Université de Picardie Jules Verne **Année 2014-2015**
B Schapira **IUT Techniques de commercialisation**

TD de statistiques groupes 1, 2, 4.
Corrigé de l'Évaluation 2.

NOM : Prénom : Groupe : .

Les « Supports de cours » et les calculatrices sont autorisées.
Chaque réponse doit être conclue par une **phrase**.
Les étapes des raisonnements doivent être présentes.

Exercice 1. On s'intéresse à la note en mathématiques des étudiants d'IUT TC de deuxième année ces quinze dernières années. Cette note est en moyenne de 10, avec un écart-type de 4.

1) Etudie-t-on ici une variable quantitative ou qualitative ? Quelle est cette variable ?

Il s'agit d'une variable quantitative, la note de mathématiques d'un étudiant.

2) Quelle est la probabilité que dans un groupe de 25 étudiants de TC2, cette moyenne soit inférieure à 5 ?

La moyenne des notes sur la population des étudiants vaut $\mu = 10$, avec un écart-type de $\sigma = 4$. Sur un échantillon aléatoire de $n = 25$ étudiants, la variable $Z = \frac{\bar{X}-\mu}{\sigma(\bar{X})} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-10}{4/5}$ suit une loi normale centrée réduite. On écrit donc

$$P(\bar{X} \leq 5) = P(\frac{\bar{X}-10}{4/5} \leq \frac{5-10}{4/5}) = P(Z \leq -6,25) = P(Z \geq 6,25) = 1 - P(Z \leq 6,25)\,.$$

La table de la loi normale donne $P(Z \leq 6,25) \simeq 1$ d'où $P(\bar{X} \leq 5) \simeq 0$.

La probabilité que dans un groupe de 25 étudiants de TC2 cette moyenne soit inférieure à 5 vaut quasiment zéro.

Exercice 2. Un responsable de production teste un lot de 100 ampoules d'une certaine marque. Il observe sur cet échantillon que la durée de vie des ampoules est en moyenne de 120 heures avec un écart-type de 3 heures.

1) Etudie-t-on ici une variable quantitative ou qualitative ? Quelle est cette variable ?

Il s'agit d'une variable quantitative, la durée de vie d'une ampoule.

2) Donner une estimation de la durée de vie moyenne de ces ampoules à l'aide d'un intervalle de confiance au niveau de confiance de 95%.

La population étudiée est celle des ampoules de la marque. La variable est la durée de vie, qui a une durée de vie moyenne μ. Sur un échantillon de taille $n = 100$, on observe une durée de vie moyenne $m = 120$ avec un écart-type $s = 3$/ L'estimation ponctuelle donne $\hat{\mu} = m = 120h$. l'estimation par intervalle de confiance au niveau de $1-\alpha = 95\%$ donne un intervalle de confiance

$$IC_{95\%}(\mu) = \left[m \pm Z_{0,975}.\frac{s}{\sqrt{n-1}}\right] = \left[120 - 1,96 \times \frac{3}{\sqrt{99}}; 120 + 1,96 \times \frac{3}{\sqrt{99}}\right] = [119,;120,5909 \simeq [119,59;120,60]$$

Ainsi il y a 95% de chances que μ appartienne à l'intervalle $IC_{95\%}(\mu) = [119,59;120,60]$.

3) Le fabricant prétend que la durée de vie moyenne de ces ampoules est de 122 heures. Le responsable de production peut-il accepter cette hypothèse, compte tenu de ce qu'il a observé ?

Première étape : On teste l'hypothèse $H_0 : \mu = 122$ heures contre l'hypothèse $H_1 : \mu \neq 122$ heures. C'est un test de conformité de la moyenne.

Deuxième étape : on choisit le seuil de signification du test à $1-\alpha = 95\%$

Troisième étape : on fait les calculs en supposant que H_0 est vraie. Léchantillon est grand et l'écart-type inconnu, donc on a $S_{th} = Z_{0,975} = 1,96$ et

$$S_{cal} = \frac{m-\mu}{s/\sqrt{n-1}} = \frac{120-122}{3/\sqrt{99}} = -2\frac{\sqrt{99}}{3} \simeq -6,63\,.$$

Quatrième étape : On a $|S_{cal}| = +6,63 > 1,96$ donc on rejette H_0, avec une probabilité 5% de se tromper (erreur de première espèce).

Exercice 3. Un responsable de production hésite entre deux fournisseurs d'ampoules A et B. Il essaie des lots de 100 ampoules de chaque marque. Il observe que sur ces échantillons, la durée de vie des ampoules A est en moyenne de 120 heures avec un écart-type de 3 heures, tandis que la durée de vie des ampoules B est en moyenne de 119 heures avec un écart-type de 2heures.

1) Etudie-t-on ici une variable quantitative ou qualitative ? Quelle est cette variable ?

Il s'agit d'une variable quantitative, la durée de vie d'une ampoule.

2) Peut-il affirmer que les ampoules ont la même durée de vie moyenne ?

Il s'agit d'un test d'homogénéité de moyennes. On regarde deux populations, les ampoules A et les ampoules B. Sur la première population, on note μ_1 la durée de vie moyenne, σ_1 son écart-type, et sur la deuxième population μ_2 la durée de vie moyenne, σ_2 son écart-type. L'échantillon de la première population a une taille $n_1 = 100$, la durée de vie moyenne sur l''echantillon vaut $m_1 = 120$ heures avec un écart-type de $s_1 = 3$ heures tandis que sur l'échantillon de la deuxième population, on a $n_2 = 100$, $m_2 = 119$ et $s_2 = 2$.

Première étape On veut tester l'hypothèse $H_0 : \mu_1 = \mu_2$ contre l'hypothèse alternative $H_1 : \mu_1 \neq \mu_2$

Deuxième étape : on choisit le seuil de signification du test à $1 - \alpha = 95\%$

Troisième étape : on fait les calculs en supposant que H_0 est vraie. Léchantillon est grand et l'écart-type inconnu, donc on a $S_{th} = Z_{0,975} = 1,96$ et

$$S_{cal} = \frac{m_1 - m_2}{\sqrt{\frac{(s_1)^2}{n_1-1} + \frac{(s_2)^2}{n_2-1}}} = \frac{120 - 119}{\sqrt{\frac{9+4}{99}}} = \sqrt{\frac{99}{13}} = 2,7593$$

Quatrième étape : On a $|S_{cal}| \simeq +2,76 > 1,96$ donc on rejette H_0, avec une probabilité 5% de se tromper (erreur de première espèce).

3) Sur ces échantillons, il observe que parmi les ampoules A il y a une proportion de 2% d'ampoules défectueuses, tandis que parmi les ampoules B, la proportion d'ampoules défectueuses est de 1,5%. Etudie-t-on maintenant une variable quantitative ou qualitative ?

On étudie maintenant une variable qualitative : une ampoule est-elle en bon état ou défectueuse.

4) Peut-il affirmer que les proportions d'ampoules A et B défectueuses sont identiques ?

Il s'agit maintenant d'un test d'homogénéité de la proportion. On étudie les deux populations d'ampoules A et B. On note π_1 la proportion d'ampoules A qui sont défectueuses et π_2 la proportion d'ampoules B qui sont défectueuses. L'échantillon d'ampoules A a une taille $n_1 = 100$, et la proportion d'ampoules défectueuses vaut $p_1 = 2\% = 0,02$ tandis que le deuxième échantillon, d'ampoules B, a une taille $n_2 = 100$ avec $p_2 = 1,5\% = 0,015$.

Première étape : On souhaite tester l'hypothèse $H_0 : \pi_1 = \pi_2$ contre l'hypothèse alternative $H_1 : \pi_1 \neq \pi_2$.

Deuxième étape : on choisit le seuil de signification du test à $1 - \alpha = 95\%$

Troisième étape : on fait les calculs en supposant que H_0 est vraie. L'échantillon est grand, on a $S_{th} = 1,96$ et

$$S_{cal} = \frac{p_1 - p_2}{\sqrt{f(1-f)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{avec} \quad f = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}.$$

On calcule $f = \frac{2+1,5}{200} = 0,0175$ d'où $S_{cal} = \frac{0,005}{\sqrt{0,0175 \times 0,9825 \times \frac{2}{100}}} = \frac{0,005 \times 10}{\sqrt{0,0175 \times 0,9825 \times 2}} \simeq 0,2696$

Quatrième étape : On a $|S_{cal}| < S_{th}$ donc on ne peut pas rejeter H_0 au seuil de 95%. Mais on ne connait pas la probabilité de se tromper.

Exercice 4. La proviseure d'un lycée amiénois s'intéresse à la réussite de ses élèves à un bac blanc de mathématiques Elle observe les résultats suivants, suivant le genre des élèves.

	Entre 0 et $4,9$	Entre 5 et $9,9$	Entre 10et $14,9$	Entre 15et 20	Total
Filles	4	12	21	10	47
Garçons	6	8	24	7	45
Total	10	20	45	17	92

Peut-elle affirmer que les résultats des filles et des garçons sont similaires ?
Vous détaillerez les calculs intermédiaires, par exemple sous forme d'un tableau.

On va faire un test d'homogénéité du χ^2.
Première étape H_0 : les garçons et les filles ont des résultats similaires. H_1 leurs résultats ne sont pas similaires.

Deuxième étape : on choisit le seuil de signification du test à $1-\alpha = 95\%$

Troisième étape : on choisit le seuil de signification du test à $1-\alpha = 95\%$
Le tableau ci-dessus représente les O_i, valeurs observées. On calcule T_i en multipliant le total de la ligne et de la colonne correspondant à la case, et on divise par l'effectif total, soit encore

	Entre 0 et $4,9$	Entre 5 et $9,9$	Entre 10et $14,9$	Entre 15et 20	
Filles	$\frac{10\times 47}{92} \simeq 5,11$	$\frac{20\times 47}{92} \simeq 10,22$	$\frac{45\times 47}{92} \simeq 22.99$	$\frac{17\times 47}{92} \simeq 8.68$	
Garçons	$\frac{10\times 45}{92} \simeq 4.89$	$\frac{20\times 45}{92} \simeq 9.78$	$\frac{45\times 45}{92} \simeq 22.01$	$\frac{17\times 45}{92} \simeq 8.31$	

Et dans un dernier tableau, on calcule les quantités $\frac{(O_i-T_i)^2}{T_i}$ puis on fait la somme de touts les cases.

	Entre 0 et $4,9$	Entre 5 et $9,9$	Entre 10et $14,9$	Entre 15et 20
Filles	$\frac{(4-5,11)^2}{5,11} \simeq 0.24$	$\frac{(12-10,22)^2}{10,22} \simeq 0.31$	$\frac{(21-22.99)^2}{22,99} \simeq 0.17$	$\frac{(10-8,68)^2}{8,68} \simeq 0.20$
Garçons	$\frac{(6-4,89)^2}{4,89} \simeq 0.25$	$\frac{(8-9,78)^2}{9,78} \simeq 0.32$	$\frac{(24-22.08)^2}{22.08} \simeq 0.17$	$\frac{(7-8.31)^2}{8.31} \simeq 0.21$

On a donc $S_{cal} = 0,24+0,31+0,17+0,20+0,25+0,32+0,17+0,21 = 1,87$.
Pour le calcul de S_{th}, on observe que cela fera intervenir un χ^2 à $\nu = (2-1).(4-1) = 3$ degrés de liberté. On a $\chi^2_{0,05}(3) = 7.81$

Quatrième étape : $|S_{cal}| = 1,87 \leq S_{th} = 7,81$ donc on ne rejette pas H_0 au seuil de risque de 5%, mais on ne connait pas la probabilité de se tromper (erreur de deuxième espèce).