CORRECTION DE L'EXAMEN BLANC DE STATISTIQUES PHARMACIE

Exercice 1

1) moyenne μ (ou $E(X)$) estimée par la moyenne de l'échantillon $m = \frac{\sum x_i}{n}$.

dans p_1 $\boxed{m_1 = 0{,}35}$

dans p_2 $\boxed{m_2 = 0{,}30}$

variance $\sigma^2(x)$ estimée par les n observations

$$s^2 = \frac{1}{n-1}\left[\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right]$$

$$s_1^2 = \frac{1}{12}\left(1{,}6072 - \frac{(4{,}55)^2}{13}\right) = 0{,}001225 \text{ soit } \boxed{s_1 = 0{,}035}$$

$$s_2^2 = \frac{1}{12}\left(1{,}1777 - \frac{3{,}9^2}{13}\right) = 0{,}000645 \text{ soit } \boxed{s_2 = 0{,}0253}$$

2) Nous disposons de deux échantillons indépendants de taille $n_1 = 13$ et $n_2 = 13$. La comparaison des variances de X dans les 2 populations p_1 et p_2 sera effectuée par un test de FISCHER :

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

variable de décision : $F = \frac{s_1^2}{s_2^2} = \frac{0{,}035^2}{0{,}0253^2} = 1{,}91$.

valeur seuil $F(\alpha = 0{,}05 \quad \theta_1 = 12 . \ \theta_2 = 12) = 3{,}28$
(on cherche dans la table à 2,5%)

$F_{calculé} < F_{seuil}$ non rejet de H_0 : on ne rejette pas l'hypothèse d'égalité des variances de X dans p_1 et p_2.

2) 3) Pour savoir si la maladie a un effet sur la concentration sanguine de l'urée (X), on teste les hypothèses :

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

La variable X est supposée distribuée selon une loi de Gauss dans les 2 populations.

De plus, on a montré que les variances ne sont pas statistiquement différentes. On peut donc comparer <u>les moyennes</u> sur <u>deux échantillons indépendants de petit effectif</u> ($n < 30$) par un <u>test t</u> de STUDENT.

On estime $s^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \dfrac{12 s_1^2 + 12 s_2^2}{24} = 0{,}000935$

$$t = \frac{m_1 - m_2}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{0{,}35-0{,}30}{0{,}0305\sqrt{2/13}} = 4{,}17.$$

valeur seuil $t(\alpha = 0{,}01 \quad \theta = 24) = 2{,}797$

$t > t_{seuil}$ au seuil $\alpha = 0{,}01 \Rightarrow$ rejet de H_0. On en déduit que la maladie a un effet sur la concentration sanguine de l'urine.

<u>Exercice 2</u>

A)

	Femmes jeunes	Femmes âgées	
précoce	56 (47)	38 (47)	94
tardive	144 (153)	162 (153)	306
	200	200	400.

$T_{ij} = \dfrac{O_{i.} \times O_{.j}}{n}$. $\quad$ ex : $\dfrac{94 \times 200}{400} = 47$

2 échantillons indépendants

$\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}$

$$\chi^2_c = \frac{(56-47)^2}{47} + \frac{(38-47)^2}{47} + \frac{(144-153)^2}{153} + \frac{(162-153)^2}{153} = 4,5$$

χ^2 théorique $= 3,84$ à 5% pour $\nu = (2-1)(2-1) = 1$ ddl

$\chi^2_c > \chi^2_{théor.}$ à 5% $\Rightarrow$ différence significative.

rejet de H_0.

le pourcentage de pubertés tardives diffèrent entre les femmes jeunes et les femmes âgées, au 5%.

B)

	40-44	45-49	50-54	55-59	
précoce	30 23,5	26 23,5	22 23,5	16 23,5	94
tardive	70 76,5	74 76,5	78 76,5	84 76,5	306.
	100	100	100	100	400

$T_{ij} = \frac{O_{i.} \times O_{.j}}{n}$ ex: $\frac{94 \times 100}{400} = 23,5$.

$$\chi^2_c = \frac{(30-23,5)^2 + (26-23,5)^2 + (22-23,5)^2 + (16-23,5)^2}{23,5}$$

$$+ \frac{(70-76,5)^2 + (74-76,5)^2 + (78-76,5)^2 + (84-76,5)^2}{76,5} = 5,95$$

$\nu = (4-1)(2-1) = 3$ ddl

χ^2 théorique $= 7,81$ à 5% et 3 ddl

$\chi^2_c < \chi^2$ théorique $\Rightarrow$ non significatif. non rejet de H_0.

c) Ces résultats semblent en contradiction avec ceux de A). le test du χ^2 n'est pas le meilleur test à utiliser ici car il ne tient pas compte de l'ordre des classes. (ce qui est toujours réalisé lorsqu'il n'existe que 2 classes).

4)

Exercice 3

1) X prend les valeurs 1, 5. $I = \{1;5\}$

Y prend les valeurs −4, 2, 7. $J = \{-4;2;7\}$

X est une V.A discrète à valeurs dans I telle que

$p(X=1) = p(X=1, Y=-4) + p(X=1;Y=2) + p(X=1;Y=7) = \frac{1}{8}+\frac{1}{4}+\frac{1}{8} = \boxed{\frac{1}{2}}$

$p(X=5) = \sum_i p(X=5; Y=j) = \frac{1}{4}+\frac{1}{8}+\frac{1}{8} = \boxed{\frac{1}{2}}$

De même, Y est une V.A discrète à valeurs dans J telle que

$p(Y=-4) = \sum_i p(X=i, Y=-4) = 1/8 + 1/4 = \boxed{3/8}$

$p(Y=2) = \sum_i p(X=i, Y=2) = 1/4 + 1/8 = \boxed{3/8}$

$p(Y=7) = \sum_i p(X=i, Y=7) = 1/8 + 1/8 = \boxed{1/4}$

2) $E(X) = \sum i p(X=i) = 1 \times 1/2 + 5 \times 1/2 = \boxed{3}$.

$E(Y) = \sum j \cdot p(X=j) = -4\times\frac{3}{8} + 2\times\frac{3}{8} + 7\times\frac{1}{4} = \boxed{1}$.

3) $Cov(X,Y) = E(XY) - E(X)E(Y)$

$E(XY) = \sum\sum_{.} ij\, p(X=i . Y=j) = -\frac{4}{8} + \frac{2}{4} + \frac{7}{8} - \frac{20}{4} + \frac{10}{8} + \frac{35}{8} = \boxed{\frac{3}{2}}$

d'où $Cov(X,Y) = \frac{3}{2} - 3\times 1 = \boxed{-\frac{3}{2}}$

$\sigma(X) = \sqrt{Var(X)} = \sqrt{E(X^2) - [E(X)]^2}$ avec $E(X^2) = \sum i^2 p(X=1) = 13$

donc : $\sigma(X) = \sqrt{13 - 3^2} = \boxed{2}$

de même $E(Y^2) = (-4)^2 \times 3/8 + 2^2 \times 3/8 + 7^2 \times \frac{1}{4} = \frac{48}{8} + \frac{12}{8} + \frac{49}{4} = \frac{79}{4}$

donc : $\sigma(Y) = \sqrt{79/4 - 1} = \sqrt{\frac{75}{4}} = \boxed{4{,}33}$

$r_{X,Y} = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)} = \frac{-3/2}{2\times\sqrt{\frac{75}{4}}} = \frac{-3/2}{\sqrt{75}} \simeq \boxed{-0{,}173}$.