

Durée de l'épreuve : 1 heure 30 mn.
Aucun document n'est autorisé. Seule la calculette (sans sa documentation) est autorisée.
Les différents exercices (encadrés) sont indépendants.
Le barème donné à titre indicatif est sur 60 ; la note finale sera donnée sur 20.
La dernière feuille (présentation du dossier PSYCHO, formules et tables) peut être détachée et conservée.
Indiquer les réponses exclusivement sur ce document. Ne rien écrire dans la marge gauche.

Type des variables (8 points)

Voir la présentation du dossier PSYCHO sur la dernière feuille.

1/ Indiquer le type de chacune des variables du dossier PSYCHO,
(Nominale / Ordinale / Numérique) :

- Nombre d'années d'études de psychologie (ANPSY) :
- Indice de charge mentale (CHARG) :
- Sexe (SEXE) :
- Réponse au questionnaire (QUEST) :
- Série du bac (BAC) :
- Année de naissance (ANNAI) :

2/ Indiquer, parmi les variables numériques précédentes, le nom d'une variable de rapport.
Justifier votre réponse :

Dossier PSYCHO : Liaison ANPSY x CHARG (18 points)

Voir la présentation du dossier PSYCHO sur la dernière feuille.

Objectif : On se demande si le nombre d'année d'études (ANPSY) contribue à faire diminuer la charge mentale induite par la lecture du texte sur la psychanalyse (CHARG).

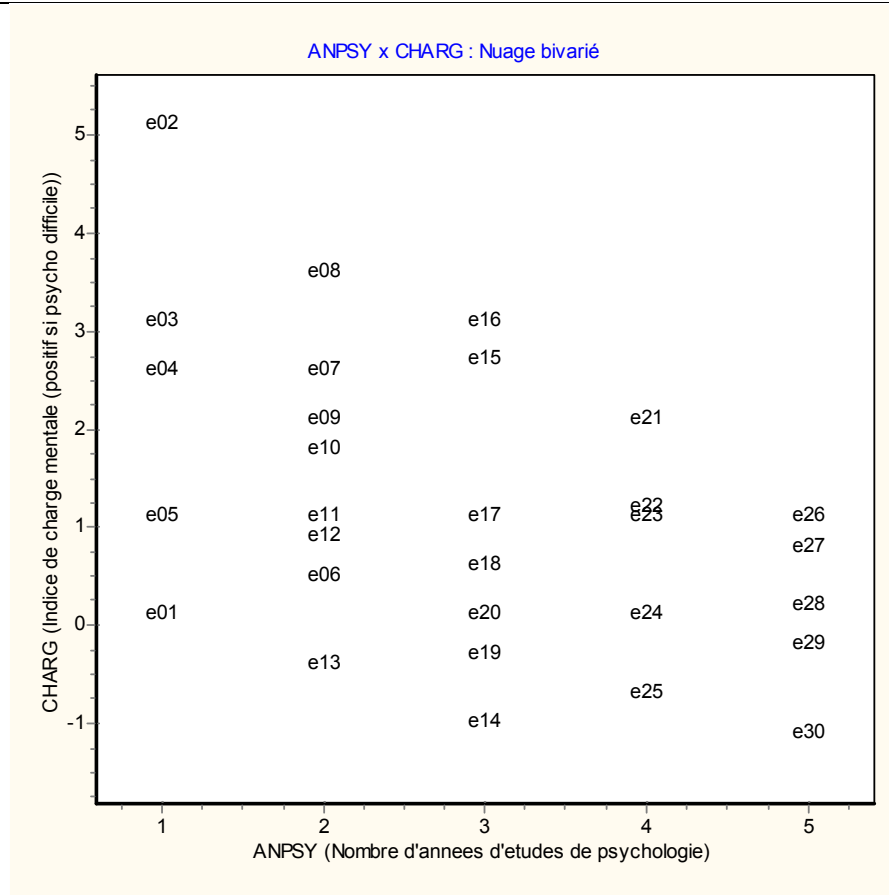
A. Analyse descriptive (11 points)

L'analyse séparée de chacune de ces deux variables donne les indications suivantes qui pourront être utilisées pour la suite de l'analyse :

	ANPSY	CHARG
Moy	2.90	1.07
Var	1.76	2.09

1/ L'analyse de la liaison entre les deux variables commence par l'analyse du graphe de corrélation (cf. page suivante).

a/ Quelles sont les coordonnées du point moyen ou centre de gravité du nuage ? On note G ce point. Reporter sa position sur le graphe.



b/ Indiquer pourquoi, pour analyser la liaison entre deux variables numériques, il est nécessaire d'analyser le graphe de corrélation (indiquer deux raisons) :

3/ La valeur du coefficient de corrélation linéaire de Bravais-Pearson est -0.48.
Indiquer ce que signifie le signe négatif de ce coefficient :

4/ On se demande si le nombre d'années d'études de psychologie (variable ANPSY) permet de prédire, au moins en partie, la charge mentale induite par la lecture du texte de psychologie (variable CHARG).

Un logiciel nous indique l'équation d'une droite qui permet de prédire la charge mentale en fonction du nombre d'années d'études de psychologie.

Cette équation est la suivante : **CHARG = -0.52 × ANPSY + 2.57**

a/ Indiquer pourquoi, compte tenu des résultats précédents, le coefficient -0.52 est nécessairement négatif.

b/ Interpréter la valeur du coefficient +2.57

c/ Tracer la droite sur le graphe de corrélation (cf. page précédente) après avoir reporté ci-dessous les calculs effectués :

d/ Calculer la charge mentale prédite par l'équation pour l'étudiant e01 :

e/ Calculer l'erreur de prédiction pour ce même étudiant :

B. Analyse inférentielle (7 points)

On se demande s'il existe bien un lien entre le nombre d'études de psychologie et la charge mentale induite par la lecture du texte de psychanalyse, dans la population d'où provient l'échantillon. On cherche donc à préciser quelle peut être la valeur de la corrélation dans la population parente. On suppose que cet échantillon a été constitué par tirage au hasard dans cette population.

1/ Préciser la nature de la population parente :

2/ Les tests présentés au cours de ce semestre servent à tester ce qui est appelé habituellement une hypothèse nulle (notée H_0). Formuler l'hypothèse nulle correspondant à la situation analysée ici :

3/ Pour tester cette hypothèse nulle, on met en œuvre un test T de Student. Calculer la valeur de cette statistique et les degrés de liberté correspondants (indiquer les formules utilisées et le développement des calculs) :

$$t =$$

$$ddl =$$

4/ Un logiciel nous indique $p = 0.8\%$ et $p/2 = 0.4\%$ pour ce test. En déduire le résultat du test, en deux étapes :

a/ Test significatif ou non significatif ?

b/ Rejet ou non rejet de l'hypothèse nulle ?

Dossier PSYCHO : Liaison SEXE x QUEST (22 points)

Voir la présentation du dossier PSYCHO sur la dernière feuille.

Objectif : On se demande s'il existe des différences entre les garçons et les filles (SEXE) dans les réponses à la question sur les raisons pour lesquelles les gens vont voir un psychologue (QUEST).

On observe la répartition suivante des réponses :

	ANORM	AIDE	ORIEN
G	6	2	0
F	8	8	6

A. Analyse descriptive (10 points)

1/ Pourquoi est-il nécessaire de calculer les effectifs d'indépendance (effectifs théoriques) pour analyser la liaison entre deux variables nominales ?

2/ Calculer l'effectif théorique pour le couple (F, ANORM) afin de retrouver sa valeur (10.3) avec une meilleure précision. Indiquer la formule, les calculs effectués et présenter le résultat arrondi à deux décimales :

	ANORM	AIDE	ORIEN
G	3.7	2.7	1.6
F	10.3	7.3	4.4

3/ Construire ci-dessous le tableau des attractions-répulsions :

	ANORM	AIDE	ORIEN
G			
F			

4/ Commenter ce tableau des attractions-répulsions :

On calcule ensuite les contributions absolues au Φ^2 (cf. tableau ci-dessous). Ces valeurs seront utilisées pour les différentes analyses qui suivent.

	ANORM	AIDE	ORIEN
G	0.0459	0.0056	0.0533
F	0.0167	0.002	0.0194
			0.1429

5/ Indiquer le nom de deux autres statistiques permettant de mesurer, case par case, l'ampleur des écarts à l'indépendance :

6/ On souhaite se prononcer sur l'importance de la liaison globale entre ces deux variables, dans l'échantillon.

a/ Indiquer le nom de la statistique permettant de se prononcer sur cette question :

b/ Calculer la valeur de cette statistique après avoir indiqué la formule utilisée :

c/ Concluez brièvement sur l'importance de la liaison observée :

B. Analyse inférentielle (12 points)

On souhaite évaluer s'il existe un lien entre ces deux variables dans la population parente. On suppose que cet échantillon a été constitué par tirage au hasard dans cette population

1/ Indiquer le nom de la statistique de test appropriée et calculer sa valeur à partir des éléments fournis dans les tableaux précédents :

2/ Calculer le nombre de degrés de liberté correspondant à ce test (indiquer la formule et le développement des calculs) :

3/ Utiliser la table appropriée pour déterminer si le test est significatif ou non. Justifier votre réponse.

4/ Rédiger une conclusion détaillée en deux parties, descriptive puis inférentielle, portant uniquement sur la liaison globale entre ces deux variables :

a/ Conclusion descriptive

b/ Conclusion inférentielle

Dossier PSYCHO – Liaison BAC x CHARG (9 points)

Voir la présentation du dossier PSYCHO sur la dernière feuille.

Objectif : On se demande si la charge mentale induite par la lecture du texte de psychanalyse (CHARG) varie selon la série du bac de la personne (BAC).

Afin de comparer les groupes, on a les premiers résultats suivants qui serviront de base à l'analyse et aux calculs suivants :

	Effectif	Moy	Var
L	10	1.01	1.93
S	8	0.79	2.24
ES	6	0.80	0.84
Autres	6	1.82	2.63
Total	30	1.07	2.09

1/ On trouve $V_{inter} = 0.15$. et $V_{intra} = 1.94$. Indiquer le développement des calculs (sans effectuer) qui permettraient de retrouver la valeur de V_{intra} :

$$V_{intra} =$$

2/ Indiquer la propriété de décomposition de la variance :

3/ Rappeler la formule du rapport de corrélation (noté r^2 ou Eta^2) et calculer sa valeur :

4/ Que peut-on en conclure sur l'importance de la liaison entre les deux variables dans l'échantillon ?

5/ Indiquer quelles configurations des moyennes des groupes et/ou des scores à l'intérieur des groupes auraient conduit à :

- un rapport de corrélation égal à 0 (0%)

- un rapport de corrélation égal à 1 (100%)?

Formules (3 points)

Démontrer que, si deux variables X et Y sont des variables numériques centrées-réduites (ou "scores Z "), la formule du coefficient de corrélation de Bravais-Pearson (cf. formulaire) se simplifie et peut s'écrire :

$$r = \frac{\sum (X \cdot Y)}{n}$$

Dossier PSYCHO

Source : inspiré d'un exemple de J.-F. Richard, dans Ghiglione, Richard (1994) – Cours de Psychologie, Paris : Dunod-CNED, Tome 2, 465-594.

On interroge un groupe de 30 étudiants parmi l'ensemble des étudiants en psychologie d'une université parisienne.

On note leur sexe (variable SEXE), on leur demande leur année de naissance (variable ANNAIS), la série du bac qu'ils ont obtenu (variable BAC : L / S / ES / Autre) et le nombre d'années d'études de psychologie qu'ils ont suivies (variable ANPSY: 1 à 5).

On leur demande de bien vouloir passer une expérience visant à étudier la charge mentale induite par la lecture de deux textes différents, un article de journal et un texte sur la psychanalyse. Pour cela on introduit une tâche secondaire consistant à demander aux sujets de donner, pendant la lecture, des chiffres au hasard. On compte le nombre de chiffres énoncés par minute de lecture dans chacune des deux conditions et en calcule la différence dans le sens "Journal – Psychanalyse". On construit ainsi un indice qui mesure la charge mentale induite par la lecture du texte sur la psychanalyse (CHARG). *Un nombre positif élevé signifie une charge mentale élevée pour la lecture du texte de psychanalyse.*

Les étudiants répondaient également à un petit questionnaire sur la psychologie (variable QUEST), comprenant entre autres la question suivante : "Pourquoi, à votre avis, les gens vont-ils voir les psychologues ?". Ils avaient le choix entre 3 réponses :

- parce qu'ils se sentent dans un état anormal (ANORM)
- par besoin d'aide, de conseil, par besoin de se connaître (AIDE).
- pour des problèmes d'orientation (ORIEN).

On trouve ci-dessous le tableau des données recueillies :

	ANPSY	CHARG	SEXE	QUEST	BAC	ANNAI
e01	1	0	F	AIDE	L	1985
e02	1	5	G	AIDE	Autre	1986
e03	1	3	F	ANORM	S	1984
e04	1	2.5	F	ORIEN	ES	1985
e05	1	1	F	ANORM	L	1985
e06	2	0.4	G	ANORM	S	1986
e07	2	2.5	F	ORIEN	Autre	1986
e08	2	3.5	F	ORIEN	L	1985
e09	2	2	F	AIDE	L	1980
e10	2	1.7	F	AIDE	Autre	1985
e11	2	1	F	ANORM	ES	1984
e12	2	0.8	F	ORIEN	S	1985
e13	2	-0.5	F	ANORM	S	1984
e14	3	-1.1	G	ANORM	L	1975
e15	3	2.6	G	AIDE	S	1983
e16	3	3	F	ORIEN	L	1982
e17	3	1	F	ORIEN	L	1983
e18	3	0.5	F	ANORM	Autre	1983
e19	3	-0.4	F	AIDE	ES	1982
e20	3	0	G	ANORM	ES	1975
e21	4	2	F	AIDE	S	1980
e22	4	1.1	G	ANORM	Autre	1982
e23	4	1	F	ANORM	ES	1981
e24	4	0	F	AIDE	L	1983
e25	4	-0.8	F	ANORM	S	1983
e26	5	1	F	AIDE	L	1982
e27	5	0.7	F	ANORM	ES	1980
e28	5	0.1	G	ANORM	Autre	1970
e29	5	-0.3	F	AIDE	L	1972
e30	5	-1.2	G	ANORM	S	1979

- Extrait de la table des valeurs critiques de la variable χ^2 (Khi^2) :

ddl	α	.05 (5%)	.01 (1%)	.001 (0.1%)
1		3.84	6.63	10.83
2		5.99	9.21	13.82
3		7.81	11.34	16.27
4		9.49	13.28	18.47
5		11.07	15.09	20.52

- Extraits de la table des valeurs critiques de la variable T de Student :

ddl	$\alpha/2$.025 (2.5%)	.005 (0.5%)	.0005 (0.05%)
	α	.05 (5%)	.01 (1%)	.001 (0.1%)
3		3.182	5.841	12.924
4		2.776	4.604	8.610
5		2.571	4.032	6.869
(...)		(...)	(...)	(...)
25		2.060	2.787	3.725
26		2.056	2.779	3.707
27		2.052	2.771	3.690
28		2.048	2.763	3.674
29		2.045	2.756	3.659
30		2.042	2.750	3.646

- Extrait de la table des valeurs critiques de la variable F de Fisher-Snedecor :

Valeurs critiques pour $\alpha = .05, .01, .001$

Ddl1	!	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ddl2	α															
10	.05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.84
	.01	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71	4.65	4.60	4.56
	.001	21.0	14.9	12.6	11.3	10.5	9.93	9.52	9.20	8.96	8.75	8.59	8.45	8.32	8.22	8.13
11	.05	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72
	.01	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40	4.34	4.29	4.25
	.001	19.7	13.8	11.6	10.3	9.58	9.05	8.66	8.35	8.12	7.92	7.76	7.63	7.51	7.41	7.32
12	.05	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62
	.01	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16	4.10	4.05	4.01
	.001	18.6	13.0	10.8	9.63	8.89	8.38	8.00	7.71	7.48	7.29	7.14	7.00	6.89	6.79	6.71

- Quelques formules...

(1) $t_{obs} = \sqrt{n-2} \times \sqrt{\frac{V_{inter}}{V_{intra}}}$

(2) $t_{obs} = \sqrt{n-2} \times \frac{r}{\sqrt{1-r^2}}$

(3) $F = \frac{(n-K)}{(K-1)} \times \frac{V_{inter}}{V_{intra}}$ avec (4) $ddl = K - 1$ et (5) $ddl2 = n - K$

(6) $r = \frac{Cov(X,Y)}{Ety(X) \cdot Ety(Y)}$

(7) $Cov(X,Y) = \frac{\sum(x,y)}{n} - \bar{x} \bar{y}$

(8) $a = \frac{Cov(X,Y)}{Var(X)} = r \times \frac{Ety(Y)}{Ety(X)}$

(9) $b = \bar{y} - a \bar{x}$

(10) $Txl_{jk} = \frac{n_{jk} - \hat{n}_{jk}}{\hat{n}_{jk}}$