

Chapitre 18

Test paramétrique de comparaison de deux moyennes estimées sur deux échantillons appariés

José LABARERE, PU-PH

Arnaud Seigneurin, MCU-PH, Bastien Boussat, MCU-PH, Alexandre Bellier, AHU, Patrice François, PU-PH

- Objectifs
- Introduction
- Test Z de l'écart réduit
- Test t de Student

Objectifs

Comparaison de deux moyennes estimées sur deux échantillons appariés :

- Identifier une situation requérant ce type de test
- Sélectionner le test pertinent
- Connaitre et vérifier les conditions d'application de ce test

Plan

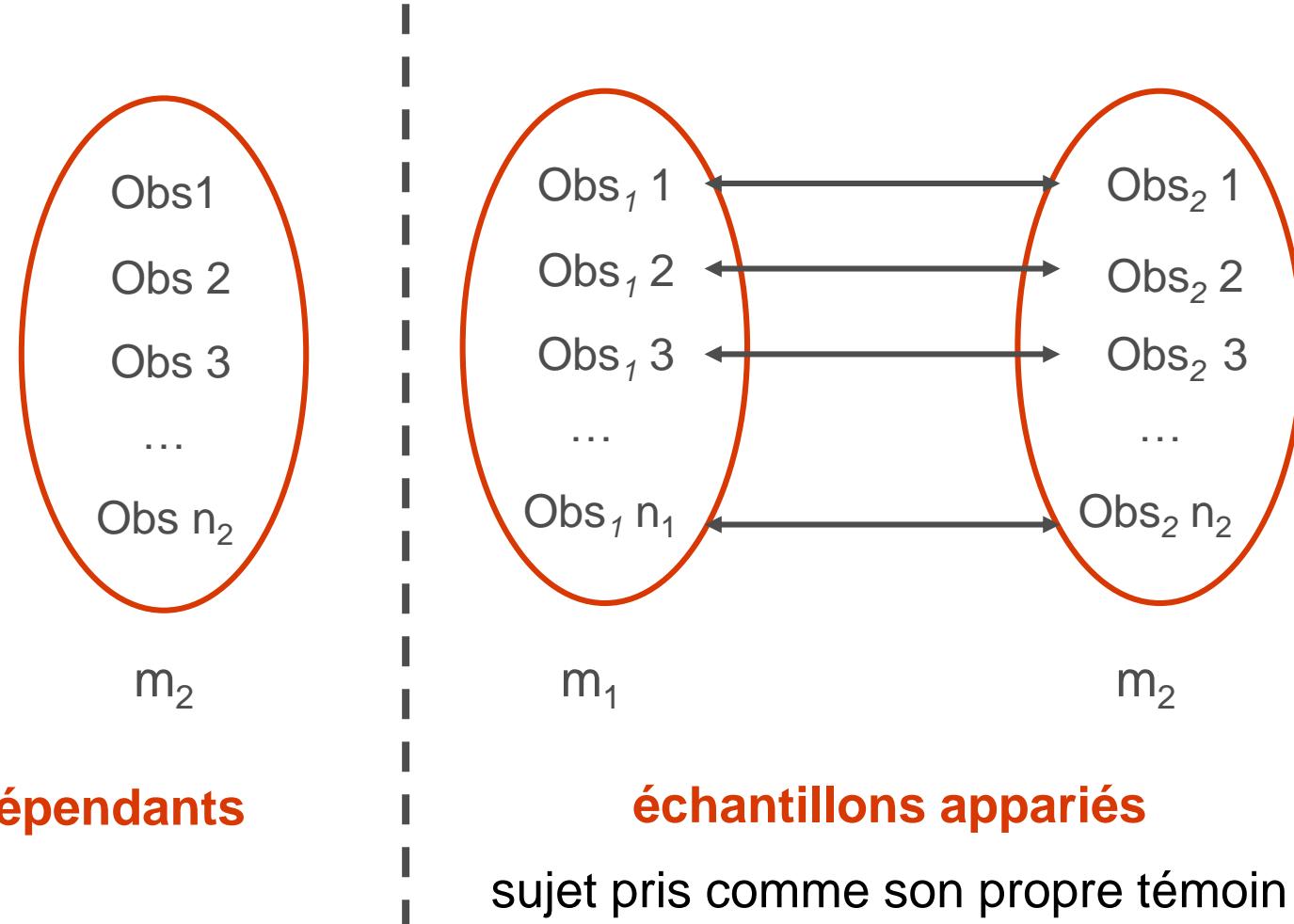
- Objectifs
- Introduction
- Test Z de l'écart réduit
- Test t de Student

Comparaison de 2 moyennes estimées



échantillons indépendants

$(n_1 = n_2 \text{ ou } n_1 \neq n_2)$

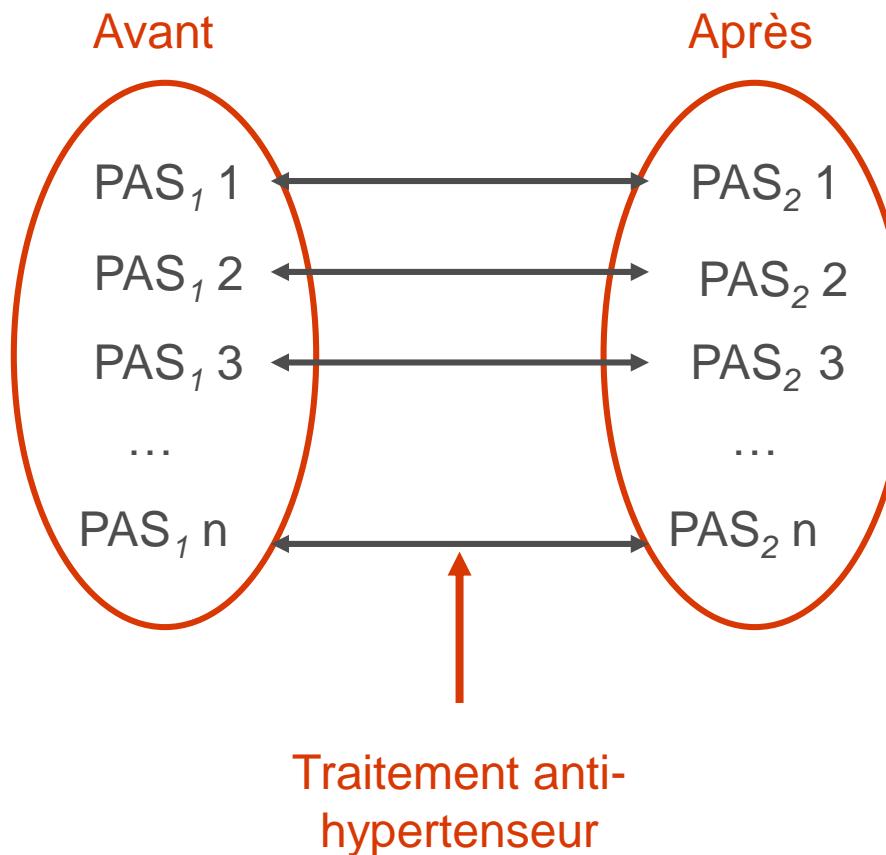


échantillons appariés

sujet pris comme son propre témoin
2 membres d'une fratrie...

$(n_1 = n_2 = n)$

Comparaison de 2 moyennes estimées à partir de 2 échantillons appariés



$$H_0 : m_{PAS_{\text{Avant}}} = m_{PAS_{\text{Après}}}$$

Les mesures PAS_1 et PAS_2 du sujet 1 ne sont pas indépendantes

Les 2 mesures ont été effectuées sur le même sujet : si PAS_1 était très élevée, il est probable que PAS_2 restera élevée (mais moins que PAS_1 si le traitement est efficace)

Le test doit prendre en compte cette dépendance des observations PAS_1 et PAS_2

(En revanche, les mesures PAS_2 du sujet 1 et PAS_2 du sujet 2 sont indépendantes)

• Z pour échantillons indépendants

$$m_1 - m_2$$

$$\text{var}(m_1 - m_2) = \text{var}(m_1) + \text{var}(m_2) - 2 \text{cov}(m_1, m_2) = \text{var}(m_1) + \text{var}(m_2)$$

$$Z = \frac{m_1 - m_2}{\sqrt{\text{var}(m_1) + \text{var}(m_2)}}$$

• Z pour échantillons appariés

$$m_1 - m_2$$

$$\text{var}(m_1 - m_2) = \text{var}(m_1) + \text{var}(m_2) - \mathbf{2 \text{cov}(m_1, m_2)}$$

$$Z = \frac{m_1 - m_2}{\sqrt{\text{var}(m_1) + \text{var}(m_2) - 2 \text{cov}(m_1, m_2)}}$$

Z_{apparié} > Z_{indépendant} → gain de puissance statistique

Plan

- Objectifs
- Introduction
- Test Z de l'écart réduit
- Test t de Student

Test Z de comparaisons de 2 moyennes estimées pour échantillons appariés

$$Z = \frac{m_1 - m_2}{\sqrt{\text{var}(m_1) + \text{var}(m_2) - 2 \text{cov}(m_1, m_2)}}$$



$$\frac{S_1^2}{n_1}$$

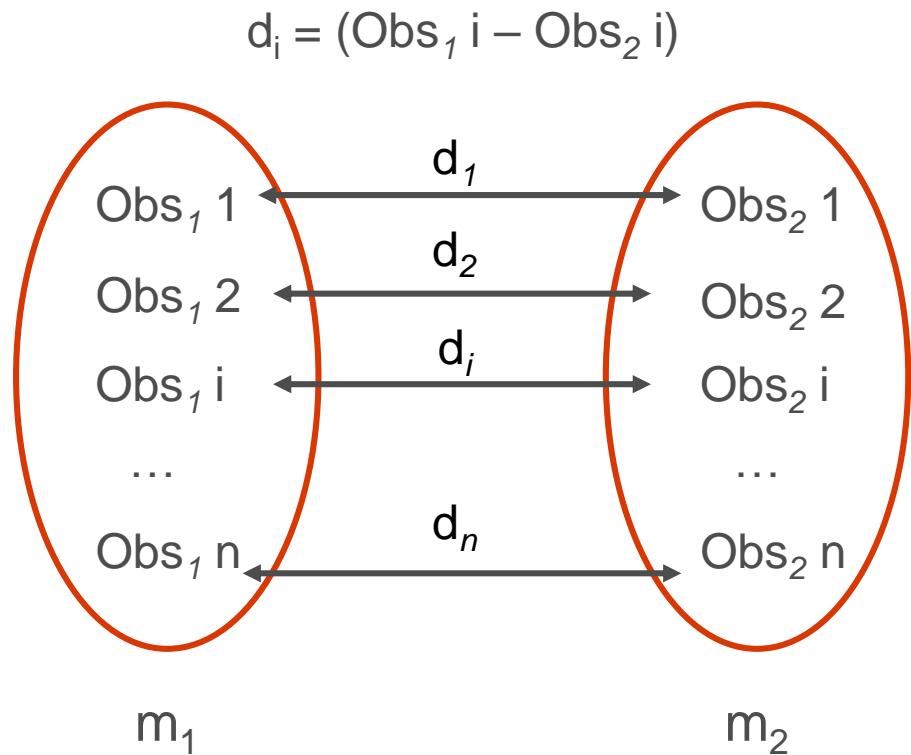
$$\frac{S_2^2}{n_2}$$

ne peut pas être estimée directement car on ne dispose que d'une mesure de m_1 et une mesure de m_2

→ Il faut estimer $\text{var}(m_1 - m_2)$ d'une autre façon

Test Z de comparaisons de 2 moyennes estimées pour échantillons appariés

L'unité d'analyse devient la différence d_i entre l'observation 1 et l'observation 2 pour chaque sujet



$$m_d = \frac{\sum_{i=1}^n d_i}{n}$$

$$s_d^2 = \frac{\sum_{i=1}^n (d_i - m_d)^2}{(n - 1)}$$

Test Z de comparaisons de 2 moyennes estimées pour échantillons appariés

$$Z = \frac{m_1 - m_2}{\sqrt{\text{var}(m_1) + \text{var}(m_2) - 2 \text{cov}(m_1, m_2)}} = \frac{m_d}{\sqrt{\text{var}(m_d)}}$$

$$(m_1 - m_2) = m_d, \text{ avec } m_d = \frac{\sum_{i=1}^n d_i}{n}$$

$$\text{var}(m_d) = s_d^2 / n, \text{ avec } s_d^2 = \frac{\sum_{i=1}^n (d_i - m_d)^2}{(n - 1)}$$

Rappel : La moyenne m d'une variable quantitative X calculée sur un échantillon d'effectif n est elle-même une variable aléatoire de loi normale $N(\mu, \sigma/\sqrt{n})$, quelle que soit la loi de distribution de la variable X si l'effectif de l'échantillon $n \geq 30$ (Chapitre 11, Théorème central limite)

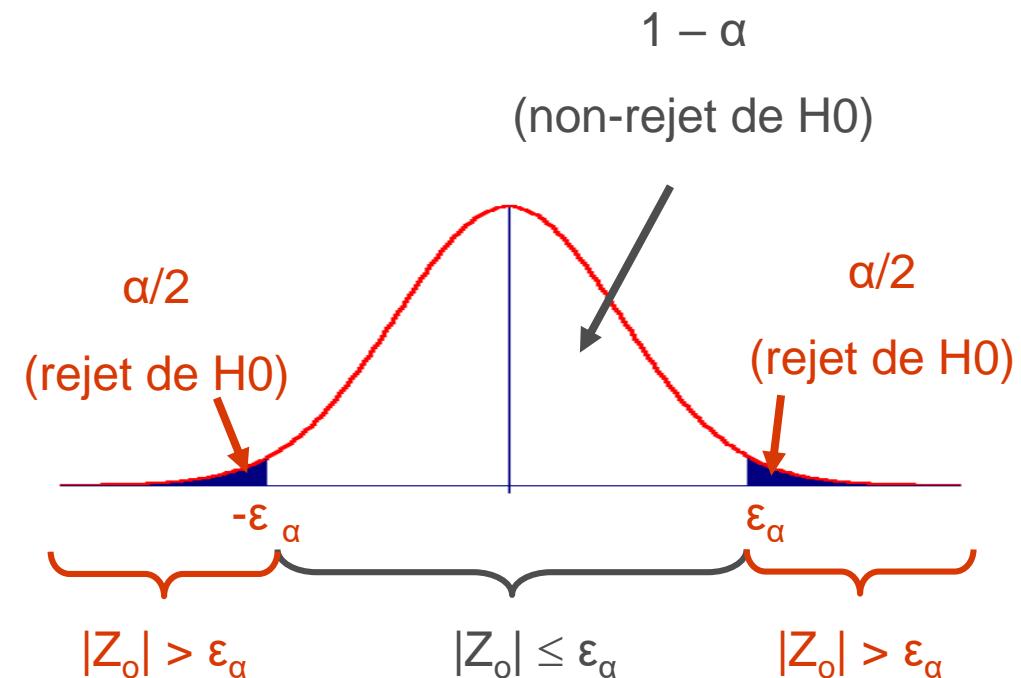
Test Z de comparaisons de 2 moyennes estimées pour échantillons appariés

$$H_0 : \mu_d = 0 (\mu_1 = \mu_2)$$

$$H_1 : \mu_d \neq 0 (\mu_1 \neq \mu_2)$$

Si $n \geq 30$ paires

$$Z = \frac{m_d}{\sqrt{\frac{s_d^2}{n}}} \rightarrow N(0,1)$$



- Objectifs
- Introduction
- Test Z de l'écart réduit
- Test t de Student

Test t de Student de comparaisons de 2 moyennes estimées pour échantillons appariés

$H_0 : \mu_d = 0 (\mu_1 = \mu_2)$

$H_1 : \mu_d \neq 0 (\mu_1 \neq \mu_2)$

Si la distribution des différences individuelles est **normale** :

$$t = \frac{m_d}{\sqrt{\frac{s_d^2}{n}}} \rightarrow t_{(n-1) ddl}$$

Messages clés

m_1	m_1	effectif	test	conditions
estimée	estimée	$n \geq 30$	Z	-
(échantillons appariés)		n	$t_{(n - 1) \text{ ddl}}$	Normalité de d_i

Glossaire

- **Covariance** : variance conjointe de 2 variables X et Y

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^N (X_i - \mu_X)(Y_i - \mu_Y)}{N}$$

Cas particulier : $X = Y$

$$\text{cov}(X, X) = \frac{\sum_{i=1}^N (X_i - \mu_X)(X_i - \mu_X)}{N} = \frac{\sum_{i=1}^N (X_i - \mu_X)^2}{N} = \text{var}(X)$$

X et Y indépendantes

cas particulier Y constant quelle que soit la valeur de X

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^N (X_i - \mu_X)(Y_i - \mu_Y)}{N} = 0$$

0 car $Y_i = \text{constante} = \mu_Y$

Mentions légales

L'ensemble de ce document relève des législations française et internationale sur le droit d'auteur et la propriété intellectuelle. Tous les droits de reproduction de tout ou partie sont réservés pour les textes ainsi que pour l'ensemble des documents iconographiques, photographiques, vidéos et sonores.

Ce document est interdit à la vente ou à la location. Sa diffusion, duplication, mise à disposition du public (sous quelque forme ou support que ce soit), mise en réseau, partielles ou totales, sont strictement réservées à l'Université Grenoble Alpes (UGA).

L'utilisation de ce document est strictement réservée à l'usage privé des étudiants inscrits à l'Université Grenoble Alpes (UGA), et non destinée à une utilisation collective, gratuite ou payante.