TEST DU SIGNE

INRA Antibes

Âge élevé à Antibes ?

Tirage 13-échantillon à Antibes :

$\Longrightarrow$ 48 41 33 60 58 27 48 40 43 32 57 51 45

Réordonné :

$\Longrightarrow$ 27 32 33 40 43 44 45 48 48 51 57 58 60

I.C. de la médiane m_A ?

I.C. DE LA MÉDIANE m_A

Fixer $\gamma = 0{,}90$ (niveau de confiance)

TABLE 2.3 $\Rightarrow c = 4$

27 32 33 [40 43 44 45 48 48 51] 57 58 60

I.C. $= [40 \,;\, 51]$

$$\Pr\{\left[X_{(4)}, X^{(4)}\right] \ni m_A\} = 0{,}908$$

$m_0 = 41{,}2$

Conclusion ?

On a décidé $m_A = m_0$: **DÉCISION CORRECTE ?**

Décision :		Réalité :
		$m_A = m_0$
$m_A = m_0$	$\nearrow$ $\searrow$	
		$m_A \neq m_0 \;\Rightarrow$ erreur de décision
		$m_A \neq m_0$
$m_A \neq m_0$	$\nearrow$ $\searrow$	
		$m_A = m_0 \;\Rightarrow$ erreur de décision

Décision		Réalité	(décision)	Erreur
$m_A = m_0$		$m_A = m_0$	correcte	
$m_A = m_0$		$m_A \neq m_0$	$\boxed{\text{incorrecte}}$	2^{e} espèce
$m_A \neq m_0$		$m_A \neq m_0$	correcte	
$m_A \neq m_0$		$m_A = m_0$	$\boxed{\text{incorrecte}}$	1re espèce

$$H_0 : m_A = m_0$$

$$H_1 : m_A \neq m_0$$

test statistique $\Longleftrightarrow$ règle de décision

		La vraie valeur inconnue est m_A :	
		$H_0 : m_A = m_0$	$H_1 : m_A \neq m_0$
Décision	$m_A = m_0$ (H_0)	pas d'erreur	erreur de 2^{e} espèce
	$m_A \neq m_0$ (H_1)	erreur de 1^{re} espèce	pas d'erreur

COMMENT PRÉCISER LA VALEUR D'UN TEST ?

$\Longrightarrow$ Étudier les risques d'erreur :

Pr{erreur de 1$^{\text{re}}$ espèce} = **risque de 1$^{\text{re}}$ espèce** = α

Pr{erreur de 2$^{\text{e}}$ espèce} = **risque de 2$^{\text{e}}$ espèce** = β

		État de la nature : $H_0 : m_A = m_0$	**État de la nature :** $H_1 : m_A \neq m_0$
Décision	H_0 : $m_A = m_0$ $[X_{(4)}\ ;\ X^{(4)}] \ni m_0$	$\gamma = 0{,}908$	? β ? = $\mathrm{Pr}_{m_A}\{[X_{(4)}\ ;\ X^{(4)}] \ni m_0\}$ risque de 2$^{\text{e}}$ espèce
	H_1 : $m_A \neq m_0$ $[X_{(4)}\ ;\ X^{(4)}] \not\ni m_0$	$\alpha = 1 - \gamma = 0{,}092$ risque de 1$^{\text{re}}$ espèce	? $1-\beta$?

État de la nature : *inconnu*, mais : .... Décision : *connue*.

TEST BILATÉRAL, TEST UNILATÉRAL

$H_0 : m_A = m_0$

$$H_1 : m_A \neq m_0 \iff \underbrace{m_A < m_0 \ \text{ ou } \ m_A > m_0}_{\text{alternative bilatérale}}$$

Test bilatéral.

On suppose que si $m_A \neq m_0$, c'est $m_A > m_0$.

On exclut la possibilité $m_A < m_0$,

avant le recueil des données.

$H_0 : m_A = m_0$

$H_1 : m_A > m_0 \Longrightarrow$ alternative unilatérale

Test unilatéral.

TEST UNILATÉRAL

$H_0 : m_A = m_0$

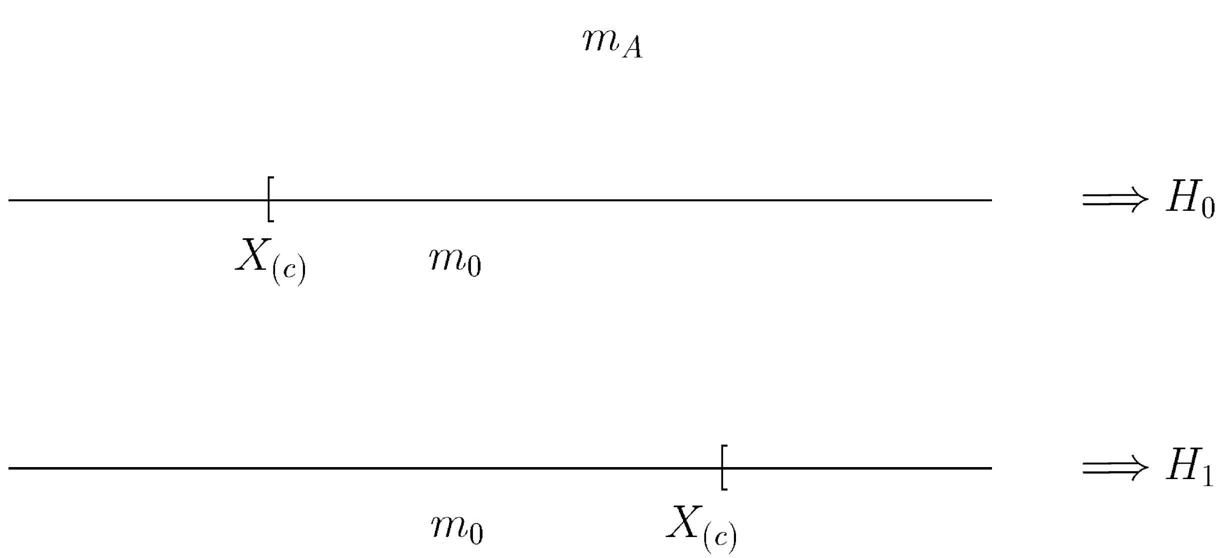
$H_1 : m_A > m_0 \Longrightarrow$ alternative unilatérale.

Règle de décision :

si m_0 en dessous de l'I.C. de $m_A \Longrightarrow$ rejet.

$$\Longrightarrow \text{I.C.} = [X_{(c)}\ ;\ +\infty[$$

I.C. unilatéral.



TEST DU SIGNE : MÉTHODE DIRECTE

$H_0 : m_A = m_0$

$H_1 : m_A > m_0$

Y_0 = nombre de valeurs $< m_0$ (statistique).

Loi de Y_0 sous H_0 | Loi de Y_0 sous H_1

H_0 ou H_1 ? On a **une** valeur de Y_0 :

$$Y_0 \text{ petit} \Longrightarrow H_1 \text{ plus probable}$$

CONDITIONS D'UTILISATION :

Test du signe : 1 échantillon (indépendance)

Test de Wilcoxon : 1 échantillon
distribution symétrique

Test de Student : 1 échantillon
distribution normale
($\Longrightarrow$ symétrique)

TEST DE STUDENT

$\mu_0 = 42$: moyenne INRA.

μ_A (inconnue) = moyenne d'Antibes.

$H_0 : \mu_A = \mu_0$

$H_1 : \mu_A > \mu_0$

$$\text{Statistique : } T_0 = \frac{\overline{X} - \mu_0}{\widehat{\sigma}/\sqrt{n}} = \frac{\overline{X} - 42}{\widehat{\sigma}/\sqrt{n}}$$

Sous $H_0, T_0 \sim t_{n-1}$.

Sous $H_1, T_0 \sim$ loi décalée à droite.

Règle de décision : si $T_0 \leq t_c \Longrightarrow H_0$.

TEST DE STUDENT BILATÉRAL

Test unilatéral :

$H_0 : \mu_A = \mu_0$
$H_1 : \mu_A > \mu_0$

$$T_0 = \frac{\overline{X} - \mu_0}{\widehat{\sigma}/\sqrt{n}} = \frac{\overline{X} - 42}{\widehat{\sigma}/\sqrt{n}}$$

Rejet si T_0 trop grand.

Réalité :

Décision :	H_0	H_1
H_0	$\gamma = 1 - \alpha$	β
H_1	α	$1 - \beta$

loi de T_0
sous H_0

$$\alpha' = 5\,\%$$

I.C. unilatéral   région de rejet

Test bilatéral :

$H_0 : \mu_A = \mu_0$

$H_1 : \mu_A \neq \mu_0$

$$T_0 = \frac{\overline{X} - \mu_0}{\widehat{\sigma}/\sqrt{n}} = \frac{\overline{X} - 42}{\widehat{\sigma}/\sqrt{n}}$$

Rejet si T_0 trop petit ou trop grand.

loi de T_0
sous H_0

$\alpha/2$ $\alpha = 5\,\%$ $\alpha/2$

I.C. bilatéral

TEST DE WILCOXON

Rappel :

$H_0 : m_A = m_0$

$H_1 : m_A \neq m_0$

$$\boxed{\text{Statistique } Y_Z^0 = \left(\text{nombre de } \frac{X_i+X_j}{2} \; < \; m_0\right)}$$

Si $[c\ ;\ n-c+1] \ni Y_Z^0$, H_0 acceptée,

si $[c\ ;\ n-c+1] \not\ni Y_Z^0$, H_0 rejetée.

TEST DE WICOXON DIRECT :
SOMME DES RANGS

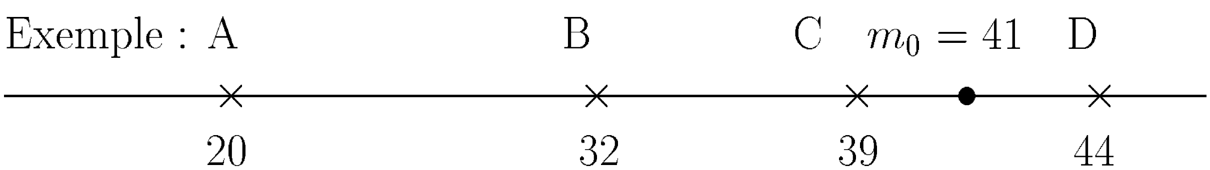
Autre façon de calculer Y_Z^0 :

$\rightarrow X_i - m_0$,

$\rightarrow$ rang de $|X_i - m_0|$,

$\rightarrow$ somme des rangs des différences négatives.

Exemple :



	A	B	C	D
X_i	20	32	39	44
$X_i - m_0$	-21	-9	-2	3
rang de $\lvert X_i - m_0 \rvert$	4	3	1	2
signe de $(X_i - m_0)$	$-$	$-$	$-$	$+$

$$Y_Z^0 = 4 + 3 + 1 = 8 = \left(\text{nombre de } \frac{X_i + X_j}{2} < m_0\right)$$

NIVEAU DESCRIPTIF ou P-VARIABLE

Problème :

La moyenne d'âge des agents d'Antibes est-elle supérieure ou égale à celle de l'ensemble de la population INRA ?

Identifier :
- la population,
- la variable,
- le paramètre.

Préciser :
- la distribution *a priori*,
- les hypothèses H_0 et H_1.

Choix de la statistique

Choix du niveau de confiance

Domaine de rejet $\longleftrightarrow$ Valeur critique

APPLICATION

Soit un 13-échantillon de la population des agents d'Antibes.

Extrait de la table de Student (12 ddl).

Probabilité de l'extrémité de la loi		Valeur critique
0,5		0,0
0,25		0,69
0,10		1,36
0,05		1,78
$\vert$		$\vert$
α'	$\dashrightarrow$	$t_{\alpha'}$
$\vert$		$\vert$
P'_{T_0}	$\dashleftarrow$	T_0

Si $T_0 > t_{\alpha'}$, H_0 est rejetée.

$$\Updownarrow$$

Si $P'_{T_0} < \alpha'$, H_0 est rejetée.

NIVEAU DESCRIPTIF ou P-VARIABLE

$$P'_{T_0} = 1 - F_0(T_0)$$

$$\boxed{\begin{array}{ll} \text{Règle :} & \text{si } P'_{T_0} < \alpha', \text{ on rejette } H_0, \\ & \text{si } P'_{T_0} > \alpha', \text{ on ne rejette pas } H_0. \end{array}}$$

P'_{T_0} est une statistique descriptive

$\downarrow$

qui **ne doit pas** être utilisée pour

choisir le risque d'erreur *a posteriori* !

NIVEAU DESCRIPTIF ou P-VARIABLE

Problème :

L'âge moyen des agents de Montpellier est-il différent de celui de l'ensemble des agents ?

$$H_0 : m_M = m_0 \quad \text{contre} \quad H_1 : m_M \neq m_0$$

$$\text{Si } T_O < 0, \qquad P_{T_0} = F_0(T_0) + [1 - F_0(-T_0)]$$

$$\text{Si } T_O > 0, \qquad P_{T_0} = [1 - F_0(T_0)] + F_0(-T_0)$$

$$\forall\ T_0, \quad P_{T_0} = F_0(-|T_0|) + [1 - F_0(|T_0|)]$$

BILAN SUR LES TESTS

* Identifier :

- la population,
- la variable d'étude,
- le paramètre de la loi à étudier.

* Préciser :

- la distribution *a priori* de la variable,
- les hypothèses H_0 et H_1.

* Choisir une **statistique** S :

de **loi connue sous** H_0	dont la forme de la loi est souvent connue sous H_1

BILAN SUR LES TESTS (suite et fin)

* Choisir un **niveau de confiance** γ.

$\Downarrow$

* Définir le **domaine de rejet** de H_0.

$\Downarrow$

* Calculer la **valeur critique** s_α,
pour un échantillon de taille n.

* Obtenir un n-**échantillon** de mesures.

* Calculer la statistique S sur les mesures de cet échantillon.

si $S > s_\alpha$	si $S < s_\alpha$
$\downarrow$	$\downarrow$
rejet de H_0	**non rejet** de H_0