# ***TD de Bio statistiques*** : ***EXERCICES SUPPLEMENTAIRES SUR LES TESTS D'HYPOTHESE***

***Exercice 1*** :

Le coût moyen d'une implantation relevé sur internet est de 200 €. Une enquête auprès de 80 femmes implantées donne les résultats suivants :

| Coût en € | ] 150 ; 180] | ] 180 ; 190] | ] 190 ; 200] | ] 200 ; 210] | ] 210 ; 250] |
|---|---|---|---|---|---|
| Effectif | 4 | 15 | 30 | 25 | 6 |

A partir de cet échantillon peut-on considérer au seuil de 5 % que le coût moyen d'une implantation est inférieur à 200 € ?

***Solution***

**Test de conformité** : l'échantillon a donné $\bar{x}$ = 197.375 ; s = 13.5472684 et n = 80.

*Formulation des hypothèses* ; test unilatéral gauche (égal contre inférieur).

Hypothèse nulle $H_0$ : $\bar{x} = \mu = 200$, contre

Hypothèse alternative $H_1$ : $\bar{x} < \mu = 200$ €.

*Le seuil critique* : le niveau de signification est $\alpha = 0.05 = P$ (rejeter $H_0/H_0$ vraie) ;

Comme n = 80 c'est un grand échantillon. $\Rightarrow$ La table n° 1 donne pour un test unilatéral droit $u_{0.05} = 1.645$ et pour un test unilatéral gauche $-u_{0.05} = -1.645$

Si la statistique de test To $\in [-1.645 ; +\infty] \Rightarrow$ on ne peut pas rejeter $H_0$, sinon on la rejette.

*Calcul de la statistique de test* : $To = \frac{\sqrt{80}\,(197.375 - 200)}{13.5472684} = -1.733095785 \approx -1.733$

*Décision* : comme $To = -1.733 \notin [-1.645 ; +\infty]$ alors on ne peut pas accepter $H_0$.

***Conclusion*** : Au seuil de 5%, on considère que le coût moyen d'implantation n'est pas conforme à la valeur vraie de 200 €, donc le coût moyen est inférieur à la valeur théorique.

## ***Exercice 2*** :

La tablette tactile s'impose de nos jours dans les foyers en général ! Les résultats de deux sondages dans notre pays effectués sur 50 personnes au cours de l'année 2016, et sur 120 personnes au cours de l'année 2017 indiquent que 26 puis 87 personnes possèdent une tablette tactile.

En 2018 le budget annuel moyen consacré par habitant dans notre pays aux nouvelles technologies de communication se répartissait en 140 € pour l'investissement et 100 € pour le fonctionnement.

Pour le sondage 2018 qui a été effectué, le tableau suivant donne la répartition du nombre de personnes en fonction du budget consacré à l'investissement et au fonctionnement dans les nouvelles technologies.

| Euros € | ] 0 ; 20] | ] 20 ; 40] | ] 40 ; 80] | ] 80 ; 150] | ]150 ; 250] | ]250 ; 500] |
|---|---|---|---|---|---|---|
| Investissement | 5 | 12 | 32 | 40 | 25 | 16 |
| Fonctionnement | 9 | 28 | 37 | 30 | 21 | 5 |

1) Peut-on au seuil de 5 % conclure à l'homogénéité des possesseurs des tables tactiles en 2016 et 2017 ?
2) Peut-on au seuil de 5 % conclure à la conformité du budget de l'investissement à la moyenne théorique de 140 €.
3) Peut-on au seuil de 5 % conclure à la conformité du budget de fonctionnement à la moyenne théorique de 100 €.
4) Peut-on au seuil de 5 %, considérer que la distribution du budget de l'investissement suit une loi normale N (130 ; $100^2$) ?
5) Peut-on au seuil de 5 %, considérer que la distribution du budget de fonctionnement suit une loi normale N (90 ; $75^2$) ?

***Solution*** :

**1) Test d'homogénéité** :

*Formulation des hypothèses* : test bilatéral où $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$.

*Le seuil critique* : α = 0.05 = P (rejeter $H_0$ / $H_0$ vraie). La table n° 1 donne $u_{0.05} = 1.96$

Si la statistique de test To ∈ [– 1.96 ; + 1.96] ⇒ on ne peut pas rejeter $H_0$.

Les deux fréquences observées : $f_1 = \frac{26}{50} = 0.52$ et $f_2 = \frac{87}{120} = 0.725 \Rightarrow f_0 = \frac{26+87}{50+120} = 0.6647059$

$\Rightarrow \sigma_c = \sqrt{f_0(1- f_0)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{\frac{113}{170} * \frac{57}{170}\left(\frac{1}{50}+\frac{1}{120}\right)} = 0.079465123 \Rightarrow$

*La statistique de test observé* : To $= \frac{0.52 - 0.725}{0.079465123} = -2.579748099$

*Décision* : comme la valeur du test : To = – 2,580 ∈ [– 1.96 ; + 1.96] ⇒ on rejette $H_0$ ⇒ Il y a une différence significative entre les 2 proportions de 2016 et de 2017.

**2) Test de conformité pour le budget de l'investissement.**

L'échantillon donne les paramètres : $\bar{x} = 137.9230769$ ; s = 105.3800516 ; n = 130.

*Formulation des hypothèses* ; test bilatéral.

Hypothèse nulle $H_0$ : $\bar{x} = \mu = 140$, contre

Hypothèse alternative $H_1$ : $\bar{x} \neq \mu = 140$.

*Le seuil critique* : le niveau de signification est α = 0.05 = P (rejeter $H_0$/$H_0$ vraie) ; grand échantillon car n = 130 ⇒ $u_{0.05} = 1.96$ tiré de la table n° 2 de la loi normale.

*La règle de décision sera* : si la statistique de test observé To ∈ [– 1.96 ; + 1.96] (dans ce cas, on ne peut pas rejeter $H_0$). Dans tous les autres cas, on peut la rejeter.

*Calcul de la statistique de test* : To $= \frac{\sqrt{130}\,(137.9230769 - 140)}{105.3800516} = -0.224715839 \approx -0.225$

*Décision* : comme To = – 0.225 ∈ [– 1.96 ; + 1.96] alors on ne peut pas rejeter $H_0$.

***Conclusion*** : Au seuil de 5%, la moyenne de l'investissement est conforme à la valeur vraie.

**3) Test de conformité pour le budget de fonctionnement.**

L'échantillon donne les paramètres : $\bar{x} = 97.5$ ; s = 81.80912271 ; n = 130.

*Formulation des hypothèses* ; test bilatéral.

Hypothèse nulle $H_0$ : $\{\bar{x} = \mu = 100\}$ ; contre Hypothèse alternative $H_1$ : $\{\bar{x} \neq \mu = 100\}$.

*Le seuil critique* : le niveau de signification est α = 0.05 = P (rejeter $H_0/H_0$ vraie).

Grand échantillon car n = 130 ⇒ $u_{0.05}$ = 1.96 tiré de la table n° 2 de la loi normale.

*La règle de décision sera* : si la statistique de test To ∈ [– 1,96 ; + 1.96] (dans ce cas, on ne peut pas rejeter $H_0$). Dans tous les autres cas, on peut la rejeter.

*Calcul de la statistique de test observé* : To = $\frac{\sqrt{130}\,(97.5 - 100)}{81.80912271}$ = – 0.348425514 ≈ **– 0.348**

*Décision* : comme To = **– 0.348** ∈ [– 1.96 ; + 1.96] alors on ne peut pas rejeter $H_0$. On conclut que la moyenne du budget de fonctionnement est conforme à la valeur vraie de 100 €.

**4) Test d'adéquation à la loi normale N (130 ; 100²) pour le budget de l'Investissement.**

*Formulation des hypothèses* :

Hypothèse nulle : $\mathbf{H_0}$ = {les observations sont issues d'une N (130 ; 100²)} **contre**

Hypothèse alternative $\mathbf{H_1}$ = {les observations ne sont pas issues d'une N (130 ; 100²)}

*Le seuil critique* : α = 0.05 = P [rejeter $H_0$ / $H_0$ vraie] = P $\left(\chi^2_{(5)} \geq SC\right)$ = 0.05, d'où SC = 11.07

*Calcul de la statistique de test observé* du Khi-Deux To = $\sum_{i=1}^{k} \frac{(n_i - n\,p_i)^2}{n\,p_i}$

On a $\mathbf{p_i}$ = P [X ∈ classe i] = P [$y_{i-1}$ < X < $y_i$] = P (N (μ = 130 ; σ = 100) ∈ [$y_{i-1}$ ; $y_i$]) = P$\left(N(0\,;1) \in [\frac{y_{i-1} - \mu}{\sigma} ; \frac{y_i - \mu}{\sigma}]\right)$ = P [$z_{i-1}$ < N (0 ; 1) < $z_i$] = $\boldsymbol{\varphi(z_i) - \varphi(z_{i-1})}$.

On prend dans le tableau suivant $z_i = \frac{y_i - \mu}{\sigma} = \frac{y_i - 130}{100}$.

To = $\chi^2_{(\text{observé})} = \sum_{i=1}^{6} \frac{(O_i - E\,t_i)^2}{E\,t_i} = \frac{(5-5.057)^2}{5.057} + \frac{(12-6.292)^2}{6.292} + \ldots + \frac{(16-14.963)^2}{14.963}$ = 24.43134

| $y_i$ | – ∞ | 0 | 20 | 40 | 80 | 150 | 250 | 500 | + ∞ |
|---|---|---|---|---|---|---|---|---|---|
| $z_i$ | – ∞ | – 1.3 | – 1.1 | – 0.9 | – 0.5 | 0.2 | 1.2 | 3.7 | + ∞ |
| $\varphi(z_i)$ | 0 | 0.0968 | 0.1357 | 0.1841 | 0.3085 | 0.5793 | 0.8849 | 0.9999 | 1 |

Les valeurs $z_i$ obtenues de la variable aléatoire Z = $\frac{Y - \mu}{\sigma}$ ; φ ($z_i$) lue de la table n° 1.

| $\mathbf{p_i}$ | 0.0968 | 0.0389 | 0.0484 | 0.1244 | 0.2708 | 0.3056 | 0.1150 | 0.0001 | 1 |
|---|---|---|---|---|---|---|---|---|---|
| $n\,p_i$ | 12.584 | 5.057 | 6.292 | 16.172 | 35.204 | 39.728 | 14.95 | 0.013 | 130 |

Vu la taille de certaines classes inférieures à 5, nous devons faire des regroupements.

| $n\,p_i$ | 12.584 | 5.057 | 6.292 | 16.172 | 35.204 | 39.728 | 14.963 | 130 |
|---|---|---|---|---|---|---|---|---|
| $n_i$ | 0 | 5 | 12 | 32 | 40 | 25 | 16 | 130 |
| To | 0 | 0.0006 | 5.1782 | 13.0673 | 0.65338 | 5.45998 | 0.07187 | 24.43134 |

*Décision* : comme To = $\chi^2_{(\text{observé})}$ = 24.43 > SC = $\chi^2_{(5)}$ = 11.07 ⇒ on ne peut pas accepter $H_0$.

***En conclusion***, le budget de l'investissement ne suit pas une loi normale.

**5) Test d'adéquation à la loi normale N (90 ; $75^2$) pour le budget de fonctionnement**

*Formulation des hypothèses* :

Hypothèse nulle : $\mathbf{H_0}$ = {les observations sont issues d'une N (90 ; $75^2$)} **contre**

Hypothèse alternative $\mathbf{H_1}$ = {les observations ne sont pas issues d'une N (90 ; $75^2$)}

*Le seuil critique* : $\alpha = 0.05$ = P [rejeter $H_0$ / $H_0$ vraie] = $P\left(\chi^2_{(k-p-1)} \geq SC\right) = 0.05$

Où le coefficient p = 0 car la loi normale N (90 ; $75^2$) est donnée et k c'est le nombre de classes dont l'effectif théorique est ≥ 5 d'où SC = $\chi^2_{(4)}$ = 9,49

*Calcul de la statistique de test observé* du Khi-Deux To = $\sum_{i=1}^{k} \frac{(n_i - n\,p_i)^2}{n\,p_i}$

On a $\mathbf{p_i}$ = P [X ∈ classe i] = P [$y_{i-1} < X < y_i$] = P $\left(N (\mu = 90 ; \sigma = 75) \in [y_{i-1} ; y_i]\right)$ =

$P\left(N(0\,;1) \in [\frac{y_{i-1}-\mu}{\sigma} ; \frac{y_i-\mu}{\sigma}]\right)$ = P [$z_{i-1}$ < N (0 ; 1) < $z_i$] = $\boldsymbol{\varphi(z_i) - \varphi(z_{i-1})}$.

On prend dans le tableau suivant $z_i = \frac{y_i - \mu}{\sigma} = \frac{y_i - 90}{75}$.

To = $\sum_{i=1}^{6} \frac{(O_i - E\,t_i)^2}{E\,t_i} = \frac{(9-7.943)^2}{7.943} + \frac{(28-9.776)^2}{9.776} + \ldots + \frac{(26-27.547)^2}{27.547}$ = 38.1318

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $y_i$ | $-\infty$ | 0 | 20 | 40 | 80 | 150 | 250 | 500 | $+\infty$ |
| $z_i$ | $-\infty$ | – 1.2 | – 0.93 | – 0.67 | – 0.13 | 0.8 | 2.13 | 5.47 | $+\infty$ |
| $\varphi(z_i)$ | 0 | 0.1151 | 0.1762 | 0.2514 | 0.483 | 0.7881 | 0.9834 | 1 | 1 |

Les valeurs $z_i$ obtenues de la variable aléatoire Z = $\frac{Y-\mu}{\sigma}$ ; $\varphi$ ($z_i$) lue de la table n° 1.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $\mathbf{p_i}$ | 0.1151 | 0.0611 | 0.0752 | 0.2316 | 0.3051 | 0.1953 | 0.0166 | 0 | 1 |
| n $p_i$ | 14.963 | 7.943 | 9.776 | 30.108 | 39.663 | 25.389 | 2.158 | 0 | 130 |

Vu la taille de certaines classes inférieures à 5, nous devons faire des regroupements.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| n $p_i$ | 14.963 | 7.943 | 9.776 | 30.108 | 39.663 | 27.547 | 0 | 130 |
| $n_i$ | 0 | 9 | 28 | 37 | 30 | 26 | 0 | 130 |
| To | 0 | 0.14066 | 33.9724 | 1.57764 | 2.35417 | 0.08688 | | 38.1318 |

*Décision* : comme To = 39.13 > SC = $\chi^2_{(4)}$= 9.49 ⇒ on ne peut pas accepter l'hypothèse nulle.

***En conclusion***, le fonctionnement ne suit pas une loi normale.

## ***Exercice 3*** :

Lors de la réunion annuelle des associations virtuelle des chirurgiens de l'Est qui se tient à la capitale, un représentant du ministère éditant les journaux et magazines relatifs à la santé était invité. Il s'est livré à une petite enquête pour connaître le temps nécessaire à un chirurgien pour compléter une certaine intervention. Les résultats obtenus sont consignés dans le tableau suivant :

| Temps nécessaire en minutes | [50-60[ | [60-70[ | [70-90[ | [90-100[ | [100-120] |
|---|---|---|---|---|---|
| Effectif | 2 | 8 | 25 | 10 | 5 |

**1) a)** Par quelles valeurs ponctuelles peut-on estimer la moyenne et la variance du temps nécessaire à l'intervention ?

**b)** Dans quel intervalle se situe la moyenne de ce temps avec un coefficient de confiance de 0,98 ?

**c)** Quelle aurait dû être la taille de l'échantillon interrogé si on avait désiré un intervalle de 8 minutes d'amplitude (on supposera que ce nouvel échantillon a les mêmes caractéristiques que le précédent) ?

**2)** En effectuant son enquête le représentant a aussi demandé aux membres de cette association si le temps qu'ils réservent à l'intervention était largement suffisant. 32 des personnes interrogées lui ont répondu que oui.

**a)** Estimer par un intervalle de confiance à 95 % la proportion des chirurgiens qui se contentent du temps réservé à l'intervention.

**b)** Tester l'hypothèse que plus de la moitié des chirurgiens sont content de ce temps avec un risque de 2 %.

**3)** Tout en menant son enquête il a discrètement relevé l'âge approximatif des personnes interrogées. A partir du tableau suivant, il aimerait bien savoir s'il existe une relation entre l'âge du chirurgien et la préférence du temps alloué à l'intervention. L'erreur sur la conclusion qu'il prendra doit être de 5%.

| Age<br>préférence du temps | Moins de 30 ans | De 30 à 45 ans | De 45 à 60 ans | Plus de 60 ans |
|---|---|---|---|---|
| OUI | 4 | 14 | 10 | 8 |
| NON | 8 | 5 | 6 | 5 |

**4)** D'après les enquêtes faites par l'éditeur du "Guide du chirurgien", on sait que 60 % des chirurgiens achètent le Guide. Sur les 60 personnes interrogées par notre enquêteur, quelle est la probabilité que plus de la moitié soit en possession de ce Guide.

***Solution***

**1) Estimation des caractéristiques de la population**

a) La moyenne $\mu$ et la variance $\sigma$ de la population sont estimées par la moyenne $\bar{x} = 82,6$ et la variance $s^2 = 193,10204$ de l'échantillon $\Rightarrow s = 13.896116$ et $n = 50$.

b) *Intervalle de confiance* d'une moyenne d'une population quelconque de variance inconnue :
Pour n = 50, on a un grand échantillon et la variable $Z = \sqrt{n} * \left(\frac{\bar{X}-\mu}{S}\right)$ suit approximativement une loi normale centrée et réduite N(0,1).
Equation de départ : $P\left[-u_\alpha < Z = \sqrt{n} * \left(\frac{\bar{X}-\mu}{S}\right) < u_\alpha\right] = 1 - \alpha = 0.98 \Rightarrow \alpha = 0.02$
$P[-u_\alpha < N(0 ; 1) < u_\alpha] = 0,98 \Leftrightarrow u_\alpha = 2,326$ : Valeur lue dans la table n° 2
$P\left[-u_\alpha < \sqrt{n}\left(\frac{\bar{X}-\mu}{S}\right) < u_\alpha\right] = 0.98 \Leftrightarrow P\left(\bar{X} - u_\alpha \frac{S}{\sqrt{n}} < \mu < \bar{X} + u_\alpha \frac{S}{\sqrt{n}}\right) = 0.98$

D'où un intervalle de confiance $\bar{x} \pm u_\alpha \frac{S}{\sqrt{n}} = [82.6 \pm 4.57107281] =$ ***[78.029 ; 87.171]***

c) *La taille n de l'échantillon* donnant l'amplitude A = 8

Supposons n > 30. Amplitude = 8 = $2u_\alpha \frac{s}{\sqrt{n}}$ $\Rightarrow$ n = $\left(\frac{s\,u_\alpha}{4}\right)^2 = \left(\frac{13.896116 * 2.326}{4}\right)^2$ = 65.3 ; d'où **n = 66** (ce qui confirme la supposition et l'utilisation de la loi normale).

***<u>2) Intervalle de confiance et test</u>***

a) Intervalle de confiance d'une proportion. La fréquence observée étant f = $\frac{32}{50}$ = 0.64

*<u>Pour n = 50</u>*, c'est un grand échantillon. La variable aléatoire Z = $\sqrt{n} * \left(\frac{F - p}{\sqrt{p(1-p)}}\right)$ suit approximativement une loi normale centrée et réduite N(0,1).

*Equation de départ* : P $\left(-u_\alpha < \sqrt{n}\left(\frac{F - p}{\sqrt{p(1-p)}}\right) < u_\alpha\right)$ = 1 – α = 0.95 ⇒ α = 0.05

P [– $u_\alpha$ < N (0 ; 1) < $u_\alpha$] = 0,95 ⇔ $u_\alpha$ =1,96

P $\left(-u_\alpha < \sqrt{n}\left(\frac{F - p}{\sqrt{p(1-p)}}\right) < u_\alpha\right)$ = 0.95 ⇔ P $\left(F - u_\alpha\sqrt{\frac{p(1-p)}{n}} < p < F + u_\alpha\sqrt{\frac{p(1-p)}{n}}\right)$ = 0.95

D'où un intervalle de confiance [f ± $u_\alpha \sqrt{\frac{p(1-p)}{n}}$] qui ne convient pas car il contient le paramètre p à encadrer; il peut être remplacé par une bonne approximation, à savoir f, d'où un intervalle de confiance [f ± $u_\alpha \sqrt{\frac{f(1-f)}{n}}$] = [0.64 ± 0.133049] = ***[0.507 ; 0.773]***

b) Test d'hypothèse de conformité sur une proportion. Test unilatéral droit (= contre >).

La fréquence de l'échantillon f = 0.64 et $u_\alpha$ > 0

*<u>Formulation des hypothèses</u>* : $H_0$ = {p = $p_0$ = 0,5} contre $H_1$ = {p > $p_0$}

*<u>Pour n = 50</u>*, la variable aléatoire F = $\sqrt{n}\left(\frac{F - p}{\sqrt{p(1-p)}}\right)$ suit approximativement une loi normale centrée et réduite N(0,1).

*<u>Le seuil critique</u>* : α = 0.02 = P [Rejeter $H_0$ / $H_0$ vraie] ⇒ $u_\alpha$ = 2.055

Si la statistique observée To = $\sqrt{n}\left(\frac{F - p}{\sqrt{p(1-p)}}\right)$ < $u_\alpha$ = 2.055 on "accepte" $H_0$ sinon on la rejette

*<u>Calcul de la statistique de test</u>* To = $\sqrt{n}\ \frac{f - p_0}{\sqrt{p_0(1-p_0)}}$

D'où To = $\sqrt{n}\ \frac{f - p_0}{\sqrt{p_0(1-p_0)}} = \sqrt{50}\ \frac{0.64 - 0.5}{\sqrt{0.5(1-0.5)}}$ = 1.9799

*<u>Décision</u>* : comme To = 1.9799 < $u_\alpha$ = 2.055, l'expérience ne permet pas de rejeter $H_0$

**3) Test d'indépendance**

*<u>Formulation des hypothèses</u>*: $H_0$ = {les 2 caractères sont indépendants} contre
$H_1$ = {les 2 caractères ne sont pas indépendants}

*<u>Le seuil critique</u>* : α = 0.05 ; il y a 2 lignes et 3 colonnes dans la table de contingence ⇒

Le nombre de degrés de liberté = k = ddl = (l–1) * (c–1) = 1 * 3 = 3 ⇒

$$\chi^2_{(k-1\,;\,0.05)} = \chi^2_{(2\,;\,0.05)} = 5.991$$

Si on a : To = $\chi^2_{(\text{observé})} < \chi^2_{(k-1\,;\,0.05)}$ ⇔ on "accepte" $H_0$ et sinon on la rejette.

*Calcul de la statistique de test observé* To = $\chi^2_{(observé)} = \sum_{i,j} \frac{(n_{ij} - n\, p_{ij})^2}{n\, p_{ij}}$

| | < 30 ans | | entre 30 et 45 | | entre 45 et 60 | | > 60 ans | | Total |
|---|---|---|---|---|---|---|---|---|---|
| OUI | 4 | 7.2 | 14 | 11.4 | 10 | 9.6 | 8 | 7.8 | 36 |
| NON | 8 | 4.8 | 5 | 7.6 | 6 | 6.4 | 5 | 5.2 | 24 |
| Total | 12 | | 19 | | 16 | | 13 | | 60 |

To = $\chi^2_{(observé)} = \frac{(4-7.2)^2}{7.2} + \frac{(14-11.4)^2}{11.4} + \ldots\ldots + \frac{(5-5.2)^2}{5.2}$ = **5,0925** et $\chi^2_{(2)}$ = 5,991

*Décision* : comme **To = 5.09** = $\chi^2_{(observé)} < \chi^2_{(2)}$ **= 5.99** ⇒ on accepte $H_0$ au seuil de 5%

***En conclusion*** les deux caractères (âge du chirurgien et la préférence du temps alloué à l'intervention) sont indépendants.

**4)** Quelle est la probabilité que plus de la moitié des chirurgiens soit en possession du Guide ?

P (F > 0.5) = P [$\sqrt{n}\left(\frac{F-p}{\sqrt{p(1-p)}}\right) > \sqrt{n}\left(\frac{0.5-p}{\sqrt{p(1-p)}}\right)$] ; où n = 60 ; p = 60 %.

Or pour n = 60, la variable aléatoire F = $\sqrt{n}\left(\frac{F-p}{\sqrt{p(1-p)}}\right)$ est approximativement une N(0,1).

Donc P (F > 0.5) = P (N (0 ; 1) > $\sqrt{n}\left(\frac{F-p}{\sqrt{p(1-p)}}\right) = \sqrt{60}\left(\frac{0.5-0.6}{\sqrt{0.6(1-0.6)}}\right)$ = – 1.58) = 0.9429 ⇒
La probabilité que plus de la moitié soit en possession de ce Guide est égale à **94.29 %**

## ***Exercice 4***

Un fabriquant de compléments alimentaires a prétendu que l'un des produits fabriqués par son entreprise agit positivement sur la mémoire des adolescents en une durée ne dépassant pas 3 mois, et par conséquent, il peut augmenter le taux de réussite des scolarisés, surtout pour les matières scolaires dépendant de la mémoire. Pour vérifier les affirmations de ce fabriquant, on a réalisé l'étude statistique suivante : on a choisi deux échantillons aléatoires d'élèves d'un même collège, les 80 élèves du premier échantillon ont pris ce complément alimentaire pendant 3 mois avant l'examen de l'histoire et les 60 élèves du deuxième échantillon ont pris un placebo pendant la même durée.

Les résultats de l'examen sont représentés dans le tableau suivant :

| | Complément alimentaire | placebo |
|---|---|---|
| Succès | 40 | 40 |
| Echec | 40 | 20 |

D'après cette étude, ce compliment alimentaire agit-il positivement sur le taux de réussite des élèves du collège dans l'examen de l'histoire ? On prend un risque de 5 %.

***Solution***

Généralement, dans des problèmes de ce genre on peut les résoudre par deux types de tests d'hypothèse :

- Test d'homogénéité du KHI-DEUX (s'il n'y a pas d'équivalence).
- Test d'homogénéité pour comparer deux proportions $p_1$ et $p_2$ (pour connaître la forte proportion).

***Solution***

**1) Test d'homogénéité du KHI-DEUX**

- *Formulation des hypothèses* :

  ***H₀*** : « L'effet du complément alimentaire et du placébo sont équivalents » ; **contre**

  ***H₁*** : « L'effet du complément alimentaire et du placébo ne sont pas équivalents ».

- *Calcul de la statistique de test* : To = $\chi^2_{(\boldsymbol{observé})}$

| | Complément alimentaire Observé | Attendu | Placebo Observé | Attendu | Total |
|---|---|---|---|---|---|
| Succès | 40 | 45.7 | 40 | 34.3 | 80 |
| Echec | 40 | 34.3 | 20 | 25.7 | 60 |
| Total | 80 | | 60 | | 140 |

$$\chi^2_{(observé)} = \sum_{i=1}^{4} \frac{(E\,O_i - E\,t_i)^2}{E\,t_i} = \frac{(40-45.7)^2}{45.7} + \frac{(40-34.3)^2}{34.3} + \frac{(40-34.3)^2}{34.3} + \frac{(20-25.7)^2}{25.7} =$$

$$= 0.71094 + 0.94723 + 0.94723 + 1.26420 = 3.869603895 \cong \mathbf{3.87}$$

- *Le seuil critique* : Le tableau de contingence a 2 lignes et 2 colonnes donc le nombre de degré de liberté = 1 ⇒ $\chi^2_{(\mathbf{1})}$ **= 3.84** ; si To > $\chi^2_{(\mathbf{1})}$ alors on ne peut pas accepter $H_0$.

- *Décision* : comme on a $\chi^2_{(\boldsymbol{observé})}$ **= 3.87 = To** > $\chi^2_{(\mathbf{1})}$ **= 3.84** alors il y a **Rejet de $H_0$.**

***En conclusion*** l'effet du complément alimentaire et l'effet du placébo sur la mémoire des adolescents ne sont pas équivalents.

**2) Test d'homogénéité** : comparaison de deux proportions $p_1$ et $p_2$.

*Formulation des hypothèses* : test unilatéral gauche (égal contre inférieur)

Hypothèse nulle : $H_0 = \{p_1 = p_2\}$ contre

Hypothèse alternative : $H_1 = \{p_1 < p_2\}$.

*Le seuil critique* : α = 0.05 = P (rejeter $H_0$ / $H_0$ vraie). La table n° 1 donne $u_{0.05} = 1.645$ et pour un test unilatéral gauche la valeur de $u_{0.05}$ est négative.

Si la statistique de test To ∈ [– ∞ ; – 1.645] ⇒ on ne peut pas accepter $H_0$.

Les deux échantillons $n_1 = 80$ (complément alimentaire) et $n_2 = 60$ (placébo) donnent les fréquences observées : $f_1 = \frac{40}{80} = 0.5$ et $f_2 = \frac{40}{60} = 0.667$ ⇒ $f_0 = \frac{40+40}{80+60} = 0.571428571$ ⇒

$$\sigma_c = \sqrt{f_0(1-f_0)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} = \sqrt{\frac{8}{14} * \frac{6}{14}\left(\frac{1}{80}+\frac{1}{60}\right)} = 0.084515425 \Rightarrow$$

*La statistique de test observé* : To = $\frac{f_1 - f_2}{\sigma_c} = \frac{0.5-0.667}{0.084515425} = -1.972026598 \cong -1.97$

*Décision* : comme la valeur du test : To = – 1,97 ∈ [– ∞ ; – 1.645] ⇒

On ne peut pas accepter $H_0$ ⇒ Il y a une différence significative entre les 2 proportions.

***En conclusion*** la prétention du fabricant est totalement fausse. Au contraire, le produit fabriqué agit négativement sur la mémoire des adolescents.