Estimation and Hypothesis Testing 1 | *Les Estimateurs et tests d'hypothèses 1*

Yabo Gwladys Vidogbena

26 June/juin 2024

Estimation | *L'estimation*

Hypothesis Testing | *Les tests d'hypothèses*

Block Randomization | *Randomisation par bloc (ou stratifiée)*

Estimation and Hypothesis Testing | *L'estimation et tests d'hypothèses*

- We have randomly assigned treatment and collected our outcome data.
- Now we use that data for:
  - Estimation: produce an estimate of the true treatment effect
  - Hypothesis testing: assess how consistent the results are with there being no effect.

- Le traitement a été assigné en façon aléatoire et nous avons mésuré les résultats.
- Nous utilisons maintenant ces données pour:
  - Estimation : produire une estimation du véritable effet du traitement.
  - Test d'hypothèse : évaluer la cohérence des résultats sans aucun effet.

Estimation | *L'estimation*

Estimation | *L'estimation*

- Remember that there is a *true* ATE but we can't observe it because of the fundamental problem of causal inference. This is our target, our **estimand**.
  - For example, the ATE.
- We use our data to make an educated guess, our **estimate**.
  - $\widehat{ATE}$
- If we run the experiment again, different units will be assigned to treatment, and our estimate will likely be different.

- Rapellez qu'il y a un *vrai* ATE, mais nous ne pouvons pas l'obsérver à cause de problème fondamentale d'inférence causale. C'est notre cible, notre **paramètre**.
  - Par exemple, l'ATE.
- Nous utilisons nos données pour faire une supposition éclairée, notre **estimation**.
  - $\widehat{ATE}$
- Si nous renouvelons l'éxpérience, diffèrentes unités seraient assigné au traitement, et notre estimation sera probablement différente.

Estimators | *Estimateurs*

- The procedure we apply to our data to produce this estimate is our **estimator**.
- There are many possible estimators for the same estimand.
- We will introduce several estimators that are commonly used to analyze experiments.

- *L'estimateur* est comment on devine la valeur du paramètre à partir des données dont on dispose (les données observées).
- Il y a plusiers estimateurs possibles pour le même paramètre.
- Nous présenterons plusiers estimateurs courrament utilisés.

Estimators | *Estimateurs*

- In general, we prefer estimators that are:
  - Unbiased: If we run the experiment many times, each estimate might be a little too high or low, but it will be correct on average.
  - Precise: The estimates are close to the estimand.

- En général, nous préférons les estimateurs qui sont:
  - Non biaisé: Cela signifie que si vous deviez exécuter l'expérience plusieurs fois, l'estimation peut parfois être trop élevée ou trop faible, mais elle sera correcte en moyenne.
  - Précis : Les estimations sont proches du paramètre.

General Principle: Analyze as you randomize | *Un principe général : Analysez comme vous randomisez*

- This means follow the design of the experiment.
- Compare groups that are created by random assignment.

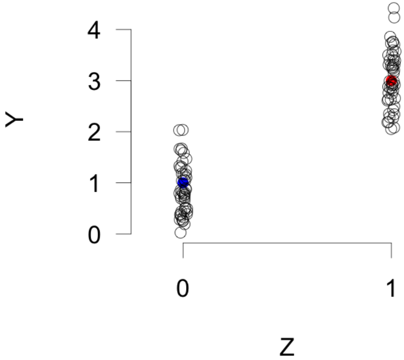
- Cela signifie suivre la concéption de l'éxpérience.
- Comparez les groupes qui sont crées par l'assignation aléatoire.

Estimator 1: difference-in-means | *Estimateur 1 : la différence des moyennes*

- Our *estimand* is the ATE.
- Random assignment to treatment or control.
- All units have the same probability of treatment assignment.
- The simplest *estimator* for the ATE is the **difference-in-means**: take the average outcome for the treatment group and subtract the average outcome for the control group.

- Notre *paramètre* est l'ATE.
- L'assignation aléatoire à traitement ou contrôle.
- Toutes les unités ont la même probabilité de recevoir le traitement.
- L'estimateur de l'ATE le plus simple est **la différence des moyennes** : soutraitez le moyen des unités assigné au contrôle du moyen des unités assigné au traitement.

Estimator 1: Difference-in-means | *Estimateur 1 : La différence des moyennes*



Estimator 1: Difference-in-means | *Estimateur 1 : La différence des moyennes*

Unit	Z_i	Y_i	$Y_i(1)$	$Y_i(0)$
a	1	5	5	
b	1	4	4	
c	1	2	2	
d	1	1	1	
e	0	1		1
f	0	1		1
g	0	0		0
h	0	2		2

Estimator 1: Difference-in-means | *Estimateur 1 : La différence des moyennes*

Unit	Z_i	Y_i	$Y_i(1)$	$Y_i(0)$
a	1	5	5	
b	1	4	4	
c	1	2	2	
d	1	1	1	
e	0	1		1
f	0	1		1
g	0	0		0
h	0	2		2

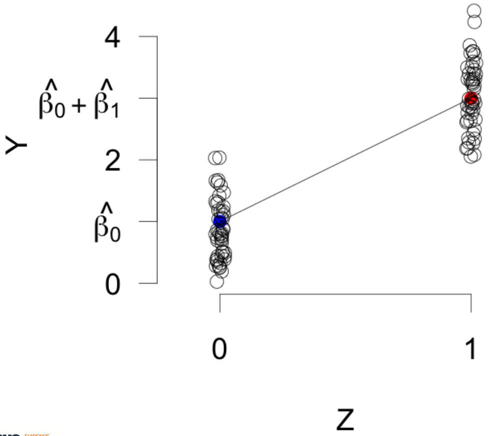
$$\frac{5+4+2+1}{4} - \frac{1+1+0+2}{4} = 3 - 1 = 2$$

Estimator 1: Difference-in-means | *Estimateur 1 : La différence des moyennes*

```
# mean(Y[treatment==1]) - mean(Y[treatment==0])

# library(estimatr)
# difference_in_means(Y ~ treatment)
```

Estimator 2: Linear Regression | *Estimateur 2 : La régression linéaire*



$$Y_i = \beta_0 + \beta_1 Z_i + \epsilon_i$$

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 Z_i + e_i$$

Estimator 2: Linear Regression | *Estimateur 2 : La régression linéaire*

$$Y_i = \beta_0 + \beta_1 Z_i + e_i$$

- With this simple experiment, we can also use a linear regression. It will produce exactly the same estimate ($\hat{\beta}_1$) of the ATE (β_1) as the difference-in-means estimator.
- $\hat{\beta}_0$ is the average outcome in the control group.

- Pour cet éxpérience simple, nous pouvon également utiliser la régression linéaire. Ça produira exactement le même estimation ($\hat{\beta}_1$) du ATE (β_1) que l'estimateur de la différence de moyennes.
- $\hat{\beta}_0$ est le résultat moyen des unités assignéss au contrôle.

Estimator 2: Linear Regression | *Estimateur 2 : La régression linéaire*

```
# lm(Y ~ treatment)
```

Hypothesis Testing | *Les tests d'hypothèses*

Hypothesis Testing | *Les tests d'hypothèses*

- Let's say that the truth is that a medicine has no effect on height. But all the short people were assigned to the medicine and all the tall people to control. If we apply the difference in means, it looks like the medicine made people shorter!
- Warning: We can get an estimate that is not zero even when there is no effect!

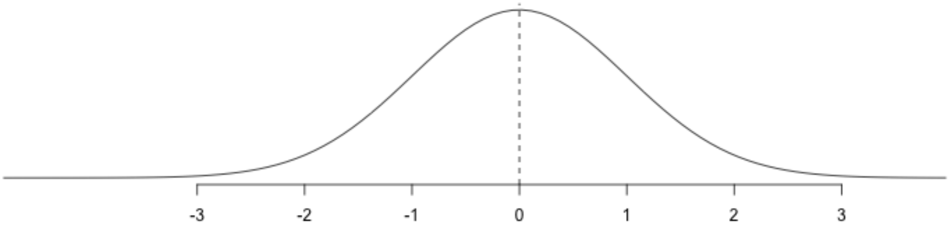
- Supposons qu'un médicament n'ait aucun effet sur la taille. Mais toutes les personnes de petite taille ont été assignées le médicament et les personnes de grande taille ont été assignées au contrôle. Si on utilise la différence de moyennes, on dirait que les médicaments ont rendu les gens plus petits!
- Avertissement : On peut obtenir une estimation que n'est pas nulle même s'il n'y a aucun effet!

Hypothesis Testing | *Les tests d'hypothèses*

- Are we confident that our non-zero estimate reflects a truly non-zero estimand (truth)?
- Start with a claim we might reject (a null hypothesis). We will use the null hypothesis that the true ATE is 0.
- But remember that we can get $\widehat{ATE}$ that is not 0, just by chance.

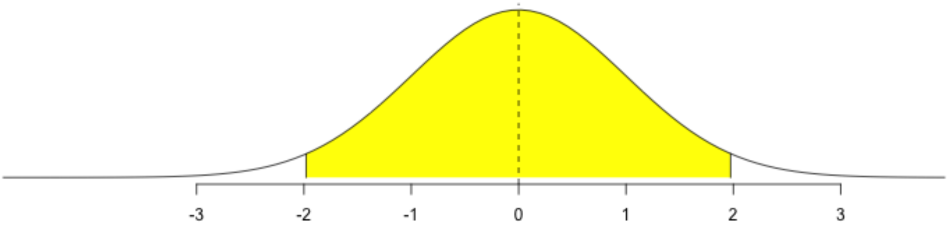
- Sommes-nous convaincus que notre estimation non nulle reflète un paramètre véritablement non nulle (la vérité) ?
- Commencer par une affirmation que nous pourrions rejeter (une hypothèse nulle). On utilisera l'hypothèse nulle que le vrai ATE est 0.
- Mais rappelez-vous que nous pouvons obtenir un $\widehat{ATE}$ différent de 0, juste par hasard.

Hypothesis Testing | *Les tests d'hypothèses*



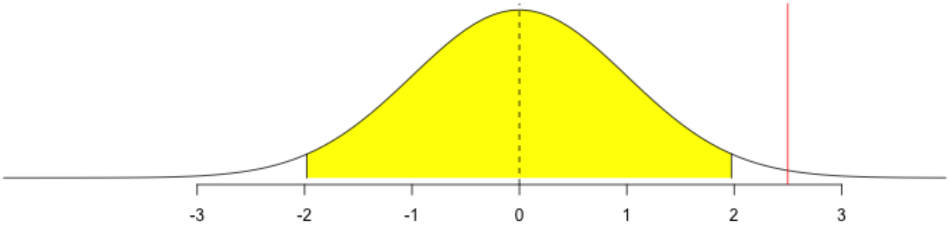
- Distribution of possible $\widehat{ATE}$ if the null hypothesis is true
- *Distribution des* $\widehat{ATE}$ possibles si l'hypothèse nulle est vraie

Hypothesis Testing | *Les tests d'hypothèses*



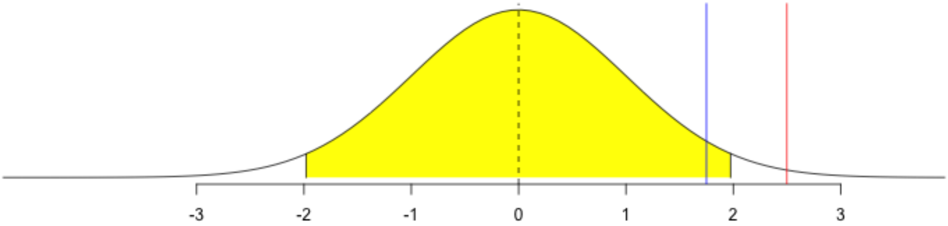
- Rejection and non-rejection regions for a two-sided alternative hypothesis at $\alpha = 0.05$
- Régions de rejet et de non-rejet pour une hypothèse alternative bilatérale à $\alpha = 0,05$

Hypothesis Testing | *Les tests d'hypothèses*



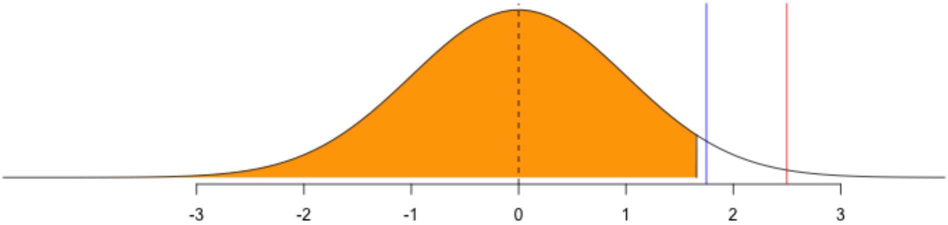
- $\widehat{ATE}$ falls in the rejection region $\rightarrow$ reject the null hypothesis
- $\widehat{ATE}$ se situe dans la région de rejet $\rightarrow$ rejetez l'hypothèse nulle

Hypothesis Testing | *Les tests d'hypothèses*



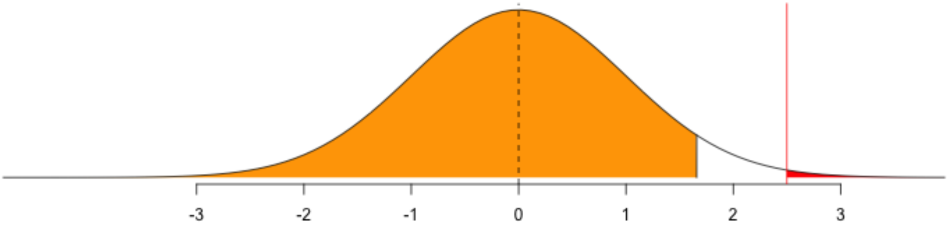
- $\widehat{ATE}$ falls outside the rejection region $\rightarrow$ do not reject the null hypothesis
- $\widehat{ATE}$ se situe en dehors de la région de rejet $\rightarrow$ ne rejetez pas l'hypothèse nulle

Hypothesis Testing | *Les tests d'hypothèses*



- Rejection and non-rejection regions for a one-sided alternative hypothesis at $\alpha = 0.05$
- Régions de rejet et de non-rejet pour une hypothèse alternative unilatérale à $\alpha = 0,05$

p-value | p-valeur



- p-value: For a one-sided test, the probability of seeing a test statistic as large as or larger than the test statistic calculated from observed data when the null hypothesis is true.
- We use the $\widehat{ATE}$ as our test statistic, but we don't need to.

- p-valeur : Pour un test d'hypothèse unilateral, la probabilité de voir une statistique de test aussi grande ou plus grande que la statistique de test calculée à partir des données observées lorsque l'hypothèse nulle est vraie.
- Nous utilisons $\widehat{ATE}$ comme statistique de test, mais ce n'est pas nécessaire.

Hypothesis Testing with Linear Regression | *Les tests d'hypothèses avec la régression linéaire*

- There are many ways to do hypothesis testing. We are going to take the simplest and most popular approach that uses regression.
- Use linear regression to calculate a p-value (two-sided test): the probability that we could have obtained a particular estimate (or greater in absolute value) by chance when the null hypothesis is true.

- Il existe de nombreuses façons de tester des hypothèses. Nous allons faire l'apprôche la plus simple et la plus populaire: la régression.
- Utiliser la régression linéaire pour calculer une p-valeur (test bilateral) : la probabilité de voir une estimation (ou une plus grande en valeur absolue) par hasard ou l'hypothèse nulle est vrai.

Hypothesis Testing with Linear Regression | *Les tests d'hypothèses avec la régression linéaire*

- Compare this *p*-value to a standard we have set in advance (an α level).
- If the *p*-value is smaller than or equal to the α level, we reject the null hypothesis of no effect.
- If the *p*-value is greater than the α level, we fail to reject the null hypothesis of no effect.

- Comparez cette *p*-valeur à une norme que nous avons fixée à l'avance (un niveau α).
- Si la *p*-valeur est plus petite ou égale au niveau α, nous rejetons l'hypothèse nulle d'aucun effet.
- Si la *p*-valeur est plus grande que le niveau α, nous ne parvenons pas à rejeter l'hypothèse nulle d'aucun effet.

Hypothesis Testing with Linear Regression | *Les tests d'hypothèses avec la régression linéaire*

Table 4: **Attitudes toward Early Forced Marriage, 2-3 Weeks After Exposure**

	Reject Forced Marriage				Reject Early Forced Marriage			
	Reject FM		Reject FM (18+)		Money		Misbehaving	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EFM Treat	0.093***	0.095***	0.088***	0.092***	0.048***	0.044***	0.013	0.013
Standard Error	0.027	0.020	0.025	0.017	0.014	0.015	0.027	0.027
RI *p*-value	0.001	<0.001	0.001	<0.001	0.001	0.003	0.318	0.318

Block Randomization | *Randomisation par bloc (ou stratifiée)*

Block Randomization | *Randomisation par bloc*

- Block randomization is like doing a separate experiment in each block.
- We present 2 estimators for block randomization. Others are also available.

- Randomisation par bloc est comme faire une expérience distincte dans chaque bloc.
- Nous presentons 2 estimateurs pour randomisation par bloc. D'autres sont également disponibles.

Estimator 1: Blocked Difference-in-Means | *Estimateur 1 : La différence des moyennes par bloc*

- Calculate the $\widehat{ATE_j}$ for each block using difference in means. j indicates which block.
- The $\widehat{ATE}$ is the average of the block-level $\widehat{ATE_j}$ weighted by block size N_j/N.
- You can use this estimator even when the probability of treatment assignment is different across blocks.

- Calculez $\widehat{ATE_j}$ pour chaque bloc en utilisant la différence des moyennes.
- $\widehat{ATE}$ est la moyenne pondérée de $\widehat{ATE_j}$ pondérée par la taille du bloc N_j/N.
- Nous pouvons utiliser cette estimateur sinon la probabilité d'assignation du traitement diffère selon les blocs.

Estimator 1 : Blocked Difference-in-Means | *Estimateur 1 : La différence des moyennes par bloc*

Unit	Block	Z_i	Y_i	$Y_i(1)$	$Y_i(0)$
a	Q	0	4		4
b	Q	1	3	3	
c	Q	0	2		2
d	R	1	3	3	
e	R	0	0		0
f	R	0	2		2
g	S	1	4	4	
h	S	0	0		0
i	S	0	2		2

$$\widehat{ATE}_Q = 3 - \frac{4+2}{2} = 0$$

$$\widehat{ATE}_R = 3 - \frac{0+2}{2} = 2$$

$$\widehat{ATE}_S = 4 - \frac{0+2}{2} = 3$$

$$\begin{aligned}\widehat{ATE} &= \frac{N_Q}{N}\widehat{ATE}_Q + \frac{N_R}{N}\widehat{ATE}_R + \frac{N_S}{N}\widehat{ATE}_S \\ &= \frac{3}{9} * 0 + \frac{3}{9} * 2 + \frac{3}{9} * 3 = \frac{5}{3}\end{aligned}$$

Estimator 1: Blocked Difference-in-Means | *Estimateur 1 : La différence des moyennes par bloc*

```r
library(estimatr)
# difference_in_means(Y ~ t, blocks = block_variable)
Y <- experiment_data$State.of.the.world
t <- experiment_data$treatment_received
difference_in_means(Y ~ t, blocks = experiment_data$Female)
```

```
## Design:  Blocked
##            Estimate Std. Error   t value  Pr(>|t|)  CI Lower CI Upper DF
## ttreatment     0.35  0.7989665 0.4380659 0.6662783 -1.322256 2.022256 19
```

Estimator 2: Linear Regression with Block Fixed Effects | *Estimateur 2 : La régression linéaire avec effets fixes par bloc*

$$Y_{ij} = \beta_0 + \beta_1 Z_{ij} + \gamma_A BlockA_{ij} + \gamma_B BlockB_{ij} + \dots + \epsilon_{ij}$$

- You can use linear regression with block fixed effects, applying weights to each observation.
- The weight is the inverse of the proportion of subjects in the same block who were assigned to the same condition.

- Nous pouvons ensuite utiliser la régression linéaire avec des effets fixes en bloc, en appliquant des pondérations à chaque observation.
- Le poids est l'inverse de la proportion de sujets d'un même bloc qui ont été assignés à la même condition.

$$w_{ij} = \frac{d_i}{p_{ij}} + \frac{1 - d_i}{1 - p_{ij}}, \text{ where } p_{ij} \equiv \frac{m_j}{N_j}$$

Block Randomization | *Randomisation par bloc*

```r
# library(estimatr)
# lm_robust(Y ~ treatment + as.factor(block_variable),
#                        weights=weight_variable)
```