Boute (2021) on Empirical risk assessment of maintenance costs under full-service contracts.

Pricing via risk classification

Actuarial fairness via multi-type features



$$\pi_i = E[F|\boldsymbol{x}_i] \times E[S|\boldsymbol{x}_i]$$

Henckaerts, Antonio et al. (2020, NAAJ) on Boosting insights in insurance tariff plans with tree-based machine learning methods.

Devriendt, Antonio et al. (2021, Insurance: Mathematics and Economics) on Sparse regression with multi-type regularized feature modeling.

Feature engineering: new types of features!

James B.

He drives 100 000 km during one year.

His road type composition is $(15\ 000, 15\ 000, 50\ 000, 20\ 000)$ or $(0.15, 0.15, 0.5, 0.2)$.

Eugène from Man Bijt Hond

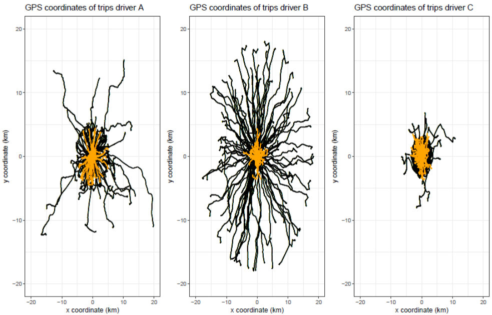
He drives 1 000 km during one year.

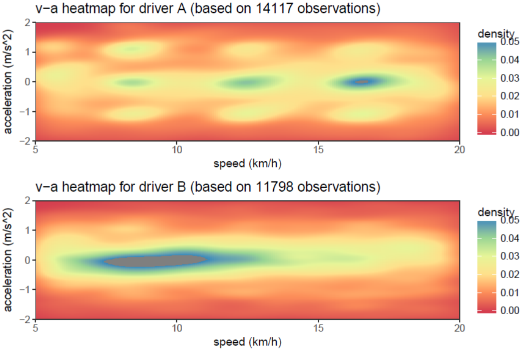
His road type composition is $(500, 300, 200, 0)$ or $(0.5, 0.3, 0.2, 0)$.

Verbelen, Antonio & Claeskens (2018, JRSS: Series C). Unravelling the predictive power of telematics data in car insurance pricing.

Henckaerts & Antonio (2021, in progress). The added value of dynamically updating motor insurance prices with telematics collected driving behavior data.

Feature engineering: new types of features!





Pictures reproducing Wüthrich (2017, European Actuarial Journal) on Covariate selection from telematics car driving data, using data from AXA Kaggle competition.

Insurance regulation and ethics

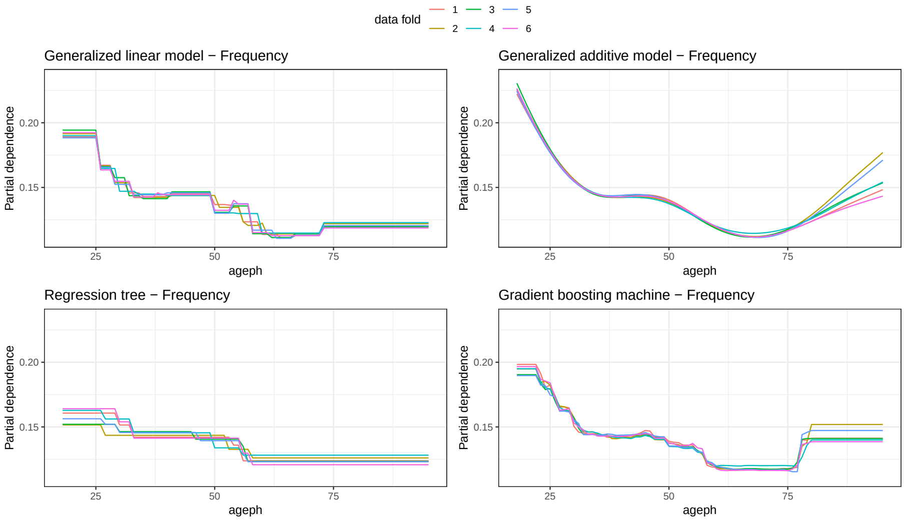
Looking under the hood

- Interpretation is **simple with mainstream actuarial models** (think: Generalized Linear Models or GLMs), e.g.

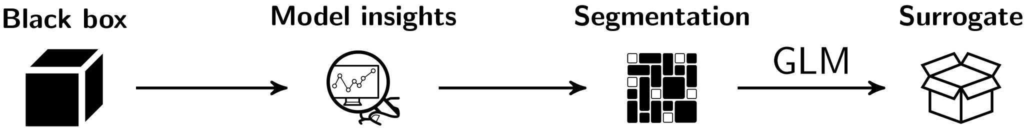
$$E[F|\boldsymbol{x}_i] = \exp(\beta_0) \cdot \exp(\beta_1 \cdot \mathcal{I}(\texttt{age\_bin\_1}) + \ldots + \beta_p \cdot \mathcal{I}(\texttt{geo\_cluster\_p})).$$

- But (way) more complicated with machine learning methods, e.g.
  - Gradient Boosting Machines (lots of small trees, fitted sequentially)
  - Neural Networks (tons of weights and bias parameters, layered structure).
- **Interpretation toolbox** demonstrated in Henckaerts, Antonio et al. (2020, NAAJ).

Looking under the hood



Explainable surrogate models



Two roads or pathways to explainable AI (XAI):

- **after the event**: use interpretation tools to (better) understand decision process in black box model
- **by design**: develop and use transparent white box model, e.g. as a global surrogate for a more complex black box model.

Henckaerts, Antonio et al. (2021). When stakes are high: balancing accuracy and transparency with Model-Agnostic Interpretable Data-driven suRRogates.

Discrimination-free insurance pricing

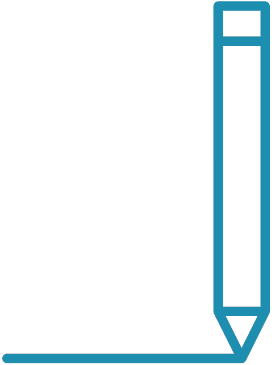
- Active and substantial literature on **fairness in machine learning**.

- Actuarial framework recently proposed in Lindholm et al. (2020), following Pope & Sydnor (2011, AEJ: Economic Policy).

- Three ways to determine (technical) insurance prices with 'sensitive or protected' features (in $\boldsymbol{d}_i$):

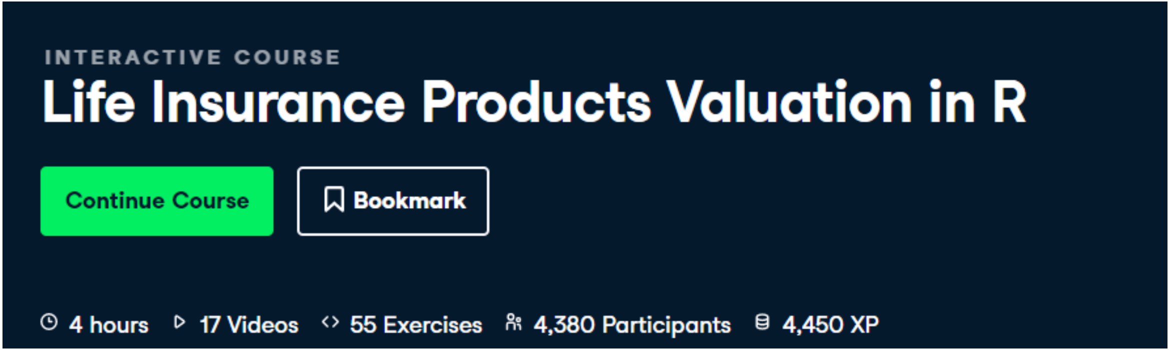
  - best-estimate (use $\{\boldsymbol{x}_i, \boldsymbol{d}_i\}$)
  - unawareness (use $\boldsymbol{x}_i$, but not $\boldsymbol{d}_i$) $\rightarrow$ proxy effects!
  - discrimination-free (use $\{\boldsymbol{x}_i, \boldsymbol{d}_i\}$, then average over $\mathbb{P}(\boldsymbol{D})$).

MSc thesis by Elien Baeten (2021, MAFE program KU Leuven) on How to avoid discrimination in insurance pricing: a probabilistic and learning toolbox.



That's a wrap!

For references, acknowledgements, see https://katrienantonio.github.io.



Online course with DataCamp on **Valuation of Life Insurance Products in R**, designed by Katrien Antonio & Roel Verbelen (target: BSc and MSc students, lifelong learning)

(in 2022, Jan 11 & 12) Junior College by Katrien Antonio and Kristien Smedts on **Kansrekening en financieel waarderen: de actuaris aan het werk**

(in 2021-2022, concept) bijhorende lesmodule Junior College STEM en Economie (FEB en dept. Wiskunde, KU Leuven).

Thank you for your attention!