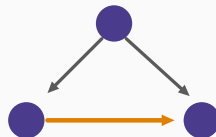


# Inférence causale et machine learning

## Chapitre 2 : Le cadre des résultats potentiels



La notation

La décomposition du biais

Les hypothèses

Ce qu'il faut retenir

## La notation

---

# Deux mondes pour chaque individu

## Résultats potentiels

Pour chaque individu  $i$  :

- $Y_i(1)$  — ce qu'il obtiendrait **s'il était traité**;
- $Y_i(0)$  — ce qu'il obtiendrait **s'il ne l'était pas**.

L'effet causal individuel est

$$\tau_i = Y_i(1) - Y_i(0)$$

## Le problème fondamental

On n'observe **jamais** les deux. Si  $T_i = 1$ , on voit  $Y_i(1)$ ; sinon,  $Y_i(0)$ .

$$Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0)$$

L'effet individuel n'est jamais observable. C'est un problème de données manquantes, et il est irréparable au niveau individuel.

### Intuition

Elle sépare ce qu'on **veut** de ce qu'on **voit**. Écrire  $Y_i(1) - Y_i(0)$  définit l'objet d'intérêt sans rien supposer sur les données.

C'est le geste décisif : **d'abord dire ce qu'on cherche, ensuite examiner si les données permettent de l'atteindre**. La démarche inverse — estimer d'abord, interpréter après — est la source de la plupart des erreurs.

## Les effets moyens

| Paramètre | Définition                  | Population      |
|-----------|-----------------------------|-----------------|
| ATE       | $E[Y(1) - Y(0)]$            | tout le monde   |
| ATT       | $E[Y(1) - Y(0) \mid T = 1]$ | les traités     |
| ATU       | $E[Y(1) - Y(0) \mid T = 0]$ | les non traités |

## Lequel intéresse le décideur

L'ATT répond à : le programme a-t-il servi à ceux qui l'ont suivi ?

L'ATE répond à : faudrait-il l'étendre à tous ?

Ce sont deux questions distinctes, et les deux réponses diffèrent dès que l'effet varie d'un individu à l'autre. **Annoncer un chiffre sans préciser lequel des trois est incomplet.**

## La décomposition du biais

---

## La décomposition

$$\underbrace{E[Y \mid T=1] - E[Y \mid T=0]}_{\text{différence observée}} = \underbrace{ATT}_{\text{ce qu'on cherche}} + \underbrace{E[Y(0) \mid T=1] - E[Y(0) \mid T=0]}_{\text{biais de sélection}}$$

## Le biais de sélection

C'est l'écart entre traités et non traités **en l'absence de traitement**.

Il vaut zéro si les deux groupes sont comparables au départ. Sinon, la différence observée mélange l'effet et l'écart préexistant.



### Exemple

| Individu | $Y(0)$ | $Y(1)$ | Effet |
|----------|--------|--------|-------|
| 1        | 10     | 14     | +4    |
| 2        | 20     | 22     | +2    |
| 3        | 30     | 36     | +6    |
| 4        | 40     | 44     | +4    |

$$ATE = \frac{4 + 2 + 6 + 4}{4} = 4$$

# Le programme sélectionne les plus favorisés

## Exemple

Supposons que les individus 3 et 4 – ceux dont  $Y(0)$  est le plus élevé – soient traités.

|                            | Calcul        | Valeur    |
|----------------------------|---------------|-----------|
| $E[Y \mid T=1]$            | $(36 + 44)/2$ | 40        |
| $E[Y \mid T=0]$            | $(10 + 20)/2$ | 15        |
| <b>Différence observée</b> |               | <b>25</b> |
| ATT                        | $(6 + 4)/2$   | <b>5</b>  |
| Biais de sélection         | $35 - 15$     | <b>20</b> |

$$25 = 5 + 20$$

La comparaison naïve annonce 25 pour un effet réel de 5. Les quatre cinquièmes du chiffre sont du biais.

## Intuition

Ici, les meilleurs ont été traités : le biais **gonfle** l'effet.

Dans l'exemple du chapitre 1, la formation ciblait les peu qualifiés : le biais **inversait** le signe.

**On ne peut pas corriger un biais dont on ignore le sens.** D'où la nécessité d'un argument d'identification, et non d'un ajustement approximatif.

## Les hypothèses

---

## Indépendance conditionnelle

$$(Y(0), Y(1)) \perp T \mid X$$

À caractéristiques  $X$  égales, l'affectation au traitement est indépendante des résultats potentiels.

On l'appelle aussi *ignorabilité*, ou hypothèse d'absence de confondant non observé.

## Support commun

$$0 < P(T = 1 \mid X) < 1$$

Pour chaque profil  $X$ , il doit exister des traités **et** des non traités. Sans cela, aucune comparaison n'est possible pour ce profil.

## Définition

Deux exigences souvent oubliées :

1. **Pas d'interférence** — le traitement d'un individu n'affecte pas le résultat d'un autre.
2. **Un seul traitement** — la version reçue est la même pour tous.

### Application en gestion

Une formation qui améliore l'employabilité de ses bénéficiaires peut **réduire** les chances des autres candidats : les emplois ne se multiplient pas.

L'effet mesuré sur les traités surestime alors l'effet d'une généralisation — le gain individuel était en partie un déplacement.

C'est l'**effet d'équilibre général**, et c'est la limite la plus sérieuse de tout ce cours. Aucune des méthodes des chapitres 5 à 11 ne le traite.

## Intuition

L'indépendance conditionnelle porte sur des grandeurs **non observées** : les  $Y(0)$  des traités.

Aucun test ne la valide. Elle se défend par l'argument, le contexte institutionnel, et des vérifications indirectes.

C'est pourquoi une bonne étude causale consacre plus de place à justifier ses hypothèses qu'à présenter ses coefficients.



Ce qu'il faut retenir

---

## Cinq points

1.  $Y_i(1)$  et  $Y_i(0)$  définissent l'effet; on n'en observe **jamais** qu'un.
2. **ATE**, **ATT**, **ATU** diffèrent dès que l'effet est hétérogène.
3. Différence observée = **ATT** + **biais de sélection** — ici  $25 = 5 + 20$ .
4. Identifier exige **indépendance conditionnelle**, **support commun** et **SUTVA**.
5. Ces hypothèses ne se **testent pas** : elles se défendent.

## La suite

L'indépendance conditionnelle exige de savoir **quelles** variables conditionner. Le chapitre 3 donne l'outil qui répond — et montre que certains contrôles sont nuisibles.