

SYS865 Inférence statistique avec programmation R

Ornwipa Thamsuwan

21 février 2024

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Recap

**Puissance
statistique**

**Détermination de
la taille
d'échantillon**

**D'autre option du
logiciel : G*Power**

Recap

- ▶ Objectifs de l'étude
- ▶ Formulation d'hypothèse
- ▶ Collecte des données : plan d'échantillonnage
- ▶ Analyse des données : **choix du test statistique**
- ▶ Décision et interprétation

- ▶ Objectifs de l'étude
- ▶ Formulation d'hypothèse
- ▶ Collecte des données : plan d'échantillonnage
- ▶ Analyse des données : **choix du test statistique**
- ▶ Décision et interprétation

Comment choisir le test statistique ?

Les conditions pour les choix du test

- ▶ Test de normalité (Shapiro-Wilk)

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Les conditions pour les choix du test

- ▶ Test de normalité (Shapiro-Wilk)
- ▶ Test de homogénéité des variances
 - ▶ Test F pour les données **normalement distribuées**
 - ▶ Test de Levene

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Les conditions pour les choix du test

- ▶ Test de normalité (Shapiro-Wilk)
- ▶ Test de homogénéité des variances
 - ▶ Test F pour les données **normalement distribuées**
 - ▶ Test de Levene

Les paramètres à étudier

- ▶ La moyenne d'un échantillon ou la différence entre deux échantillons appariés
 - ▶ Test t avec paired=TRUE pour les données **normalement distribuées**
 - ▶ Test de Wilcoxon pour échantillons appariés

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Les conditions pour les choix du test

- ▶ Test de normalité (Shapiro-Wilk)
- ▶ Test de homogénéité des variances
 - ▶ Test F pour les données **normalement distribuées**
 - ▶ Test de Levene

Les paramètres à étudier

- ▶ La moyenne d'un échantillon ou la différence entre deux échantillons appariés
 - ▶ Test t avec paired=TRUE pour les données **normalement distribuées**
 - ▶ Test de Wilcoxon pour échantillons appariés
- ▶ Les moyennes des deux échantillons indépendants
 - ▶ Test t par défaut pour les données **normalement distribuées**
 - ▶ Test de rang somme de Wilcoxon

Recap

Puissance
statistique

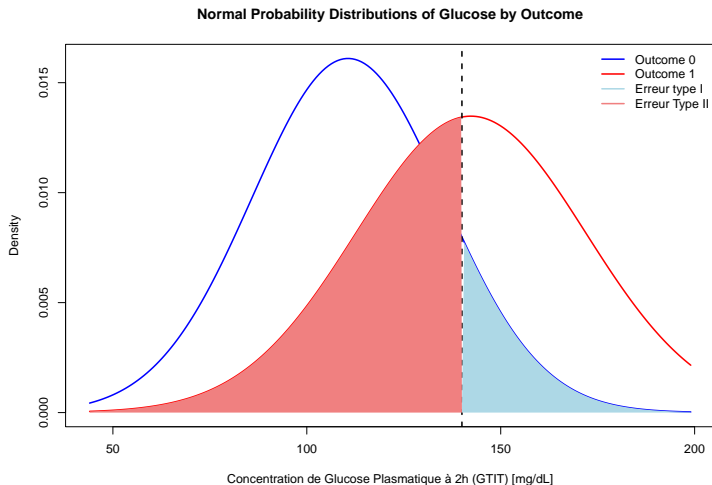
Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Erreur statistique (retour)

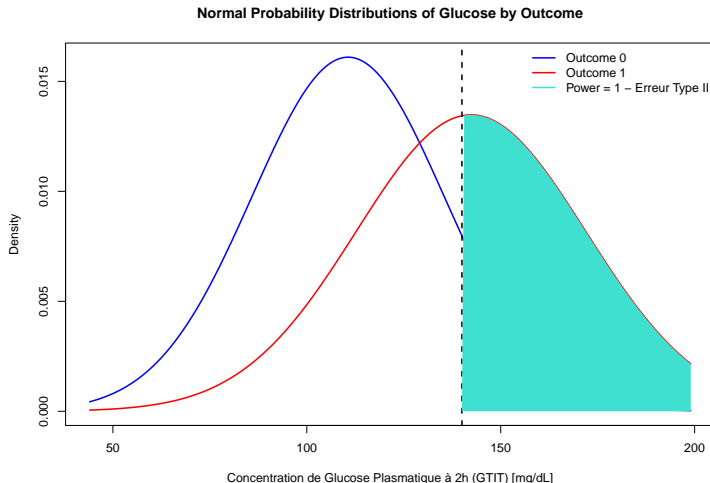
Erreur type I (α) : Rejeter H_0 quand H_0 est vraie.

Erreur type II (β) : Ne pas rejeter H_0 quand H_0 est fausse.



Erreur statistique (retour)

“Power” ($1 - \beta$) : Correctement rejeter H_0 quand H_0 est fausse, ou quand H_1 est vraie.



Recap

**Puissance
statistique**

**Détermination de
la taille
d'échantillon**

**D'autre option du
logiciel : G*Power**

Puissance statistique

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

		Reality	
		Positive	Negative
Study findings	Positive	True positive (Power) $(1 - \beta)$	False positive Type I error (α)
	Negative	False negative Type II error (β)	True negative

Figure 1: Tableau de contingence

1. Définir les hypothèses et décider du type de test t
 - ▶ **Hypothèse nulle (H_0)**: Il n'existe aucune différence entre les groupes que vous comparez.
 - ▶ **Hypothèse alternative (H_1)**: Il existe une différence significative.
 - ▶ Décidez si l'on réalise un test t unilatéral (si l'on attend une différence dans une direction spécifique) ou bilatéral (en cas où les différences dans les deux directions sont pertinentes).

1. Définir les hypothèses et décider du type de test t
 - ▶ **Hypothèse nulle (H_0)**: Il n'existe aucune différence entre les groupes que vous comparez.
 - ▶ **Hypothèse alternative (H_1)**: Il existe une différence significative.
 - ▶ Décidez si l'on réalise un test t unilatéral (si l'on attend une différence dans une direction spécifique) ou bilatéral (en cas où les différences dans les deux directions sont pertinentes).
2. Choisir un niveau de signification (α). Typiquement, $\alpha = 0.05$.

1. Définir les hypothèses et décider du type de test t
 - ▶ **Hypothèse nulle (H_0)**: Il n'existe aucune différence entre les groupes que vous comparez.
 - ▶ **Hypothèse alternative (H_1)**: Il existe une différence significative.
 - ▶ Décidez si l'on réalise un test t unilatéral (si l'on attend une différence dans une direction spécifique) ou bilatéral (en cas où les différences dans les deux directions sont pertinentes).
2. Choisir un niveau de signification (α). Typiquement, $\alpha = 0.05$.
3. Supposer que la taille de l'échantillon (n) pour chaque groupe est donnée.

4. Déterminer la taille de l'effet

La taille de l'effet de Cohen's d est définie comme:

$$d = \frac{\mu_1 - \mu_2}{SD_{\text{poolisée}}}$$

Où μ_1 et μ_2 sont les moyennes des deux groupes, et $SD_{\text{poolisée}}$ est l'écart-type poolisé, calculé en fonction des écarts-types des deux groupes.

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

4. Déterminer la taille de l'effet

La taille de l'effet de Cohen's d est définie comme:

$$d = \frac{\mu_1 - \mu_2}{SD_{\text{poolisée}}}$$

Où μ_1 et μ_2 sont les moyennes des deux groupes, et $SD_{\text{poolisée}}$ est l'écart-type poolisé, calculé en fonction des écarts-types des deux groupes.

$$SD_{\text{poolisée}} = \sqrt{\frac{(n_1 - 1) \times SD_1^2 + (n_2 - 1) \times SD_2^2}{n_1 + n_2 - 2}}$$

Où n_1 et n_2 sont les tailles d'échantillon des deux groupes, et SD_1 et SD_2 sont les déviations standard des deux groupes.

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

5. Calculer ou estimer la puissance $(1 - \beta)$ avec R

Recap

Puissance statistique

Détermination de la taille d'échantillon

D'autre option du logiciel : G*Power

5. Calculer ou estimer la puissance $(1 - \beta)$ avec R

Paramètres

```
effect_size <- 0.5  
n1 <- 30  
n2 <- 40  
alpha <- 0.05
```

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

5. Calculer ou estimer la puissance ($1 - \beta$) avec R

Paramètres

```
effect_size <- 0.5  
n1 <- 30  
n2 <- 40  
alpha <- 0.05
```

Calcul des degrés de liberté

```
df <- n1 + n2 - 2  
df
```

```
## [1] 68
```

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Calcul de la valeur t critique

```
t_critical <- qt(1 - alpha / 2, df)
t_critical
```

```
## [1] 1.995469
```

La fonction `qt()` fournit le quantile de la distribution t.

Calcul du paramètre de non-centralité

```
ncp <- effect_size * sqrt((n1 * n2) / (n1 + n2))  
ncp
```

```
## [1] 2.070197
```

Le paramètre de non-centralité représente à quel point l'effet réel est éloigné de l'hypothèse nulle dans le contexte d'une hypothèse alternative spécifique.

Le paramètre de non-centralité est ajusté pour les tailles d'échantillon inégales en utilisant la formule suivante :

$$ncp = \text{effect size} \times \sqrt{\frac{n_1 \times n_2}{n_1 + n_2}}$$

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Calcul de la puissance

```
power <- 1 - pt(t_critical, df, ncp)
power
```

```
## [1] 0.5322582
```

La fonction `pt()` donne la fonction de répartition
(c'est-à-dire, la probabilité cumulée) pour la distribution `t`.

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Si nous n'avons pas encore collecté de données ...

Et nous voulons connaître la taille de l'échantillon pour
atteindre une puissance statistique prédéfinie ...

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

Détermination de la taille d'échantillon

Recap

Puissance
statistique

Détermination de
la taille
d'échantillon

D'autre option du
logiciel : G*Power

D'autre option du logiciel : G*Power

D'autre option du logiciel : G*Power

Lien pour le télécharger :

<https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower>



Figure 2: G*Power