

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

Dans une étude publiée en 1995, J. Westaby examine l'influence de la culture (japonaise v/s américaine) et des conditions de la tâche (travail seul v/s en présence d'un groupe) sur la productivité et la qualité du travail.

99 sujets féminins ont participé à l'expérience : 46 sujets japonais (30 travaillant seuls, 16 travaillant en présence d'un groupe) et 53 sujets américains (33 travaillant seuls et 20 en présence d'un groupe).

Les sujets devaient accomplir une tâche de tracé de cercles à main levée. On mesurait d'une part la quantité de travail fourni (nombre de figures tracées en un temps donné) et d'autre part la qualité du travail fourni (nombre moyen d'erreurs).

1) Le tableau 1 ci-dessous est un extrait du tableau donnant la quantité de travail fourni par les sujets américains et japonais travaillant seuls. Dans ce tableau, les notations n , \bar{x} et s_c désignent respectivement l'effectif, la moyenne et l'écart type corrigé dans chacun des deux échantillons. Etudier l'influence de la culture à l'aide d'un test de comparaison de moyennes (test bilatéral au seuil de 5%).

Il s'agit ici d'un test de comparaison de moyennes sur deux groupes indépendants, avec des échantillons de taille supérieure ou égale à 30. Notons μ_1 et μ_2 les moyennes sur les deux populations parentes.

Hypothèse nulle $H_0 : \mu_1 = \mu_2$

Hypothèse alternative $H_1 : \mu_1 \neq \mu_2$ (conformément aux indications de l'énoncé, on fait ici un test bilatéral).

La statistique de test, $Z = \frac{\bar{x}_1 - \bar{x}_2}{E}$, avec $E^2 = \frac{s_{c,1}^2}{n_1} + \frac{s_{c,2}^2}{n_2}$ suit une loi normale centrée réduite.

Pour un seuil de 5%, la valeur critique déduite de la table est $z_c = 1.96$. La règle de décision est donc :

– si $-1.96 \leq z_{obs} \leq 1.96$, on retient H_0 .

– si $z_{obs} < -1.96$ ou $z_{obs} > 1.96$, on rejette H_0 et on retient H_1 .

Avec les données observées, $E^2 = \frac{34.48^2}{30} + \frac{33.39^2}{33} = 73.41$ d'où $E = 8.57$ et $z_{obs} = \frac{214.07 - 189.18}{8.57} = 2.90$.

On retient donc l'hypothèse H_1 : on a montré que la culture a une influence sur la quantité de travail fourni.

2) On soumet à l'utilitaire d'analyse d'Excel (méthode "test d'égalité des espérances : deux observations de variances égales"), les scores (quantité de travail fourni) des deux groupes de sujets américains. En spécifiant un seuil de 1%, on obtient le résultat donné dans le tableau 2.

a) Justifier le calcul du nombre de degrés de liberté et celui des valeurs critiques de la statistique t .

Il s'agit ici d'un test de comparaison de moyennes sur deux groupes indépendants. L'un des groupes ayant un effectif inférieur à 30, on va utiliser une statistique distribuée selon la loi T de

Student. Le nombre de degrés de liberté à prendre en compte est : $n_1 + n_2 - 2 = 33 + 20 - 2 = 51$, conformément aux indications d'Excel. La consultation des tables de cette loi, pour 51 ddl (ou plus exactement pour 50 et 52 ddl) et $\alpha = 1\%$ nous conduit à des valeurs critiques très voisines de celles fournies par Excel, aussi bien pour un test unilatéral ($T_c = 2.40$) que pour un test bilatéral ($T_c = 2.68$).

b) Quelle interprétation peut-on donner des pourcentages (7.30% et 14.59%) indiqués?

Ces deux pourcentages représentent les niveaux de significativité de la valeur observée de T ($t_{obs} = -1.48$), pour un test unilatéral (NivSig= 7.30%) et pour un test bilatéral (NivSig= 14.59%). Ainsi, si l'hypothèse H_0 d'homogénéité des groupes est vérifiée, la probabilité de tirer des échantillons conduisant à $T \leq -1.48$ est égale à 7.30%, celle de tirer des échantillons conduisant à $T \leq -1.48$ ou $T \geq 1.48$ est égale à 14.59%. Notez l'écriture incorrecte d'Excel dans ce cas. Une notation correcte aurait pu être : $P(|T| \geq 1.48) = 14.59\%$.

c) Quelle conclusion peut-on tirer du test effectué?

Que ce soit en raisonnant à partir des niveaux de significativité (tous deux supérieurs à 1%) ou en comparant t_{obs} aux valeurs critiques, on conclut sur H_0 : on n'a pas mis en évidence de différence de performance entre les Américains travaillant seuls et les Américains travaillant en groupe.

3) Le tableau 3 donne le nombre moyen d'erreurs produites pour les sujets des deux cultures travaillant en groupe. Utiliser un test de la médiane pour déterminer si les scores des deux groupes sont significativement différents.

Après classement des 36 observations par ordre croissant, on observe en 18^e et 19^e positions des scores de 0.62 et 0.64. La médiane de l'ensemble des observations vaut donc $M = 0.63$.

On obtient donc le tableau de contingence et le tableau d'effectifs théoriques suivants :

	Japonais	Américains		Japonais	Américains
$\leq M$	10	8	$\leq M$	8	10
$> M$	6	12	$> M$	8	10

D'où $\chi_{obs}^2 = 2 \times 0.5 + 2 \times 0.4 = 1.8$. On a ici $ddl = 1$ et donc, au seuil de 5%, $\chi_{crit}^2 = 3.84$. Comme $\chi_{obs}^2 < \chi_{crit}^2$, on conclut à l'indépendance des variables "position par rapport à la médiane" et "culture". Autrement dit, on ne peut pas refuser l'hypothèse H_0 d'égalité des médianes dans les populations parentes.

Tâche individuelle		
	Japonais	Américains
	170	214

	251	200
	213	156
		143
		228
		189
n	30	33
\bar{x}	214,07	189,18
s_c	34,48	33,29

Tableau 1: Quantité de travail fourni

Test d'égalité des espérances: deux observations de variances égales		
	Am-Indiv	Am-Gr
Moyenne	189.18	202,75
Variance	1108.28	955,57
Observations	33	20
Variance pondérée	1051.39	
Différence hypothétique des moyennes	0	
Degré de liberté	51	
Statistique t	-1.48	
P(T<=t) unilatéral	7.30%	
Valeur critique de t (unilatéral)	2.40	
P(T<=t) bilatéral	14.59%	
Valeur critique de t (bilatéral)	2.68	

Tableau 2: Quantité de travail (Américains seuls v/s en groupe)

Exercice 2

Dans une étude sur la relation entre les faits additifs et les faits multiplicatifs, on présente aux sujets une série de problèmes arithmétiques simples sur l'écran d'un ordinateur. Les sujets ont pour consigne de dire si oui ou non, le résultat présenté est correct.

Les problèmes sont de type additif ou multiplicatif et le résultat présenté peut être "vrai" (ex: $4+3=7$), "faux" (ex: $4+3=9$) ou "confusion" (ex: $4+3=12$).

Pour cette expérience, 26 sujets sont utilisés. Leur comportement est mesuré par le temps de réaction (TR) devant chaque problème. Le tableau 4 donne le temps de réaction médian (en millisecondes) des sujets pour les problèmes de type "additif-vrai" et pour les problèmes de type "additif-faux". Le tableau 5 donne les paramètres descriptifs des séries de données observées.

1) A l'aide d'un test de comparaisons de moyennes, étudier s'il existe une différence significative des temps de réaction selon le type ("vrai" ou "faux") de réponse proposée.

Il s'agit ici d'un test de comparaison de moyennes sur deux groupes appariés, avec un échantillon de taille inférieure à 30. Notons μ_1 et μ_2 les moyennes dans les deux conditions.

Hypothèse nulle $H_0: \mu_1 = \mu_2$

Hypothèse alternative $H_1: \mu_1 > \mu_2$ (l'énoncé ne nous précise pas le type de test à effectuer ; on peut choisir, par exemple, de faire un test unilatéral).

La statistique de test, $T = \frac{\bar{d}}{E}$, avec $E^2 = \frac{s_c^2}{n}$ suit une loi de Student avec 25 ddl.

Pour un seuil de 5%, la valeur critique déduite de la table est $t_c = 1.70$. La règle de décision est donc :

– si $t_{obs} \leq 1.70$, on retient H_0 .

Japonais				Américains			
0.20	0.23	0.26	0.27	0.18	0.18	0.32	0.32
0.27	0.28	0.34	0.35	0.44	0.44	0.58	0.62
0.48	0.48	0.64	0.75	0.82	0.89	1.03	1.19
0.86	0.99	0.99	1.12	1.20	1.24	1.25	1.27
				1.37	1.45	1.57	1.64

Tableau 3: Qualité de travail fourni

	Add-Vrai	Add-Faux	Diff		Add-Vrai	Add-Faux	Diff
s1	1452	1462	-10	s14	1285	1164	121
s2	1294	1240	54	s15	1331	1193	138
s3	1225	1334	-109	s16	1396	1248	148
s4	1315	1172	143	s17	1533	1483	50
s5	1498	1562	-64	s18	1361	1306	55
s6	1306	1231	75	s19	1366	1365	1
s7	1401	1430	-29	s20	1383	1245	138
s8	1343	1400	-57	s21	1320	1107	213
s9	1201	919	282	s22	1289	1348	-59
s10	1199	1124	75	s23	1214	1140	74
s11	1580	1391	189	s24	1274	1131	143
s12	1345	1053	292	s25	1410	1410	0
s13	1542	1624	-82	s26	1508	1325	183

Tableau 4: Temps de réaction

	Add-Vrai	Add-Faux	Diff
Moyenne	1360.42	1284.88	75.54
Ecart type	105.24	160.77	107.79
Ecart type corrigé	107.32	163.95	109.92

Tableau 5: Paramètres des séries “temps de réaction”

– si $t_{obs} > 1.70$, on rejette H_0 et on retient H_1 .

Avec les données observées, $E^2 = \frac{109.92^2}{26} = 464.71$ d'où $E = 21.56$ et $t_{obs} = \frac{75.54}{21.56} = 3.50$.

On retient donc l'hypothèse H_1 : on a montré que le temps de réaction dépend du type de réponse proposée.

2) Reprendre la question précédente à l'aide d'un test du signe.

Soit H_0 l'hypothèse: “dans la population parente, la fréquence des différences négatives est de 50%” et H_1 l'hypothèse: “ la fréquence des différences négatives est inférieure à 50%”.

On observe sur l'échantillon 7 différences négatives, 18 différences positives, soit 25 différences non nulles.

Sous H_0 , le nombre de différences négatives suit une loi binomiale de paramètres $N = 25$ et $p = 0.5$.

En utilisant l'approximation par une loi normale, on obtient $z_{obs} = \frac{2 \times 7 - 1 - 25}{\sqrt{25}} = 2$. Or, pour un test unilatéral au seuil de 5%, $z_{crit} = 1.64$. On retient donc l'hypothèse H_1 .

Si l'on juge l'effectif trop faible pour utiliser l'approximation par une loi normale, on peut utiliser la loi binomiale de paramètres $N = 25$ et $p = 0.5$ et raisonner en termes de niveau de significativité. On a:

$p(0) = 2.9810^{-8}$, $p(1) = 7.4510^{-7}$, $p(2) = 8.9410^{-6}$, $p(3) = 6.8510^{-5}$, $p(4) = 3.7710^{-4}$, $p(5) = 1.5810^{-3}$, $p(6) = 5.2810^{-3}$, $p(7) = 1.4310^{-2}$ et donc $P(X \leq 7) = 2.16 \times 10^{-2}$.

Ce niveau de significativité étant inférieur à 5%, l'hypothèse H_1 est retenue au seuil de 5%.

Exercice 3

Le centre national des statistiques sur l'éducation a indiqué que 47% des étudiants (américains) travaillent pour payer leurs études. Supposons qu'un échantillon de 450 étudiants ait été utilisé

pour cette étude.

Déterminer un intervalle de confiance avec un degré de confiance de 95% pour la proportion de la population des étudiants qui travaillent pour payer leurs études.

On a ici : $f = 0.47$. L'erreur type est donnée par : $E = \sqrt{\frac{f(1-f)}{n}} = 0.02353$.

Pour un degré de confiance de 95%, la valeur z_α lue dans la table de la loi normale centrée réduite est $z_\alpha = 1.96$. D'où l'intervalle de confiance :

$$0.47 - 1.96 \times 0.02353 \leq p \leq 0.47 + 1.96 \times 0.02353$$

c'est-à-dire : [42.38%; 51.61%].

Exercice 4

1) Dans les systèmes informatiques actuels, l'interface-utilisateur fait usage de nombreuses représentations icônes. A titre d'exemples, on a représenté ci-dessous dix icônes d'applications ou de documents.

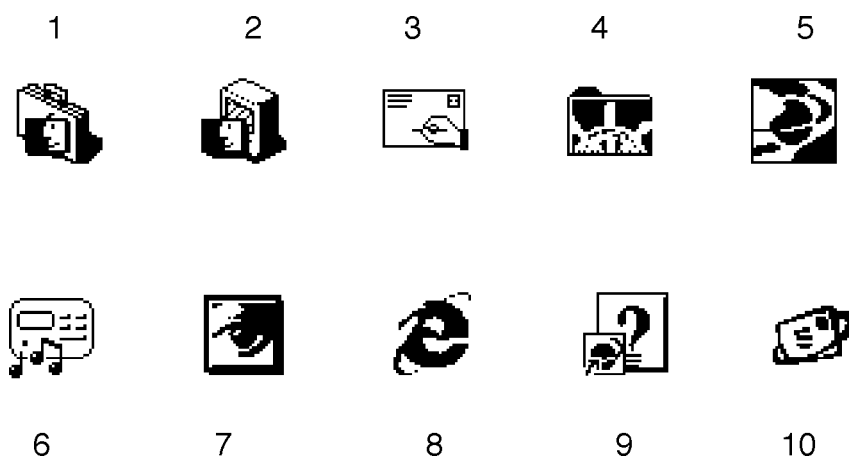
a) Donner quelques-unes des raisons pour lesquelles les concepteurs font usage de ce type de représentations.

Parmi les buts poursuivis par les concepteurs de ces représentations icônes, on peut notamment citer :

- la suggestion de métaphores par des icônes symbolisant des objets du monde réel ;
- faciliter la manipulation directe ; par exemple, poser l'icône d'un document sur celle d'une imprimante ;
- rendre "visible" ce qui est faisable sur le poste de travail ;
- limiter la charge mentale de l'utilisateur ;
- permettre des raccourcis dans l'exécution des commandes.

On pourrait également ajouter l'aspect "international" (les icônes seront les mêmes, quelle que soit la langue de la version du logiciel utilisée) et l'aspect "appropriation" de l'outil par la possibilité de personnalisation de certaines icônes.

b) Les icônes représentées ci-dessus sollicitent-elles toutes l'intuition de l'utilisateur de la même façon ?



Certaines de ces icônes s'appuient de manière évidente sur des métaphores du monde réel (N° 6 : écoute audio, N° 3 : courrier – en fait électronique –, N° 10 : courrier électronique, le "réseau" internet étant symbolisé par un anneau).

Dans d'autres cas, les modèles utilisés sont moins explicites. La barre à roue de l'icône N° 4 fait référence à un navigateur. Mais ce terme est lui-même une métaphore pour désigner un

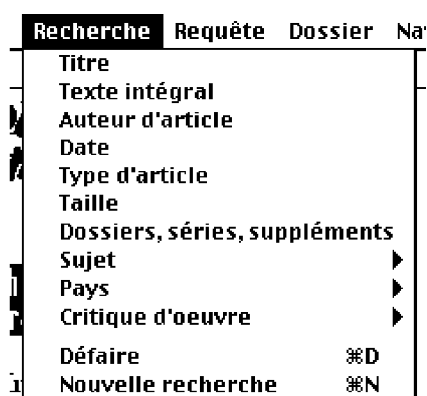
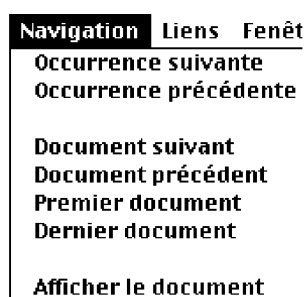
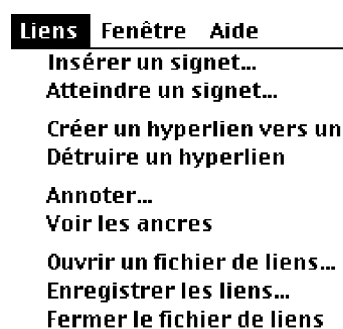
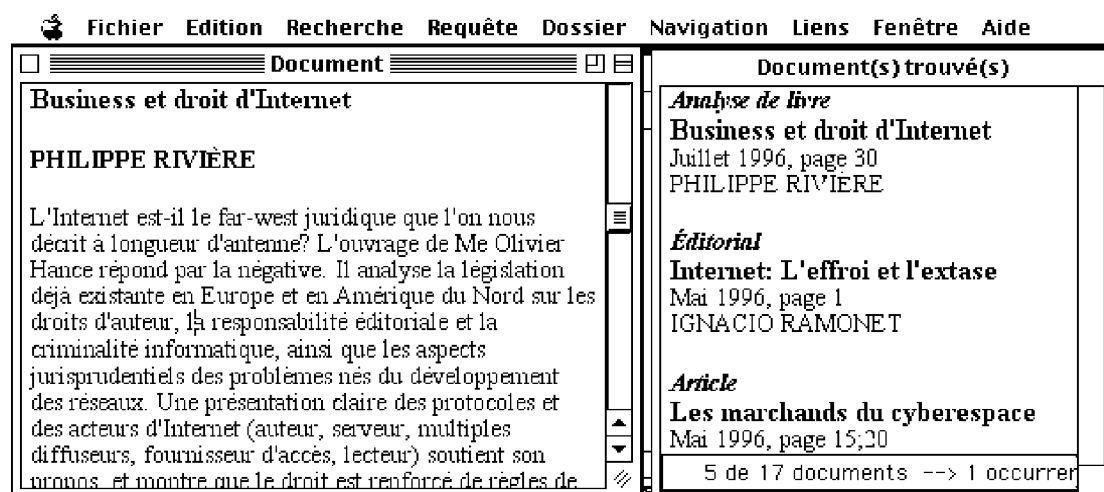
logiciel servant à la consultation de pages Web. Une icône telle que le N° 9 superpose deux représentations : le point d'interrogation, symbole de l'aide et une mini-icône rappelant l'icône N° 5 : il s'agit de l'aide concernant le logiciel désigné par le N° 5.

Les icônes N° 1, 2 et 8 font appel à des modèles assez différents de ceux décrits précédemment. La lettre "e" de l'icône N° 8 réfère à l'Explorer de Microsoft. Ce symbolisme est compris dans la mesure où il fait l'objet d'une large médiatisation. De la même façon, les icônes N° 1 et 2 sont construites autour d'un logo très utilisé par Apple, mais leur compréhension dépend étroitement de la connaissance que l'on peut avoir de cette marque.

Ainsi, ces icônes font toutes référence à un modèle, mais ces modèles sont variés : métaphores simples du monde réel, symbolismes plus ou moins fixés par l'usage, logos choisis par les marques, etc.

2) Le "Monde Diplomatique" publie régulièrement des archives d'anciens numéros sur CD-ROM. On a représenté ci-dessous une copie de l'écran et de quelques menus du logiciel de navigation du CD-ROM couvrant la période 1989-1996.

Parmi les fonctionnalités généralement rencontrées dans les logiciels de navigation, quelles sont celles qui semblent être présentes dans le logiciel utilisé ?



On retrouve dans ce logiciel pratiquement toutes les fonctions rencontrées dans les logiciels de navigation : liens hypertextes, table de matières et index (plusieurs) réactifs, historique de la navigation, annotations et signets et des fonctions de recherche très étendues (par titre, auteur, genre, et en texte intégral).

On notera particulièrement :

- la richesse des fonctions de recherche ;
- la possibilité pour l'utilisateur de construire son propre réseau de liens (créer/détruire un hyperlien) ;
- la présence simultanée d'une structure hiérarchique (table des matières, index) et d'une structure réseau.

En revanche, la navigation s'effectue uniquement en mode texte. On ne retrouve pas ici l'aspect multimodal propre aux documents multimédias.