Statistiques

– UE III : Statistiques

Statistique inférentielle

Semaine : n°2 (du 12/09/16 au 16/09/16) ***Date :*** 15/09/2016	***Heure : de*** 8h00 à 9h00	***Professeur :*** Pr. LEMDANI
Binôme : n°41		***Correcteur :*** n°42
Remarques du professeur : *Rien à signaler*		

PLAN DU COURS

III) Tests :

C) Deux échantillons :

1) Test de type « observé/observé » :

D) Séries appariées :

1) Plans en séries appariées :

2) Comparaison de proportions en séries appariées : test de MAC NEMAR :

III) Tests :

C) Deux échantillons :

1) Test de type « observé/observé » :

COMPARAISON DE MOYENNES ET DE VARIANCES OBSERVÉES (SUITE) :

Comparaison des variances :

Variable de décision :

$$F = \frac{S_A^2}{S_B^2}$$

avec S_A^2 = plus grande des variances entre S_1^2 et S_2^2

Conditions d'utilisation du test F : elles sont identiques à celles du test t c'est à dire :

- X suit une loi normale (X ~ N{0 ; 1}),
- on a deux grands échantillons ($n \geq 30$).

Sous K_0, F ~ $F_{n_B-1}^{n_A-1}$

Sous H_0, t ~ $S_{t\nu}$ avec :

- $\nu = n_1 + n_2 - 2$ si non rejet de K_0
- $\nu = \dfrac{(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2})^2}{\frac{1}{n_1-1}(\frac{S_1^2}{n_1})^2+\frac{1}{n_2-1}(\frac{S_2^2}{n_2})^2}$ si rejet de K_0

Exemple : comparaison des dosages de transaminases de 2 groupes de patients, soumis à 2 variantes d'un traitement pouvant perturber les fonctions hépatiques :

Variante 1 : 22 ; 18 ; 25 ; 14 ; 16 ; 22 ; 17 ; 19 ; 30 ; 17 ; 23 ; 17

Variante 2 : 31 ; 35 ; 32 ; 30 ; 28 ; 26 ; 27 ; 19 ; 25 ; 20 ; 18 ; 31 ; 27 ; 29 ; 24

X = « dosage de transaminases » (variable quantitative continue)

μ_1 = moyenne de X (variante 1) ; μ_2 = moyenne de X (variante 2)

$n_1 = 12$ et $n_2 = 15$

$H_0 : \{\mu_1 = \mu_2\}$ contre $H_1 : \{\mu_1 \neq \mu_2\}$

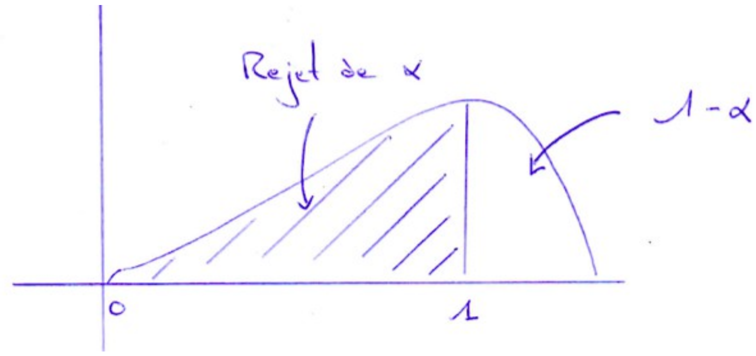
Ici on a affaire à deux échantillons → (deux) test(s) de normalité pour utiliser un test t.

H'_0 : {$X \sim N(\mu ;\sigma)$} contre H'_1 : {$X \sim/\sim N(\mu ;\sigma)$}

Calcul des variables de décisions : W1 = 0,923 et W2 = 0,956

La table de **SHAPIRO et WILK** correspond à :

- Une courbe comprise entre 0 et 1,
- W calculé est un pourcentage qui donne la proportion de normalité :
  - Si W est proche de 1 : normalité,
  - Sinon, pas de normalité.



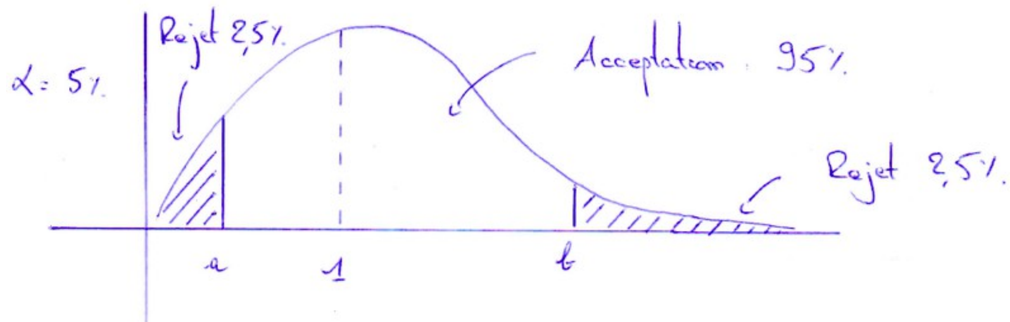
Zones de non – rejet respectives ($\alpha = 0,05$) : [0,859 ; 1] et [0,881 ; 1]

→ **Non rejet** de H'_0 au seuil de 5% pour les deux populations par conséquent on peut utiliser le test t.

<u>Comparaison des variances</u> : K_0 : {$\sigma1 = \sigma2$} contre K_1 : {$\sigma1 \neq \sigma2$}

<u>Calculs</u> : $x_1 = 20,0$ et $s_1 \approx 4,533$; $x_2 = 26,8$ et $s_2 \approx 4,945$

$$F = \frac{S_2^2}{S_1^2}$$ $\sim F^{14}_{11}$ sous K_0 → demi zone de rejet ($\alpha = 0,05$) :]3,36 ; + ∞[



<u>Calcul (suite)</u> : $F_c \approx 1,19$ → non rejet de K_0 au seuil de 5%

$$S^2 = \frac{(11.4,533^2)+(14.4,945^2)}{25} \approx 22,735$$

Sous H_0, $t \sim S_{t\upsilon}$ avec : $\nu = 12 + 15 - 2 = 25$

Zone d'acceptation ($\alpha = 0{,}05$) : **[-2,060 ; +2,060]**

$$t_c = \frac{20 - 26{,}8}{\sqrt{\frac{22{,}735}{12}} + \sqrt{\frac{22{,}735}{15}}} \approx -3{,}682$$

→ **Conclusion** : rejet de H0 au seuil de 5%.

La **p-value** vaut **0,001.**

D) Séries appariées :

1) Plans en séries appariées :

Variable X observée sur **un** échantillon à deux reprises : essais croisés, stabilité d'un dosage, détection d'une substance par deux méthodes différentes (concordance/discordance), …

Comparaison de moyennes en séries appariées :

But : comparer les moyennes de X (**quantitatives**) entre les deux « situations ».

Exemple : essai croisé traitement 1 X traitement 2

X observée après traitement 1 (X_1, moyenne μ_1) et traitement 2 (X_2, moyenne μ_2)

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

D = X2 – X1 → $H_0 : \{\mu_D = 0\}$ versus $H_1 : \{\mu_D \neq 0\}$ (test de type « observé/théorique »)

$$t = \frac{\bar{d}}{s_D / \sqrt{n}}$$

Cette variable $t \sim S_{n-1}$ sous H_0 si : $D \sim N(\mu_D ; \sigma_D)$ ou si n est grand c'est à dire $n \geq 30$.

2) Comparaison de proportions en séries appariées : test de MAC NEMAR

Variable : X binaire (2 valeurs) observée à 2 reprises sur un même échantillon.

Exemple : détection de la présence de salmonelles par analyse sérologique et bactériologique

Bactériologique → X prend la valeur {+ ; -}

		Méthode B	
Méthode A		+	-
	+	A → p + +	B → p + -
	-	C → p - +	D → p - -