

Introduction à la Causalité pour l'actuaire

Un guide pour la modélisation

Fabien FAIVRE - MACIF



SOMMAIRE



01 - La causalité une matière nouvelle pour l'actuaire

02 - Comment formaliser la causalité ?

03 - Pourquoi s'intéresser à la causalité (en tant qu'actuaire, en environnement incertain)?

04 - Pour aller plus loin

01

La causalité, une matière nouvelle pour l'actuaire

01 Décider en univers incertain : au coeur du métier d'actuaire



Éclairer les risques, tracer l'avenir

DÉCOUVRIR L'INSTITUT

DEVENIR ACTUAIRE

SE PERFECTIONNER

CONTRIBUER À L'INSTITUT

SE DOCUMENTER

ESPACE EMPLOI

Accueil > Devenir Actuaire > Qu'est-ce qu'un actuaire ?

Qu'est-ce qu'un actuaire ?

Issu d'une formation de haut niveau en mathématiques appliquées mais aussi en droit, économie et comptabilité, un actuaire est un expert de l'évaluation, de la modélisation et de la gestion des risques.

Les actuaires exercent une grande variété de métiers dans de nombreux secteurs d'activité, publics et privés, tels que l'assurance, la réassurance, les mutuelles, la prévoyance, les retraites, le contrôle, la banque et gestion d'actifs, l'industrie, le conseil, l'audit, la recherche académique ou en entreprise et l'enseignement.

Commission ERM

Objectif : Veille de l'actualité ERM, échanges sur les problématiques et définition de règles de bonnes pratiques ERM / conformité.

Présidente : Isabelle Delestrade

Vice-Présidents : Nicolas Marescaux et Pierre Valade

Nombre de membres : 16

GT rattachés :

- Captives
- Cartographie des risques
- Gestion et allocation du capital
- Anticiper en univers incertain



01 La causalité ne fait pas partie du cursus de formation des actuaires



=> Cette présentation se veut surtout une introduction

1ère année L3 Licence

2ème année Master M1

3ème année Master M2

- Méthodes probabilistes en actuariat
- Mathématiques financières et marchés financiers
- Informatique actuarielle
- Outils mathématiques pour l'actuaire
- Actuariat et opérations d'assurance
- Comptabilité financière
- Microéconomie
- Macroéconomie
- Droit civil
- Anglais

Cours complémentaires DU

- Mathématiques et probabilités
- Bases de finance
- Informatique (Python, Excel et VBA, C++, intro à R)
- Economie
- Droit et comptabilité
- Actuariat-Gestion

Stage 1 mois

Fiches bibliographiques
Fiches de cours 2023-2024

Stage 3 m

Possibilité l'étranger
Fiches bibliographiques
Fiches de cours 2023-2024

Matière	ECTS	
Actuariat de l'assurance non-vie	4	
Actuariat de l'assurance-vie	4	
Foundations of Risk management	3	18+0
Risk Theory	2	15+6
Théorie microéconomique appliquée à l'assurance	3	18+0

INSTITUT DE STATISTIQUE DE SORBONNE UNIVERSITÉ

Programme Master 2

Master 2

SEMESTRE 3

UE SCIENCES DES DONNÉES (9 ECTS)

- Machine Learning
- Big Data et Assurance

UE ASSURANCE II (6 ECTS)

- Comptabilité en assurance
- Prévoyance sociale et retraite
- Réassurance

UE RISK MANAGEMENT (6 ECTS)

- Théorie des valeurs extrêmes
- ALM
- Analyse de survie
- Gestion quantitative des risques

UE ENVIRONNEMENT ACTUARIEL (6 ECTS)
4 cours au choix parmi :

- Data visualisation
- MCEV
- Block-chain
- Compléments non-vie (sur les 2 semestres)
- Solvabilité II

CAMB

FORMATION PROGRAMME

UE fondamentales S4

- Règlementation en assurance : Solvency II
- Gestion actif-passif d'une société d'assurance
- Retraite et engagements sociaux
- Principe de l'assurance IARD
- Méthodes numériques en finance
- Apprentissage statistique et Monte-Carlo accéléré pour le calcul du SCR en assurance vie
- Prévoyance et santé

UE complémentaires S4

- Machine Learning
- Démographie et tables de mortalité
- Séries temporelles et applications actuarielles
- Gestion de données massives avec Hadoop et Spark en assurance

UE facultatives S4

- Deep learning avec Python
- Data project

Bloc mémoire

- Mémoire

01 La reconnaissance académique des approches causales

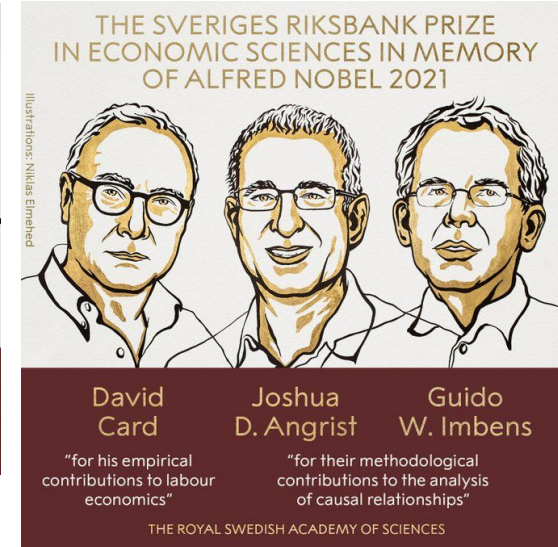


prix Alan Turing
2011



Judea Pearl

prix Nobel économie
2019, 2021

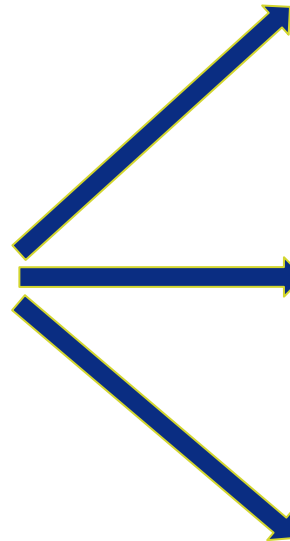


01 Pourquoi s'intéresser à la causalité ?

en tant qu'actuaire, en environnement incertain



La **causalité** comme un **guide mental** pour la modélisation



Pour mieux appréhender les apports et limites des **nouvelles technologies** (LLMs, Agents etc.)

Pour disposer des bons outils lorsqu'on cherche à **agir sur un événement**

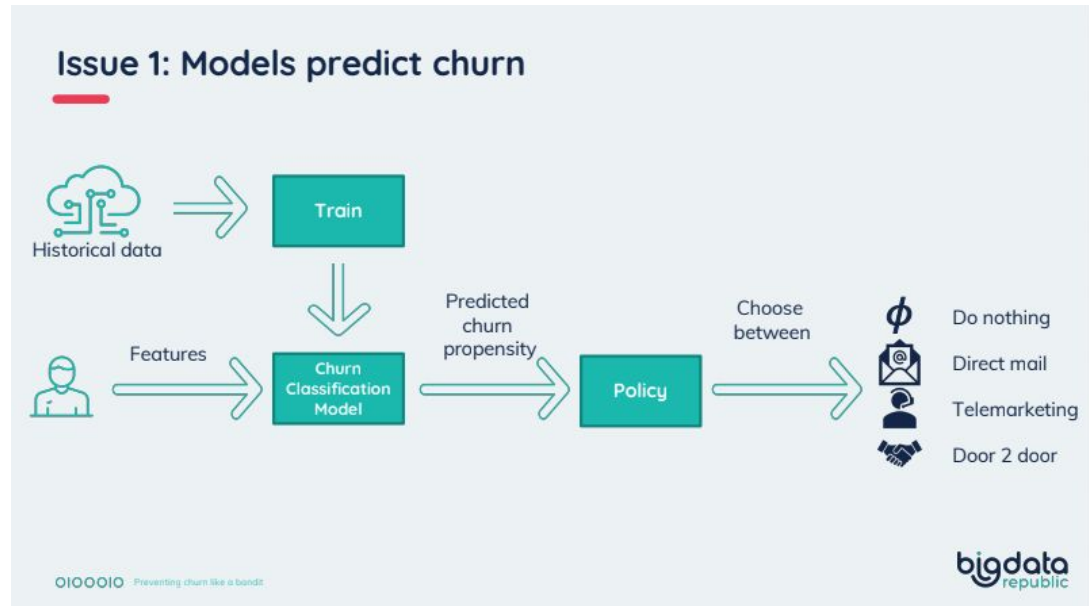
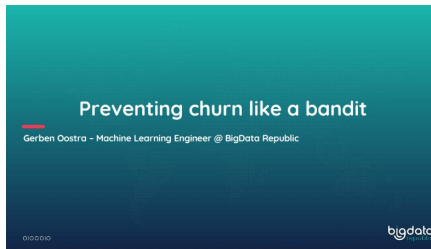
Pour mieux **communiquer** avec des experts d'autres corps de métiers / entre actuaires



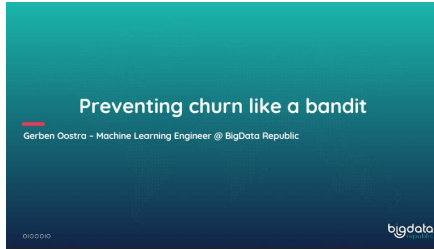
01 Comment j'ai découvert la causalité ?



Découverte des sujets **marketing** à l'occasion de la création d'une Direction Réponse Besoin Sociétaires et Innovation regroupant notamment des activités Marketing, Offre et Data.



01 Comment j'ai découvert la causalité ?



Les approches prédictives
ne sont pas le bon outil
dans ce cas !

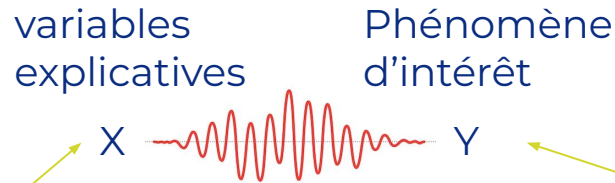


source : <https://bigdatarepublic.nl/articles/preventing-churn-bandit/>

01 Pourquoi s'intéresser à la causalité ?



Principalement pour identifier des relations de causes à effet de manière à pouvoir **agir sur les causes** afin de **favoriser un effet désiré**



Si je veux agir sur les causes pour influencer Y, les méthodes **causales** sont nécessaires

Si je veux agir sur la base d'une anticipation d'un phénomène sans essayer de le modifier, les méthodes **prédictives** suffisent (mais la causalité peut aider)

02

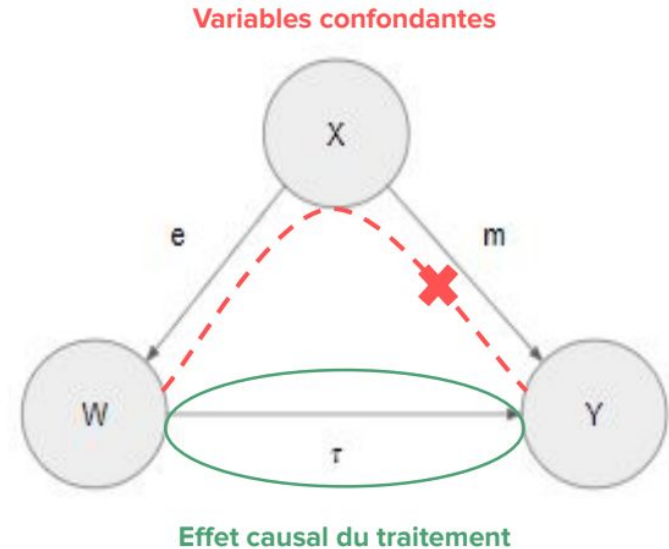
Comment modéliser la causalité ?

02 La causalité invite à réfléchir sur le processus qui a généré les données

La corrélation seule ne suffit pas pour agir



“You are smarter than your data. Data do not understand causes and effects; humans do.” *Judea Pearl, The Book of Why*



02 Techniquement, un formalisme assez simple



L'objet de base de la causalité est un **SCM** (Structural Causal Model) :

- un choix de variables **endogènes** qu'on choisit de modéliser
- une variable aléatoire dite **exogène** pour chaque variable endogène
- les variables exogènes sont **indépendantes** entre elles
- une série d'assignations au travers de **fonctions déterministes**.

$$\begin{aligned} T &\leftarrow f_t(u_t) \\ Y &\leftarrow f_y(T, u_y) \end{aligned}$$

$$u_t \perp u_y$$

02 Techniquement, un formalisme assez simple

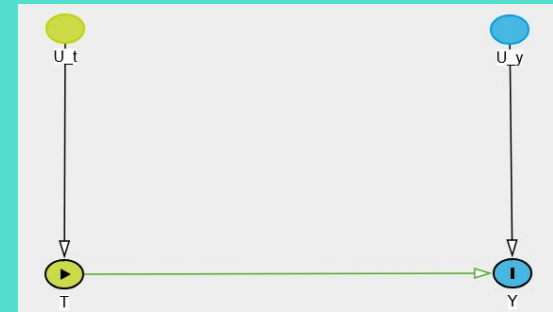


De ce SCM nous pouvons notamment déduire une relation entre ses variables (**le graphe causal**) sous la forme d'un Graph Acyclique Orienté (Directed Acyclic Graph ou **DAG** en anglais):

- Le choix des variables à modéliser ou non et de la structure peut se faire par mobilisation de connaissances AVANT d'avoir vu les données
- Néanmoins le choix d'une structure a des implications dont certaines sont vérifiables sur des jeux de données

$$T \leftarrow f_t(u_t)$$
$$Y \leftarrow f_y(T, u_y)$$

$$u_t \perp u_y$$



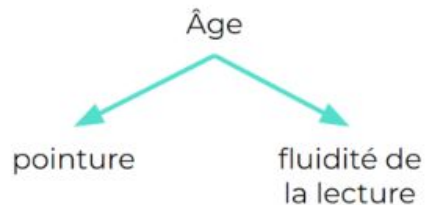
02 Un graph causal admet 3 structures de base

Qui ont des conséquences vérifiables sur les données

Chaîne /
médiation



Bifurcation



Facteur
de
collision



population
générale :

$alarme \perp feu$

$pointure \perp fluence$

$beauté \perp talent$

population
spécifique :

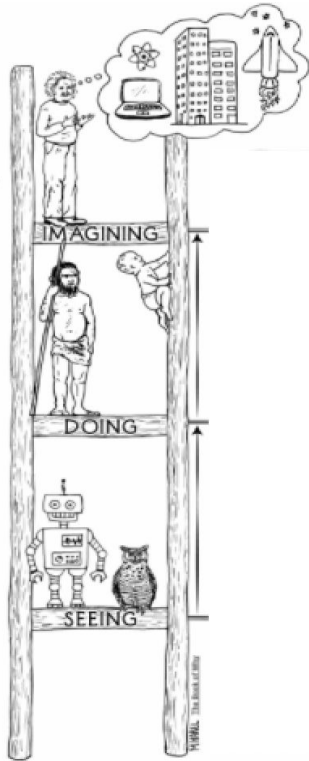
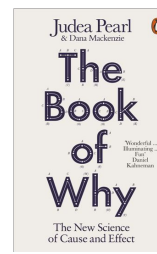
$alarme \perp feu \mid fumée$

$pointure \perp fluence \mid âge$

$beauté \perp talent \mid célébrité$

02 Que peut on faire quand on a un SCM ?

L'échelle de causalité de Pearl



Typologie de questions	Outil mathématique	Nature de données
<p>3. Contrefactuel - Penser l'existant modifié</p> <p>Si j'avais agi différemment, quel aurait été le résultat? Si je n'avais pas pris l'aspirine, est-ce que j'aurais toujours mal à la tête ?</p>	$P(y_x x', y')$	
<p>2. Intervention - Agir sur le monde</p> <p>Que serait Y, si je faisais X ? Si je prends de l'aspirine, est-ce que ma douleur s'arrêtera ?</p>	$P(y do(x) z)$	Expérimentales / Observationnelles
<p>1. Observation - Etre passif</p> <p>Dans quelle mesure, observer X change ma croyance sur Y? Est ce que le symptôme X est lié / associé à la maladie Y ?</p>	$P(y x)$	Observationnelles

02 L'inférence causale

L'inférence causale propose une grammaire et une série de méthodes pour :

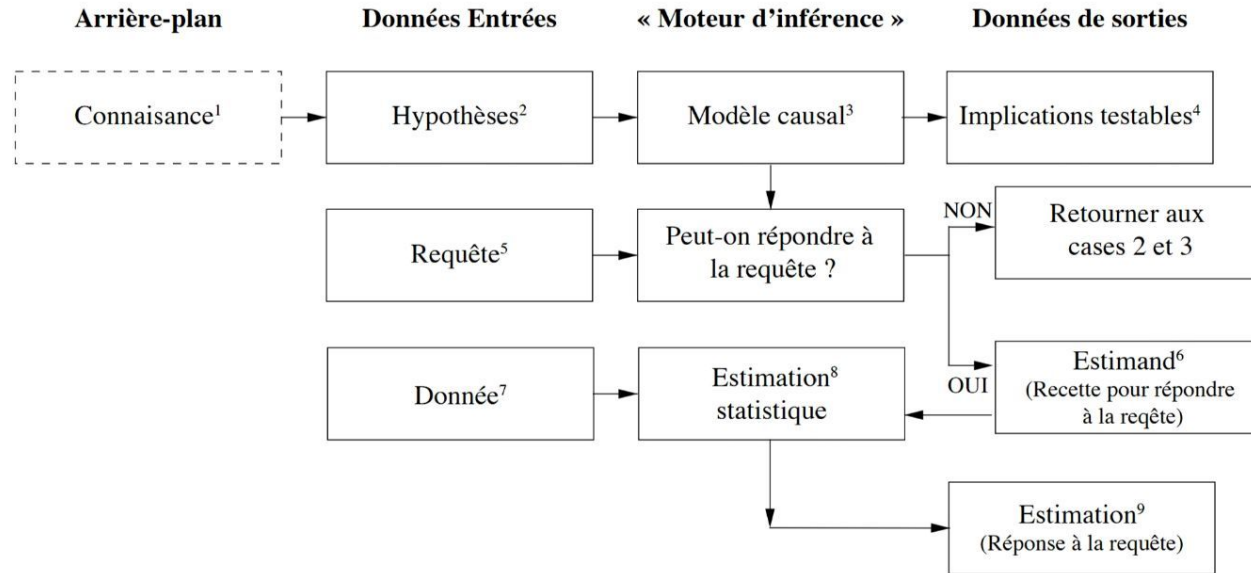
- estimer l'effet d'une action (une “**intervention**”) dans le cadre d'une **expérimentation**
- dans certains cas, **anticiper le résultat d'une expérimentation** à partir de données **observationnelles**

=> Particulièrement utile lorsqu'il n'est pas possible pour des raisons éthiques ou pratiques de forcer les paramètres d'un système

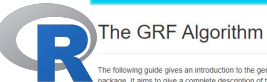
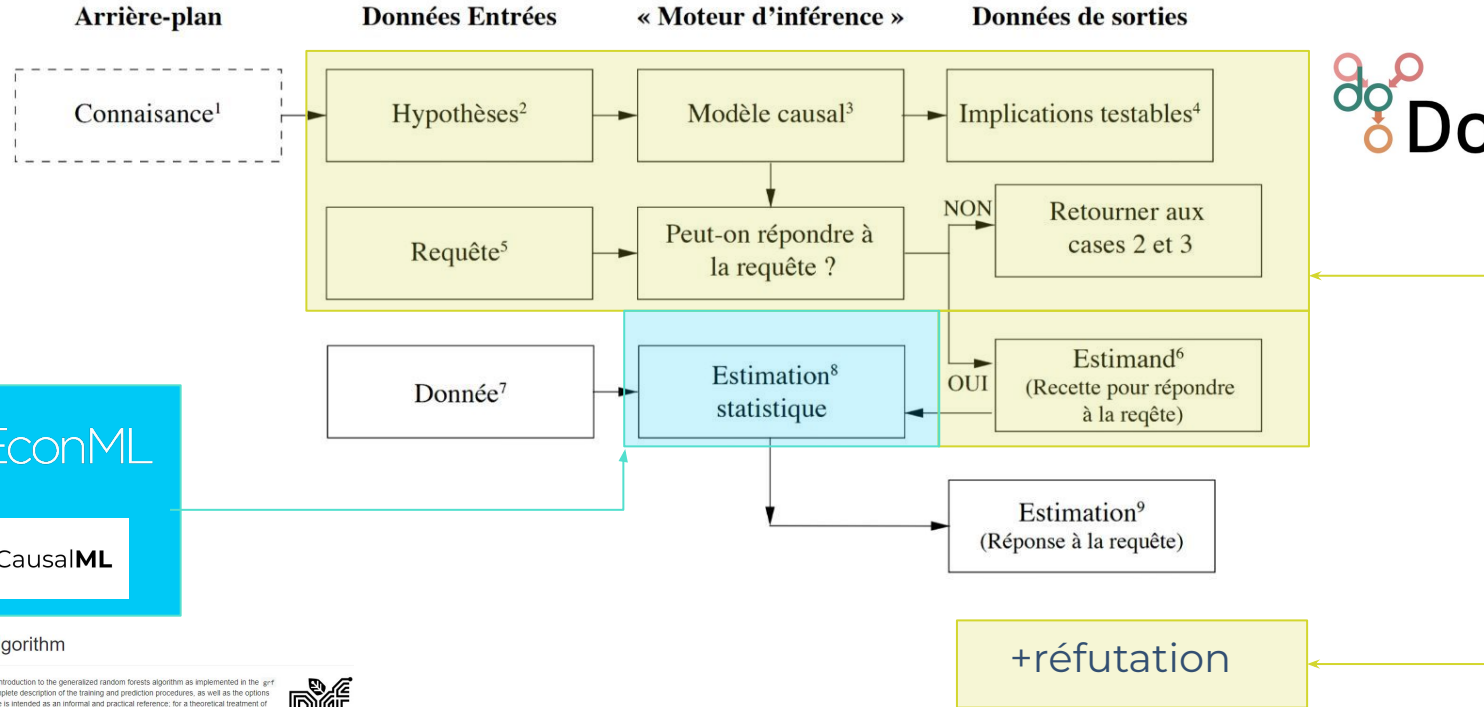
Ex : “Est-ce que le tabagisme favorise le cancer ?”

- là où une expérimentation ne permet d'identifier qu'un effet moyen de l'intervention, certaines méthodes permettent d'inférer des variations individuelles au traitement, ce qui ouvre la voie à une logique de **personnalisation**

02 Processus d'analyse : les données ne sont plus le point de départ !



02 Outillage



The following guide gives an introduction to the generalized random forests algorithm as implemented in the `grf` package. It aims to give a complete description of the training and prediction procedures, as well as the options available for tuning. This guide is intended as an informal and practical reference; for a theoretical treatment of GRF, please consult the "Generalized Random Forests" paper.

GRF extends the idea of a classic random forest to allow for estimating other statistical quantities besides the expected outcome. Each forest type, for example `quantile_forest`, trains a random forest targeted at a particular problem, like quantile estimation. The most common use of GRF is in estimating treatment effects through the function `causal_forest`.



03

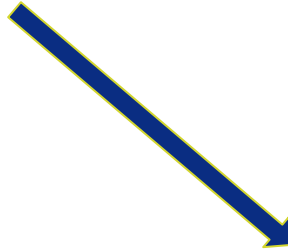
Pourquoi s'intéresser à la causalité en tant qu'actuaire, en environnement incertain ?

03 Pourquoi s'intéresser à la causalité ?

en tant qu'actuaire, en environnement incertain



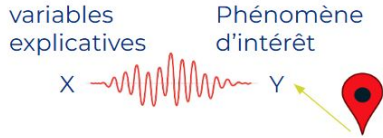
La **causalité** comme
un **guide mental**
pour la modélisation



Pour mieux **communiquer** avec
des experts d'autres corps de
métiers / entre actuaires

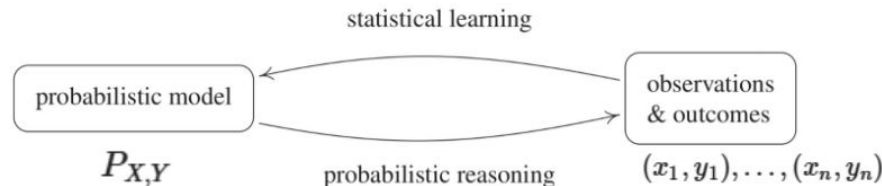


03 L'actuaire est souvent confronté à des besoins de modélisation prédictive

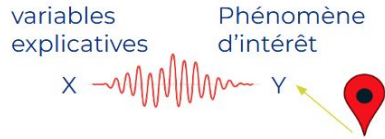


- **Pour faire des prédictions sur une grandeur qui impacte les activités de la compagnie, mais sur laquelle cette dernière n'a que peu d'influence :**
 - quelle sera la valeur du CAC40 à la fin de l'année ?
 - quels sont les scénarios probables de tempêtes sur l'année ?
 - est-ce que cette commune sera reconnue en état de CatNat au titre de la sécheresse ?
- **Et établir des stratégies de gestion des risques qui n'influent pas le phénomène de départ :**
 - quelle stratégie de composition du portefeuille d'actifs ?
 - quelle structure de réassurance mettre en place ?
 - quel niveau de prudence retenir dans les provisions ?

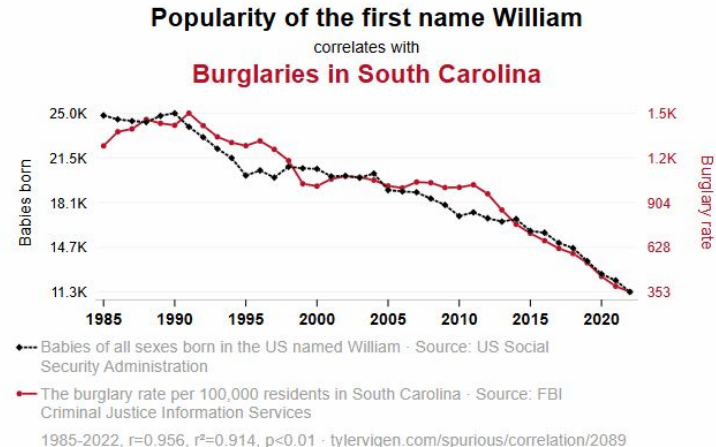
- **Outils :**



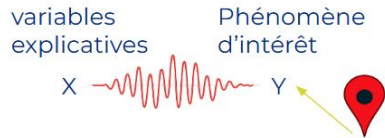
03 Processus de modélisation



- **processus (actuaire junior) :**
 - récupération des données disponibles
 - modèle statistique
 - sélection du modèle qui présente le meilleur ajustement aux données de test
- **=> le modèle ajusté risque de tirer partie de relations fortuites (mais qui peuvent être stables dans le temps)**
 - garanties en cas de changement de régime ?
 - risque de surinterprétation du rôle d'une variable (“**mythologie**” autour du modèle)

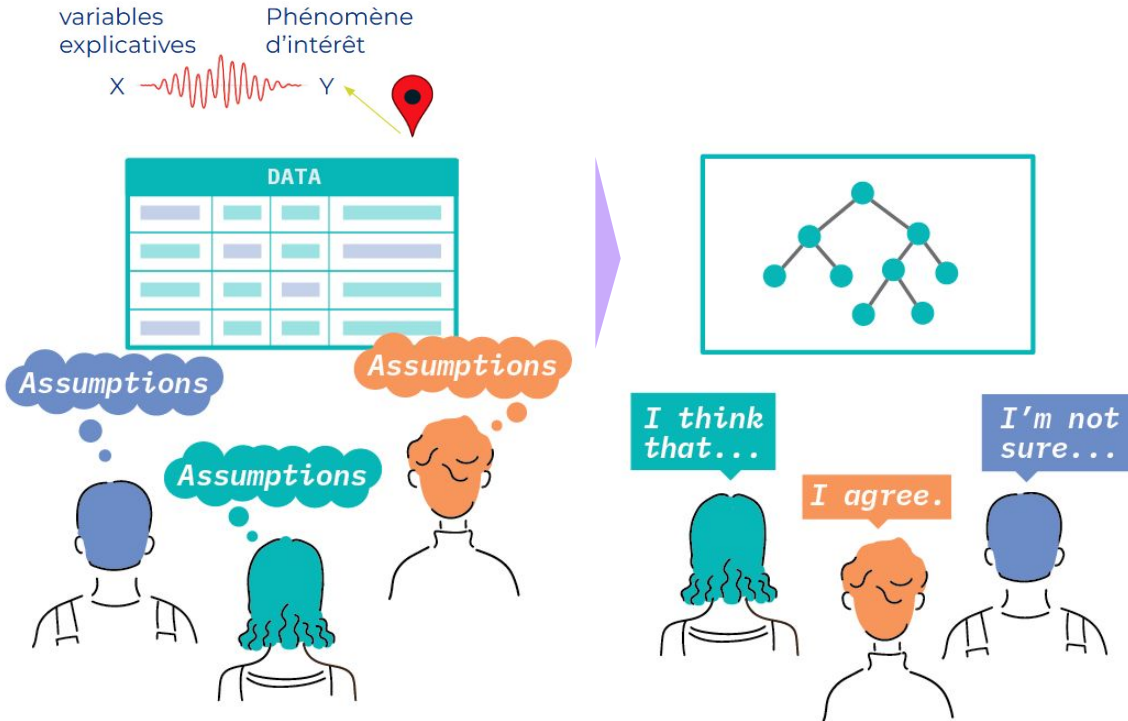


03 Pourquoi cette situation se retrouve peu en pratique ?



- **Le processus suivi par un actuaire senior diffère**
- **processus (actuaire senior) :**
 - précision du besoin
 - *identification de données a priori pertinentes (en liaison avec des experts)*
 - sélection d'un modèle apportant de la valeur tout en intégrant d'autres contraintes (robustesse, maintenabilité etc)
- **Nous mobilisons déjà un modèle causal du monde (souvent implicite) pour justifier nos choix de variables a priori**
 - Nous faisons déjà implicitement des hypothèses, comment les rendre explicites ?
 - Comment échanger sur ces hypothèses lorsque le consensus n'est pas garanti / complètement compris (*environnement incertain*) ?

03 Raison 1 : le graph causal comme guide a priori de modélisation

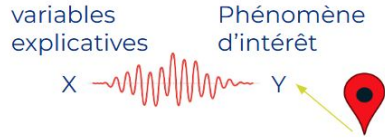


We all make causal assumptions when looking at data.

Making a causal graph makes those assumptions explicit.

- Le graph causal peut être discuté avec des experts
- Particulièrement adapté à la modélisation de process
- La construction du graph devrait démarrer simplement et ne rajouter des variables qu'au besoin
- Permet d'identifier des variables non observées mais qui seraient utiles

03 Exemple d'utilisation : anticipation des arrêts CatNat sécheresse



Provisionnement CatNat RGA -
Adaptation des modèles de charge ultime à la circulaire du 6 mai 2024

Benoit Chaput,
(Macif)



Fabien Faivre,
(Macif)



Luigia Ripani,
(Axiorable)



21 novembre 2024

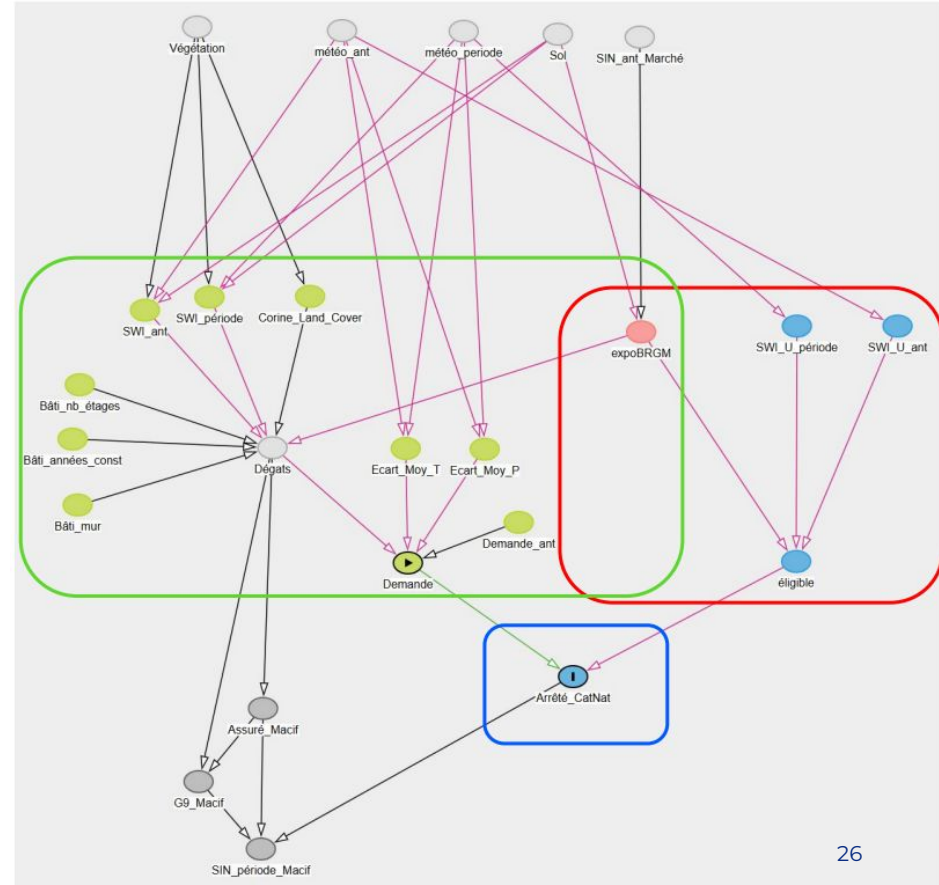
1

En s'inspirant des **approches causales**, il est possible de représenter sous formes d'un graphe notre compréhension du phénomène.

Étape implicite dans la plupart des projets

3 composants clefs se détachent :

- un modèle comportemental de dépôt de dossier par la mairie
- un modèle de règles d'éligibilité
- la réalité des deux modèles précédents n'étant observables qu'à la publication des décisions d'arrêt CatNat

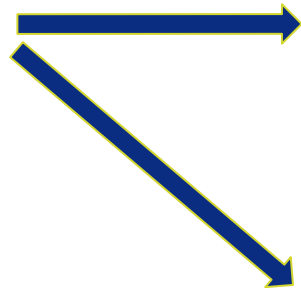


03 Pourquoi s'intéresser à la causalité ?

en tant qu'actuaire, en environnement incertain



La **causalité** comme un **guide mental** pour la modélisation



Pour disposer des bons outils lorsqu'on cherche à **agir sur un événement**



Pour mieux **communiquer** avec des experts d'autres corps de métiers / entre actuaires



03 La prévention quantitative : prochain terrain de jeu des actuaires ?



Sous le Sapin, un PNACC-3 Ambitieux, mais des Lacunes à Comblir



Thierry Langreney

Président des Ateliers du Futur, ONG Action Climatque | EFRAG
TIAP member | UNEP TIP member







Cependant, plusieurs de nos propositions clés restent sans suite à ce stade :

1. Maximiser le retour sur subventions octroyées par le FPRNM en hiérarchisant les impacts et en adoptant une démarche proactive vis à vis des acteurs clé (les collectivités des zones les plus exposées). Ceci est d'autant plus critique que le budget de l'Etat est sous pression et les événements climatiques à la hausse...





03 Comment mesurer l'impact d'un traitement ?



	X	T	Y
	X_a	1	1
	X_b	1	0
	X_b	0	0
	X_a	0	1





03 Comment mesurer l'impact d'un traitement ?



	X	T	Y = Y(T)	Y(1)	Y(0)
	X_a	1	1	1	
	X_b	1	0	0	
	X_b	0	0		0
	X_a	0	1		1





03 Comment mesurer l'impact d'un traitement ?



	X	T	Y = Y(T)	Y(1)	Y(0)
	X_a	1	1	1	?
	X_b	1	0	0	?
	X_b	0	0	?	0
	X_a	0	1	?	1





03 Comment mesurer l'impact d'un traitement ?



	X	T	Y = Y(T)	Y(1)	Y(0)	Y(1) - Y(0)
	X_a	1	1	1	?	?
	X_b	1	0	0	?	?
	X_b	0	0	?	0	?
	X_a	0	1	?	1	?

03 Comment mesurer l'impact d'un traitement ?







	X	T	Y = Y(T)	Y(1)	Y(0)	Y(1) - Y(0)
	X_a	1	1	1	?	?
	X_b	1	0	0	?	?
	X_b	0	0	?	0	?
	X_a	0	1	?	1	?

ITE $\tau_i = Y_i |do(T = t_1) - Y_i |do(T = t_0)$

03 Comment mesurer l'impact d'un traitement ?



$$ATE = E[\tau_i]$$

	X	T	Y = Y(T)	Y(1)	Y(0)	Y(1) - Y(0)
	X_a	1	1	1	?	?
	X_b	1	0	0	?	?
	X_b	0	0	?	0	?
	X_a	0	1	?	1	?

ITE $\tau_i = Y_i |do(T = t_1) - Y_i |do(T = t_0)$

03 Comment mesurer l'impact d'un traitement ?



$$E[Y | T=1] - E[Y | T=0] \stackrel{?}{=} E[Y(1) - Y(0)]$$

La plupart du temps, **non** !

Nécessite

- 1- Echangeabilité
- 2- Positivité
- 3- Non confusion
- 4- Absence d'interférence

	X	T	Y = Y(T)	Y(1)	Y(0)	Y(1) - Y(0)
	X_a	1	1	1	?	?
	X_b	1	0	0	?	?
	X_b	0	0	?	0	?
	X_a	0	1	?	1	?

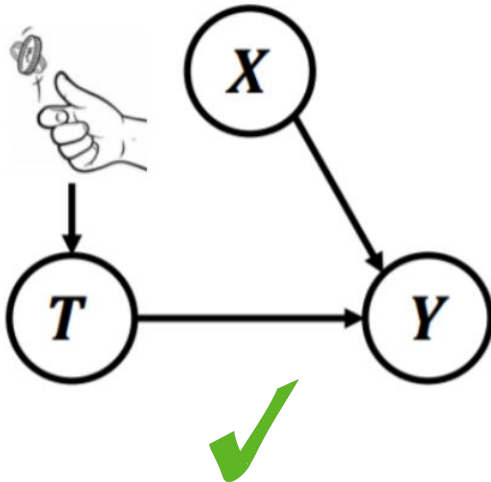
03 Comment mesurer l'impact d'un traitement ?



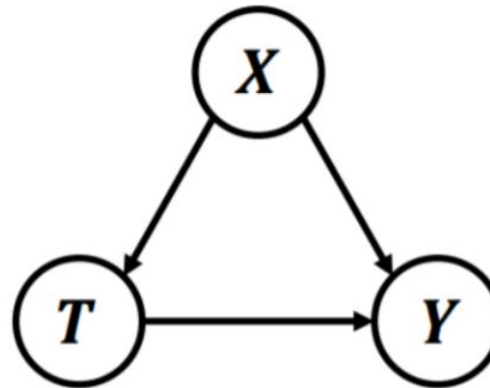
$$E[Y | T=1] - E[Y | T=0] = E[Y(1) - Y(0)]$$

La plupart du temps, **non** !

Données
expérimentales



Données
observationnelles



? hypothèses à vérifier / graph

Nécessite

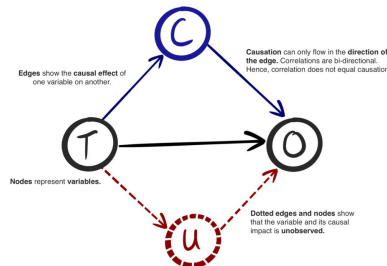
- 1- Echangeabilité
- 2- Positivité
- 3- Non confusion
- 4- Absence d'interférence

03 Inférence Causale + Machine Learning = Causal ML



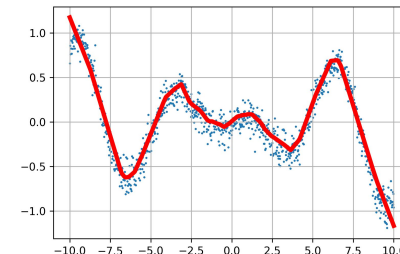
Utilisation du cadre causal pour

- **expliquer nos hypothèses** sur le phénomène analysé
- **identifier les variables sur lesquelles conditionner** ou non dans un exercice de régression (pour estimer un effet causal ou éviter les corrélations fallacieuses)



Utilisation du ML

- Pour estimer la forme et l'intensité des relations
- **En lien avec le cadre causal**



03 A l'aide des méthodes de Causal ML il est possible d'aller vers de la personnalisation en estimant le CATE (Conditional Average Treatment Effect)



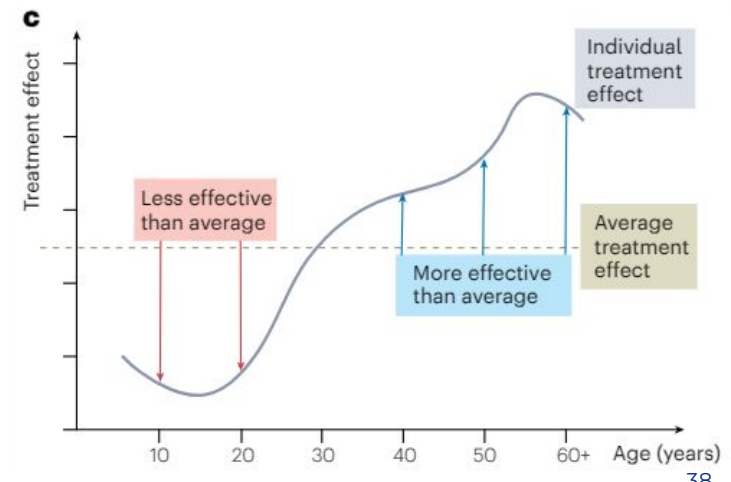
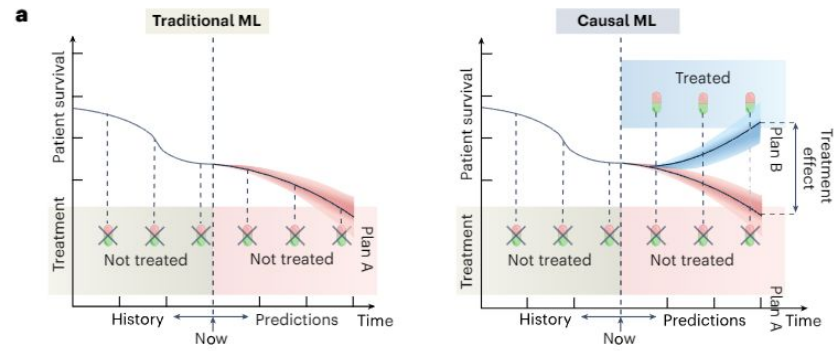
b

Traditional ML				Causal ML				
Patient	Covariates	Treatment	Patient outcome	Patient	Covariates	Treatment	Patient outcome	
							If not treated	If treated
1	Age, sex, etc.	0	-1.0	1	Age, sex, etc.	0	-1.0	
2	↓	1	2.3	2	↓	1		2.3
3	↓	1	0.3	3	↓	1		0.3

Patient	Covariates	Treatment	Patient outcome	Patient	Covariates	Potential outcomes		Treatment effect	
						If not treated	If treated	If treated	If not treated
1	Age, sex, etc.	1	?	1	Age, sex, etc.	?	?	?	?
2	↓	0	?	2	↓	?	?	?	?

□ Missing observations ? Prediction targets

$$CATE = E[Y_{1i} - Y_{0i} | X = x]$$



source : https://www.nature.com/articles/s41591-024-02902-1.epdf?sharino_token=BHCH9LTMdPwdfTcmLTVNqN0IAiWel9inR3Z0Tv0N0aZozK20IAXuHdNNUVLZW9GQdhrFtrUWyzISNnk8WZvL8hy9SX3V7uRnT4nguv7VCnVcoZ7sq134RckCdr7J0ILZosiTLuLcunlOs-n8Doto7A5seeAvrKtwAKMk3D

03 Comment valider les résultats ?

Approche naïve : essayer d'appliquer les mêmes pratiques qu'en ML



Evaluating uplift models

Predicted treatment effect

Actual treatment effect

$$\frac{1}{n} \sum_i^n (\hat{\tau}(\mathbf{x}_i) - \tau(\mathbf{x}_i))^2$$

- **Problème : $\tau(\mathbf{x}_i)$ n'est pas observable**

[1] Shalit, U., Johansson, F.D. and Sontag, D., 2017, July. Estimating individual treatment effect: generalization bounds and algorithms. In International Conference on Machine Learning (pp. 3076-3085). PMLR.
[2] Saito, Y. and Yasui, S., 2020, November. Counterfactual Cross-Validation: Stable Model Selection Procedure for Causal Inference Models. In International Conference on Machine Learning (pp. 8398-8407). PMLR.

03 Méthode par regroupement de populations



Evaluating uplift models: uplift per segment

Le modèle de **CATE** sert à **ordonner** les populations

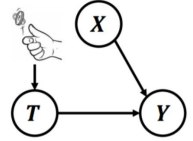


High predicted CATE

Low predicted CATE

[1] Radcliffe, N.J., 2007. Using control groups to target on predicted lift: Building and assessing uplift models. Direct Marketing Analytics Journal, 1(3), pp.14-21.
[2] Gutierrez, P. and Gérardy, J.Y., 2017, July. Causal inference and uplift modelling: A review of the literature. In International Conference on Predictive Applications and APIs (pp. 1-13). PMLR.

03 Méthode par regroupement de populations



Evaluating uplift models: uplift per segment

Sample means

$$E[Y|T=1] - E[Y|T=0]$$

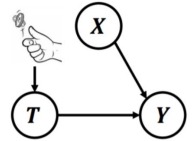


High predicted CATE

Low predicted CATE

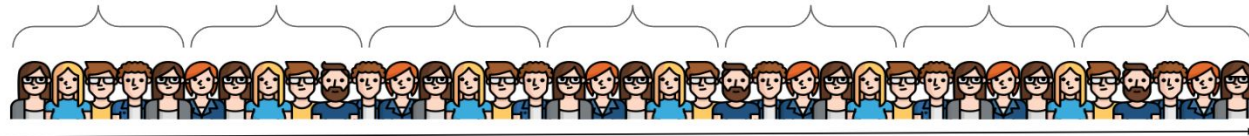
- [1] Radcliffe, N.J., 2007. Using control groups to target on predicted lift: Building and assessing uplift models. Direct Marketing Analytics Journal, 1(3), pp.14-21.
- [2] Gutierrez, P. and Gérardy, J.Y., 2017, July. Causal inference and uplift modelling: A review of the literature. In International Conference on Predictive Applications and APIs (pp. 1-13). PMLR.
- [3] Devriendt, Floris, et al. "Learning to rank for uplift modeling." *IEEE Transactions on Knowledge and Data Engineering* 34.10 (2020): 4888-4904.

03 Méthode par regroupement de populations



Evaluating uplift models: uplift per segment

$$E[Y|T=1] - E[Y|T=0]$$

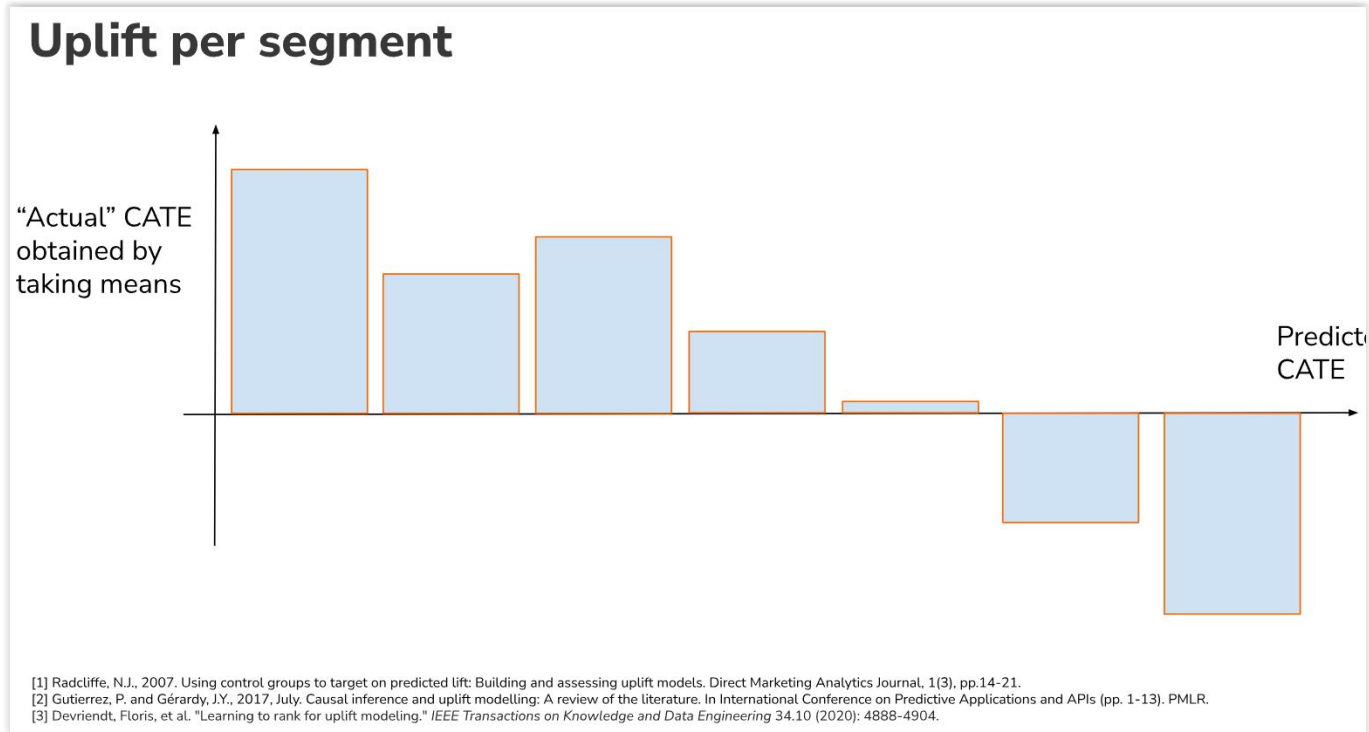
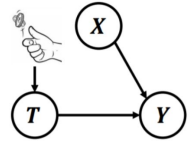


High predicted CATE

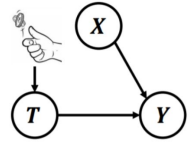
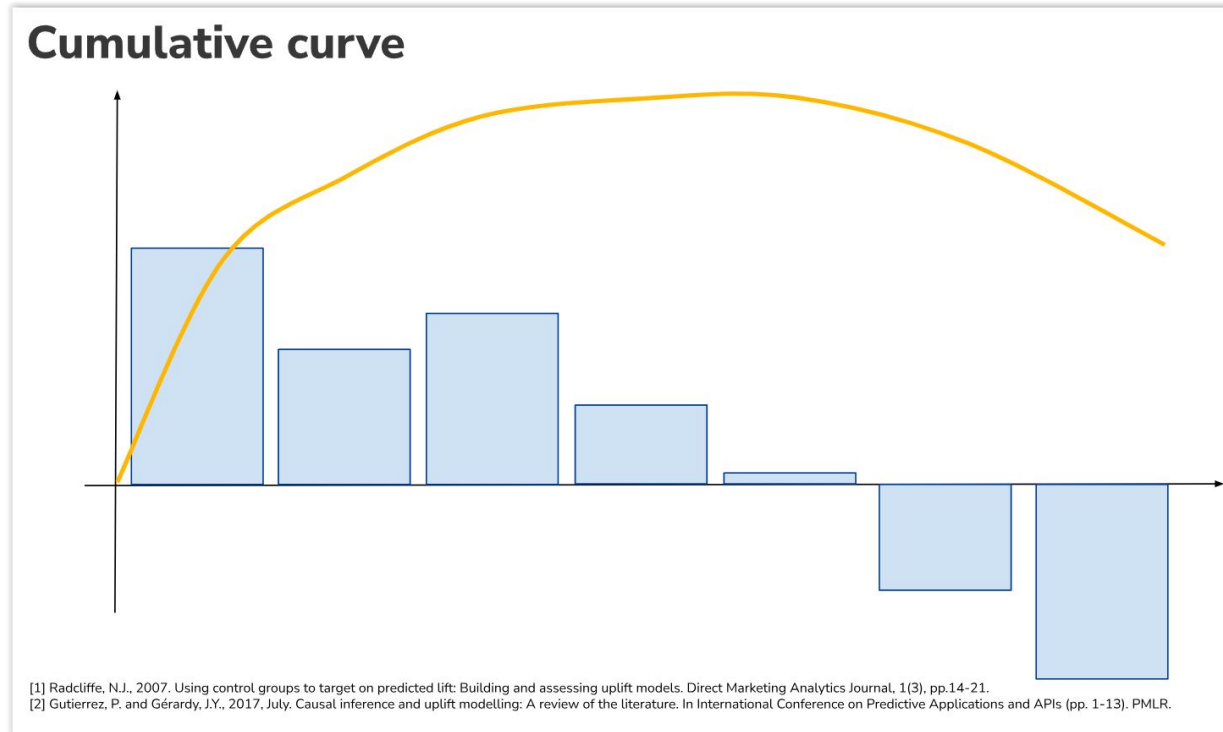
Low predicted CATE

- [1] Radcliffe, N.J., 2007. Using control groups to target on predicted lift: Building and assessing uplift models. *Direct Marketing Analytics Journal*, 1(3), pp.14-21.
- [2] Gutierrez, P. and Gérardy, J.Y., 2017, July. Causal inference and uplift modelling: A review of the literature. In *International Conference on Predictive Applications and APIs* (pp. 1-13). PMLR.
- [3] Devriendt, Floris, et al. "Learning to rank for uplift modeling." *IEEE Transactions on Knowledge and Data Engineering* 34.10 (2020): 4888-4904.

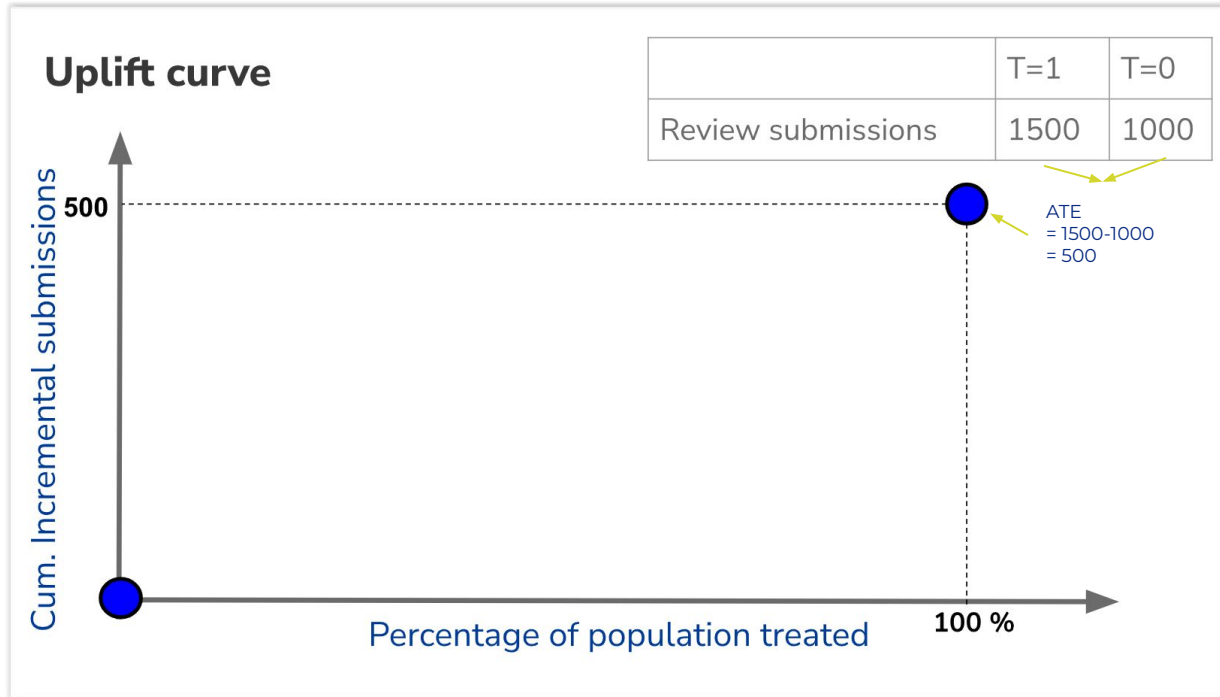
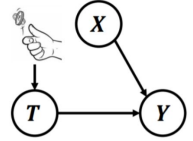
03 Méthode par regroupement de populations



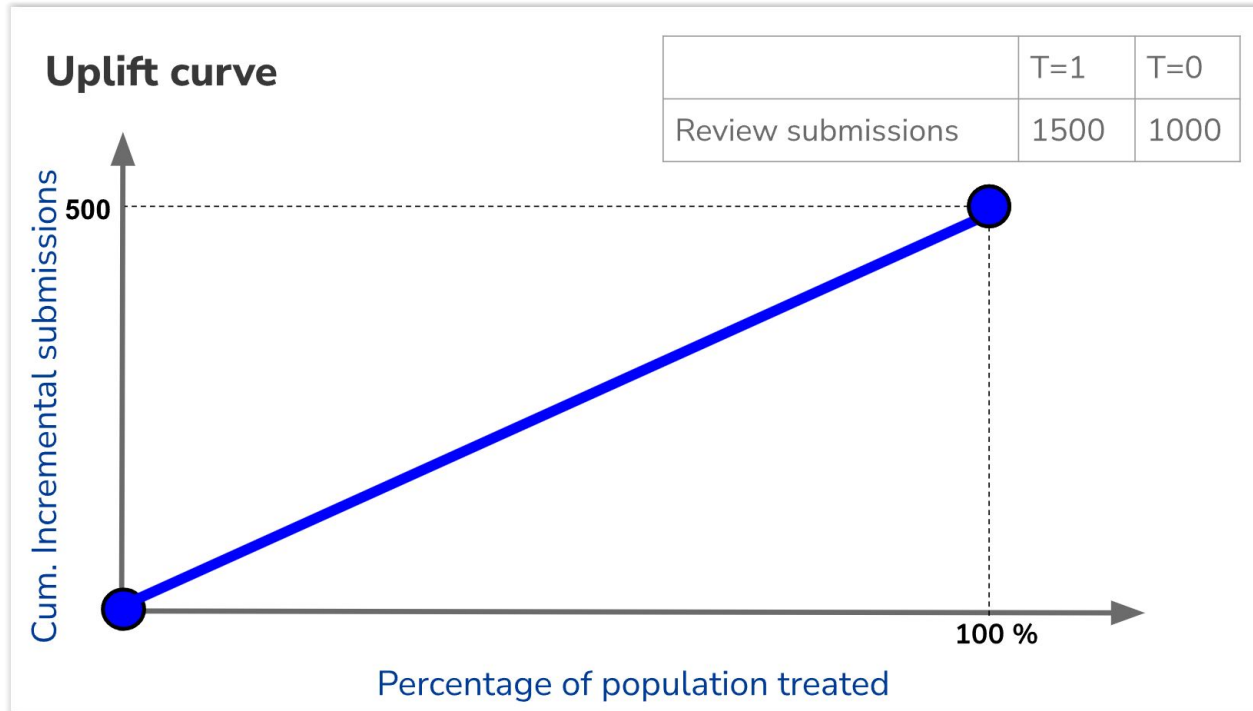
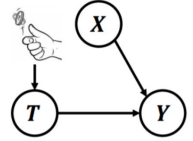
03 Méthode par regroupement de populations



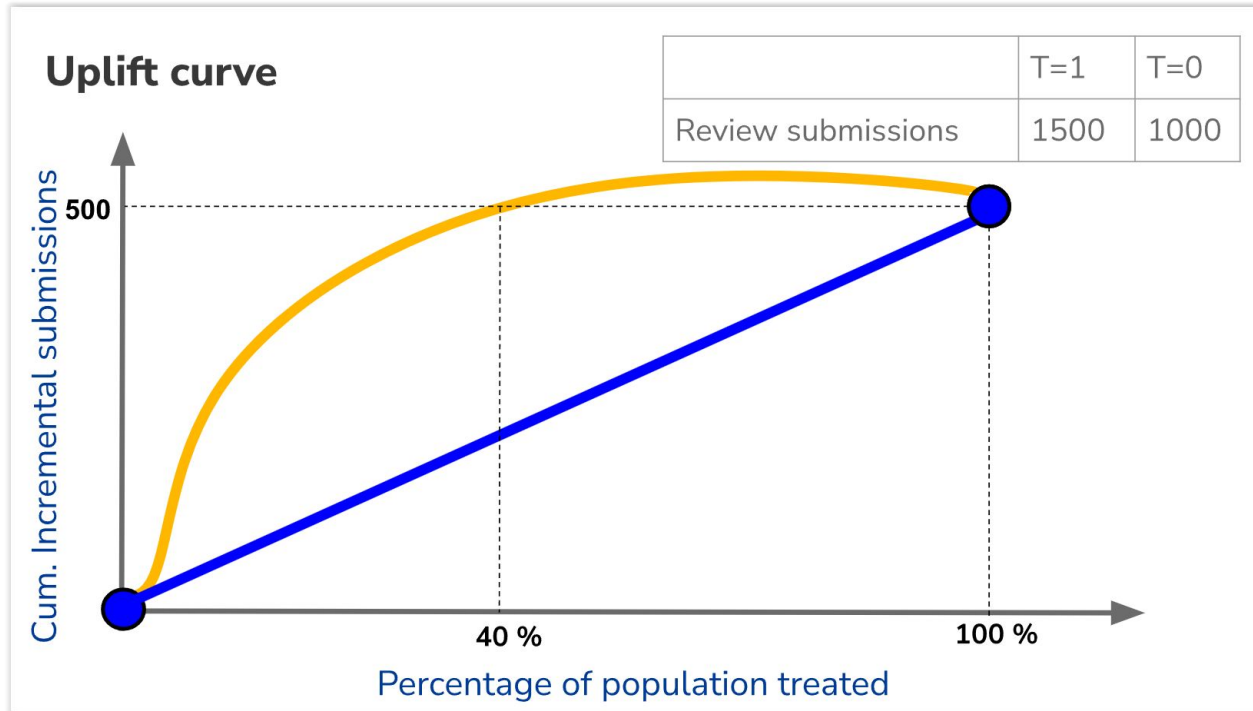
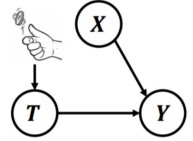
03 Uplift curve



03 Uplift curve

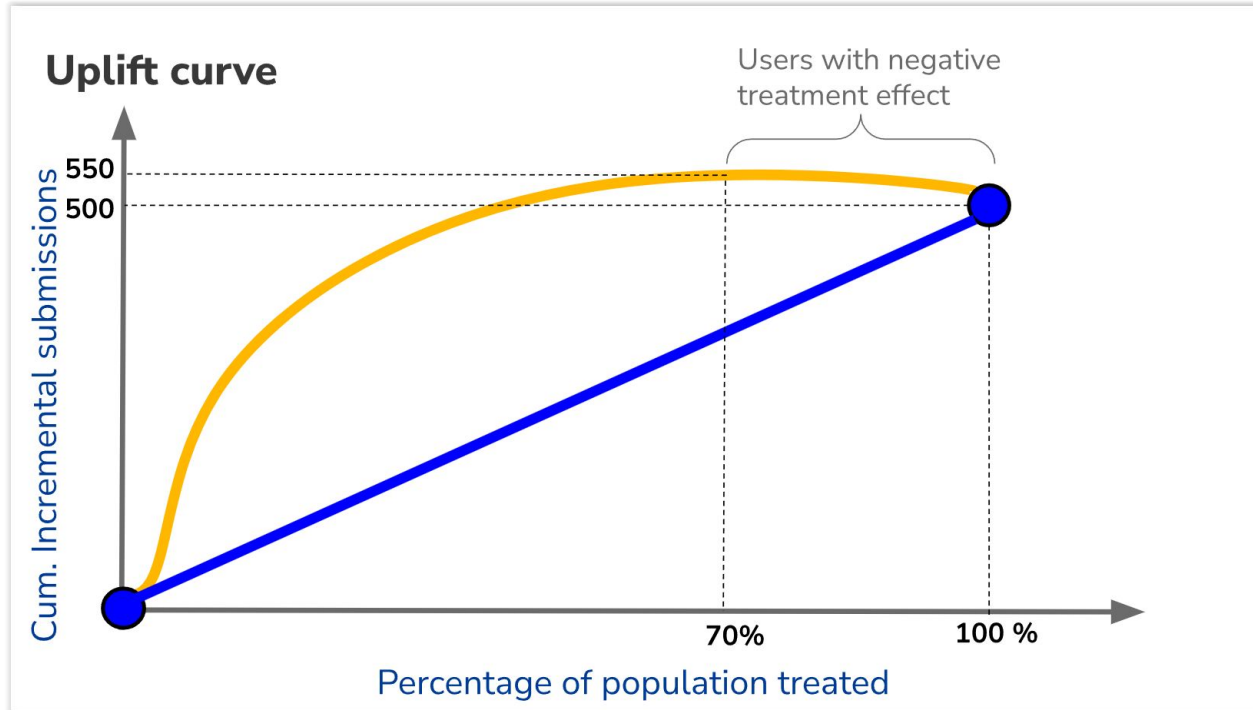
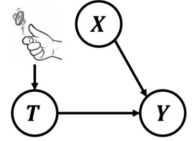


03 Uplift curve

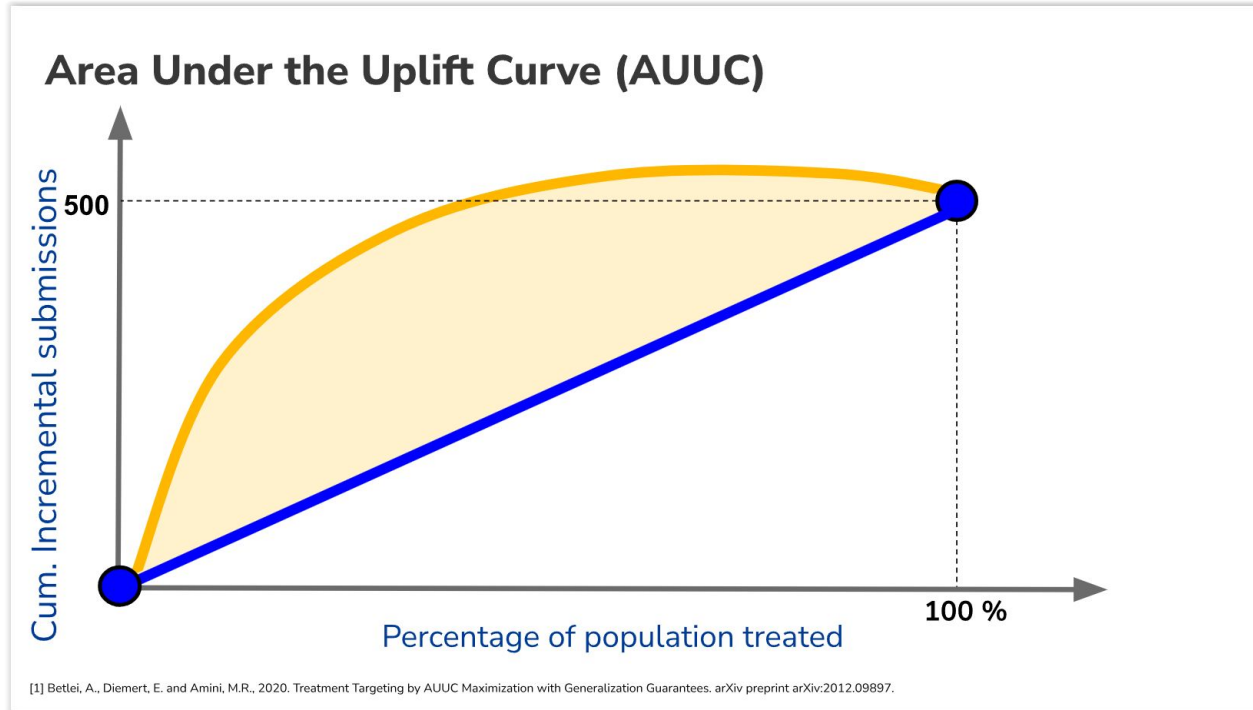
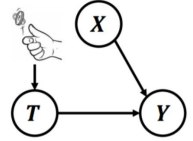


source : <https://sites.google.com/view/uplift-modeling-cikm23?pli=1>

03 Uplift curve



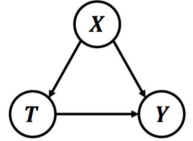
03 Area Under the Uplift curve



03 Quelle adaptation dans le cas de données observationnelles ?



- Il n'est **plus possible** de comparer directement les valeurs observées sur les jambes traitées et contrôle
- Garder en tête que le modèle de CATE appris sur les données de train sert à **ordonnancer** une priorisation de traitements
- L'impact en termes d'effet du traitement sur le jeu de test pour la proportion à traiter peut être **estimé via une mesure doublement robuste** (CAIPW ou TMLE)

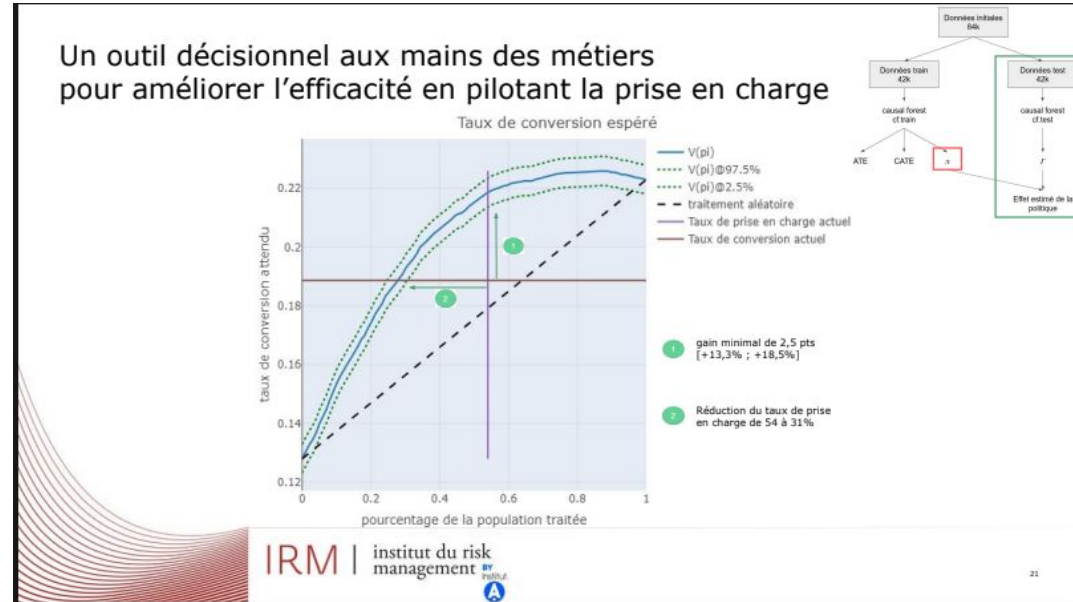
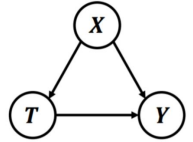
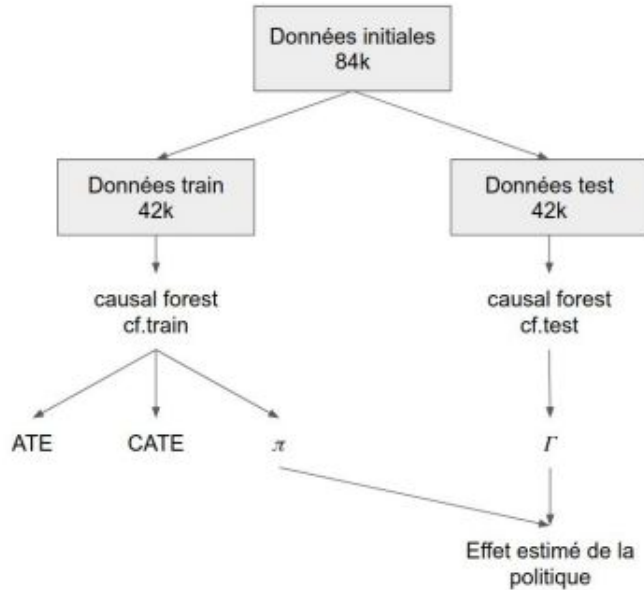


Athey & Wager 2020 : <https://arxiv.org/pdf/1702.02896>

Bookdown : <https://bookdown.org/stanfordgsbsilab/ml-ci-tutorial/policy-learning-i---binary-treatment.html#parametric-policies>

03 Quelle adaptation dans le cas de données observationnelles ?

Exemple de mise en oeuvre

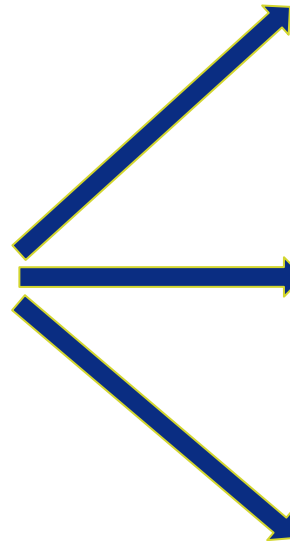


03 Pourquoi s'intéresser à la causalité ?

en tant qu'actuaire, en environnement incertain



La **causalité** comme un **guide mental** pour la modélisation



Pour mieux appréhender les apports et limites des **nouvelles technologies** (LLMs, Agents etc.)

Pour disposer des bons outils lorsqu'on cherche à **agir sur un événement**

Pour mieux **communiquer** avec des experts d'autres corps de métiers / entre actuaires

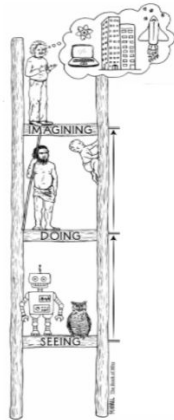
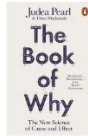


03 La causalité pour comprendre les possibilités et limites des LLMs



02 Que peut on faire quand on a un SCM ?

L'échelle de causalité de Pearl



Typologie de questions	Outil mathématique	Nature de données
3. Contrefactuel - Penser l'existant modifié Si j'avais agi différemment, quel aurait été le résultat? Si je n'avais pas pris l'aspirine, est-ce que j'aurais toujours mal à la tête ?	$P(y x', y')$	
2. Intervention - Agir sur le monde Que serait Y, si je faisais X ? Si je prends de l'aspirine, est-ce que ma douleur s'arrêtera ?	$P(y do(x), z)$	Expérimentales / Observationnelles
1. Observation - Etre passif Dans quelle mesure, observer X change ma croyance sur Y? Est ce que le symptôme X est lié / associé à la maladie Y ?	$P(y x)$	Observationnelles

http://bayes.cs.ucla.edu/jp_home.html

15

Les LLMs **ne sont pas soumis** à l'échelle de causalité de Pearl !

Les LLMs ont accès à une **nouvelle source de données** : des textes écrits par des humains qui contiennent des modèles causaux du monde !

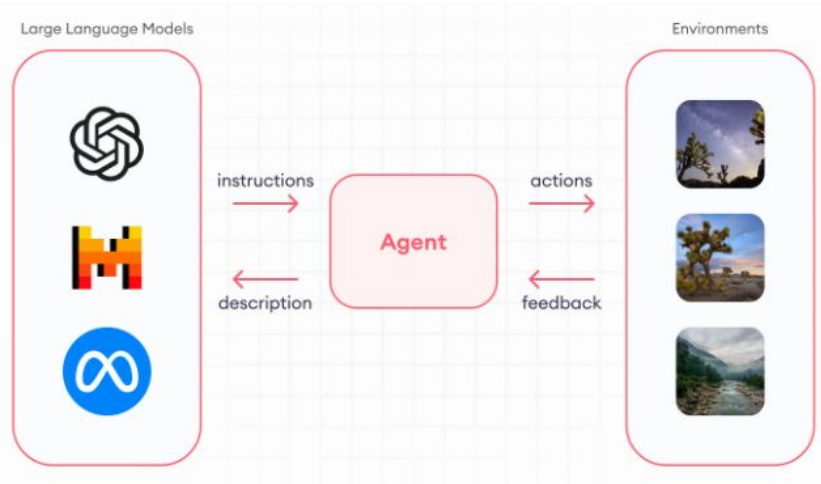
“Les LLMs sont une **soupe de rumeurs** à propos de **modèles causaux du monde**”

J. Pearl

source :

<https://www.youtube.com/watch?v=vqKJ9pUQ6Q8&t=1059s>

03 La causalité pour comprendre les possibilités et limites des agents



source : <https://www.superannotate.com/blog/llm-agents>

Les Agents sont soumis au risque de confusion



Les Agents robustes, c'est-à-dire soumis à suffisamment de dérives de distribution, apprennent un modèle causal (approché) du monde



Comments: ICLR 2024 (oral). Updated agents section, new corollary

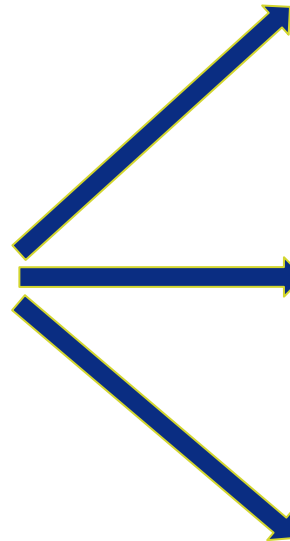
03 Pourquoi s'intéresser à la causalité ?

en tant qu'actuaire, en environnement incertain



+ root cause analysis, fairness, attribution, dérives...

La **causalité** comme un **guide mental** pour la modélisation



Pour mieux appréhender les apports et limites des **nouvelles technologies** (LLMs, Agents etc.)



Pour disposer des bons outils lorsqu'on cherche à **agir sur un événement**



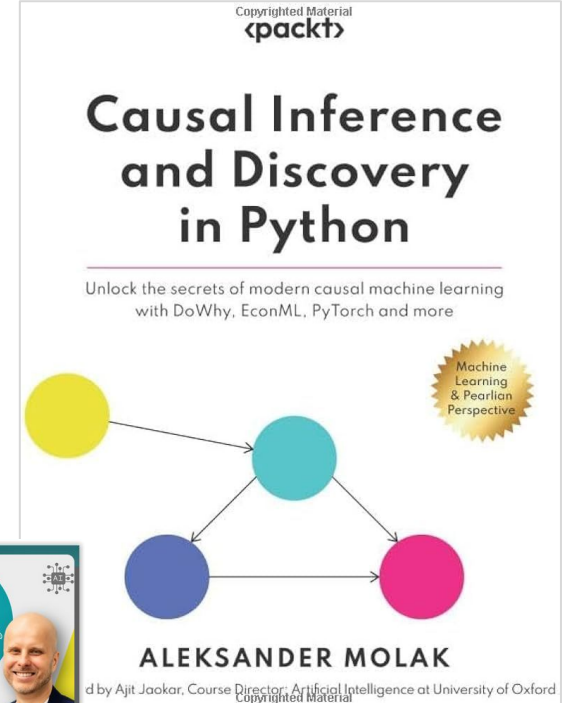
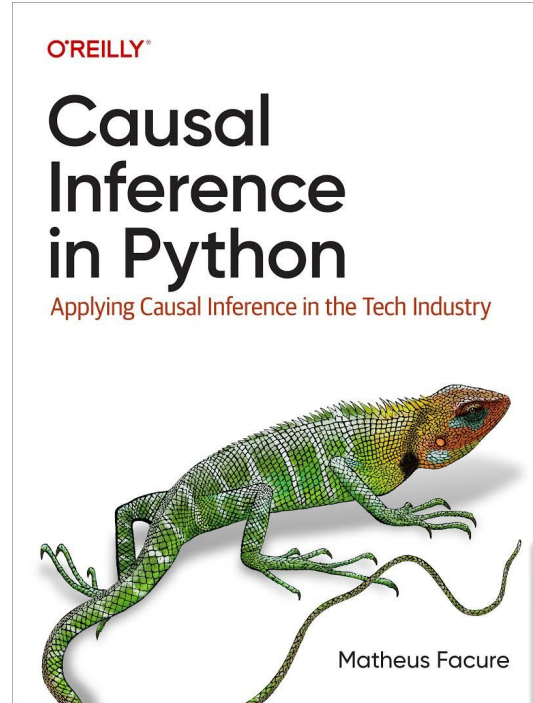
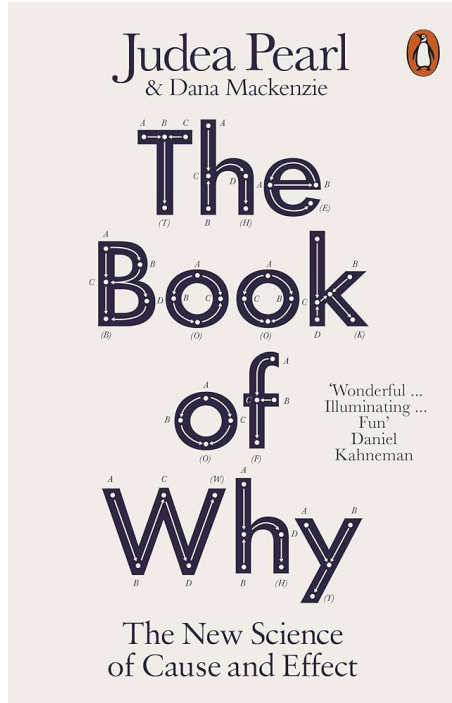
Pour mieux **communiquer** (avec des experts d'autres corps de métiers / entre actuaires)



04

Pour aller plus loin

04 Pour aller plus loin



04 Pour aller plus loin



datacraft*

- [Alessandro Leite](#), chercheur à l'INRIA
- [Audrey Poinso](#), actuellement en thèse sur le thème de la causalité
- [Fabien Faivre](#), Responsable R&D et Innovation Data à la MACIF
- [Marianne Clause](#), Professeure Université Lorraine
- [Myriam Tami](#), Enseignante-chercheuse en statistical learning, Université Paris Saclay, CentraleSupélec
- [Georges Oppenheim](#), ancien Professeur Université Paris Saclay

14

February
Tue,
2023

ATELIER – LES APPORTS DE L'APPROCHE CAUSALE : EXPLORATION SUR DES CAS D'USAGE CONCRETS

Atelier animé par le groupe de travail 'Causalité'

16

March
Thu,
2023

ATELIER – CAUSAL DISCOVERY : PARTONS À LA DÉCOUVERTE DU GRAPHE CAUSAL

Atelier animé par le groupe de travail 'Causalité'.

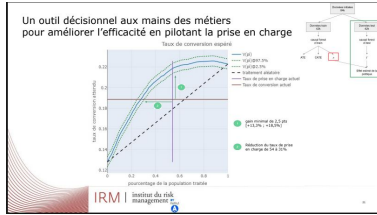
26

June
Wed,
2024

ATELIER – CAUSALITÉ : L'INFÉRENCE CAUSALE

Atelier animé par Marianne Clause, professeur, Université de Lorraine, Fabien Faivre, Responsable R&D et Innovation Data, MACIF et Georges Oppenheim, Professeur, Université Paris-Est Marne-la-Vallée

04 Pour aller plus loin



Améliorer la performance d'un processus marketing grâce à l'uplift modelling et aux modèles causaux (24/05/2022)

<https://www.institutdesactuaires.com/agenda/dsa-certificat-management-en-data-science-pour-l-actuariat-3154>



Encadrement de l'IA et renouveau des questions de discrimination dans les travaux actuariels (17/11/2022)

https://www.youtube.com/watch?v=sq_F0haA02i&list=PL8lXSworDvbYVbTnS5cPP0opmzvO2UGfy&index=28



Provisionnement CatNat RGA -
Adaptation des modèles de charge
ultime à la circulaire du 6 mai 2024

Benoit Chaput,
(Macif)



Fabien Favre,
(Macif)



Luigia Ripani
(Avonabile)



avonabile

21 novembre 2024

1

Provisionnement CatNat RGA - Adaptation des modèles de charge ultime à la circulaire du 6 mai 2024 (21/11/2024)

https://www.institutdesactuaires.com/global/aene/link.php?doc_id=19234&fq=1



Merci !

CONTACT :

Fabien FAIVRE

Responsable R&D et Innovation Data
Data Office
Direction Réponse Besoins des Sociétaires et Innovation

ffavre@macif.fr