Taille d'un effet – Puissance d'un test

Exemple introductif

On se place dans une situation de comparaison de deux groupes indépendants, avec une VD numérique.

Ces deux groupes sont issus de deux populations et l'on estime que :

– La moyenne de la VD dans la population 1 est 100.
– La moyenne de la VD dans la population 2 est 105.

On se pose des questions telles que :

– La différence entre les deux groupes est-elle grande, facile à mettre en évidence ou au contraire petite, difficile à mettre en évidence ?

– Est-on sûr de pouvoir mettre en évidence une différence entre les deux groupes ?

– Si les effectifs des échantillons sont $n_1 = 25$ et $n_2 = 30$, pourra-t-on mettre en évidence une différence entre les deux groupes ?

– Quelles tailles d'échantillons doit-on choisir pour avoir au moins 80% de chances de mettre en évidence une différence entre les deux groupes ?

Rappel concernant les tests statistiques : erreurs de type I et II

		Hypothèse vraie	
		H_0	H_1
Hypothèse	H_0	$1-\alpha$	β
retenue	H_1	α	$1-\beta$

• α : seuil de significativité. C'est aussi la probabilité de rejeter H_0 alors que H_0 est vraie (risque de première espèce ou risque de commettre une erreur de type I) : risque de voir une différence là où il n'y en a pas.

• β : risque de seconde espèce. C'est la probabilité d'accepter H_0 alors que H_0 est fausse (risque de commettre une erreur de type II) : risque de ne pas mettre en évidence une différence qui, pourtant, existe.

• $1-\beta$: probabilité de détecter correctement un cas où H_0 doit être rejetée. **Puissance du test**.

Comment influer sur la puissance d'un test ?

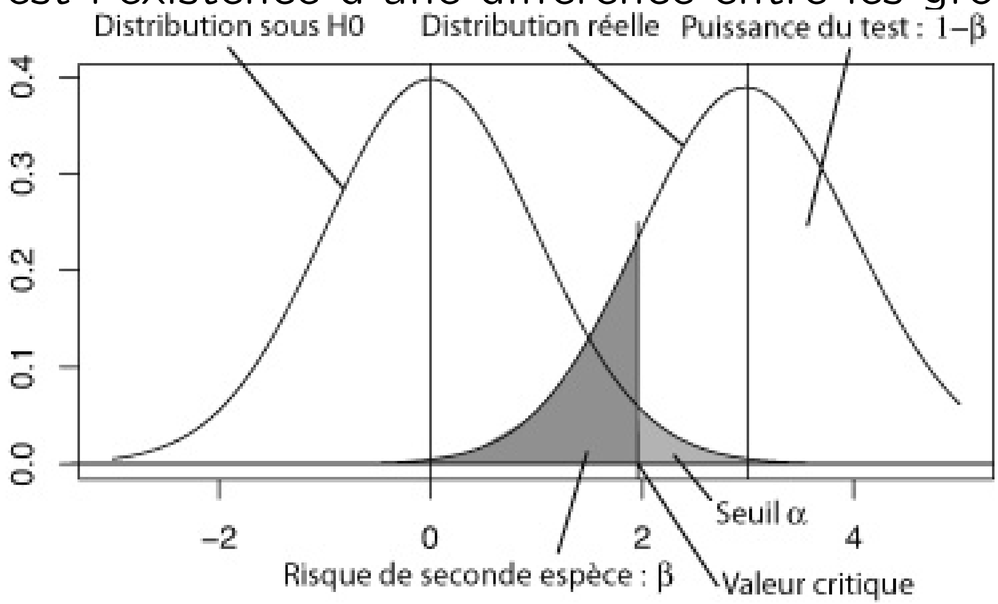
– Changement du seuil

– Choix d'une autre variable dépendante

– Taille des échantillons

Risque β et puissance du test $1 - \beta$

Illustration graphique dans le cas où l'hypothèse *vraie* est l'existence d'une différence entre les groupes.



Sous l'hypothèse nulle H_0, la distribution de la statistique de test est celle représentée par la courbe de gauche.

Mais, la *vraie* distribution de la statistique de test est représentée par la courbe de droite.

La probabilité que le test nous permette de conclure correctement est de $1 - \beta$.

Taille d'un effet - Effet calibré de Cohen

Deux groupes appareillés - Variable numérique

Moyenne des différences individuelles dans la population : δ.
Ecart type des différences individuelles : σ.

La **taille de l'effet** est définie par :

$$d = \frac{\mid \delta \mid}{\sigma}$$

Estimation de la taille de l'effet à partir d'un échantillon : **effet calibré** observé sur l'échantillon :

$\overline{d}$: moyenne observée sur l'échantillon
s_c : écart type corrigé observé sur l'échantillon.

L'effet calibré est défini par :

$$\text{EC} = \frac{\mid \overline{d} \mid}{s_c}$$

Critère psychométrique :

- EC $< 1/3$: effet faible
- $1/3 \leq$ EC $\leq 2/3$: effet moyen
- EC $> 2/3$: effet important.

Deux groupes indépendants - Variable numérique

Variable dépendante X définie sur deux populations.
Moyennes (sur les populations) : μ_1 et μ_2.
Ecart type (le même pour les deux populations) : σ.

Taille de l'effet :

$$d = \frac{|\mu_1 - \mu_2|}{\sigma}$$

Effet calibré

Paramètres évalués à partir de deux échantillons de tailles n_1 et n_2 :

Moyennes : $\overline{x}_1$, $\overline{x}_2$
Ecarts types (non corrigés) : s_1 et s_2
Ecart type corrigé pondéré s défini par : $s^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$

Effet calibré : $\text{EC} = \dfrac{|\overline{x}_1 - \overline{x}_2|}{s}$

Deux groupes indépendants - Variable dichotomique

Paramètres évalués à partir de deux échantillons de tailles n_1 et n_2 :

Fréquences observées : f_1, f_2
Fréquence pondérée moyenne : f, définie en général par : $f = \dfrac{n_1 f_1 + n_2 f_2}{n_1 + n2}$

Effet calibré : $\text{EC} = \dfrac{|f_1 - f_2|}{\sqrt{f(1-f)}}$

Estimer la puissance d'un test

Exemple

– On se place dans une situation de groupes appareillés.
– On connaît une évaluation de la taille de l'effet.
– Pour une taille d'échantillon donnée, et un seuil donné, quelle estimation peut-on faire de la puissance du test (autrement dit : quelles chances a-t-on de tirer un échantillon permettant de conclure sur H_1) ?

Par exemple :
– on estime que $d = 0.33$
– on choisit $n = 40$ et $\alpha = 5\%$

On remarque que la statistique de test s'écrit :

$$t = \text{EC}\ \sqrt{n}$$

On calcule : $\Delta = d\sqrt{n} = 0.33\sqrt{40} = 2.09$

Par lecture de la table, on a : Puissance = 0.56

Estimer la taille d'échantillon requise

Deux groupes appareillés - Variable numérique

– On se place dans une situation de groupes appareillés
– On connaît une évaluation de la taille de l'effet d
– On se donne un seuil donné et une puissance de test
– Quelle taille minimale d'échantillon faut-il choisir ?

On remarque que la statistique de test s'écrit :

$$t = \text{EC}\ \sqrt{n}$$

On introduit : $\Delta = d\sqrt{n}$

Par lecture de la table, on obtient la valeur Δ requise.

On calcule ensuite n de façon que : $\Delta = d\sqrt{n}$.

Exemple : Dans la situation précédente ($d = 0.33$ et $\alpha = 5\%$), quelle taille d'échantillon faut-il choisir pour avoir au moins 4 chances sur 5 de conclure sur H_1 ?

On veut que la puissance du test soit au moins 80%.
Lecture de la table : $\Delta = 2.80$
L'équation : $0.33\sqrt{n} = 2.80$ donne : $n = 72$.

La taille minimale d'échantillon est donc de 72.

Deux groupes indépendants équilibrés

Variable numérique ou dichotomique.

Soit n la taille commune de chacun des deux échantillons. La statistique de test s'écrit alors :

$$t = \text{EC} \sqrt{\frac{n}{2}}$$

On raisonne comme précédemment, mais on a alors :

$$\Delta = d\sqrt{\frac{n}{2}}$$

Deux groupes indépendants non équilibrés

Dans ce cas : $\Delta = d\sqrt{\frac{n_h}{2}}$

où n_h est la *moyenne harmonique* des deux tailles d'échantillons définie par :

$$\frac{2}{n_h} = \frac{1}{n_1} + \frac{1}{n_2} \quad \text{ou} \quad n_h = \frac{2n_1n_2}{n_1 + n_2}$$