

# Causal Inference | *Inférence causale*

Yabo Gwladys Vidogbena

24 June/juin 2024

Causal effects | *Les effets causals*

Learning the average causal effect from an experiment | *Comment apprendre l'effet causal moyen d'une expérience aléatoire*

Core assumptions | *Hypothèses clés*

Randomized experiments vs observational studies | *Études aléatoires vs. études observationnelles*



## Causal effects | *Les effets causals*

# What can we learn from an experiment? | *Qu'apprendre d'une expérience?*

The **average causal effect** of the treatment  $D$  on an outcome  $Y$  for the units that you have in your study.

**L'effet causal moyen** du traitement  $D$  sur un résultat  $Y$  pour pour les unités de votre étude.

## What do we mean by “ $D$ causes $Y$ ”? | *Comment interpréter “ $D$ cause $Y$ ”?*

- ▶ We are taking a counterfactual approach.
- ▶ “ $D$  causes  $Y$ ” is a claim about what didn’t happen.
  - ▶ “If  $D$  had not occurred, then  $Y$  would not have occurred.”
  - ▶ “With  $D$ , the probability of  $Y$  is higher than would be without  $D$ .”
- ▶ Nous prenons une approche contrefactuelle.
- ▶ “ $D$  cause  $Y$ ” est une affirmation sur ce qui n’a pas eu lieu.
  - ▶ “Si  $D$  n’avait pas été donné, il n’y aurait pas de  $Y$ .”
  - ▶ “Avec  $D$ , la probabilité de  $Y$  est plus élevée qu’elle ne le serait sans  $D$ .”

## What do we mean by “ $D$ causes $Y$ ”? | *Comment interpréter “ $D$ cause $Y$ ”?*

- ▶ “ $D$  causes  $Y$ ” requires a *context*.
  - ▶ Small classrooms improve test scores but require experienced teachers and funding.
- ▶ “ $D$  causes  $Y$ ” doesn’t mean “ $W$  does not cause  $Y$ .”
- ▶ “ $D$  cause  $Y$ ” nécessite un *contexte*.
  - ▶ Les petites salles de classe améliorent les résultats des examens mais nécessitent des enseignants expérimentés et un financement adapté.
- ▶ “ $D$  cause  $Y$ ” n’implique pas nécessairement que  $W$  ne cause pas  $Y$ .

## Key subtleties | *Les subtilités clés*

- ▶ We must define what the treatment means and what it means to not have the treatment.
- ▶ The units in your study might or might not be representative of a larger population.
- ▶ Nous devons définir la condition de traitement et de contrôle pour définir un effet causal.
- ▶ Les unités de votre étude peuvent ou non être représentatives d'une population plus large.

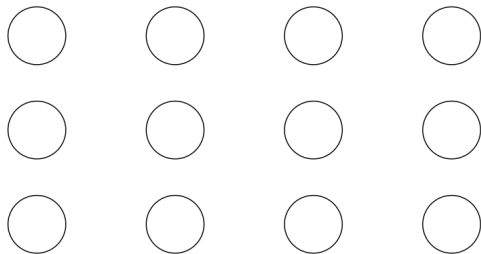


Learning the average causal effect from an experiment |  
*Comment apprendre l'effet causal moyen d'une expérience  
aléatoire*

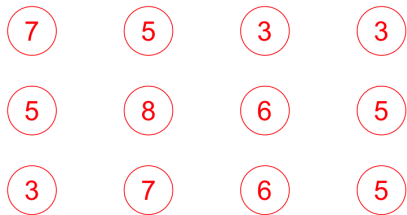
## Potential outcomes | Résultats potentiels

- ▶ For each unit we assume that there are two *post-treatment* outcomes:  $Y_i(1)$  and  $Y_i(0)$ .
  - ▶  $Y_i(1)$  is the outcome that *would* obtain *if* the unit received the treatment ( $D_i = 1$ ).
  - ▶  $Y_i(0)$  is the outcome that *would* obtain *if* the unit received the control ( $D_i = 0$ ).
- ▶ Pour chaque unité, nous supposons qu'il y a deux résultats *post-traitement* :  $Y_i(1)$  et  $Y_i(0)$ .
  - ▶  $Y_i(1)$  est le résultat qu'on *obtiendrait* si l'unité recevait le traitement ( $D_i = 1$ ).
  - ▶  $Y_i(0)$  est le résultat qu'on *obtiendrait* si l'unité recevait le contrôle ( $D_i = 0$ ).

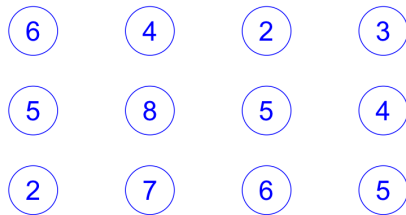
We have 12 units | *Nous avons 12 unités*



Each unit has both  $Y_i(1)$  and  $Y_i(0)$  | *Chaque unité a à la fois  $Y_i(1)$  et  $Y_i(0)$*



true average of  $Y(1) = 5.25$

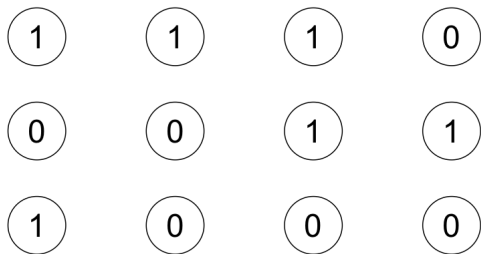


true average of  $Y(0) = 4.75$

So each unit has a treatment effect  $\tau_i = Y_i(1) - Y_i(0)$  | Alors, pour chaque unité, il y a un effet de traitement  $\tau_i = Y_i(1) - Y_i(0)$

1	0	1	0
0	1	0	1
1	0	1	0

Average Treatment Effect / *L'effet moyen du traitement* (ATE) =  
 $\bar{\tau}_i = \overline{Y_i(1) - Y_i(0)}$



true average treatment effect = 0.5

# Fundamental problem of causal inference | *Le problème fondamental de l'inférence causale*

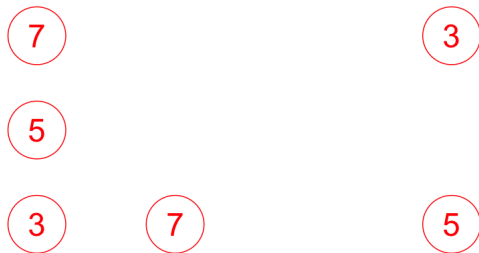
- ▶ **We can't measure the individual-level causal effect**, because we can't observe both  $Y_i(1)$  and  $Y_i(0)$  at the same time.
- ▶ But we can still estimate the **average treatment effect** with a randomized experiment.
- ▶ **Nous ne pouvons pas mesurer l'effet causal au niveau individuel**, car nous ne pouvons pas observer à la fois  $Y_i(1)$  et  $Y_i(0)$  pour un même individu.
- ▶ Mais nous pouvons encore estimer **l'effet moyen du traitement** avec une expérience aléatoire.

Let's go back to the  $Y_i(1)$  | *Revenons à  $Y_i(1)$*

7	5	3	3
5	8	6	5
3	7	6	5



We can take a random sample of these  $Y_i(1)$  | *Sélectionnons un échantillon aléatoire de  $Y_i(1)$*



average  $Y(1)$  of sample #1 = 5

We can take another random sample of these  $Y_i(1)$  | *Sélectionnons un autre échantillon aléatoire de  $Y_i(1)$*



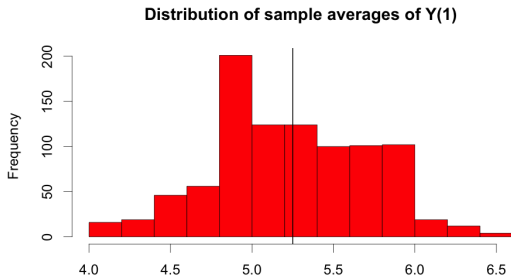
average  $Y(1)$  of sample #2 = 5.5

And another! | *Et un autre!*



average  $Y(1)$  of sample #3 = 5.67

## Average of sample averages | *La moyenne des moyennes de l'échantillon*



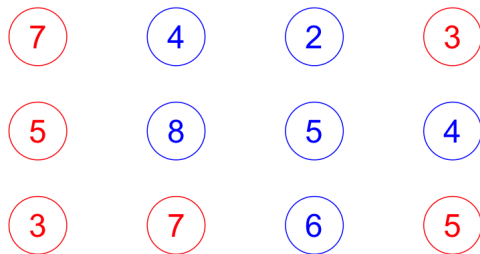
- ▶ A particular sample average will not be the true population average.
- ▶ But the average of all the sample averages will be the true population average!

- ▶ Un échantillon particulier ne nous donnera probablement pas la vraie valeur.
- ▶ Mais la moyenne de toutes les moyennes possibles est la vraie valeur!

## A random assignment | *Une assignation aléatoire*

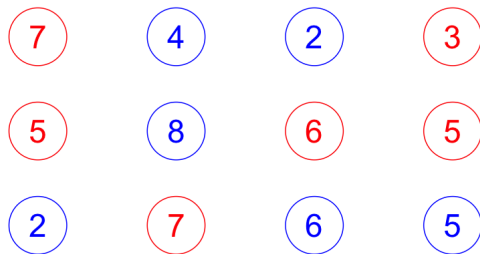
7	4	2	3
5	8	5	4
3	7	6	5

A random assignment produces an estimate | *Une assignation aléatoire produit une estimation*



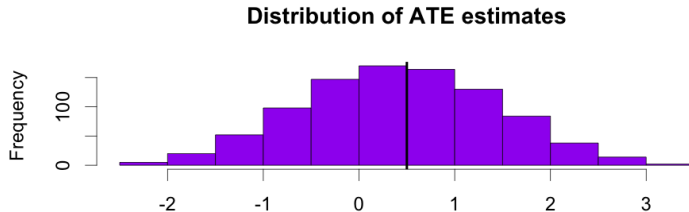
$$5 - 4.83 = 0.17$$

A different random assignment produces another estimate | *Une autre assignation aléatoire génère une autre estimation*



$$5.5 - 4.5 = 1$$

## Randomization is powerful | *Randomiser est puissant*



- ▶ A particular random assignment will probably not give us the true average treatment effect.
- ▶ But the average of the estimates produced by all possible randomizations will be the true average treatment effect!
- ▶ Une assignation aléatoire particulière ne nous donnera probablement pas la vraie valeur de l'effet causal moyen de traitement.
- ▶ Mais la moyenne de toutes les estimations possibles est la vraie valeur.



## Randomization is powerful | *Randomiser est puissant*

Average treatment effect / l'effet moyen du traitement (ATE):

$$\begin{aligned}ATE &= \bar{\tau}_i = \overline{Y_i(1) - Y_i(0)} \\ &= \overline{Y_i(1)} - \overline{Y_i(0)}\end{aligned}$$

1. The average of a difference equals the difference of averages.

1. La moyenne des différences est égale à la différence des moyennes.

## Randomization is powerful | *Randomiser est puissant*

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

2. For the treated units, we observe  $Y_i = Y_i(1)$ .
3. For the control units, we observe  $Y_i = Y_i(0)$ .

2. Pour les unités traitées, nous observons  $Y_i = Y_i(1)$ .
3. Pour les unités au contrôle, nous observons  $Y_i = Y_i(0)$ .

## Randomization is powerful | *Randomiser est puissant*

$$E_R[\overline{Y_i} | D_i = 1] = \overline{Y_i(1)}$$

$$E_R[\overline{Y_i} | D_i = 0] = \overline{Y_i(0)}$$

4. Because the units were randomly assigned to treatment, the average of the  $Y_i = Y_i(1)$  for the treated units is a good estimate of the average of  $Y_i(1)$  for all the units. The same logic applies to the units randomly assigned to control.

4. Parce que les unités ont été assignées de manière aléatoire au traitement, la moyenne de  $Y_i = Y_i(1)$  pour les unités traitées est une bonne estimation de la moyenne de  $Y_i(1)$  pour toutes nos unités. La même logique s'applique pour les unités assignées aléatoirement au contrôle.

## Randomization is powerful | *Randomiser est puissant*

$$ATE = \overline{Y_i(1)} - \overline{Y_i(0)}$$

$$\widehat{ATE} = \overline{(Y_i|D_i = 1)} - \overline{(Y_i|D_i = 0)}$$

5. So we plug in these estimates for  $\overline{Y_i(1)}$  and  $\overline{Y_i(0)}$  to estimate the ATE.

5. Donc nous utilisons ces estimations de  $\overline{Y_i(1)}$  et  $\overline{Y_i(0)}$  pour estimer l'ATE.

## Randomization is powerful | *Randomiser est puissant*

$$ATE = \bar{\tau}_i = \overline{Y_i(1) - Y_i(0)} = \overline{Y_i(1)} - \overline{Y_i(0)}$$

$$E_R[\overline{Y_i} | D_i = 1] = \overline{Y_i(1)}$$

$$E_R[\overline{Y_i} | D_i = 0] = \overline{Y_i(0)}$$

- ▶  $\overline{Y}$  means the average over a set of observations.
- ▶  $E_R[\ ]$  means the average over repeated operations.
- ▶ We may not get exactly  $\overline{Y_i(1)}$  from a particular randomization, but we can expect to get  $\overline{Y_i(1)}$  over repeated experiments.
- ▶  $\overline{Y}$  signifie “moyenne sur un ensemble d’observations”.
- ▶  $E_R[\ ]$  signifie “moyenne sur des opérations répétées”.
- ▶ Nous ne pourrions pas obtenir pas exactement  $\overline{Y_i(1)}$  d’une assignation aléatoire particulière, mais nous pouvons nous attendre à  $\overline{Y_i(1)}$  au cours d’expériences répétées.

## Core assumptions | *Hypothèses clés*

## Core assumptions | *Hypothèses clés*

We have random assignment of treatment.  
We need two core assumptions for randomized experiments:

1. SUTVA (Stable unit treatment value assumption).
2. Excludability.

Nous avons l'assignation aléatoire du traitement.

Nous avons besoin de deux hypothèses clés pour les expériences aléatoires:

1. L'hypothèse SUTVA (Hypothèse stable de la valeur de traitement d'une unité).
2. La restriction d'exclusion.

## Definition: Random assignment of treatment | *Le définition: L'assignation aléatoire du traitement*

- ▶ Random assignment means that each observation has a *known* probability of assignment to experimental conditions *between* 0 and 1.
- ▶ No unit in the experimental sample is assigned to treatment with certainty (probability = 1) or to control with certainty (probability = 0).
- ▶ L'assignation aléatoire signifie que chaque observation a une probabilité *connue* d'assignation au traitement comprise entre 0 et 1.
- ▶ Aucune unité de l'échantillon expérimental n'est assignée au traitement ou au contrôle avec certitude.



## Definition: Random assignment of treatment | *Le définition: L'assignation aléatoire du traitement*

- ▶ It doesn't mean haphazard or uncontrolled!
- ▶ Careful: The condition assigned is not always the same as the condition received.
- ▶ Cela ne signifie ni désordonné ni incontrôlé!
- ▶ Attention ! La condition assignée n'est pas toujours la même que la condition reçue.

## Core assumption 1: SUTVA, part 1 | *Hypothèse clé 1: SUTVA, partie 1*

1.1 No interference – A subject's potential outcome reflects only whether that subject receives the treatment himself/herself. It is not affected by how treatments happen to be allocated to other subjects.

- ▶ A classic violation is the case of vaccines and their spillover effects.

1.1 La non-interférence – Le résultat potentiel d'un sujet reflète uniquement la réception du traitement par ce sujet. Il n'est pas affecté par le statut du traitement des autres sujets.

- ▶ Une violation classique est le cas des vaccins et de leurs effets de contamination.

## Core assumption 1: SUTVA, part 2 | *Hypothèse clé 1: SUTVA, partie 2*

### 1.2 No hidden variations of the treatment

- Say treatment is taking a vaccine, but there are two kinds of vaccines and they have different ingredients.

### 1.2 Pas de variations cachées du traitement

- Disons que le traitement consiste à prendre un vaccin, mais il existe deux types de vaccins et ils ont des compositions différentes.

## Core assumption 2: Excludability | *Hypothèse clé 2: La restriction d'exclusion*

2. Treatment assignment has no effect on outcomes except through its effect on whether treatment was received.
  - ▶ Treatment and control groups should be treated the same, except for the actual treatment.
2. L'assignation au traitement n'a aucun effet sur les résultats, sauf à travers l'effet de recevoir ou non le traitement.
  - ▶ Important de maintenir la symétrie entre les groupes de traitement et de contrôle, excepté le statut du traitement.

## Randomized experiments vs observational studies | *Études aléatoires vs. études observationnelles*

## Observational studies | *Études observationnelles*

- ▶ Treatment is not randomly assigned. It is observed, but not manipulated.
- ▶ Differences in outcomes might be due to underlying differences (selection bias) or the treatment (plus noise). It's often very hard to tell.
- ▶ Le traitement n'est pas assigné de manière aléatoire. Il est observé, mais pas manipulé.
- ▶ Les différences de résultats peuvent être dues à des différences sous-jacentes (biais de sélection) ou au traitement (plus le bruit). C'est souvent très difficile à dire.

## Randomized studies | *Études aléatoires*

- ▶ Randomize treatment, then go measure outcomes.
  - ▶ Randomization creates two groups that are similar except that one will receive the treatment and the other will not. Differences before receiving treatment are only due to chance.
  - ▶ We can attribute differences in outcomes to the treatment (plus chance) instead of other factors.
- ▶ Le traitement est assigné de manière aléatoire, ensuite on mesure les résultats.
  - ▶ L'assignation aléatoire crée deux groupes d'unités qui sont similaires excepté qu'un recevra le traitement et l'autre non. Les différences entre les groupes avant de recevoir le traitement ne sont dues qu'au hasard.
  - ▶ Nous pouvons attribuer les différences de résultats au traitement (en plus du hasard) plutôt qu'à d'autres facteurs.

## Internal validity | *La validité interne*

- ▶ Randomized studies have internal validity – high confidence that we have learned the causal effect of a treatment on an outcome – from randomization.
  - ▶ Observational studies need additional strong assumptions to try to deal with selection bias.
- ▶ La randomisation apporte une validité interne élevée à une étude – nous avons appris l'effet causal du traitement sur le résultat avec un certain niveau de confiance – grâce à l'assignation aléatoire.
  - ▶ Les études observationnelles ont besoin d'autres hypothèses fortes pour tenter de traiter le biais de sélection.



## External Validity (Generalizability) | *La validité externe (Généralisabilité)*

- ▶ The finding from a particular study in one particular place and at one particular time may not hold in other settings (i.e., not have external validity).
- ▶ This is a general concern for observational studies as well, not just a concern for randomized studies.
- ▶ Le résultat d'une étude spécifique à un endroit et moment particuliers peut ne pas être valable dans d'autres contextes (i.e. la validité externe).
- ▶ C'est une préoccupation générale, pas seulement une préoccupation pour les études aléatoires.

- ▶ EGAP Methods Guide on Randomization  
(<https://egap.org/resource/10-things-to-know-about-randomization/>)
- ▶ EGAP's Metaketa Initiative works to accumulate knowledge by pre-planning a meta-analysis of multiple studies that have high internal validity due to randomization.
- ▶ Guide des méthodes EGAP sur la randomisation  
(<https://egap.org/fr/resource/10-choses-a-savoir-sur-la-randomisation/>)
- ▶ L'initiative Metaketa d'EGAP vise à accumuler des connaissances en pré-planifiant une méta-analyse de plusieurs études qui ont une validité interne élevée en raison de la randomisation.