L2 - Psychologie 2019-2020 - Trois exercices sur ajustement de proportion et de moyenne

Octobre 2019

Exercice 1. Une étude est menée pour tester l'efficacité d'une campagne de vaccination. Un chercheur a constaté qu'avant la campagne, 45% des foyers acceptaient la vaccination. Un mois après la campagne, sur 200 foyers interrogés, 112 acceptent la vaccination.

1. Peut-on en conclure, au risque d'erreur $\alpha = 5\%$, que la campagne de vaccination a significativement changé les comportements des personnes vis a vis de la vaccination ?
2. Calculer la puissance du test $\eta\,(.57)$.

Exercice 2. Le tableau suivant représente le nombre d'heures x passées devant un écran par jour pour un échantillon de jeunes adultes :

Nombres d'heures	$4-5$	$5-6$	$6-7$	$7-8$	$8-9$	$9-10$
Effectifs	3	6	10	12	11	5

1. Calculer la moyenne et l'écart type de l'échantillon.
2. Au vu de cet échantillon peut-on considérer au niveau $\alpha = 0.05$ que la durée moyenne passée devant un écran dans la population des jeunes adultes est inférieure à 8 heures ?
3. Calculer la puissance $\eta\,(7.6)$.

Exercice 3. Une mesure de la taille en centimètres a été réalisée sur un échantillon de 21 garçons de 5 à 15 ans. On obtient une moyenne de 128 et un écart-type de 3.1. On suppose que la taille suit une loi normale. Peut-on dire, avec un risque d'erreur de 5%, que la taille moyenne des garçons de 5 à 15 ans est inférieure à 130 ?

Corrigé de l'Exercice 1. Une étude est menée pour tester l'efficacité d'une campagne de vaccination. Un chercheur a constaté qu'avant la campagne, 45% des foyers acceptaient la vaccination. Un mois après la campagne, sur 200 foyers interrogés, 112 acceptent la vaccination.

1. Peut-on en conclure, au risque d'erreur $\alpha = 5\%$, que la campagne de vaccination a significativement changé les comportements des personnes vis a vis de la vaccination ?
On procède a un test *bilatéral* d'ajustement de proportion.
a) Hypothèses. On désigne par p la proportion de foyers qui acceptent la vaccination.$H_0 : p = p_0 = 0.45.H_1 : p \neq p_0$.
b) Statistique du test. On observe que $n = 200 > 30$, $n \times p_0 = 90 > 5$ et $n \times q_0 = 110 > 5$. Donc, si on désigne par P_n la proportion de foyers qui acceptent la vaccination, on a, sous l'hypothèse nulle : $P_n \hookrightarrow \mathcal{N}\left(p_0, \sqrt{\frac{p_0 q_0}{n}}\right)$, c'est à dire $P_n \hookrightarrow \mathcal{N}(0.45, 0.035)$.
c) Région critique K_α au niveau $\alpha = 0.05$. Il s'agit d'un test *bilatéral*. Donc on détermine la valeur $p_{1,\alpha}$ telle que $\mathbb{P}[P_n \leq p_{1,\alpha}] = \frac{\alpha}{2} = 0.025$ et la valeur $p_{2,\alpha}$ telle que $\mathbb{P}[P_n \geq p_{2,\alpha}] = \frac{\alpha}{2} = 0.025$. On trouve, avec la fonction Invnorm (ou une variante), et éventuellement, sur Casio, Tail=CENTRAL, que $p_{1,\alpha} = 0.381$ et $p_{2,\alpha} = 0.519$. Donc $K_\alpha = \{P_n \leq 0.381\} \cup \{P_n \geq 0.519\}$.
d) Décision du test. On observe que $p_e = \dfrac{112}{200} = 0.560 \in K_\alpha$. Donc, au niveau $\alpha = 0.05$,on accepte l'hypothèse H_1 : la campagne de vaccination a significativement changé le comportement des personnes vis a vis de la vaccination.
2. Calculer la puissance du test $\eta(0.57)$.
On a : $\eta(0.57) = \mathbb{P}\left[K_\alpha; \mathcal{N}\left(0.57, \sqrt{\frac{0.57 \times (1 - 0.57)}{200}}\right)\right] = \mathbb{P}[\{P_n \leq 0.381\} \cup \{P_n \geq 0.519\}; \mathcal{N}(0.57, 0.035)]$.
Ce nombre est égal à $1 - \mathbb{P}[0.0381 \leq P_n \leq 0.519]$. Sur la calculatrice, on trouve $\mathbb{P}[0.0381 \leq P_n \leq 0.519]$ on donnant :
$$\text{LOWER} = 0.381,\ \text{UPPER} = 0.519,\ \sigma = 0.035,\ \mu = 0.381,$$
ce qui donne $\mathbb{P}[0.0381 \leq P_n \leq 0.519] = 0.0721$. On obtient $\eta(0.57) = 1 - 0.0721 = 0.928$. Ainsi, si la proportion de foyers qui acceptent la vaccination est de 57%, le test a 92.8% de chances d'accepter H_1 (à juste titre).

Corrigé de l'Exercice 2. Le tableau suivant représente le nombre d'heures x passées devant un écran par jour pour un échantillon de jeunes adultes :

Nombre d'heures	$4-5$	$5-6$	$6-7$	$7-8$	$8-9$	$9-10$
Effectifs	3	6	10	12	11	5

1. Calculer la moyenne et l'écart type de l'échantillon. Sur la calculatrice, on rentre les centres de classe dans la colonne L_1, par exemple avec la commande SEQ(X,X,4.5,9.5,1), puis les effectifs dans la colonne L_2. On calcule ensuite moyenne et écart-type :
CASIO : en mettant les bons Settings dans Casio : 1Var XList : List1, 1Var Freq : List2, puis commande 1VAR
TEXAS : 1-Var Stats L1,L2.
On trouve $m_e = 7.287, s_e = 1.384$
2. Au vu de cet échantillon peut-on considérer au niveau $\alpha = 0.05$ que la durée moyenne passée devant un écran dans la population des jeunes adultes est inférieure à 8 heures ? On procède à un test *unilatéral* d'ajustement d'une moyenne.
Hypothèses. On désigne par m la moyenne du nombre d'heures passées par jour devant un écran par les jeunes adultes.
$H_1 : m = m_0 = 8$.
$H_2 : m < m_0$.
Statistique du test. $n = 47 > 30$, grand échantillon. On sait alors que, sous l'hypothèse nulle, la moyenne aléatoire M_n d'heures passées devant un écran par les jeunes adultes vérifie :
$$M_n \hookrightarrow \mathcal{N}\left(m_0, \frac{s_e}{\sqrt{n-1}}\right), \text{ c'est-à-dire } M_n \hookrightarrow \mathcal{N}(8, 0.204).$$
Région critique au niveau $\alpha = 0.05$. On cherche m_α tel que $\mathbb{P}[M_n \leq m_\alpha] = 0.05$. On trouve $m_\alpha = 7.665$, c'est à dire $K_\alpha = \{M_n \leq 7.665\}$.
Décision du test. On voit que $m_e = 7.287 < m_\alpha$, donc $m_e \in K_\alpha$. Au niveau $\alpha = 0.05$, on accepte l'hypothèse H_1.
3. Calculer la puissance $\eta(7.6)$. On a $\eta(7.6) = \mathbb{P}\left[K_\alpha; \mathcal{N}\left(7.6, \frac{s_e}{\sqrt{n-1}}\right)\right] = \mathbb{P}[\{M_n \leq 7.665; \mathcal{N}(7.6, 0.204)\}] = 0.624$: si le nombre moyen d'heures passées devant un écran par les jeunes adultes est de 7.6 heures, notre test a 62.4% de chances d'accepter l'hypothèse H_1.

Correction de l'Exercice 3. Une mesure de la taille en centimètres a été réalisée sur un échantillon de 21 garçons de 5 à 15 ans. On obtient une moyenne de 128 et un écart-type de 3.1. On suppose que la taille suit une loi normale. Peut-on dire que la taille moyenne des garçons de 5 à 15 ans est inférieure à 130 ? On procède à un test unilatéral d'ajustement de moyenne.

Hypothèses. On désigne par m la taille moyenne des garçons de 5 à 15 ans.

$H_1 : m = m_0 = 130.$

$H_2 : m < 130.$

Statistique du test. Puisque $n = 21 < 30$, il s'agit d'un petit échantillon. Puisqu'on a supposé que la taille suit une loi normale, on sait que la moyenne aléatoire M_n de la taille des enfants de 5 à 15 ans, ainsi que l'écart-type aléatoire S_n, vérifient :

$$T_n = \frac{M_n - m_0}{\frac{S_n}{\sqrt{n-1}}} \hookrightarrow \mathrm{St}\,(n-1) \text{ c'est à dire } T_n = \frac{M_n - 130}{\frac{S_n}{\sqrt{20}}} \hookrightarrow \mathrm{St}\,(20)\,,$$

où St(20) est la loi de Student à 20 ddl.

Région critique K_α au niveau $\alpha = 0.05$. On utilise la table de Student, pour chercher la valeur t_α telle que $\mathbb{P}\,[T_n \leq t_\alpha] = 0.05$. Puisque $0.05 < 0.5$, le nombre t_α est négatif, est donc n'apparait pas sur la table. On cherche donc, par symétrie, le nombre t'_α tel que $\mathbb{P}\,[T_n \geq t'_\alpha] = 0.05$. A l'intersection de la colonne $\alpha = 0.05$ et de la ligne ddl $= 20$, on trouve $t'_\alpha = 1.7247$. Donc $t_\alpha = -1.7247$ et $K_\alpha = \{T_n \leq -1.7247\}$.

Décision du test. On voit que

$$t_\mathrm{e} = \frac{m_\mathrm{e} - 130}{\frac{s_\mathrm{e}}{\sqrt{n-1}}} = \frac{128 - 130}{\frac{3.1}{\sqrt{20}}} = -2.885,$$

donc $t_\mathrm{e} < t_\alpha = -2.7247$, c'est à dire $t_\mathrm{e} \in K_\alpha$. On en déduit, au niveau $\alpha = 0.05$, qu'on peut accepter l'hypothèse H_1.