Science des données I : module 9

Test bilatéral *versus* unilatéral

Philippe Grosjean & Guyliann Engels

Université de Mons, Belgique
Laboratoire d'Écologie numérique

https://wp.sciviews.org
sdd@sciviews.org

De quoi allons-nous parler ?

Le test t de Student et son homologue non paramétrique, le test de Wilcoxon-Mann-Whitney proposent plusieurs variantes.

Nous nous focalisons sur les variantes bi- et unilatérales :

- au niveau de la **définition des hypothèses**
- au niveau du **positionnement des zones de rejet et non rejet** de H_0

Le test d'hypothèse t de Student

Définition de l'hypothèse nulle :

- Comparaison des moyennes entre deux **populations distinctes** *p1* et *p2* (test indépendant, différence des moyennes)

$$H_0 : \overline{x_{p1}} - \overline{x_{p2}} = cste$$

- Comparaison de la moyenne entre deux mesures x_a et x_b réalisées à chaque fois sur le **même individu** (test apparié, moyenne des différences)

$$H_0 : \overline{x_a - x_b} = cste$$

- Comparaison de la moyenne **d'une seule population** à une valeur constante de référence (test univarié)

$$H_0 : \overline{x} = cste$$

UMONS

Variantes bilatérale et unilatérales

Exemple du test indépendant où la moyenne de la variable x est comparée entre deux populations *p1* et *p2*.

- L'hypothèse nulle est **toujours la même**:

$$H_0 : \overline{x_{p1}} - \overline{x_{p2}} = cste$$

- **Test bilatéral :** l'hypothèse alternative peut être > ou < à *cste*

$$H_1 : \overline{x_{p1}} - \overline{x_{p2}} \neq cste$$

- **Test unilatéral à gauche :** l'hypothèse alternative est uniquement < *cste*

$$H_1 : \overline{x_{p1}} - \overline{x_{p2}} < cste$$

- **Test unilatéral à droite :** l'hypothèse alternative est uniquement > *cste*

$$H_1 : \overline{x_{p1}} - \overline{x_{p2}} < cste$$

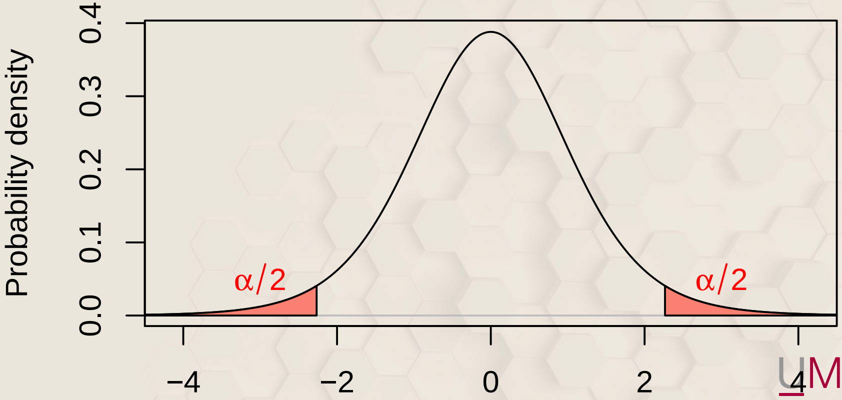
UMONS

Aires de rejet et non rejet en bilatéral

Exemple : distribution t réduite pour $n = 10$ (moyenne = 0, écart type = 1, ddl = 9)

$$H_0 : \overline{x_{p1}} - \overline{x_{p2}} = 0$$
$$H_1 : \overline{x_{p1}} - \overline{x_{p2}} \neq 0$$

On répartit α en **deux moitiés** aux deux extrémités de la distribution.

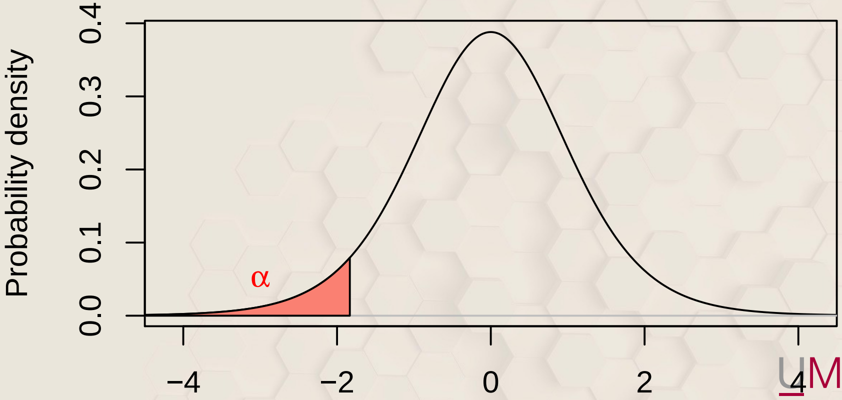


Aires de rejet et non rejet en unilatéral à gauche

Exemple : distribution t réduite pour $n = 10$ (moyenne = 0, écart type = 1, ddl = 9)

$$H_0 : \overline{x_{p1}} - \overline{x_{p2}} = 0$$
$$H_1 : \overline{x_{p1}} - \overline{x_{p2}} < 0$$

On répartit α uniquement à la **gauche** de la distribution.

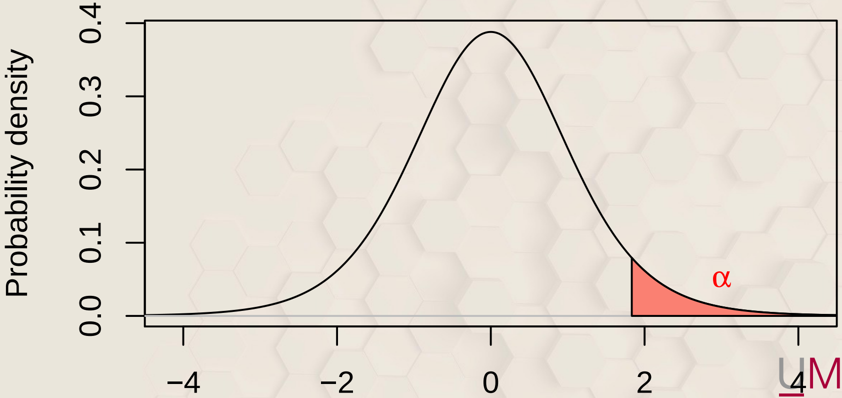


Aires de rejet et non rejet en uniliatéral à droite

Exemple : distribution t réduite pour $n = 10$ (moyenne = 0, écart type = 1, ddl = 9)

$$H_0 : \overline{x_{p1}} - \overline{x_{p2}} = 0$$
$$H_1 : \overline{x_{p1}} - \overline{x_{p2}} > 0$$

On répartit α uniquement à la **droite** de la distribution.



Choix de H_1

- Le choix se fait en fonction d'information **connue *a priori*** du phénomène (valeur impossible d'un côté ou de l'autre)

- Si pas d'information on utilise **un test bilatéral**

- Si test unilatéral, attention au calcul en fonction de **l'ordre des niveaux de la variable qualitative** (*p1 - p2*), voir fonction `levels()` pour afficher ces niveaux

UMONS

Infos pratiques

- Les mini-présentations sont disponibles sur https://github.com/BioDataScience-Course/sdd_lessons et sont enregistrées dans Teams

- **A vous la parole maintenant !**

UMONS