L2 - Psychologie 2019-2020 - Corrigé de l'Exercices 6 du TD 5

Exercice 6. Un examen psychotechnique contient deux types de tests $\mathcal{T}_1$ et $\mathcal{T}_2$. Les difficultés des deux tests sont de natures différentes. Pour le test $\mathcal{T}_1$ un groupe de 11 adultes choisis dans une classe socio- professionnelle $\mathcal{C}_1$ a donné les résultats suivants (normalité des variables étudiées admise) :

$\mathcal{T}_1$: x_1	118	123	127	119	134	127	129	131	128	118	134

Un autre groupe de 11 adultes choisis dans une autre classe socio-professionnelle $\mathcal{C}_2$ a donné les résultats suivants pour les deux tests :

$\mathcal{T}_1 : x_1$	135	131	125	132	126	125	128	134	138	127	132
$\mathcal{T}_2 : x_2$	131	129	128	133	131	135	119	125	127	121	123

1. Tester si la classe $\mathcal{C}_2$ donne de meilleurs résultats pour le test $\mathcal{T}_1$ que la classe $\mathcal{C}_1$.

Les deux classes $\mathcal{C}_1$ et $\mathcal{C}_2$ ne sont pas reliées. Il s'agit donc d'un test (orienté) de *comparaison de moyennes dans le cas de deux petits échantillons indépendants* (bien qu'ils aient la même taille), qu'il faut faire précéder d'un test (bilatéral) de *comparaison des écart-types.*

A) TEST DE COMPARAISON DES ÉCART-TYPES.

Hypothèses. On désigne par σ_1 et σ_2 les écart-types des résultats des classes $\mathcal{C}_1$ et $\mathcal{C}_2$ à la tâche $\mathcal{T}_1$. ATTENTION : nous parlons bien des écart-types des résultats généraux des deux classes $\mathcal{C}_1$ et $\mathcal{C}_2$, et pas des ***écart-types expérimentaux*** obtenus sur les deux échantillons de 11 adultes choisis dans chaque classe.

$$H_0 : \sigma_1 = \sigma_2 \,; \ H_1 : \sigma_1 \neq \sigma_1 .$$

Statistique du test. Pour déterminer la variable aléatoire F, on calcule les écart-types corrigés des deux échantillons. Celui qui est le plus grand déterminera la variance corrigée aléatoire qu'il faut mettre au numérateur de F.

On travaille avec des échantillons de taille $n_1 = n_2 = 11$. Avec la calculatrice, on trouve : $s_{1,\mathrm{e}} = 5.670$, $\widehat{s}_{1,\mathrm{e}} = 5.947$, $s_{2,\mathrm{e}} = 4.180$, $\widehat{s}_{2,\mathrm{e}} = 4.384$. Donc $\widehat{s}_{1,\mathrm{e}} > \widehat{s}_{2,\mathrm{e}}$. Ainsi, sous l'hypothèse nulle H_0, on a :

$$F = \frac{\left(\widehat{S}_{n_1}\right)^2}{\left(\widehat{S}_{n_2}\right)^2} = \frac{\widehat{V}_{n_1}}{\widehat{V}_{n_2}} \hookrightarrow \mathrm{FS}\left(n_1 - 1, n_2 - 1\right) = \mathrm{FS}\left(10, 10\right),$$

où $\widehat{S}_{n_1}$ et $\widehat{S}_{n_2}$ représentent les écart-types aléatoires corrigés des deux classes $\mathcal{C}_1$ et $\mathcal{C}_2$, et $\widehat{V}_{n_1}$ et $\widehat{V}_{n_2}$ les variables aléatoires corrigées des classes $\mathcal{C}_1$ et $\mathcal{C}_2$.

Région critique K_α au niveau $\alpha = 0.05$. On fait un test bilatéral, donc on travaille avec les tables de la loi de Fisher-Snedecor pour $\alpha = 0.025$. Les tables ne donne que "la moitié supérieure" de la région critique bilatérale, qu'on note $K_\alpha^+ = \{F \geq f_\alpha\}$. Pour la colonne $\nu_1 = 10$ et la ligne $\nu_2 = 10$, on trouve $f_\alpha = 3.72$.

$$K_\alpha = \{F \geq 3.72\} .$$

Décision du test. La valeur expérimentale est

$$f_{\mathrm{e}} = \frac{\left(\widehat{s}_{1,\mathrm{e}}\right)^2}{\left(\widehat{s}_{2,\mathrm{e}}\right)^2} = \frac{5.947^2}{4.384^2} = 1.840 .$$

On voit que $f_{\mathrm{e}} < f_\alpha$. Donc $f_{\mathrm{e}} \notin K_\alpha$. On en déduit que, au niveau $\alpha = 0.05$, on accepte l'égalité des écart-types : $\sigma_1 = \sigma_2$.

B) TEST DE COMPARAISON DES MOYENNES.

On procède au test de comparaison de deux moyennes dans le cas de deux ***petits échantillons indépendants***, pour deux classes pour lesquelles on suppose $\sigma_1 = \sigma_2$ (grâce à l'exercice précédent).

Hypothèses. On désigne par μ_1 et par μ_2 les moyennes des scores des classes $\mathcal{C}_1$ et $\mathcal{C}_2$ pour la tâche $\mathcal{T}_1$.

$$H_0 \ : \mu_1 = \mu_2 \, ; \ H_1 \, : \, \mu_2 > \mu_1 .$$

Statistique du test. On a $n_1 = n_2 = 11$. Puisque on peut supposer $\sigma_1 = \sigma_2$, on a, sous l'hypothèse nulle H_0 :

$$\frac{M_{n_1} - M_{n_2}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \hookrightarrow \mathrm{St}\,(n_1 + n_2 - 2)\,,$$

où M_{n_1} et M_{n_2} sont les moyennes aléatoires des résultats des classe $\mathcal{C}_1$ et $\mathcal{C}_2$ à la tâche $\mathcal{T}_1$, et

$$s = \sqrt{\frac{n_1\,(s_{1,e})^2 + n_2\,(s_{2,e})^2}{n_1 + n_2 - 2}} = \sqrt{\frac{11 \times 5.947^2 + 11 \times 4.384^2}{20}} = 5.224 \text{ et } s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 5.224\sqrt{\frac{1}{11} + \frac{1}{11}} = 2.227\ .$$

Donc :

$$\frac{M_{n_1} - M_{n_2}}{2.227} \hookrightarrow \mathrm{St}\,(20)\ .$$

Région critique K_α au niveau $\alpha = 0.05$. L'hypothèse H_1 est vérifiée pour les valeurs $M_{n_1} < M_{n_2}$, c'est à dire $M_{n_1} - M_{n_2} < 0$. La région critique est donc de la forme $K_\alpha = \{T \leq t_\alpha\}$. La table de Student pour $\alpha = 0.05$ et 20 ddl donne le nombre 1.7247. Donc $t_\alpha = -1.7247$:

$$K_\alpha = \{T \leq -1.7247\}\,.$$

Décision du test. Un calcul montre que les moyennes expérimentales sont $m_{1,\mathrm{e}} = 126.182$ et $m_{2,\mathrm{e}} = 130.273$. Le résultat expérimental est donc :

$$t_\mathrm{e} = \frac{m_{1,\mathrm{e}} - m_{2,\mathrm{e}}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{126.182 - 130.273}{2.227} = -1.836\ .$$

On voit que $t_\mathrm{e} < t_\alpha$. Donc $t_\mathrm{e} \in K_\alpha$: au niveau $\alpha = 0.05$, on accepte l'hypothèse H_1, qui affirme que la classe $\mathcal{C}_2$ a en moyenne de meilleurs résultats que la classe $\mathcal{C}_1$.

2. Comparer les résultats des deux tests pour la classe $\mathcal{C}_2$. Puisqu'on compare les résultats pour ***la même classe*** $\mathcal{C}_2$, il s'agit cette fois d'un ***test de comparaison de deux échantillons appariés***. Ce test n'est pas orienté : on demande de comparer les moyennes, sans suggérer que l'une est plus grande que l'autre.

Fixons les notations. On désigne maintenant, comme il est suggéré dans l'énoncé, par X_1 la variable aléatoire qui représente les scores de la classe $\mathcal{C}_2$ à la tâche $\mathcal{T}_1$ et par X_2 la variable aléatoire qui représente les scores de la classe $\mathcal{C}_2$ à la tâche $\mathcal{T}_2$. On pose également $D = X_1 - X_2$. Enfin, on désigne par m_1 la moyenne des scores de la classe $\mathcal{C}_2$ à la tâche $\mathcal{T}_1$ et par m_2 la moyenne des scores de la classe $\mathcal{C}_2$ à la tâche $\mathcal{T}_2$.

Hypothèses. $H_0 \ : m_1 = m_2\,;\ H_1 \ : m_1 \neq m_1$.
Statistique du test. Puisque nous travaillons avec deux petits échantillons appariés, de taille $n = 11$, on a, sous l'hypothèse nulle H_0 :

$$T = \frac{M_n\,(D)}{S_n\,(D)\,/\sqrt{n-1}} \hookrightarrow \mathrm{St}\,(n-1) = \mathrm{St}\,(10),$$

où $S_n\,(D)$ et $M_n\,(D)$ désignent les écart-type et moyenne aléatoires de la variable D.
Région critique K_α au niveau $\alpha = 0.05$. La table de Student pour 10 ddl et $\alpha = 0.025$ (puisqu'on fait un test bilatéral) donne le nombre $t_\alpha = 2.2281$. Donc

$$K_\alpha = \{T \leq -2.2281\} \cup \{T \geq 2.2281\}\,.$$

Décision du test. Les valeurs expérimentales de la variable D sont $4, 2, -3, -1, -5, -10, 9, 9, 11, 6, 9$. Sa moyenne expérimentale est $m_\mathrm{e} = 2.8181$ et son écart-type expérimental est 6.5201. La valeur expérimentale du test est donc :

$$t_\mathrm{e} = \frac{m_\mathrm{e}}{s_\mathrm{e}/\sqrt{n-1}} = \frac{2.8181}{6.5201/\sqrt{10}} = 1.369\,.$$

Donc $t_\mathrm{e} \notin K_\alpha$. Au niveau $\alpha = 0.05$, on conserve l'hypothèse nulle, la classe $\mathcal{C}_2$ a les mêmes résultats moyens aux deux tests.