Estimation and Hypothesis Testing 2 | *Les estimateurs et les tests d'hypothèses 2*

Christelle Zozoungbo

27 June/juin 2024

EVIDENCE IN GOVERNANCE AND POLITICS

Covariate Adjustment | *Ajustement des covariables*

Cluster Randomization | *Randomisation par grappe*

Experiments with Multiple Arms | *Les éxperiences avec plusiers bras*

Factorial Design | *La concéption factorielle*

A Quick Reminder | *Un pétit rappel*

- Remember: Analyze as you randomize
- We prefer estimators that are unbiased and have greater precision

- N'oubliez pas : Analysez comme vous randomisez
- Nous préférons les estimateurs non biaisés et plus précis.

Covariate Adjustment | *Ajustement des covariables*

Estimator: Linear regression with covariates | Estimateur : La régression linéaire avec des covariables

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

- Including a **pre-treatment covariate** X that is *predictive of* the outcome variable in our regression model is called *covariate adjustment*.
- For example: pre-treatment measure of the outcome.
- Careful: This can bias our estimates, but improve their precision.

- L'inclusion d'une **covariable pré-traitement** X qui est *prédictive* de la variable de résultat dans notre modèle de régression est appelée *ajustement des covariables*.
- Par exemple : un mesure du résultat avant le traitement.
- Attention : Cela peut biaiser nos estimations, mais améliorer leur précision.

Estimator: Linear regression with covariates | *Estimateur : La régression linéaire avec des covariables*

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

- The coefficient on the treatment variable (β_1) is again our ATE.
- The coefficient on the covariate (γ) is *not* the causal effect of that variable.

- Le coefficient sur la variable de traitement (β_1) est encore notre ATE.
- Le coefficient de la covariable (γ) n'est *pas* l'effet causal de cette variable.

Estimator: Linear regression with covariates | *Estimateur : La régression linéaire avec des covariables*

	Reject FM	
	(1)	(2)
EFM Treat	0.093***	0.095***
Standard Error	0.027	0.020
RI *p*-value	0.001	<0.001
Hypothesis	+	+
Control Mean	0.82	0.82
Control SD	0.16	0.16
DV Range	[0-1]	[0-1]
Blocked FE	Yes	Yes
Controls	No	16
Adj-R^2	0.09	0.23
Observations	998	998

Note: * p < .1, ** p < 0.05, and *** p < 0.01

Estimator: Linear regression with covariates | *Estimateur : La régression linéaire avec des covariables*

```r
library(estimatr)
# lm_robust(Y ~ treatment + Language + Gender)
```

Cluster Randomization | *Randomisation par grappe*

Estimator: Regression with cluster-robust standard errors | *Estimateur : La régression avec des erreurs types robustes au niveau du cluster*

$$Y_{ic} = \beta_0 + \beta_1 Z_c + e_{ic}$$
$$Y_{ic} = \beta_0 + \beta_1 Z_c + \beta_2 X_{ic} + e_{ic}$$

- Our analysis has to take into account the fact that treatment is assigned at the cluster level with *cluster-robust standard errors*.
- β_1 is the ATE of the treatment at the individual level.
- We can also do covariate adjustment at the same time.

- Notre analyse doit prendre en compte le fait que le traitement est attribué au niveau du cluster avec des *erreurs types robustes au niveau du cluster*.
- β_1 est l'ATE du traitement au niveau individuel.
- Nous pouvons également effectuer un ajustement covariable en même temps.

–>

Cluster Randomization | *Randomisation par grappe*

```r
library(estimatr)

# lm_robust(Y ~ treatment, clusters=cluster_variable)

# lm_robust(Y ~ treatment + covariate, clusters=cluster_variable)
```

Experiments with Multiple Arms | *Les éxperiences avec plusiers bras*

Estimator 1: Difference-in-Means | *Estimateur 1 : La différence en moyennes*

Z_A only	Z_B only	Neither (control)

- We can always take the difference-in-means between any two groups.

- Nous pouvons toujours tenir compte de la différence de moyens entre deux groupes.

egap EVIDENCE IN GOVERNANCE AND POLITICS

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

$$Y_i = \alpha + \beta_A Z_{Ai} + \beta_B Z_{Bi} + e_i$$

- Regression with an indicator variable for each of the two treatment arms.
- We can also do covariate adjustment at the same time.

- Régression avec une variable indicatrice pour chacun des deux bras de traitement.
- Nous pouvons également effectuer un ajustement covariable en même temps.

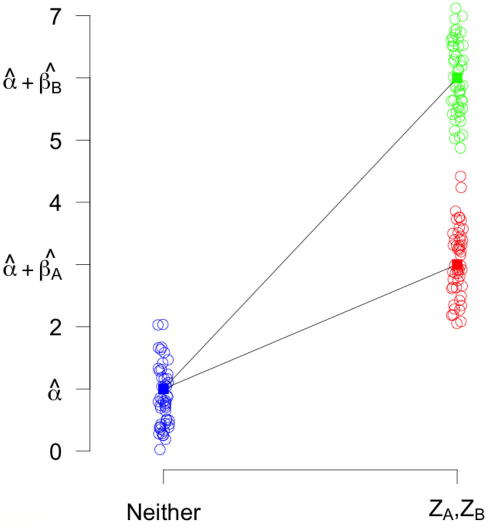
Estimator 2: Linear regression | *Estimateur 2 : La régression linéaire*

$$Y_i = \alpha + \beta_A Z_{Ai} + \beta_B Z_{Bi} + e_i$$

- β_A is the ATE of Z_A (compared with control).
- β_B is the ATE of Z_B (compared with control).

- β_A est l'ATE de Z_A (par rapport au contrôle).
- β_B est l'ATE de Z_B (par rapport au contrôle).

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



Z_A only
Z_B only
Neither (control)

$$Y_i = \alpha + \beta_A Z_{Ai} + \beta_B Z_{Bi} + e_i$$

Estimators for Multi-arm Designs | *Les estimateurs pour les éxperiences avec plusiers bras*

```r
library(estimatr)

# difference_in_means(Y ~ treatment,
#           condition1="T2",
#           condition2="T1")

# lm_robust(Y ~ as.factor(treatment))
```

Factorial Design | *La concéption factorielle*

Estimator 1: Difference-in-Means | *Estimateur 1 : La différence en moyennes*

Neither	Z_2 only
Z_1 only	Both Z_1 and Z_2

- If we have a 2*2 factorial design, we have four groups.
- We can always take the difference-in-means between any two groups.

- Si nous avons une concéption factorielle 2*2, nous avons 4 groupes.
- Nous pouvons toujours tenir compte de la différence de moyens entre deux groupes.

–>

Estimator 2: Linear Regression with an Interaction Term | *Estimateur 2 : La régression linéaire avec un terme d'interaction*

$$Y_i = \beta_0 + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{1i} * Z_{2i} + e_i$$

$$Y_i = \beta_0 + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{1i} * Z_{2i} + \beta_4 X_i + e_i$$

- Indicator variables for Z_1 and Z_2.
- We can also do covariate adjustment at the same time.

- Variables indicatrices pour Z_1 et Z_2.
- Nous pouvons également effectuer un ajustement covariable en même temps.

Estimator 2: Linear Regression with an Interaction Term | *Estimateur 2 : La régression linéaire avec un terme d'interaction*

Neither	Z_2 only
Z_1 only	Both Z_1 and Z_2

$$Y_i = \beta_0 + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{1i} * Z_{2i} + e_i$$

- β_1 is the ATE of Z_1 conditional on $Z_2 = 0$ | *l'ATE de Z_1* conditionnel à $Z_2 = 0$, $(E[Y(Z_1 = 1) - Y(Z_1 = 0)|Z_2 = 0])$
- β_2 is the ATE of Z_2 conditional on $Z_1 = 0$ | *l'ATE de Z_2* conditionnel à $Z_1 = 0$, $(E[Y(Z_2 = 1) - Y(Z_2 = 0)|Z_1 = 0])$

Estimator 2: Linear Regression with an Interaction Term | *Estimateur 2 : La régression linéaire avec un terme d'interaction*

Neither	Z_2 only
Z_1 only	Both Z_1 and Z_2

$$Y_i = \beta_0 + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{1i} * Z_{2i} + e_i$$

- $\beta_1 + \beta_3$ = ATE of Z_1 conditional on $Z_2 = 1$ | *l'ATE de* Z_1 conditionnel à $Z_2 = 1$, $(E[Y(Z_1 = 1) - Y(Z_1 = 0)|Z_2 = 1])$
- $\beta_2 + \beta_3$ = ATE of Z_2 conditional on $Z_1 = 1$ | *l'ATE de* Z_2 conditionnel à $Z_1 = 1$, $(E[Y(Z_2 = 1) - Y(Z_2 = 0)|Z_1 = 0])$
- β_3 is called the interaction effect. β_3 est appelé l'effet d'interaction.

Estimator 2: Linear Regression with an Interaction Term | *Estimateur 2 : La régression linéaire avec un terme d'interaction*

Table 4. Adoption for Parents Sampled for SAFI & Subsidy Programs

	Used Fertilizer Season 1
Panel A. 2004 Season 1 Treatments	(1)
SAFI Season 1	0.114
	(0.035)***
Starter Kit Farmer	0.059
	(0.042)
Starter Kit Farmer * Demonstration Plot School	-0.026
	(0.060)
Demonstration Plot School	0.006
	(0.314)

Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

$$\hat{Y} = \hat{\alpha} + 0.059Kit + 0.006Demo - 0.026Kit * Demo$$

Estimator 2: Linear Regression with an Interaction Term | *Estimateur 2 : La régression linéaire avec un terme d'interaction*

$$\hat{Y} = \hat{\alpha} + 0.059 Kit + 0.006 Demo - 0.026 Kit * Demo$$

Neither: $\hat{Y} = \hat{\alpha}$

Kit only: $\hat{Y} = \hat{\alpha} + 0.059$

Demo only: $\hat{Y} = \hat{\alpha} + 0.006$

Kit and Demo: $\hat{Y} = \hat{\alpha} + 0.059 + 0.006 - 0.026$

$\widehat{ATE}$ of Kit conditional on No Demo $= \hat{\alpha} + 0.059$ - $\hat{\alpha} = 0.059$

$\widehat{ATE}$ of Kit conditional on Demo $= \hat{\alpha} + 0.059 + 0.006$ - 0.026 - $(\hat{\alpha} + 0.006) =$ 0.059-0.026

Interaction effect $= 0.059$ - 0.026 - $0.059 = 0.026$

Estimator 2: Linear Regression with an Interaction Term | *Estimateur 2 : La régression linéaire avec un terme d'interaction*

```r
library(estimatr)
# lm_robust(Y ~ Z1 + Z2 + Z1*Z2)

# lm_robust(Y ~ Z1*Z2)

# lm_robust(Y ~ Z1*Z2 + covariate)
```