

Section : Psychologie - Licence 3^è année

Enseignant responsable : F.-G. Carpentier

**CORRECTION DE L'ÉPREUVE DE STATISTIQUES ET INFORMATIQUE
APPLIQUÉES À LA PSYCHOLOGIE**

N.B. Calculatrices, tables des lois statistiques et résumé de cours autorisés.

Exercice 1

Des chercheurs ont étudié l'effet d'une prise en charge psychologique chez des patients âgés victimes d'une chute. Les sujets ont notamment été évalués à l'aide de l'échelle MOS SF-36. Le MOS SF-36 (Medical outcome study short form 36 item health survey) est un questionnaire généraliste destiné à permettre le recueil du point de vue des patients sur leur propre état de santé dans divers domaines médicaux. C'est un reflet indirect du niveau de qualité de vie ressentie par les patients.

La sous-échelle BP de l'échelle SF-36 est relative aux douleurs physiques ressenties par les sujets. Cette sous-échelle a été étalonnée sur une population de référence, et le score de référence pour les personnes de 75 ans et plus est de 61.7.

L'échantillon étudié se composait de 15 sujets. Les paramètres descriptifs des scores observés sont les suivants :

- Avant rééducation, un score moyen de 48.5, avec un écart type corrigé de 25.8 ;
- Après rééducation, un score moyen de 53.7, avec un écart type corrigé de 21.4.

Les scores sur l'échelle BP mesurés avant et après rééducation sont-ils significativement différents de ceux de la population de référence ? Répondre à cette question à l'aide de tests unilatéraux, au seuil de 5%.

Soient μ_{Av} et μ_{Ap} les scores moyens avant et après rééducation dans la population d'où est tiré l'échantillon. Soient \bar{x}_{Av} et \bar{x}_{Ap} les moyennes des scores observés sur l'échantillon étudié.

Au niveau descriptif, on constate que les scores avant rééducation sont inférieurs au score de référence (61.7). Les hypothèses du test correspondant seront donc :

- $H_0 : \mu_{Av} = 61.7$
- $H_1 : \mu_{Av} < 61.7$.

La statistique de test est : $t = \frac{\bar{x}_{Av} - 61.7}{E}$ avec $E^2 = \frac{25.8^2}{15}$. Sous H_0 , cette statistique suit une loi de Student à 14 ddl. Au seuil unilatéral de 5%, on lit dans la table : $t_{lue} = 1.76$. La valeur critique est donc $t_{crit} = -1.76$ et la règle de décision est :

- Si $t_{obs} \geq -1.76$, on retient H_0 ;
- Si $t_{obs} < -1.76$, on retient H_1 .

Les calculs donnent : $E^2 = 44.376$; $E = 6.66$; $t_{obs} = \frac{48.5 - 61.7}{6.66} = -1.98$.

Comme $t_{obs} < -1.76$, l'hypothèse H_1 est retenue : avant rééducation, les sujets de la population étudiée ont un score significativement plus faible que le score de référence.

Pour le deuxième test, les hypothèses sont :

- $H_0 : \mu_{Ap} = 61.7$
- $H_1 : \mu_{Ap} < 61.7$.

La statistique de test est : $t = \frac{\bar{x}_{Ap} - 61.7}{E}$ avec $E^2 = \frac{21.4^2}{15}$. La règle de décision est la même que dans le test précédent.

Les calculs donnent : $E^2 = 30.53$; $E = 5.52$; $t_{obs} = \frac{53.7 - 61.7}{5.52} = -1.45$.

Comme $t_{obs} \geq -1.76$, l'hypothèse H_0 est retenue : après rééducation, les sujets de la population étudiée ont un score qui n'est pas significativement différent du score de référence.

Exercice 2

Les théories de la motivation sont nombreuses et plusieurs auteurs ont établi une distinction entre *motivation intrinsèque* et *motivation extrinsèque*.

Dans une étude de terrain réalisée en 2006, on demande à 25 experts économiques d'évaluer un même dossier de demande de prêt destiné à une création d'activité. Pour 13 experts, le dossier est accompagné d'une fiche décrivant le porteur du projet comme motivé intrinsèquement, pour les 12 autres, la fiche le décrit comme motivé extrinsèquement.

Les 25 experts devaient se prononcer sur l'attribution du prêt par un avis exprimé sur une échelle allant de 1 (très défavorable) à 9 (très favorable). Les résultats obtenus sont les suivants :

	Motivation intrinsèque		Motivation extrinsèque
s1	3	s14	1
s2	3	s15	1
s3	4	s16	3
s4	6	s17	4
s5	6	s18	4
s6	6	s19	5
s7	6	s20	5
s8	6	s21	5
s9	7	s22	6
s10	7	s23	7
s11	8	s24	7
s12	8	s25	8
s13	9		

1) Calculer la moyenne et l'écart type non corrigé de la variable "avis de l'expert" pour chacun des deux types de motivation.

Formons les tableaux permettant de calculer les paramètres descriptifs demandés :

	Motivation intrinsèque	x_i^2		Motivation extrinsèque	x_i^2
s1	3	9	s14	1	1
s2	3	9	s15	1	1
s3	4	16	s16	3	9
s4	6	36	s17	4	16
s5	6	36	s18	4	16
s6	6	36	s19	5	25
s7	6	36	s20	5	25
s8	6	36	s21	5	25
s9	7	49	s22	6	36
s10	7	49	s23	7	49
s11	8	64	s24	7	49
s12	8	64	s25	8	64
s13	9	81			
Tot.	79	521		56	316

Pour le groupe “motivation intrinsèque”, la moyenne vaut donc : $\bar{x}_I = \frac{79}{13} = 6.077$, la variance :

$$s_I^2 = \frac{521}{13} - \left(\frac{79}{13}\right)^2 = 3.148 \text{ et l'écart type d'échantillon : } s_I = 1.77.$$

Pour le groupe “motivation extrinsèque”, les paramètres sont respectivement : $\bar{x}_E = \frac{56}{12} = 4.667$,

$$s_E^2 = \frac{316}{12} - \left(\frac{56}{12}\right)^2 = 4.556 \text{ et } s_E = 2.134.$$

2) Faire un test bilatéral de comparaison de moyennes pour étudier si l'avis de l'expert est influencé par la motivation supposée du porteur de projet (seuil : 5%).

Soient μ_I et μ_E les moyennes de la variable “avis de l'expert” dans les populations parentes des deux groupes. Les hypothèses du test bilatéral demandé s'écrivent :

- $H_0 : \mu_I = \mu_E$
- $H_1 : \mu_I \neq \mu_E$.

La statistique de test est : $t = \frac{\bar{x}_I - \bar{x}_E}{E}$ avec $E^2 = \frac{n_I s_I^2 + n_E s_E^2}{n_I + n_E - 2} \left(\frac{1}{n_I} + \frac{1}{n_E}\right)$. Sous H_0 , cette statistique suit une loi de Student à 23 ddl. Pour un seuil $\alpha = 5\%$ bilatéral, la valeur critique est $t_{crit} = 2.07$.

La règle de décision est donc :

- Si $|t_{obs}| \leq 2.07$, on retient H_0 ;
- Si $|t_{obs}| > 2.07$, on retient H_1 .

Les calculs donnent : $E^2 = 0.666$; $E = 0.816$; $t_{obs} = \frac{6.077 - 4.667}{0.816} = 1.73$.

On conclut donc sur H_0 : on n'a pas mis en évidence de différence significative de l'avis de l'expert selon le type de motivation.

3) On a également demandé aux experts de donner sur une échelle de 1 à 9, un pronostic relatif à la réussite professionnelle du porteur de projet.

Une comparaison des moyennes des pronostics observés dans chacune des conditions a été faite à l'aide de Statistica. Elle conduit au résultat suivant :

Test t pour des Echantillons Indépendants (Feuille de données6)							
Note : Variables traitées comme des échantillons indépendants							
	Moyenne Groupe 1	Moyenne Groupe 2	valeur t	dl	p	N Actifs Groupe 1	N Actifs Groupe 2
Intrinsèque vs. Extrinsèque	5.7692	4.2500	2.464064	23	0.021639	13	12

Interpréter les résultats fournis par le logiciel.

On effectue une comparaison de moyennes analogue à la précédente, pour la variable “pronostic”. Statistica réalise un test bilatéral, et indique une p-value égale à 2.16%. Il existe donc une différence significative entre les deux groupes, au seuil traditionnel de 5%. En revanche, cette différence n’est pas significative au seuil de 1%. L’examen des moyennes nous montre en outre que le pronostic est meilleur pour le groupe “motivation intrinsèque”.

Exercice 3

Un chercheur s’est intéressé à l’articulation entre l’identification nationale et l’identification à l’Europe pour les habitants de l’Union Européenne.

Il a interrogé un échantillon composé de 182 étudiants belges francophones. Les participants ont rempli un questionnaire dans lequel l’identification nationale et l’identification européenne étaient mesurées par 8 items présentés sous forme d’échelles allant de 1 (désaccord total) à 11 (accord total).

Deux échelles d’identification, l’une nationale, l’autre européenne sont ensuite calculées en faisant les moyennes sur les 8 items concernés du questionnaire.

Les paramètres descriptifs des données observées sont les suivants :

	Moyenne	Ecart type corrigé
Identification nationale	5.70	2.04
Identification européenne	5.36	2.21
Différences individuelles	0.34	2.40

1) L’identification nationale diffère-t-elle de l’identification à l’Europe ? Préciser quel test statistique permet d’apporter une réponse à cette question.

L’étude porte ici sur deux groupes appariés et la variable considérée est numérique. On peut donc comparer l’identification nationale et l’identification à l’Europe à l’aide d’un test bilatéral de comparaison de moyennes sur deux groupes appariés.

2) Mettre en œuvre ce test et conclure au seuil de 5%.

Soient μ_N et μ_E les moyennes de la variable dépendante dans les deux conditions “Identification nationale” et “Identification européenne” pour la population parente. Les hypothèses du test bilatéral demandé s’écrivent :

- $H_0 : \mu_N = \mu_E$
- $H_1 : \mu_N \neq \mu_E$.

Soient \bar{x}_N et \bar{x}_E les moyennes observées sur l’échantillon dans les deux conditions et soit s_c l’écart type corrigé de la série des différences individuelles. La taille de l’échantillon est $n = 182$.

La statistique de test est : $Z = \frac{\bar{x}_N - \bar{x}_E}{E}$ avec $E^2 = \frac{s_c^2}{n}$. Sous H_0 , cette statistique suit une loi normale centrée réduite. Pour un seuil $\alpha = 5\%$ bilatéral, la valeur critique est $Z_{crit} = 1.96$.

La règle de décision est donc :

- Si $|Z_{obs}| \leq 1.96$, on retient H_0 ;
- Si $|Z_{obs}| > 1.96$, on retient H_1 .

Les calculs donnent : $E^2 = 0.0316$; $E = 0.1779$; $Z_{obs} = \frac{0.34}{0.1779} = 1.91$.

On conclut donc sur H_0 : on n'a pas mis en évidence de différence significative entre le score d'identification nationale et celui d'identification à l'Europe.

Remarques. Si on avait choisi de faire un test unilatéral, avec comme hypothèse alternative, un score d'identification à l'Europe inférieur au score d'identification nationale, on aurait au contraire conclu sur H_1 . Par ailleurs, on notera que la conclusion peut également être inversée si les calculs ne sont pas assez précis, par exemple si on prend $E = 0.17$.

Exercice 4

On s'intéresse aux représentations de la délinquance chez l'enfant en abordant les quatre dimensions suivantes ainsi que leurs liens :

- 1) L'évaluation des différents types de délinquances ;
- 2) Les origines des actes délinquants ;
- 3) Les sanctions jugées justes ;
- 4) Les raisons des sanctions.

L'échantillon interrogé est composé de 127 enfants âgés de 5 à 10 ans normalement scolarisés en écoles maternelles et primaires (20 garçons et 18 filles de 5 ans, 26 garçons et 21 filles de 8 ans, 24 garçons et 18 filles de 10 ans), et 83 étudiants inscrits en 1ère année de Biologie (33 hommes et 50 femmes).

Les sujets répondaient à un questionnaire dans lequel ils devaient notamment indiquer, par ordre d'importance, les causes qu'ils estimaient être à l'origine d'un acte de délinquance commis par un enfant. Les réponses proposées pouvaient se regrouper en trois types de causes : le caractère de l'enfant, le milieu social, et la façon dont les parents s'occupaient de leur enfant.

Le croisement de la réponse classée en première position et de la classe d'âge du sujet conduit au résultat suivant :

	Caractère	Milieu social	Parents
5 ans	14	12	12
8 ans	36	9	2
10 ans	9	17	16
Etudiants	5	22	56

1) Réaliser un test statistique permettant de déterminer si le choix de l'une ou l'autre cause dépend de l'âge du sujet (seuil : 5%).

Les variables étudiées sont deux variables nominales, et nous recherchons s'il existe un lien entre ces deux variables, autrement dit si elles sont dépendantes. Le test qui convient est donc le test du χ^2 sur un tableau de contingence.

Les hypothèses sont ici :

- H_0 : Les variables "Type de cause" et "Classe d'âge" sont indépendantes.
- H_1 : Les variables "Type de cause" et "Classe d'âge" sont dépendantes.

Le nombre de degrés de liberté est ici de $(3 - 1) \times (4 - 1) = 6$. Pour un seuil de 5%, on lit dans la table : $\chi_{crit}^2 = 12.59$. La règle de décision est donc la suivante :

- Si $\chi_{obs}^2 \leq 12.59$, on retient H_0 ;
- Si $\chi_{obs}^2 > 12.59$, on retient H_1 .

Les tableaux des effectifs théoriques et des contributions au χ^2 sont donnés par :

	Caractère	Milieu social	Parents
5 ans	11.58	10.86	15.56
8 ans	14.32	13.43	19.25
10 ans	12.80	12.00	17.20
Etudiants	25.30	23.71	33.99

	Caractère	Milieu social	Parents
5 ans	0.51	0.12	0.81
8 ans	32.82	1.46	15.46
10 ans	1.13	2.08	0.08
Etudiants	16.29	0.12	14.25

On obtient ainsi : $\chi_{obs}^2 = 85.14$. On conclut donc sur H_1 : les types de causes invoquées pour expliquer la délinquance dépendent de l'âge du sujet interrogé.

2) On se limite aux classes d'âge 5 ans et 10 ans (les sujets de 8 ans et les étudiants sont donc exclus de l'étude) et on réalise le test précédent à l'aide d'un logiciel de traitement statistique. On obtient 2.3263 comme valeur de la statistique de test et un niveau de significativité $p = .3125$. Que peut-on conclure au seuil de 5% ?

Lorsqu'on ne considère que les deux classes d'âge 5 ans et 10 ans, le niveau de significativité du test du χ^2 est de 31.25%. Comme cette valeur est largement supérieure au seuil de 5%, on conclut alors sur H_0 : dans ces deux classes d'âge, les types de causes invoquées pour expliquer la délinquance se répartissent de la même façon.