- Biostatistique -

Chapitre 02 : Les Tests d'hypothèse :

Lorsqu'on effectue une comparaison entre deux ou plusieurs séries de données on observe toujours une différence entre les paramètres mesurés.

Introduction

Les tests d'hypothèse sont des procédures statistiques permettant de prendre une décision sur une hypothèse concernant une population à partir d'un échantillon

Le but du test est de déterminer si cette différence observé est due au hasard ou au contraire la différence est réelle.

- Concepts de base :

= Hypothèse nulle (H_0) : Hypothèse à tester, elle représente souvent l'absence d'effet ou de différence. ex. ($\mu_1 = \mu_2$).

. Hypothèse alternative (H_1) : Hypothèse opposée à H_0, elle représente une différence ou un effet significatif. (ex: $\mu_1 <> \mu_2$; $\mu_1 \overset{<}{\underset{>}{\text{ou}}} \mu_2$).

Niveau de signification (seuil) : la probablité de rejeter H_0 alors qu'elle est vraie (erreur de type 01) ; (appelé risque α.)

- détermine si la différence est significative ou non (1% – 5% – 10%

Étapes d'un test d'hypothèse

1) : Déterminer le genre dei test (conformité ou homogénéité)

2) : Formuler les hypothèses. H_0 : H_1 {
   - H_0 : Hypothèse nulle : la différence n'est pas significative ∄
   - H_1 : Hypothèse alternative : la différence est significative ∃

3). Fixer α (souvent 5%).

4) : Sélectionner le test statistique approprié (Z ou t)

5) : Calculer statistique de test.

6) : Comparer la statistique calculée avec la statistique théorique.

7) : prendre la décision : rejeter (∃) ou Accepter (∄) H_0

(la décision relier à H_0)

Comparaison

Echantillon / Population

Comparaison entre une valeur expérimentalle et une valeur théorique (référence – norme).

⟹ Test de conformité

- Il s'intéresse :
  - moyenne
  - pourcentage.
- Question :

Est-il-conforme ?

Est-il-représentatif ?

Echantillon / Echantillon

comparaison entre deux valeur expérimentalles de 2 échantillons différents ou de la même populat

⟹ Test d'homogénéite.

- Il s'intéresse
  - moyenne
  - pourcentage
  - variance
- Question

Sont elles homogènes ?

N.B

H0 $\overline{X}_{ex} = \overline{X}_{thé}$: ∄ difference non significative : échantillon ∈ Population

H1 : $\overline{X}_{ex} <> \overline{X}_{th}$: ∃ différence significative : échantillon ∉ population

Cas de Echan / Popu

• Moyenne (bilatéral).

Test de conformité : pour comparer une moyenne expérimentalle (m_0) avec une moyenne théorique donnée (m).

1). Enoncer H_0, H_1

H_0: $m = m_0$ ∄ differenca non significative

H_1 $m <> m_0$ ∃ différence significative

2) Fixer α. (1% . 5% . 10%)

3) Statistique de test:

taille de l'échantillon

$n \geq 30$ →

$$Z_{cal} = \frac{\bar{X}_{ech} - \mu_{pop}}{\sqrt{\frac{\sigma^2_{pop}}{n}}}$$

$n < 30$

σ_{pop} connu → $$Z_{cal} = \frac{\bar{X}_{ech} - \mu_{pop}}{\sqrt{\frac{\sigma^2_{pop}}{n}}}$$

σ_{pop} inconnu ↓

$$t_{cal} = \frac{\bar{x} - \mu}{\sqrt{\frac{\hat{\sigma}^2}{n}}} = \frac{\bar{x} - \mu}{\sqrt{\frac{S^2}{n-1}}}$$

avec $\hat{\sigma}^2$ estimé $= \frac{1}{n-1} \sum (x_i - \bar{x})^2$

4). Déterminer $Z_{théo}$ ou $t_{théo}$

$Z_{théo}$ pour $\alpha = 5\%$

à partir de table de Z

$1 - \frac{\alpha}{2}$ = proba

$Z_{théo} = \dots + \dots$

$t_{théo}$ pour $\alpha = 5\%$

ddl = n − 1 : ligne ...

et $\frac{\alpha}{2}$: ... colonne

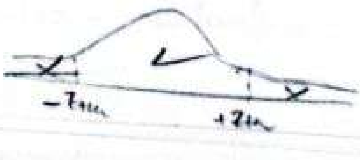
à partir de Table de student t

$t_{théo}$ = ... valeur

5) Décision $P\left(-\frac{\alpha}{2} < T > \frac{\alpha}{2}\right) = 1 - \alpha$

Si ∄ : $Z_{calculé} = Z_{théo}$. Ho Accepté.

Si ∃ $Z_{cal} \in [-Z_{théo} ; +Z_{théo}]$: Ho rejeté
H1 Accepté



• Pourcentage :

Test de conformité pour comparer une valeur expé (P_0) avec une valeur théorique donnée (P). (Pourcentage)

1) Enoncer Ho H1

Ho: $P = P_0$ ∄

H_1 : $P <> P_0$ ∃

2) Fixer α. (souvent 5%)

3) Statistique de test

$$Z_{cal} = \frac{\frac{K}{n} - P_0}{\sqrt{\frac{P_0 q_0}{n}}}$$

K : nombre de succés. 30 [illegible] 100 [illegible]

n : taille de l'echantillon

P_0 : $\frac{\% \text{ donné}}{100}$

$q_0 = 1 - P_0$

$$P\left(-\frac{\alpha}{2} \leqslant T \geqslant \frac{\alpha}{2}\right) = 1 - \alpha$$

Cas de Echan / Echan

[redacted] : (Bilatérale)

Test d'homogénéité pour comparer 2 moyennes expérimentales m_1 m_2.

1): Enoncer H0, H1

$H_0 : m_1 = m_2$

$H_1 \quad m_1 <> m_2$

2). Fixer α (1% - 5%) 10%

Souvent α = 5%

3). statistique de test.

Taille des échantillons
Variance de la population

$\sigma_{pop_1}, \sigma_{pop_2}$: connus : $n \geq 30$

$$Z_{cal} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma^2_{pop_1}}{n_1} + \frac{\sigma^2_{pop_1}}{n_2}}}$$

4)= Détermine $Z_{théo}$

$Z_{théo}$ pour α = 5%

$1 - \frac{\alpha}{2} = [\text{proba}]$

de Table de Z : on trouve $Z_{théo}$

(la m̂ méthode précédente)

5) Décision

Si $Z_{cal} \in [-Z_{théo} ; Z_{théo}]$

Ho : est rejeté - H1 est accepte

$\sigma_{pop_1}, \sigma_{pop_2}$: inconnus : $n < 30$

Test de comparaison de variance estimé

Égales

$$t_{cal} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Inégales

$n_1 . n_1 \geq 30$ → $\cancel{n_1, n_2 < 30}$

$$Z_{cal} \,①= \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} = ② \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1 - 1} + \frac{S_2^2}{n_2 - 1}}}$$

$$\hat{\sigma}^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2} \; ⊛ \text{ avec } \quad S_1^2 = \frac{1}{n_1 - 1} \sum (x_{i_1} - \overline{x}_1)^2$$

+ بتبسيط الكسر ① يعطي العلاقة ⊛ [illegible]
الكسر ②

$$\hat{S}_2^2 = \frac{1}{n_2 - 1} \sum (x_{i_2} - \overline{x}_2)^2$$

4) Détermine $T_{théo}$

T_{th} pour α = 5%

$ddl = n_1 + n_2 - 2$

de Table de Student ($\frac{\alpha}{2}$; ddl)

On trouve $t_{théo}$

5) Décision

Si $t_{cal} \in [-T_{th} ; T_{th}]$

Ho est rejeté ; H1 est Accepte

• [redacted]. (Bilatéral).

(Proportion P)

Test d'homogénété pour comparer 2 proportion expérimentale.

$P_1 . P_2$

1) : Enoncer Ho et H_1:

H_0 : $P_1 = P_2$

H_1 : $P_1 <> P_2$.

2) : Fixer α (1% . 5%, 10%)

Souvent 5%.

3). Statistique de test:

$$Z_{cal} = \frac{\frac{K_1}{n_2} - \frac{K_2}{n_2}}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$$

• Variance (Bilateral)

Test d'homogénité pour comparer 2 variances $F_1 ; F_2$

1) = Enoncer H_0, H_1

H_0 : $F_1 = F_2$

H_1 : $F_1 <> F_2$

2) : Fixer α (1% - 5% - 10%)

3) - Statistique de test :

$$F_{cal} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{\frac{n_1}{n_1-1} \cdot S_1^2}{\frac{n_2}{n_2-1} \cdot S_2^2}$$

avec : $S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

4) déterminer $F_{théo}$

colonne ... $ddl_1 = n_1 - 1$

ligne ... $ddl_2 = n_2 - 1$

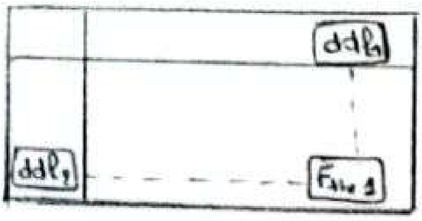
avec α (souvent 5%) :

NB : c'est α qui détermine la table de Fishe
à utiliser.

- Détermin l'intervalle :

$F_{théo\,1}$: $(\frac{\alpha}{2} ; ddl_1 ; ddl_2)$ ⇒ lire de table de Fisher ($F_{théo\,1}$)

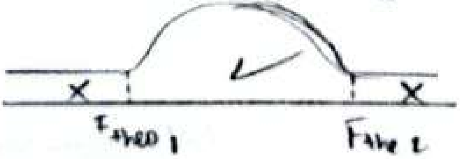
Table de Fischer selon $(\frac{\alpha}{2})$ (bilatéral)



$F_{théo\,2}$ $(1 - \frac{\alpha}{2} ; ddl_1 ; ddl_2)$ ⇒ $F_{théo\,2} = \frac{1}{F_{thé\,1}}$

5) Décision

- Ho est accepté : si $F_{cal} \in [F_{thé\,1} ; F_{thé\,2}]$ ⇒ Il y a aucune variation entre les 2 échantillons



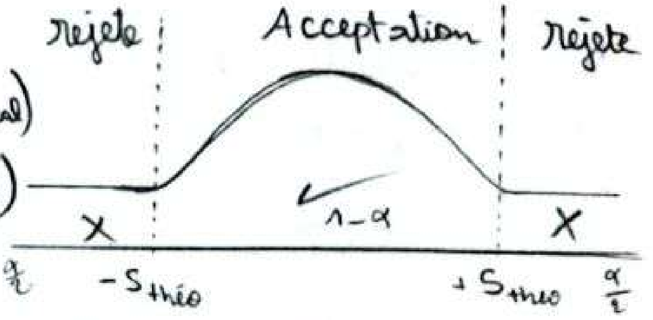
- Ho est rejeté ; H_1 est Accepté si : $F_{cal} \notin [F_{théo\,1} ; F_{théo\,2}]$.

La différence entre les types de Test (bilatéral - unilatéral)

Test bilatéral:

Test evalue si la valeur observée (S_{cal}) est différente à une valeur théorique (seuil) $S_{théo}$

$S_{cal} <> S_{théo}$:



H_0 ✓ si $S_{cal} \in [-S_{théo} ; +S_{théo}]$

Test unilatéral: (À droite - à gauche).

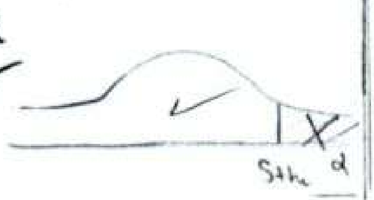
Test qui évalue si la valeur observée est supérieure ou inférieure (S_{cal}) à une valeur théorique (seuil) ($S_{théo}$)

Test à droite: $S_{cal} < S_{théo}$

- Ne dépasse pas.
- Moins de
- Au plus
- Inférieur à
- Au maximum

$S_{cal} < S_{théo} \Rightarrow H_0$ accepté ✓

$S_{cal} > S_{théo} \Rightarrow H_0$ X
H_1 ✓

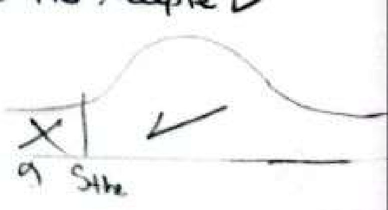


Test à gauche $S_{cal} > S_{théo}$

- Passe
- Plus de
- Au moins
- Supérieur à
- Au minimum.

$S_{cal} > S_{théo} \Rightarrow H_0$ Accepté ✓

$S_{cal} < S_{théo} \Rightarrow H_0$ X
H_1 ✓



unilatéral

moyenne / Proportion { $t_{théo}$ ~~dép~~ α
$Z_{théo}$ ~~d~~ $1-\alpha$

Variance ⇒ $F_{théo}$ dépend α

Bilatéral

moyenne / Proportion { $t_{théo}$ dépend $\frac{\alpha}{2}$
ou
$Z_{théo}$ dépend $1-\frac{\alpha}{2}$

Variance ⇒ $F_{théo}$ dépend. $\frac{\alpha}{2}$.

Echantillons dépendants (appariés ou couples).

- Les mesures sont liées ou comparées entre elles.

- On retrouve cette situation dans des études où chaque sujet est mesuré avant et après une intervention.

Ex. Le niveau de calcium dans le sang (= calcemie) avant et après un traitement chez un echantillon.

- Egalement utilisé pour des sujets appariés : comme : des jumeaux ou des personnes testées dans des conditions différentes.

Comparer la moyenne. $t_{cal} = \dfrac{\Sigma y \sqrt{N-1}}{\sqrt{N.\Sigma y^2 - (\Sigma y)^2}}$

- y : diff. absolue entre valeur avant et après.
- N : taille de l'échan

$t_{théo} = ddl = N - 1$

Echantillions indépendants :

- Les mesures des groupes sont Totalement distincts et ne pas une les unes des autres.

- comparaison entre 2 groupes sans lien :

Ex: niveau de calcium dans le sang (= calcemie) chez deux populations differentes (un groupe sportifs et un groupe sédentaire).

- Chaque individu ou mesur est autonome, sans correspondance directe directe avec l'autre.

les mêmes lois mentionnées au debut d

Résumé