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

NAME/NOM

DATE

Multiple Arm Experiments | *Les éxperiences avec plusieurs bras*

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

Cluster Randomization | *Randomisation par grappe*

Factorial Design | *La conception factorielle*

A Quick Reminder | *Un pétit rappel*

- Remember: Analyze as you randomize
- We prefer estimators that are unbiased and have greater precision
- Hypothesis testing can be simple with linear regression

- N'oubliez pas : Analysez comme vous randomisez
- Nous préférons les estimateurs non biaisés et plus précis
- Test d'hypothèse peut être simple avec la régression linéaire

Multiple Arm Experiments | *Les éxperiences avec plusieurs 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.

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

$$Y_i = \hat{\alpha} + \hat{\beta}_A Z_{Ai} + \hat{\beta}_B Z_{Bi} + e_i$$

$$Y_i = \hat{\alpha} + \hat{\beta}_A Z_{Ai} + \hat{\beta}_B Z_{Bi} + \hat{\gamma} X_i + e_i$$

- Regression with an indicator variable for each of the two treatment arms.
  - $Z_{Ai} = 1$ if unit i has treatment Z_A, 0 otherwise
  - $Z_{Bi} = 1$ if unit i has treatment Z_B, 0 otherwise
- We can also do covariate adjustment at the same time.

- Régression avec une variable indicatrice pour chacun des deux bras de traitement.
  - $Z_{Ai} = 1$ si unité i a traitement Z_A, sinon 0
  - $Z_{Bi} = 1$ si unité i a traitement Z_B, sinon 0
- Nous pouvons également effectuer un ajustement covariable en même temps.

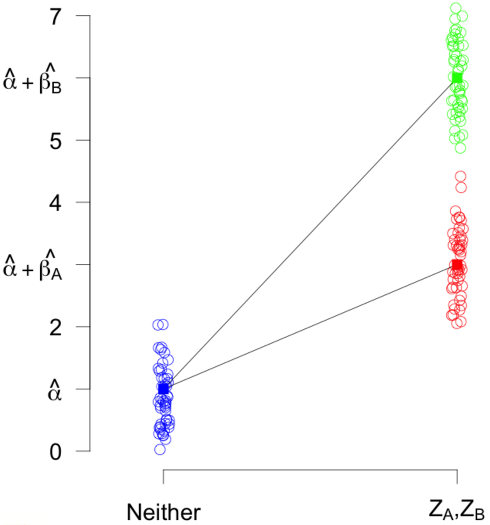
Estimator 2: Linear regression | *Estimateur 2 : La régression linéaire*

$$Y_i = \hat{\alpha} + \hat{\beta}_A Z_{Ai} + \hat{\beta}_B Z_{Bi} + e_i$$

- $\hat{\beta}_A$ is the $\widehat{ATE}$ of Z_A (compared with control).
- $\hat{\beta}_B$ is the $\widehat{ATE}$ of Z_B (compared with control).
- How do we estimate the effect of Z_B compared to Z_A?

- $\hat{\beta}_A$ est $\widehat{ATE}$ de Z_A (par rapport au contrôle).
- $\hat{\beta}_B$ est $\widehat{ATE}$ de Z_B (par rapport au contrôle).
- Comment estimer l'effet de Z_B par rapport à Z_A ?

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



Z_A only

Z_B only

Neither (control)

$Y_i = \hat{\alpha} + \hat{\beta}_A Z_{Ai} + \hat{\beta}_B Z_{Bi} + e_i$

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

```
# library(estimatr)
# difference_in_means(Y ~ treatment,
#         condition1="T1",
#         condition2="T2")
#
#
# library(car)
# M <- lm_robust(Y ~ as.factor(treatment))
# linearHypothesis(M, "T1=T2")
```

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 by 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
a	Q	0	4
b	Q	1	3
c	Q	0	2
d	R	1	3
e	R	0	0
f	R	0	2
g	S	1	4
h	S	0	0
i	S	0	2
j	S	1	4

$$\widehat{ATE}_Q = \frac{3}{1} - \frac{4+2}{2} = 0$$

$$\widehat{ATE}_R = \frac{3}{1} - \frac{0+2}{2} = 2$$

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

$$\begin{aligned}\widehat{ATE} &= \frac{N_Q}{N}\widehat{ATE}_Q + \frac{N_R}{N}\widehat{ATE}_R + \frac{N_S}{N}\widehat{ATE}_S \\ &= \frac{3}{10} * 0 + \frac{3}{10} * 2 + \frac{4}{10} * 3 \\ &= \frac{0+6+12}{10} = \frac{9}{5}\end{aligned}$$

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

```
# library(estimatr)
# difference_in_means(Y ~ t, blocks = block_variable)
```

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

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

- 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*

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

Cluster Randomization | *Randomisation par grappe*

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

$$Y_{ic} = \hat{\beta}_0 + \hat{\beta}_1 Z_c + e_{ic}$$

$$Y_{ic} = \hat{\beta}_0 + \hat{\beta}_1 Z_c + \hat{\gamma} X_{ic} + e_{ic}$$

- Our analysis has to take into account the fact that treatment is assigned at the cluster level with *cluster-robust standard errors*.
- $\hat{\beta}_1$ is the $\widehat{ATE}$ of the treatment on individual units.
- 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*.
- $\hat{\beta}_1$ est $\widehat{ATE}$ du traitement sur les unités individuelles.
- 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)
```

Factorial Design | *La conception factorielle*

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

		Z_2 = 1	Z_2 = 0
Z_1	1	Both Z_1 and Z_2	Z_1 only
	0	Z_2 only	Neither

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

- Nous utilisons un plan factoriel quand nous nous intéressons aux effets d'interaction.
- Si nous avons une conception 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*

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 Z_{1i} + \hat{\beta}_2 Z_{2i} + \hat{\beta}_3 Z_{1i} * Z_{2i} + e_i$$

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 Z_{1i} + \hat{\beta}_2 Z_{2i} + \hat{\beta}_3 Z_{1i} * Z_{2i} + \hat{\gamma} X_i + e_i$$

- 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*

	$Z_2 = 1$	$Z_2 = 0$
$Z_1 = 1$	Both Z_1 and Z_2	Z_1 only
$Z_1 = 0$	Z_2 only	Neither

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 Z_{1i} + \hat{\beta}_2 Z_{2i} + \hat{\beta}_3 Z_{1i} * Z_{2i} + e_i$$

- Estimand | Paramètre: $E[Y(Z_1 = 1)|Z_2 = 0] - E[Y(Z_1 = 0)|Z_2 = 0]$
  - $\hat{\beta}_1$ is the $\widehat{ATE}$ of Z_1 conditional on $Z_2 = 0$ | $\widehat{ATE}$ de Z_1 conditionnel à $Z_2 = 0$.

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

	$Z_2 = 1$	$Z_2 = 0$
$Z_1 = 1$	Both Z_1 and Z_2	Z_1 only
$Z_1 = 0$	Z_2 only	Neither

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 Z_{1i} + \hat{\beta}_2 Z_{2i} + \hat{\beta}_3 Z_{1i} * Z_{2i} + e_i$$

- Estimand | Paramètre: $E[Y(Z_1 = 1)|Z_2 = 1] - E[Y(Z_1 = 0)|Z_2 = 1]$
  - $\hat{\beta}_1 + \hat{\beta}_3 = \widehat{ATE}$ of Z_1 conditional on $Z_2 = 1$ | $\widehat{ATE}$ de Z_1 conditionnel à $Z_2 = 1$
- 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*

Antisocial behaviors, z-score	
Intercept	0.151
Therapy	-0.249**
	(0.088)
Cash	-0.079
	(0.091)
Cash*Therapy	-0.308**
	(0.089)

NOTE: * p<0.05; ** p<0.01; *** p< 0.001 . The table reports intent-to-treat estimates of the effect of each treatment arm, controlling for covariates and block fixed effects. Taken from Table 2 of Blattman, Jamison, and Sheridan (2017)

$$\hat{Y} = \hat{\alpha} - 0.249\,\mathit{Therapy} - 0.079\,\mathit{Cash} - 0.308\,\mathit{Therapy} * \mathit{Cash}$$

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

$$\hat{Y} = \hat{\alpha} - 0.249\,Therapy - 0.079\,Cash - 0.308\,Therapy * Cash$$

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

Cash only: $\hat{Y} = \hat{\alpha} - 0.079$

Therapy only: $\hat{Y} = \hat{\alpha} - 0.249$

Therapy and Cash: $\hat{Y} = \hat{\alpha} - 0.249 - 0.079 - 0.308$

$\widehat{ATE}$ of Cash conditional on No Therapy $= \hat{\alpha} - 0.079 - \hat{\alpha} = -0.079$

$\widehat{ATE}$ of Cash conditional on Therapy $=$
$\hat{\alpha} - 0.249 - 0.079 - 0.308 - (\hat{\alpha} - 0.249) = -0.079 - 0.308 = -0.387$

Interaction effect $= -0.079 - 0.308 - (-0.079) = -0.308$

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)
```