

Tests multiples

G. Marot-Briend
guillemette.marot@univ-lille.fr

2021-2022

Rappels sur les tests

Question fréquente :

Est-ce que deux variables sont liées ?

La réponse dépend du type des variables.

- deux variables qualitatives : test du χ^2
- deux variables quantitatives : test du coefficient de corrélation
- une quantitative, une qualitative : t-test ou ANOVA

Tout comme les intervalles de confiance, les tests sont utilisés pour généraliser des résultats dans une population à partir d'observations d'un échantillon.

Notion de risques (α, β) - Rappel

Décision	Réalité	
	H_0	H_1
non rejet H_0	conclusion correcte	risque de 2ème espèce (β)
rejet H_0	risque de 1ère espèce (α)	conclusion correcte

Décision	Réalité	
	H_0	H_1
non rejet H_0	Niveau de confiance $1 - \alpha$	β
rejet H_0	α	Puissance $1 - \beta$

Degré de signification (p-value)

Ne pas confondre niveau de signification (risque de 1ère espèce α), niveau de confiance ($1 - \alpha$) et degré de signification (p-value).

Degré de signification ou p-value $p(t)$:

pour une réalisation t d'une statistique de test T , probabilité (calculée sous l'hypothèse nulle) d'obtenir une statistique de test au moins aussi extrême que celle réellement observée.

Autrement dit, la p-value correspond à la plus petite valeur de risque α accepté pour cette réalisation.

$p\text{-value} < \alpha \Rightarrow$ on rejette H_0

$p\text{-value} \geq \alpha \Rightarrow$ on ne rejette pas H_0

Degré de signification (p-value)

Plus la réalisation de la statistique de test est grande en valeur absolue, plus la p-value est petite.

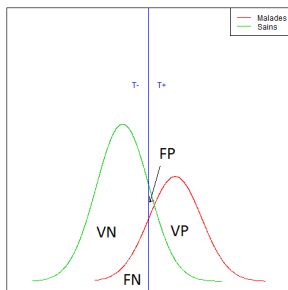
Cas bilatéral : $p(t) = P_{H_0}(|T| \geq |t|) = 2.(1 - F(|t|))$

Cas unilatéral à droite : $p(t) = P_{H_0}(T \geq t) = 1 - F(t)$

Cas unilatéral à gauche : $p(t) = P_{H_0}(T \leq t) = F(t)$

Vrais positifs et faux positifs (à un seuil donné)

	H0 rejetée	H0 non rejetée
H0 vraie	Faux Positif	Vrai Négatif
H0 fausse	Vrai Positif	Faux Négatif



Plan

1 Tests multiples

Tests multiples

Dans le cas de nombreux tests contrôlant chacun un taux d'erreur de première espèce à 5%, le nombre de faux positifs peut devenir très grand .

Ex : $m = 10000$ tests et $\alpha = 0.05 \rightarrow \mathbb{E}(FP) = 500$

Nécessité de corriger pour la **multiplicité des tests**.

- Family Wise Error Rate : rejeter à tort au moins une hypothèse nulle (e.g. Bonferroni)
- False Discovery Rate : contrôler la proportion attendue de faux positifs parmi les positifs (e.g. Benjamini Hochberg)

Family Wise Error Rate (FWER)

Definition

Probabilité d'obtenir au moins une erreur de type I (faux positif)

$$FWER = \mathbb{P}(FP \geq 1)$$

Procédure de Bonferroni

Soit chaque test est réalisé au niveau $\alpha = \alpha^*/m$ (m : nombre de tests)

soit on utilise une p-value ajustée $p_{Bonf_i} = \min(1, p_i * m)$ et $FWER \leq \alpha^*$.

Ex : pour $m = 2000$ et $\alpha^* = 0,05$; $\alpha = 2,5 \cdot 10^{-5}$.

Facile mais manque de sensibilité quand beaucoup de tests.

Family Wise Error Rate (FWER)

Justification :

$$\begin{aligned} FWER &= \mathbb{P}(FP \geq 1) = \mathbb{P}(\cup(FP_{test} \geq 1)) \\ &\leq \sum_{tests} \mathbb{P}(FP_{test} \geq 1) \\ &\leq m\alpha \end{aligned}$$

Remarque : Sidak suppose l'indépendance entre les tests et écrit :
 $FWER \leq 1 - (1 - \alpha)^m$

False Discovery Rate (FDR)

Idee : Ne pas contrôler la probabilité d'avoir au moins une erreur de type I mais la proportion d'erreurs
 \Rightarrow moins conservatif que le contrôle du FWER.

Définition

Le taux de faux positifs FDR de (Benjamini Hochberg, 1995) est la proportion attendue d'erreurs de type I parmi les hypothèses rejetées

$$\text{FDR} = \mathbb{E}(FP/P) \text{ si } P > 0 \text{ et } 0 \text{ si } P = 0$$

Prop

$$\text{FDR} \leq \text{FWER}$$

Tests multiples : éléments clefs

- Important de prendre en compte la multiplicité des tests
- le choix entre FDR et FWER dépend du coût associé aux FN et FP

Contrôler le FWER :

Avoir une grande confiance dans les gènes différentiellement exprimés (contrôle très fort). Accepter de ne pas détecter certains gènes intéressants. (manque de sensibilité \Leftrightarrow peu de DE)

Contrôler le FDR :

Accepter une proportion de FP parmi les DE. Très intéressant dans une étude exploratoire.