

CREST - GENES

Cours doctoraux 2023 – 2024

BIASES, DISCRIMINATION, AND FAIRNESS

Arthur Charpentier

Université du Québec à Montréal

| | | | | |
|----------|----------|--|--------------------|-----|
| SCHEDULE | Monday | 29th January 2024 5th February 2024 | From 9:00 to 12:15 | TBA |
| | Thursday | 1st February 2024 | From 9:00 to 12:15 | TBA |

Aims and objectives

This course will provide a state-of-the-art, on fairness and discrimination, in the context of insurance pricing (and more generally, predictive models). As explained by Avraham et al. (2014) "insurance companies are in the business of discrimination. Insurers attempt to segregate insureds into separate risk pools based on the differences in their risk profiles, first, so that different premiums can be charged to the different groups based on their differing risks and, second, to incentivize risk reduction by insureds. This is why we let insurers discriminate. There are limits, however, to the types of discrimination that are permissible for insurers. But what exactly are those limits and how are they justified?". First, we will come back to the specificities of predictive models in insurance. We will come back to the different places where a potential discrimination can intervene, by insisting on the possible biases in the data, in the models. We will present in particular the regulations in Europe and North America. In a second step, we will see how to quantify a possible discrimination, insisting on the main measures of "group-fairness", before discussing the individual approach, in particular in relation with the causal approaches. Indeed, the central question of discrimination is "would the price have been different if this person had been a man instead of a woman". We will see how to build a counterfactual allowing to quantify a possible discrimination. Finally, we will see how to correct a discrimination, insisting on the in-processing (through penalized models) and post-processing approaches (using optimal transport).

Outline

- Predictive models in insurance**
From GLMS to machine learning
Calibration and interpretability
- Data and biases**
Selection biases
Simpson's paradox
Retroaction bias
- Discrimination vs segmentation, in insurance**
Economic perspective on "efficient" discrimination
Legal perspective
- Quantifying discrimination and fairness**
Group fairness (demographic parity, equalized odds)
Individual fairness and counterfactuals

5. **Discrimination mitigation**

Pre-processing / in-processing / post-processing

Pre-requisites

Course on predictive modelling, and ideally, a course in insurance pricing.

Exam: group project

Some related literature (*suggested before the course)

Charpentier A (2023). Insurance: biases, discrimination and fairness. Springer Verlag.

Avraham, R., Logue, K. D., & Schwarcz, D. (2014). Towards a universal framework for insurance anti-discrimination laws. Connecticut Ins. Law Journal, 21, 1.

Avraham R (2017) Discrimination and insurance. In: Lippert-Rasmussen K (ed) Handbook of the Ethics of Discrimination, Routledge, pp 335–347

Barocas S, Hardt M, Narayanan A (2019) Fairness and Machine Learning. fairml-book.org

Halpern JY (2016) Actual causality. MIT Press

Pearl J, Mackenzie D (2018) The book of why: the new science of cause and effect. Basic books