

Cours 15

Mme medouer nawel

Tests d'hypothèses

Tests d'hypothèses

L'objectif de ce cours est de répondre à la problématique suivante:

On est souvent conduit à prendre une décision au sujet d'une population à partir des informations données sur un échantillon. Les tests d'hypothèses consistent à choisir entre deux hypothèses *H0 et H1*. Il y a deux types de tests : test de conformité et test d'homogénéité.

Test de conformité

a) Comparaison de la moyenne

de l'échantillon m à une moyenne vraie de la population μ . Données : n taille de l'échantillon, σ l'écart-type de la population et S l'écart-type de l'échantillon.

b) Comparaison d'une proportion calculée \tilde{P} (de l'échantillon) et une proportion vraie P (de la population).

Procédure générale d'un test

1) Détermination des hypothèses à tester H_0 et H_1 .

Remarque: les deux hypothèses doivent être contradictaires

Test bilatéral	Test Unilatéral droit	Test Unilatéral gauche
Hypothèse nulle égale Contre Hypothèse alternative inégal	Hypothèse nulle égale Contre Hypothèse alternative supérieure	Hypothèse nulle égale Contre Hypothèse alternative inférieure

Procédure générale d'un test

2) Calcul de la statistique de test observée T_0 .

3) Identification du seuil critique

- Choix du seuil de signification α (généralement donné dans l'énoncé) (sinon par convention en prend $\alpha = 0.05$)
- La lecture de la valeur du seuil critique (soit à partir de la table Normale ou la table de student).

Procédure générale d'un test

4) Décision :

- *Conclusion statistique conservation ou rejet de l'hypothèse H_0 .*
- *Conclusion pratique (répondre à la question donnée).*

Notations... Test de conformité

- ✓ n taille de l'échantillon
- ✓ \bar{m} moyenne de l'échantillon
- ✓ S écart- type de l'échantillon(corrigé)
- ✓ μ moyenne de la population
- ✓ σ écart- type de la population(*connu* ou inconnu)
- ✓ P proportion théorique
- ✓ \tilde{P} *proportion observée de l'échantillon*

Formulation des hypothèses

	<i>Test bilatéral</i>	<i>Test Unilatéral droit</i>	<i>Test Unilatéral gauche</i>
Test de conformité de la moyenne	$H_0 : m = \mu$ $H_1 : m \neq \mu$	$H_0 : m = \mu$ $H_1 : m > \mu$	$H_0 : m = \mu$ $H_1 : m < \mu$
Test de conformité de la proportion	$H_0 : \tilde{P} = P$ $H_1 : \tilde{P} \neq P$	$H_0 : \tilde{P} = P$ $H_1 : \tilde{P} > P$	$H_0 : \tilde{P} = P$ $H_1 : \tilde{P} < P$

Moyennes et fréquences

	Cas 1	Cas 2	Cas 3
Conditions	- $n \geq 30$ - σ connu	- $n \geq 30$ - σ inconnu	- $n < 30$ - σ connu -loi normale
La statistique de test T_0	$T_0 = \frac{m - \mu}{\frac{\sigma}{\sqrt{n}}}$	$T_0 = \frac{m - \mu}{\frac{S}{\sqrt{n}}}$	$T_0 = \frac{m - \mu}{\frac{\sigma}{\sqrt{n}}}$
Le seuil critique	test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$ Test unilatéral à droite $\bar{Z}_{2\alpha} = Z_{1-\alpha}$ Test unilatéral à gauche $-\bar{Z}_{2\alpha} = -Z_{1-\alpha}$	test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$ Test unilatéral à droite $\bar{Z}_{2\alpha} = Z_{1-\alpha}$ Test unilatéral à gauche $-\bar{Z}_{2\alpha} = -Z_{1-\alpha}$	test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$ Test unilatéral à droite $\bar{Z}_{2\alpha} = Z_{1-\alpha}$ Test unilatéral à gauche $-\bar{Z}_{2\alpha} = -Z_{1-\alpha}$

Moyennes et fréquences

	Cas 4	Cas 5
Conditions	<ul style="list-style-type: none"> - $n < 30$ - σ inconnu - loi normale 	<ul style="list-style-type: none"> - $n \geq 30$ -- $nP > 5$ - $n(1 - P) > 5$
La statistique de test T_0	$T_0 = \frac{m - \mu}{\frac{S}{\sqrt{n}}}$	$T_0 = \frac{\tilde{P} - P}{\sqrt{\frac{P(1-P)}{n}}}$
Le seuil critique	<ul style="list-style-type: none"> test bilatéral t_α Test unilatéral à droite $t_{2\alpha}$ Test unilatéral à gauche $-t_{2\alpha}$ 	<ul style="list-style-type: none"> test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$ Test unilatéral à droite $\bar{Z}_{2\alpha} = Z_{1-\alpha}$ Test unilatéral à gauche $-\bar{Z}_{2\alpha} = -Z_{1-\alpha}$

Test d'homogénéité

Notations...

- ✓ n_1, n_2 tailles des échantillons
- ✓ \bar{X} et \bar{Y} moyennes des échantillons
- ✓ S_X, S_Y écarts-types des échantillons
- ✓ μ_1 et μ_2 moyennes des populations
- ✓ σ_X, σ_Y écarts-types des populations(*connus* ou inconnus)
- ✓ S_C^2 **variance commune**

telle que
$$S_C^2 = \frac{(n_1-1)S_X^2 + (n_2-1)S_Y^2}{n_1+n_2-2}$$

✓

Test d'homogénéité

Notations...

- ✓ \tilde{P}_1 et \tilde{P}_2 proportions observées dans les échantillons
- ✓ P_1 et P_2 proportions théoriques
- ✓ \tilde{P} proportion commune *telle que* $\tilde{P} = \frac{n_1\tilde{P}_1 + n_2\tilde{P}_2}{n_1 + n_2}$

Formulation des hypothèses

	<i>Test bilatéral</i>	<i>Test Unilatéral droit</i>	<i>Test Unilatéral gauche</i>
Test d'homogénéité de la moyenne	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$
Test d'homogénéité de la proportion	$H_0 : P_1 = P_2$ $H_1 : P_1 \neq P_2$	$H_0 : P_1 = P_2$ $H_1 : P_1 > P_2$	$H_0 : P_1 = P_2$ $H_1 : P_1 < P_2$

Moyennes et fréquences

	Cas 1	Cas 2	Cas 3
Conditions	$-n_1 \geq 30, n_2 \geq 30$ σ_X et σ_Y connus	$-n_1 \geq 30, n_2 \geq 30$ σ_X et σ_Y inconnus	$-n_1 < 30$ et (ou) $n_2 < 30$ σ_X et σ_Y connus -loi normale
La statistique de test T_0	$T_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}\right)}}$	$T_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}\right)}}$	$T_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}\right)}}$
Le seuil critique	test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$ Test unilatéral à droite $\bar{Z}_{2\alpha} = Z_{1-\alpha}$ Test unilatéral à gauche $-\bar{Z}_{2\alpha} = -Z_{1-\alpha}$	test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$ Test unilatéral à droite $\bar{Z}_{2\alpha} = Z_{1-\alpha}$ Test unilatéral à gauche $-\bar{Z}_{2\alpha} = -Z_{1-\alpha}$	test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$ Test unilatéral à droite $\bar{Z}_{2\alpha} = Z_{1-\alpha}$ Test unilatéral à gauche $-\bar{Z}_{2\alpha} = -Z_{1-\alpha}$

Moyennes et fréquences

	Cas 4	Cas 5
Conditions	<p>$-n_1 < 30$ et (ou) $n_2 < 30$</p> <p>σ_X et σ_Y inconnus mais <u>égaux</u></p> <p>-loi normale</p>	<p>$-n_1 \geq 30$ $-n_2 \geq 30$</p> <p>-- $n_1 \tilde{P}_1 > 5$</p> <p>- $n_1 (1 - \tilde{P}_1) > 5$</p> <p>-- $n_2 \tilde{P}_2 > 5$</p> <p>- $n_2 (1 - \tilde{P}_2) > 5$</p>
La statistique de test T_0	$T_0 = \frac{\bar{X} - \bar{Y}}{S_C \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ <p>avec $S_C^2 = \frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{n_1 + n_2 - 2}$</p> <p>Appelée variance commune</p>	$T_0 = \frac{\tilde{P}_1 - \tilde{P}_2}{\sqrt{\tilde{P}(1 - \tilde{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ <p>avec $\tilde{P} = \frac{n_1 \tilde{P}_1 + n_2 \tilde{P}_2}{n_1 + n_2}$</p> <p>Appelée Proportion commune</p>
Le seuil critique	<p>test bilatéral t_α</p> <p>Test unilatéral à droite $t_{2\alpha}$</p> <p>Test unilatéral à gauche $-t_{2\alpha}$</p>	<p>test bilatéral $\bar{Z}_\alpha = Z_{1-\frac{\alpha}{2}}$</p> <p>Test unilatéral à droite</p> <p>$\bar{Z}_{2\alpha} = Z_{1-\alpha}$</p> <p>Test unilatéral à gauche</p> <p>$-\bar{Z}_\alpha = -Z_{1-\alpha}$</p>

Remarques

1. Pour identifier le seuil critique (test de conformité ou test d'homogénéité), on utilise la table normale pour les cas 1,2,3 et 5, par contre on utilise la table de Student pour le cas 4.
2. Pour un test bilatéral : le seuil critique \bar{Z}_α se lit dans la table 2 (table de la variable normale réduite) ou bien $Z_{1-\frac{\alpha}{2}}$ dans la table 1 (la table 1 de la fonction de répartition de la loi normale).

Remarques

3. Pour un test unilatéral à gauche ou à droite, le seuil critique $\bar{Z}_{2\alpha}$ se lit dans la table 2 (table de la variable normale réduite) ou bien $Z_{1-\alpha}$ dans la table 1 (la table 1 de la fonction de répartition de la loi normale).

4. L'écart-type corrigé $S = \sqrt{\frac{\sum(x_i - m)^2}{n-1}}$ de l'échantillon est donné dans l'énoncé sous le nom de « écart-type de l'échantillon ». Si l'écart-type n'est pas donné, vous devez le calculer

Remarques

- a) *Cas1* : des données brutes ou des données regroupées dans un tableau statistique On utilise la calculatrice, après la saisie des données vous choisissez $\chi\sigma n-1$.

Remarques

a) Cas2 : dans l'énoncé vous avez les deux sommes

$\sum(x_i)$ et $\sum(x_i)^2$, alors :

- calculer l'écart-type empirique de l'échantillon

$$\tilde{S} = \sqrt{\frac{\sum(x_i)^2}{n} - m^2}, \text{ telle que } m = \frac{\sum(x_i)}{n}$$

Pour identifier l'écart-type de l'échantillon S , on utilise la

$$\text{relation : } S = \sqrt{\frac{n}{n-1}} \tilde{S}$$

A retenir

Les Décisions selon le type du test (cas 1, 2,3 et 5)

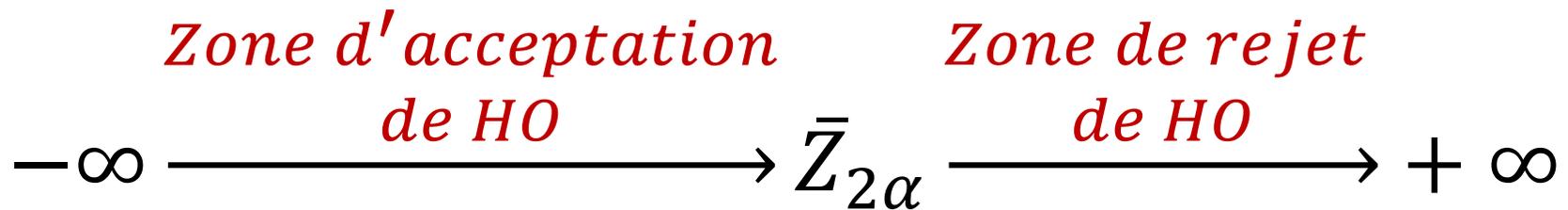
Test bilatéral : (égale contre inégale)



A retenir

Les Décisions selon le type du test (cas 1, 2,3 et 5)

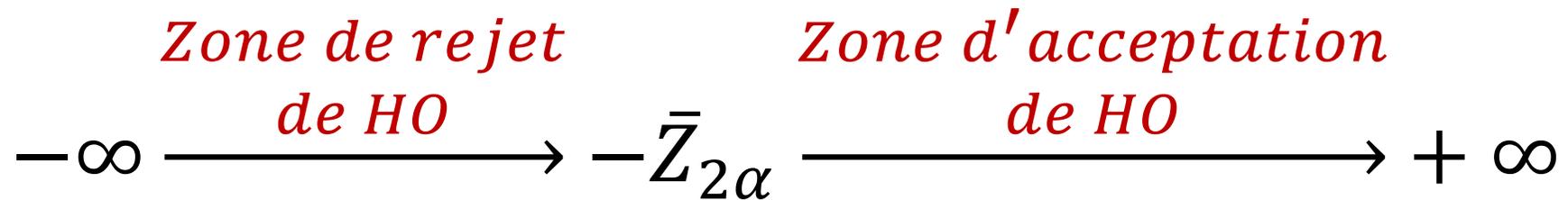
Test unilatéral à droite : (égale contre supérieure)



A retenir

Les Décisions selon le type du test (cas 1, 2,3 et 5)

Test unilatéral à gauche : (égale contre inférieure)



A retenir

La Décision selon le type du test (cas 4)

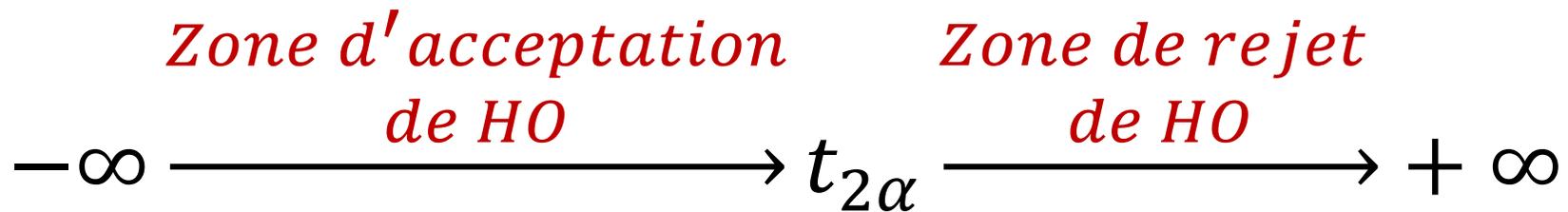
Test bilatéral : (égale contre inégale)



A retenir

La Décision selon le type du test (cas 4)

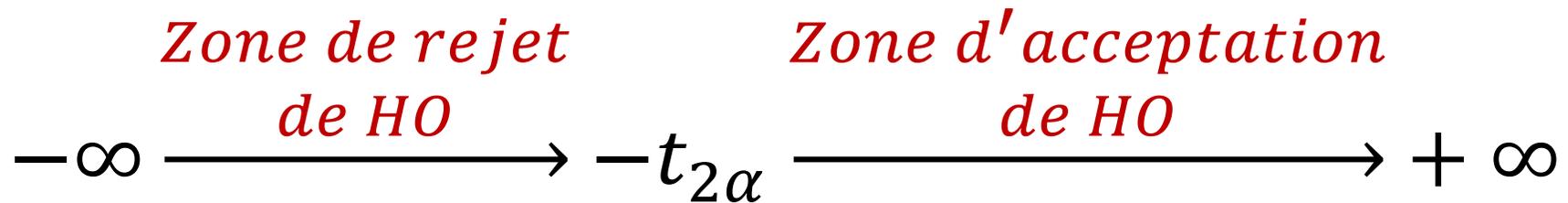
Test unilatéral à droite : (égale contre supérieure)



A retenir

Les Décisions selon le type du test (cas 4)

Test unilatéral à gauche : (égale contre inférieure)



Exemple 1

On étudie la dépendance à un médicament. Une région du cerveau, appelée VTA, contient des récepteurs (GABA-A) qui, exposés à l'enzyme anhydrase carbonique contrôlent le basculement d'un état de non-accoutumance à un état d'accoutumance.

Chez les sujets à risque, le dosage de l'activité de cet enzyme suit une loi normale d'espérance 10.7 et d'écart-type inconnu. Une série de dosages effectués sur une personne a donné : 12.9 ; 8.7 ; 9.0 ; 1.2 ; 2.7 ; 9.7 ; 9.1 ; 10.3

Cette personne peut-elle être considérée comme à risque au seuil de 5% ?

Solution:

$$m = \frac{\sum(x_i)}{n} = 7.95 \quad S = \sqrt{\frac{\sum(x_i - m)^2}{n-1}} = 3.949$$

Ou directement utiliser la calculatrice, après la saisie des données vous choisissez $\chi\sigma n-1$.

(test de conformité cas 4)

1) Choix des hypothèses : $H_0 : m = \mu$ $H_1 : m \neq \mu$

Il s'agit d'un **test bilatéral**(égale contre inégale)

2) Calcul de la statistique de test observée :

$$T_0 = \frac{m - \mu}{\frac{S}{\sqrt{n}}} = -1.9693$$

Alors on cherche le seuil critique dans la table 3 (de Student)

Suite:

3) Identification de seuil critique :

$$n < 30 \quad -\sigma \text{ inconnu} \quad \text{-loi normale}$$

Alors on cherche le seuil critique dans la table 3 (de Student)

$$\alpha = 0.05 \quad ddl = n - 1 = 7 \quad t_\alpha = 2.365$$

4) Décision :

$$-t_\alpha \leq T_0 \leq t_\alpha$$

$$-2.365 \leq T_0 \leq 2.365$$

4) Décision:

Décision statistique : La statistique de test observée se trouve dans la zone d'acceptation de H_0

Décision pratique: La personne est considérée comme étant en risque

Exemple 2

1. Pour mesurer le pH d'une solution, on utilise un pH- mètre qui affiche un résultat dont la loi est $N(\mu; 0,05)$, où μ la vraie valeur du pH de la solution. On a mesuré le pH d'une solution A par 12 mesures indépendantes et trouvé une moyenne de 7.4, et le pH d'une solution B par 10 mesures indépendantes et trouvé une moyenne de 7.5.

Peut-on considère que les deux solutions ont même pH (au seuil 1%) ?

Exemple 2

2. Pour mesurer le pH d'une solution, on utilise un nouveau pH-mètre qui affiche un résultat dont la loi est $N(\mu; \sigma)$, où μ est la vraie valeur du pH de la solution et où σ n'a pas été déterminé. On a mesuré le pH d'une solution A par 12 mesures indépendantes et trouvé une moyenne de 7.4 et un écart-type SA de 0.09, et le pH d'une solution B par 10 mesures indépendantes et trouvé une moyenne de 7.5 et un écart-type SB de 0.08.

Peut-on considérer que les deux solutions ont même pH (au seuil 1%) ?

Solution:

Question 1:

1. *Il s'agit d'un test bilatéral (égale contre inégale)*

1) *Choix des hypothèses :*

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_A \neq \mu_B$$

2) *Calcul de la statistique de test observée :*

$$T_0 = \frac{\bar{X} - \bar{Y}}{\sigma \sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} = -4.67$$

3) *Identification de seuil critique :*

$$n_A < 30 \text{ et } n_B < 30$$

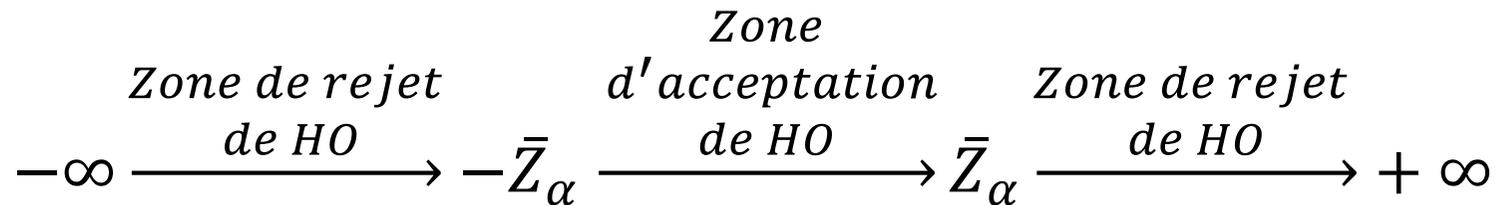
$$\sigma_X = \sigma_Y = \sigma \text{ (Connus) et Loi Normale}$$

On cherche le seuil critique dans la table 2 de la loi Normale

$$\alpha = 0.05 \quad \bar{Z}_\alpha = 2.575$$

Décision :

Test bilatéral : (égale contre inégale)



$$T_0 \notin [-\bar{Z}_\alpha ; \bar{Z}_\alpha]$$

Décision statistique : la statistique de test observée se trouve dans la zone de rejet de H_0 .

Décision pratique : On conclut que les deux solutions n'ont pas le même PH

Question 2:

1) Choix des hypothèses :

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_A \neq \mu_B$$

2) Calcul de la statistique de test observée :

$$T_0 = \frac{\bar{X} - \bar{Y}}{S_C \sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} = -0.27269 \text{ avec}$$

$$S_C^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2} = 0.7335$$

3) Identification de seuil critique :

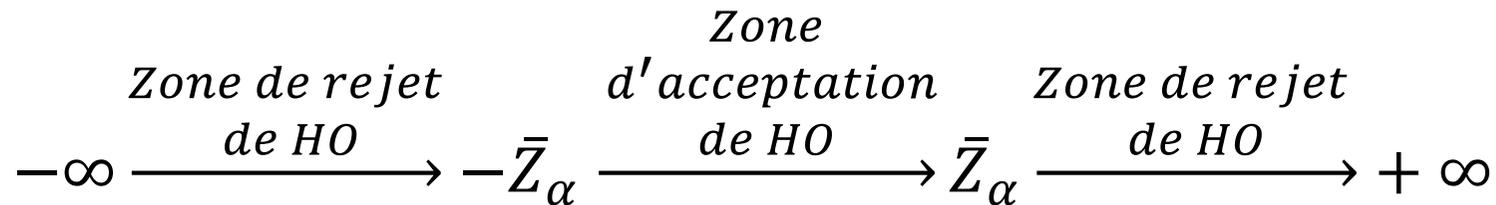
$n_A < 30$ et $n_B < 30$ σ_X, σ_Y (inconnus) mais égaux et Loi Normale

On cherche le seuil critique dans la table 3 de Student

$$\alpha = 0.01 \quad ddl = n_A + n_B - 2 = 20 \quad \bar{t}_\alpha = 2.845$$

Décision :

Test bilatéral : (égale contre inégale)



$$T_0 \in [-t_\alpha ; t_\alpha]$$

$$-2.845 < -0.272 < 2.845$$

Décision statistique : la statistique de test se trouve dans la zone d'acceptation de H_0

Décision pratique : On conclut que les deux solutions sont de mêmes PH

Exemple 3

Un laboratoire pharmaceutique garantit que son médicament destiné au traitement d'une affection cutanée est efficace dans 80% des cas. Un groupe indépendant a collecté les données de cabinets médicaux d'une région de 40 patients atteints par l'affection ont été traités par le médicament. Parmi eux, 28 ont été définitivement guéris. Pour les 12 autres patients le médicament a été sans effet, ils ont de changer de traitement .

Ces résultats sont-ils conformes à l'affirmation du laboratoire pharmaceutique% ? ($\alpha = 0.05$).

Solution:

1) Choix des hypothèses : **(test de conformité cas 5)**

$H_0: \tilde{P} = P$ l'affirmation du laboratoire est vraie

$H_1: \tilde{P} < P$ l'affirmation du laboratoire est fausse

Il s'agit d'un **test unilatéral à gauche** (égale contre inférieure)

2) Calcul de la statistique de test observée :

$$T_0 = \frac{\tilde{p} - P}{\sqrt{\frac{P(1-P)}{n}}} = 1.5811$$

3) Identification de seuil critique :

le test est valable car les trois conditions sont bien remplies :

$n > 30$ **$nP = 28 > 5$** **$n(1 - P) = 8 > 5$** Alors on cherche le seuil critique dans la table 2 (de la variable normale réduite)

Solution:

$$\alpha = 0.05 \quad \bar{Z}_{2\alpha} = 1.645$$

Décision :

Test unilatéral à gauche : (égale contre inférieure)

$$-\infty \xrightarrow[\text{de } H_0]{\text{Zone de rejet}} -\bar{Z}_{2\alpha} = -1.645 \xrightarrow[\text{de } H_0]{\text{Zone d'acceptation}} +\infty$$

$$-1.645 < -1.5811 = T_0$$

Décision statistique : la statistique de test observée se trouve dans la zone d'acceptation de H_0 .

Décision pratique : l'affirmation du laboratoire est vraie.

Des échantillons appariés

Des échantillons appariés sont des échantillons construits de façon à ce qu'ils soient composés d'individus possédant les mêmes caractéristiques. C'est le cas par exemple lorsque l'on mesure le même caractère sur les mêmes individus à deux moments différents (avant et après)

Exemple 4

Neuf malades présentant des symptômes d'anxiété reçoivent un traitement. On évalue l'état des malades avant et après traitement par un indice que le médecin traitant calcule d'après les réponses à une série de questions. Si le traitement est efficace, l'indice doit diminuer. Les valeurs de cet indice sur les neuf patients sont les suivants :

Patient	1	2	3	4	5	6	7	8	9
Avant	1.83	0.5	1.62	2.48	1.68	1.88	1.55	3.06	1.3
Après	0.88	0.65	0.59	2.05	1.06	1.29	1.06	3.14	1.29

Le traitement est-il efficace au seuil de 0.05 ?

Solution:

Patient	1	2	3	4	5	6	7	8	9
Différence	0.95	-0.15	1.03	0.43	0.62	0.59	0.49	-0.08	0.01

Différence= avant-après

$$\bar{D} = \frac{\sum D_i}{n} = 0.43222 \quad S_D = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n-1}} = 0.42845$$

Remarque : directement utiliser la touche $\chi\sigma n-1$ dans la calculatrice pour identifier S_D

1) Choix des hypothèses : **(test de conformité cas 5)**

Il s'agit d'un **test unilatéral à droite** (égale contre supérieure)

Parce que on discute la diminution de l'indice (avant- après>0)

$$H_0: \mu = 0$$

$$H_1: \mu > 0$$

Solution:

2) Calcul de la statistique de test observée :

$$T_0 = \frac{\bar{D}}{\frac{s_D}{\sqrt{n}}} = 3.0239$$

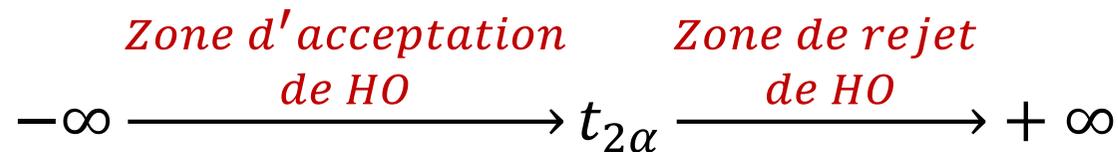
3) Identification de seuil critique :

$$\alpha = 0.05 \quad ddl = n - 1 = 8 \quad t_{2\alpha} = 1.86$$

4) Décision :

$$T_0 > t_{2\alpha} \quad 3.0239 > -1.86$$

Test unilatéral à droite : (égale contre supérieure)



Décision statistique : La statistique de test se trouve dans la zone de rejet de H_0

Décision pratique : Le traitement est efficace