Feuille d'EXERCICES III : TESTS D'HYPOTHESES (feuille 2)

Exercice 1 *Des études comparatives sur la fécondité des femmes au sein de la communauté européenne ont été menées sur des femmes de 40 ans. Pour cela, un statisticien a réalisé deux échantillons: l'un (de 100 femmes) en France, l'autre (de 80 femmes), en Belgique. Il a obtenu une moyenne de 1,78 enfants / femme et une somme de carrés d'écarts (SCE) de 427,68 (enfants / femme)2 pour l'échantillon français et une moyenne de 2,12 enfants / femme et une SCE 281,24 (enfants / femme)2 de pour l'échantillon belge. Les femmes françaises sont-elles moins fécondes que leurs homologues belges?*

$n_1 = 100$
$n_2 = 80$
$\mu_1 = 1{,}78$
$\mu_2 =$

Exercice 2 *Pour deux catégories différentes de raisins (catégorie 1 et catégorie 2), on a observé l'acidité (ph) de 7 et 11 grappes respectivement. On remarque que l'échantillon de la catégorie 1 a une acidité moyenne de 3.556 (variance 0.011) et celle de l'échantillon de la catégorie 2 est de 3.477 (variance 0.007). Tester si la différence est significative.*

Exercice 3 *Soit deux populations de chevaux : les bons sauteurs et les mauvais sauteurs. On étudie la hauteur du garrot que l'on suppose distribuée normalement dans les deux populations. Pour cela, on prélève un échantillon dans chacune des deux populations, ce qui donne les résultats suivants :*
- Bons sauteurs : $n_1 = 55$, $\bar{X}_1 = 164$, $S_1^2 = 7.4$,
- Mauvais sauteurs : $n_2 = 50$, $\bar{X}_2 = 161.5$, $S_2^2 = 5.2$.

1. *La différence des moyennes observeées est-elle significative ? Répondre par un test de niveau* $\alpha = 5\%$
2. *Même question mais en supposant que les effectifs sont :* $n_1 = 12$ *et* $n_2 = 10$.

Exercice 4 *Le tableau ci-dessous contient les nombres* N_i *d'apparition des entiers* x_i *dans les 10000 premiers chiffres de la partie décimale du nombre* π. *Tester l'hypothèse d'une répartition uniforme de ces entiers.*

x_i	0	1	2	3	4	5	6	7	8	9
N_i	968	1025	1021	974	1014	1045	1021	970	948	1014

Exercice 5 *Le tableau ci-dessous fournit la distribution du nombre X d'accidents graves par semaine à un carrefour dangereux. On demande de tester l'hypothèse que X suit une loi de Poisson de paramètre 2.*

x_i	0	1	2	3	4	5
N_i	5	10	7	4	3	1

Exercice 6 *Peut-on accepter l'hypothèse que les données suivantes proviennent d'une loi* $N(1.5, 1)$ *?*
2.51; 1.45; 0.71; 0.25; 2.35; 0.00; 2.06; 1.95; 0.41; 2.75; 0.78; 1.01; 1.75; 2.32; 1.36.

Exercice 7 *Un échantillon aléatoire de 1367 diplômes d'université, délivrés en 1984, a donné la répartition suivante :*

	Licence	*Maîtrise*	*Doctorat*
Masculin	*534*	*144*	*22*
Féminin	*515*	*141*	*11*

Le sexe et le niveau de diplôme obtenu sont-ils indépendants ?

Exercice 8 *Un traitement est administré à trois doses différentes D1, D2, D3, à un groupe de sujets atteints d'une même maladie. L'expérimentation est faite en double aveugle. On compte le nombre de guérisons pour chaque dose. Les résultats sont les suivants :*

	Sujets guéris	*Sujets non guéris*	*Total*
Dose D1	*30*	*30*	*60*
Dose D2	*42*	*35*	*77*
Dose D3	*58*	*31*	*89*
Total	*130*	*96*	*226*

L'efficacité du traitement est-elle liée à la dose utilisée ?

- TD 2 : Statistiques

Rappel : tests au niveau α

Exemple : test sur la moyenne

1) Unilatéral

$\mu = \mu_0$ contre H_1 : $\mu > \mu_0$ ← à droite

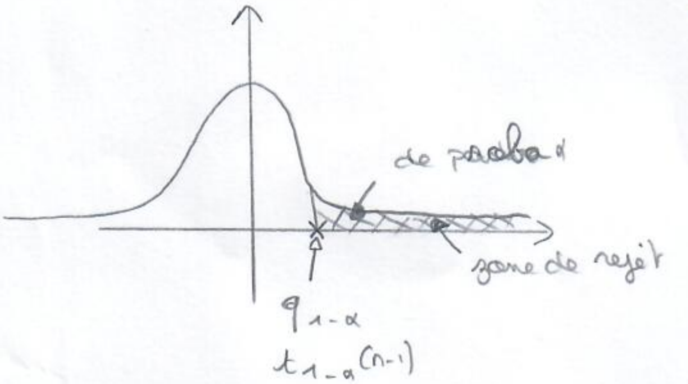
Statistique de test : si σ connu $M_\sigma = \dfrac{\overline{X_n} - \mu}{\frac{\sigma}{\sqrt{n}}} \underset{\text{sous } H_0}{\sim} \mathcal{N}(0,1)$

si σ inconnu $M_s = \dfrac{\overline{X_n} - \mu}{\frac{s}{\sqrt{n}}} \underset{\text{sous } H}{\sim} T(n-1)$

Règle de décision :

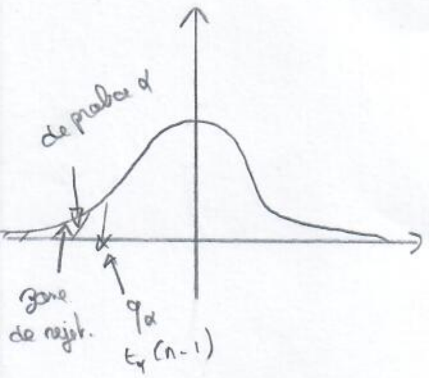
on rejette H_0 si $M_\sigma, H_0 > q_{1-\alpha}$ σ connu

si $M_s, H_0 > t_{1-\alpha}(n-1)$ σ inconnu



↙ à gauche.

H_0 : $\mu = \mu_0$ contre H_1 : $\mu < \mu_0$



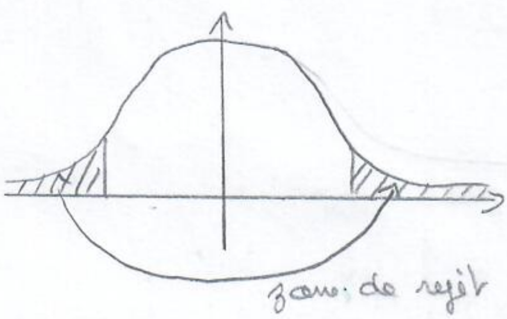
Règle de décision :

on rejette H_0 si $M_\sigma, H_0 < q_\alpha$ σ connu

si $M_s, H_0 < t_\alpha(n-1)$ σ inconnu

2) Bilatéralité

(H_0) $\mu = \mu_0$ contre $\mu \neq \mu_0$



Stat de test : M_σ et M_s

Règle de décision

on rejette (H_0) si

$M_\sigma, H_0 < q_{\frac{\alpha}{2}}$

ou $M_\sigma, H_0 > q_{1-\frac{\alpha}{2}}$

-TD 2 : Statistiques-

Règle de décision : voir rappel

<u>Exercice 1 :</u> France : $n_1 = 100$ Belgique : $n_2 = 80$

$\overline{x_n} = 1,78$ $\overline{x_n} = 2,12$

$SCE_1 = 427,68$ $SCE_2 = 281,24$

SCE : somme des carrés des écarts à la moyenne

<u>Question :</u> $\mu_1 < \mu_2$

moyenne ~~de~~ pop° des femmes françaises

Test : (H_0) $\mu_1 = \mu_2$ contre $\mu_1 < \mu_2$

↳ unilatéral à gauche

$(H_0) : \frac{\mu_1}{\mu_2} = 1$ contre $(H_1) : \frac{\mu_1}{\mu_2} < 1$.

<u>Méthode :</u>

1) Test de comparaison des variances.

$(H_0) : \sigma_1^2 = \sigma_2^2$ contre $\sigma_1^2 = \sigma_2^2$

↳ Test de Fisher

2) Test de comparaison des moyennes.

① Statistiques de test

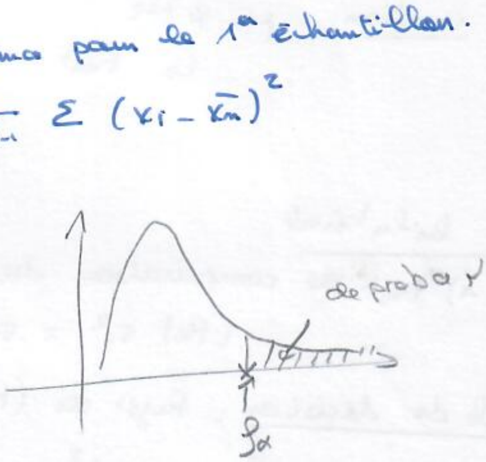
$$F = \frac{\sigma_2^2}{\sigma_1^2} \times \frac{s_1^2}{s_2^2}$$

s_1^2 : estimat° de la variance pour le 1er échantillon. $= \frac{1}{n-1} \Sigma (x_i - \overline{x_n})^2$

<u>Règle de décision</u>

On rejette (H_0)

si $F_{H_0} > f_\alpha(n_1+n_2-2)$



or $F_{H_0} = \frac{s_1^2}{s_2^2}$

$$= \frac{\frac{1}{n_1-1} SCE_1}{\frac{1}{n_2-1} SCE_2} = \frac{4,32}{3,56} \simeq 1,213$$

- $f_{0,05}(99, 78) \simeq 1,47$

Donc on accepte (H_0)

② <u>Test de comparaison</u>

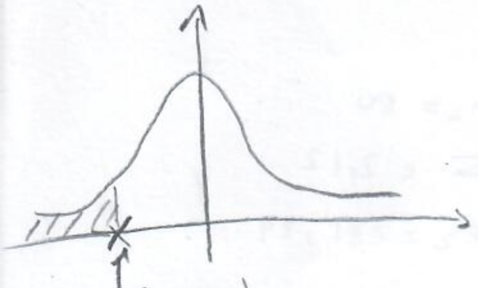
$$T = \frac{\overline{x_{n_1}} - \overline{x_{n_2}} - (\mu_1 - \mu_0)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{avec } S = \frac{(n_1-1)s_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

sous H_0, $T \sim T(n_1+n_2-2)$

$\Rightarrow T_{H_0} \sim T(n_1+n_2-2)$

Dans notre cas : test unilatéral à gauche

Règle de décision : On rejette (H_0) si $T_{H_0} < t_\alpha(n_1+n_2-2)$



$t_\alpha\,(n_1+n_2-2)$

$$T_{H_0} = \frac{\overline{X_{n_1}} - \overline{X_{n_2}}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$= \frac{\overline{x_{n_1}} - \overline{x_{n_2}}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$\overline{x_{n_1}} = 1,78$ $\quad$ $s_1^2 = 4,32$

$\overline{x_{n_2}} = 2,12$ $\quad$ $s_2^2 = 3,56$

↳ $T_{H_0} = \dots$ $\quad$ On compare T_{H_0} et $t_{0,05}(178)$ et on conclut.

$t_{0,05}(178) =$

Exercice 2 :

$n_1 = 7$ $\quad$ $n_2 = 11$

$\overline{x_{n_1}} = 3,556$ $\quad$ $\overline{x_{n_2}} = 3,477$

$s_1^2 = 0,011$ $\quad$ $s_2^2 = 0,007$

Question : $\mu_1 \neq \mu_2$

↳ test : $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$

$\frac{\mu_1}{\mu_2} = 1$ $\quad$ $\frac{\mu_1}{\mu_2} \neq 1$

Test bilatéral

1) Test de comparaison des variances

$(H_0)\ \sigma_1^2 = \sigma_2^2$ contre $(H_1)\ \sigma_1^2 \neq \sigma_2^2$

Règle de décision : Rejet de (H_0) si $F_{H_0} > f_\alpha(n_1-1, n_2-1)$

- F_{H_0} ? $= \frac{s_1^2}{s_2^2} = \frac{0,011}{0,007} \simeq 1,57$

- $f_{0,05}(6,10) \simeq 3,22$

Donc on accepte H_0

2) Test de comparaison des moyennes (bilatéral)

Règle de décision : Rejet de (H_0) si $T_{H_0} < t_{\frac{\alpha}{2}}(n_1+n_2-2)$ ou $T_{H_0} > t_{1-\frac{\alpha}{2}}(n_1+n_2-2)$

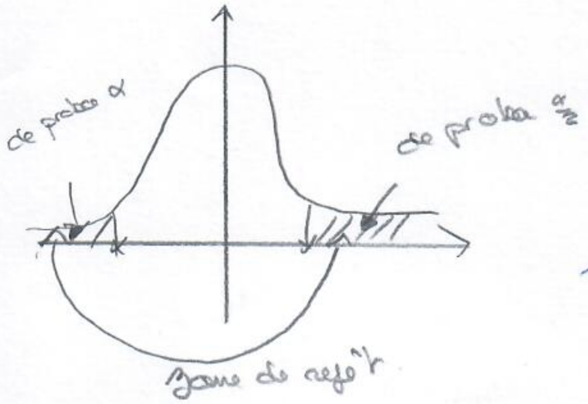
$$\frac{\overline{x_{n_1}} - \overline{x_{n_2}}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} \times \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$\overline{x_{n_1}} = 3,556$ $\quad$ $S_1^2 = 0,011$

$\overline{x_{n_2}} = 3,477$ $\quad$ $S_2^2 = 0,007$

$x_1 = 7$ $\quad$ $x_2 = 11$

$t_{1-\frac{\alpha}{2}}(n_1+n_2-2)$.



Sans reserve :

$s^2 = 0{,}085$

$\hookrightarrow s = 0{,}092$

$\hookrightarrow T_{H_0} \simeq 1{,}79$

- $t_{0{,}025}(16) = -2{,}12$

$t_{0{,}975}(16) = 2{,}12$

Observation du test : on a ni $T_{H_0} < t_{\frac{\alpha}{2}}(n_1+n_2-2)$

ni $T_{H_0} > t_{1-\frac{\alpha}{2}}(n_1+n_2-2)$

Donc on ne rejette pas (H_0).

$\hookrightarrow$ on accepte H_0.