Université de Bretagne Occidentale
U.F.R. de Lettres et Sciences Sociales
CS 93837 - 29238 BREST CEDEX 3

Année Universitaire 2003-04

Section : Psychologie - Licence

Enseignant responsable : F.G. Carpentier

Corrigé de l'épreuve partielle et ponctuelle de Statistiques et Informatique appliquées à la Psychologie

Exercice 1

Pour réaliser une expérience, on doit constituer deux groupes de sujets (un groupe expérimental et un groupe témoin). La méthode de recrutement et de sélection des sujets est la suivante :

- Tout d'abord, 60 sujets sont sélectionnés.
- Chacun des 60 sujets est ensuite affecté au hasard (à l'aide d'une pièce de monnaie) dans l'un des deux groupes.

Les deux groupes ainsi constitués ne seront vraisemblablement pas équilibrés (c'est-à-dire, ne seront de même effectif). On se pose le problème d'évaluer le minimum et le maximum de la taille du groupe expérimental que ce tirage au hasard risque de produire.

Soit X une variable statistique dichotomique définie sur une population supposée très nombreuse. La fréquence p de la modalité "succès" est 50%. On tire au hasard un échantillon de taille $n = 60$ dans cette population. Soit F la variable "fréquence de la modalité "succès" sur l'échantillon tiré".
1) Quelles sont les caractéristiques de la distribution de la variable F (moyenne $\overline{F}$, écart type S, loi suivie) ?

D'après le résultat relatif à l'échantillonnage, la variable F a pour caractéristiques :

— distribution suivant une loi normale
— moyenne $\overline{F} = 0.5$
— écart type S vérifiant : $S^2 = \frac{p(1-p)}{n} = \frac{0.25}{60}$, c'est-à-dire $S = 0.0645$.

2) Donner un intervalle centré autour de $\overline{F}$ rassemblant 95% des valeurs observées de la variable F.

Soit Z la variable normale centrée réduite associée à F. On a : $F = SZ + \overline{F} = 0.0645 \times Z + 0.5$.
D'après la table de la loi normale, $P(-z_\alpha \leq Z < z_\alpha) = 0.95$ pour $z_\alpha = 1.96$.
Pour $Z = -1.96$, on a : $F = 0.0645 \times (-1.96) + 0.5 = 0.37$.
Pour $Z = 1.96$, on a : $F = 0.0645 \times 1.96 + 0.5 = 0.63$.
L'intervalle recherché est donc : $[0.37, 0.63]$.
3) Quels sont les effectifs de la modalité "succès" correspondant aux deux bornes de l'intervalle déterminé dans la question précédente.

Rapporté à un effectif de 60 sujets, l'intervalle précédent devient : $[22, 38]$.

Exercice 2

Dans le cadre d'un programme visant à réduire le tabagisme, une organisation a mené une campagne publicitaire pour convaincre les gens d'arrêter de fumer ou de réduire leur consommation de tabac. Afin d'évaluer l'efficacité de leur campagne, les responsables ont demandé à 12 sujets de noter le nombre moyen de cigarettes fumées par jour durant la semaine précédant et la semaine qui a suivi la campagne. Les données sont les suivantes :

Sujet	1	2	3	4	5	6	7	8	9	10	11	12
Avant	45	16	20	33	30	19	33	25	26	40	28	36
Après	43	20	17	30	25	19	34	28	23	41	26	40

La campagne a-t-elle été efficace ? Répondre à cette question à l'aide d'un test unilatéral, avec un seuil de 5%.

Nous répondrons à la question posée à l'aide d'un test d'égalité de deux moyennes sur des groupes appareillés.
Soient μ_1 et μ_2 le nombre moyen de cigarettes fumées dans la population dont est issu l'échantillon (la population des fumeurs ?) respectivement avant et après la campagne publicitaire. L'efficacité de la campagne devrait se traduire par une diminution de ce nombre moyen. Les hypothèses sont donc ici :
- $H_0 : \mu_1 = \mu_2$
- $H_1 : \mu_1 > \mu_2$

La statistique de test est ici $T = \frac{\overline{d}}{E}$, avec $E^2 = \frac{s_c^2}{n}$, où $\overline{d}$ est la moyenne des différences individuelles (qui seront calculées ici dans le sens $x_{\text{Avant}} - x_{\text{Après}}$), s_c^2 est la variance corrigée de la série de ces différences et n est la taille de l'échantillon.
Cette statistique suit une loi de Student à 11 *ddl*. Pour un seuil de 5%, on obtient : $t_{crit} = 1.7959$.
La règle de décision est donc :
- Si $t_{obs} > 1.7959$ on retient l'hypothèse H_1
- Si $t_{obs} \leq 1.7959$ on retient l'hypothèse H_0

Calcul de la valeur observée de la statistique :

Sujet	1	2	3	4	5	6	7	8	9	10	11	12	Total
d_i	2	-4	3	3	5	0	-1	-3	3	-1	2	-4	5
d_i^2	4	16	9	9	25	0	1	9	9	1	4	16	103

$\overline{d} = \frac{5}{12} = 0.42 \quad ; \quad s^2 = \frac{103}{12} - \frac{25}{144} = 8.4097 \quad ; \quad s_c^2 = \frac{12}{11} 8.4097 = 9.1742$
$E^2 = \frac{9.1742}{12} = 0.7645 \quad ; \quad E = 0.87 \quad ; \quad t_{obs} = \frac{0.42}{0.87} = 0.48.$
D'après la règle de décision indiquée ci-dessus, on retient donc l'hypothèse H_0 : la campagne publicitaire n'a pas eu d'effet sur la consommation de cigarettes dans la population considérée.

Exercice 3

Des chercheurs ont réalisé une expérience visant à mettre en évidence l'effet d'une séance d'intervention motivante brève sur le comportement relatif à la consommation d'alcool.
Soixante sujets, qui ont déclaré avoir bu occasionnellement à 2 reprises ou plus au cours du mois précédant l'expérience ont été affectés au hasard soit dans un groupe contrôle, sans traitement (31 sujets) soit dans un groupe expérimental dit "groupe d'intervention brève" (29 sujets).
Le comportement des sujets est mesuré par la variable "nombre de verres bus par semaine".
Les sujets sont évalués avant l'expérience (condition de référence). Chacun des sujets du groupe d'intervention brève bénéficie d'un entretien personnalisé relatif aux problèmes liés à l'alcool.
Six semaines après l'entretien, l'ensemble des sujets est de nouveau évalué.
1) La répartition par sexe des sujets dans les deux groupes est donnée par :

	H	F
Groupe contrôle	14	17
Groupe expérimental	12	17

Etudier à l'aide d'un test approprié s'il existe une différence significative entre les deux groupes de ce point de vue.

Pour étudier s'il existe une différence significative de la répartition par sexe dans les deux groupes, plusieurs tests conviennent. On peut réaliser un test du χ^2 sur le tableau de contingence ci-dessus. On peut également faire un test d'égalité de deux proportions sur deux groupes indépendants. Par exemple, on peut comparer la proportion de H dans le groupe contrôle (14/31) à la proportion de H dans le groupe expérimental (12/29). On donne ici la solution utilisant le test du χ^2.
H_0 : Les variables sexe et affectation à l'un des groupes sont indépendantes. Autrement dit, chacun des sexes est représenté de la même façon dans les populations dont est issu chacun des groupes.
H_1 : Les variables sexe et affectation à l'un des groupes sont dépendantes.
Nous utilisons ici une statistique du χ^2, avec 1 *ddl*, et un seuil de 5%. Par lecture de la table, on obtient : $\chi^2_{crit} = 3.84$. La règle de décision est la suivante :
— Si $\chi^2_{obs} > 3.84$ on retient l'hypothèse H_1
— Si $\chi^2_{obs} \leq 3.84$ on retient l'hypothèse H_0
Les marges du tableau d'effectifs observés, le tableau d'effectifs théoriques et celui des contributions au χ^2 sont donnés par :

	H	F	Tot.
Gr. ctr.	14	17	31
Gr. exp.	12	17	29
Tot.	26	34	60

	H	F
Gr. ctr.	13.43	17.57
Gr. exp.	12.57	16.43

	H	F
Gr. ctr.	0.0242	0.0185
Gr. exp.	0.0258	0.0198

On obtient donc $\chi^2_{obs} = 0.0883$, et c'est donc l'hypothèse H_0 qui est retenue : les deux groupes ne présentent pas de différence significative du point de vue de la répartition par sexe.

2) Le tableau 1 ci-dessous est un extrait du tableau donnant le nombre de verres bus par semaine dans la condition de référence. Dans ce tableau, les notations n, $\overline{x}$ et s désignent respectivement l'effectif, la moyenne et l'écart type (non corrigé) dans chacun des deux échantillons.

Condition de référence		
	Int. brève	Contrôle
	8	6
	………	
	………	
	12	22
	22	5
		10
		2
n	29	31
$\overline{x}$	17.79	18.81
s	8.33	12.27

TABLE 1 – Condition de référence

Existe-t-il une différence significative de comportement des deux groupes dans cette condition ? Répondre à cette question à l'aide d'un test statistique bilatéral, avec un seuil fixé à 5%.

Nous répondrons à la question posée à l'aide d'un test d'égalité de deux moyennes sur des groupes indépendants.
Soient μ_{IB} et μ_C les moyennes de la variable dépendante dans les deux populations parentes. Les hypothèses sont ici :

— H_0 : $\mu_{IB} = \mu_C$
— H_1 : $\mu_{IB} \neq \mu_C$

La statistique de test est ici $T = \dfrac{\overline{x}_{IB} - \overline{x}_C}{E}$, avec $E^2 = \dfrac{n_{IB}s_{IB}^2 + n_C s_C^2}{n_{IB} + n_C - 2}\left(\dfrac{1}{n_{IB}} + \dfrac{1}{n_C}\right)$, où $\overline{x}_{IB}$, $\overline{x}_C$, s_{IB}^2, s_C^2, n_{IB} et n_C sont respectivement les moyennes, les écarts types de la variable dépendante et les effectifs de chacun des deux groupes.
Cette statistique suit une loi de Student à $n_{IB} + n_C - 2 = 58$ ddl. Pour un seuil de 5%, on obtient : $t_{crit} = 20017$. La règle de décision est donc :

— Si $|t_{obs}| > 2.0017$ on retient l'hypothèse H_1
— Si $|t_{obs}| \leq 2.0017$ on retient l'hypothèse H_0

Les calculs donnent ici :
$E^2 = \dfrac{29 \times 8.33^2 + 31 \times 12.27^2}{58}\left(\dfrac{1}{29} + \dfrac{1}{31}\right) = 7.6860 \quad ; \quad E = 2.77$
$t_{obs} = \dfrac{17.79 - 18.81}{2.77} = -0.3679$
Ici, $|t_{obs}| = 0.3679$ et donc $|t_{obs}| \leq 2.0017$. L'hypothèse H_0 est donc ici retenue : en condition de référence, il n'existe pas de différence significative entre les deux groupes.

3) Les valeurs de la variable observées lors de l'étude de suivi ont été traitées à l'aide de l'utilitaire d'analyse d'Excel. Les résultats produits par Excel sont indiqués dans le tableau 2.

	Intervention brève	Contrôle
Moyenne	11.72	15,77
Variance	50.3498	61.2473
Observations	29	31
Variance pondérée	55.99	
Différence hypothétique des moyennes	0.00	
Degré de liberté	58	
Statistique t	-2.10	
P(T<=t) unilatéral	2.03%	
Valeur critique de t (unilatéral)	1.67	
P(T<=t) bilatéral	4.05%	
Valeur critique de t (bilatéral)	2.00	

Table 2 – Test d'égalité des espérances : deux observations de variances égales

Utiliser ces résultats pour réaliser un test statistique unilatéral au seuil de 5 %.

Les données observées lors de l'étude de suivi ont été traitées à l'aide d'un test d'égalité de deux moyennes sur des groupes indépendants, analogue à celui réalisé à la question précédente. Du tableau Excel, on retiendra essentiellement les éléments suivants : $t_{obs} = -2.10$, et le niveau de significativité de cette valeur, pour un test unilatéral est de 2.03%.
Au seuil de 5%, on retient donc l'hypothèse H_1 : il existe une différence significative de comportement des deux groupes lors de l'étude de suivi, la variable dépendante prenant des valeurs significativement moins élevées dans le groupe "Intervention brève".

4) Quelle conclusion les auteurs peuvent-ils tirer de leur expérience ?

Les deux premiers tests montrent que l'équivalence des deux groupes de sujets, d'une part du point de vue de la composition par sexe, d'autre part du point de vue du comportement initial par rapport à l'alcool.
Le troisième test montre que le groupe qui a bénéficié de l'intervention brève a, six semaines après cette intervention, un comportement différent de celui du groupe contrôle.
L'intervention brève semble donc avoir un effet sur le comportement des sujets relativement à l'alcool.