

Séance 25

Des probabilités aux scores : une approche historique et critique

22 avril 2020 (reprogrammée du 18 décembre 2019)

Présentation

Les scores sont partout. Que l'on accepte ou non l'idée d'une révolution apportée par le big data, force est de constater que les nouveaux modèles servent à classer, trier, choisir et organiser le monde grâce à l'attribution de scores. Là où auparavant on se contentait de catégoriser les personnes dans des classes d'équivalence assez larges, les techniques actuelles cherchent à appliquer à chacun un chiffre, dont on voit bien que la dimension à la fois probabiliste et individuelle l'écarte d'une catégorisation de type classique. Le sens possible ou supposé de ces scores n'a pourtant rien d'évident, même si la question est rarement posée.

Cette étude est une tentative de mise à plat d'une interprétation des scores tels qu'ils sont calculés dans les grandes familles de modèles. Nous partirons des débats historiques sur l'interprétation des probabilités, débats qui se sont tus dès les années cinquante, sans pour autant avoir fourni une réponse consensuelle. De ces interprétations aucune ne s'applique parfaitement aux scores actuels, sauf peut-être la propension proposée par Karl Popper dans le contexte d'une polémique de mécanique quantique. Après avoir mis en avant le glissement des probabilités aux scores (comme propension ?) dans les modèles de data science, nous essaierons de montrer les limites de cette dernière interprétation dans ses applications potentielles aux phénomènes sociaux.

Intervenant : Laurence Barry actuaire IA, PhD, Chaire PARI

*

La chaire PARI (programme sur l'appréhension des risques et des incertitudes), portée par l'ENSAE et Sciences Po, a pour objectif d'identifier et comprendre (i) le champ de pertinence de nos outils d'appréhension des risques, et (ii) leurs conditions d'émergence et d'utilisation. Créée début 2015, elle organise un séminaire de recherche mensuel de 2h pour présenter et échanger autour de ses travaux et des thématiques connexes. Le deuxième cycle de la chaire porte sur les enjeux du big data pour l'assurance.