

MEDI

Durée de l'épreuve : 2 heures

Aucun document n'est autorisé. Seule la calculatrice (sans sa documentation) est autorisée.

Attention : les exercices (encadrés) sont indépendants. Le barème donné à titre indicatif est sur 60 ; la note finale sera donnée sur 20.

La page 8 (tables et formules) peut être détachée et conservée. **Indiquer les réponses exclusivement sur ce document (ne pas écrire dans la marge).**

Dossier Stress&Tumeur (14 points)

Source : Visintainer, M.A., Volpicelli, J.R., & Seligman, M. (1982). Tumor Rejection in Rats after inescapable or escapable Shock. *Science*, 216 (3), 437-438.

Lors d'une recherche médicale effectuée sur 93 rats, échantillonnés au hasard, une équipe de chercheurs a soumis un premier groupe de 30 rats à des chocs électriques qu'ils ne pouvaient éviter. Un deuxième groupe de 30 rats avait la possibilité d'éviter ces chocs en soulevant une trappe qui leur permettait de passer dans une autre cage. Un troisième groupe de 33 animaux ne reçoit aucun choc. Ces 93 animaux ayant tous subi l'implantation d'une tumeur, les auteurs ont dénombré le nombre de rejets spontanés des tumeurs pour chacun des groupes. Le tableau ci-dessous résume les observations recueillies selon les deux variables étudiées chez ces rats : type de réaction de l'organisme vis à vis de la *Tumeur* (rejet/non rejet) et *Niveau de stress* subi (choc inévitable/choc évitable/pas de choc).

Tableau 1 : Effectifs observés

<i>Niveau stress</i> <i>Tumeur</i>	Choc inévitable	Choc évitable	Pas de choc	Total
Rejet	8	19	18	45
Non-rejet	22	11	15	48
Total	30	30	33	93

Question initiale : existe-t-il une relation entre le niveau de stress subi et la capacité de l'organisme à se défendre contre certaines tumeurs ?

A - Analyse descriptive : analyse locale du tableau de contingence

1) On se demande tout d'abord, avant l'analyse de la liaison, quelle est la réaction à la Tumeur la plus fréquemment observée, indépendamment du stress subi (tous les niveaux de stress confondus).

a) Calculer les pourcentages pertinents pour répondre à cette question (les résultats seront présentés arrondis à l'entier le plus proche) :

b) Commenter les résultats obtenus :

2) Le tableau des effectifs théoriques est le suivant :

Tableau 2 : Effectifs théoriques

<i>Niveau stress</i> <i>Tumeur</i>	Choc inévitable	Choc évitable	Pas de choc
Rejet	14.52	14.52	15.97
Non-rejet	15.48	15.48	17.03

a) Comment a-t-on obtenu la valeur 15.97 (en haut à droite de ce tableau 2) ? Donner la formule utilisée ainsi que le détail des calculs.

b) Donner la signification de cet effectif théorique (15.97) en langage naturel :

3) Le tableau des taux de liaison est le suivant :

Tableau 3 : Taux de liaison

<i>Niveau stress</i> <i>Tumeur</i>	Choc inévitable	Choc évitable	Pas de choc
Rejet	-0.45	0.31	0.13
Non-rejet	0.42	-0.29	-0.12

a) Comment a-t-on obtenu la valeur 0.13 (en haut à droite de ce tableau 3) ? Donner la formule utilisée ainsi que le détail des calculs :

b) Représenter ci-après le graphe des attractions :

c) Commenter ce graphe des attractions :

B – Analyse descriptive : analyse globale de la liaison entre les deux variables

1) On trouve $\Phi^2 = 0.0952$. Interpréter cette valeur :

2) Rédiger une conclusion descriptive :

C- Analyse inférentielle.

On se demande si on peut généraliser les résultats obtenus sur ces 93 rats à l'ensemble des rats ayant les mêmes caractéristiques que celles observées sur cet échantillon. Pour ce faire, on va utiliser un test statistique.

1) Indiquer ci-après le nom du test à utiliser, sa formule, ainsi que le nombre de degré de liberté (ddl) qui lui est associé :

2) Mise en œuvre du test.

a) Développer le calcul de ce test :

b) Donner le résultat du test :

c) Rédiger une conclusion inférentielle :

Dossier Mémoire&Sémantique (16 points)

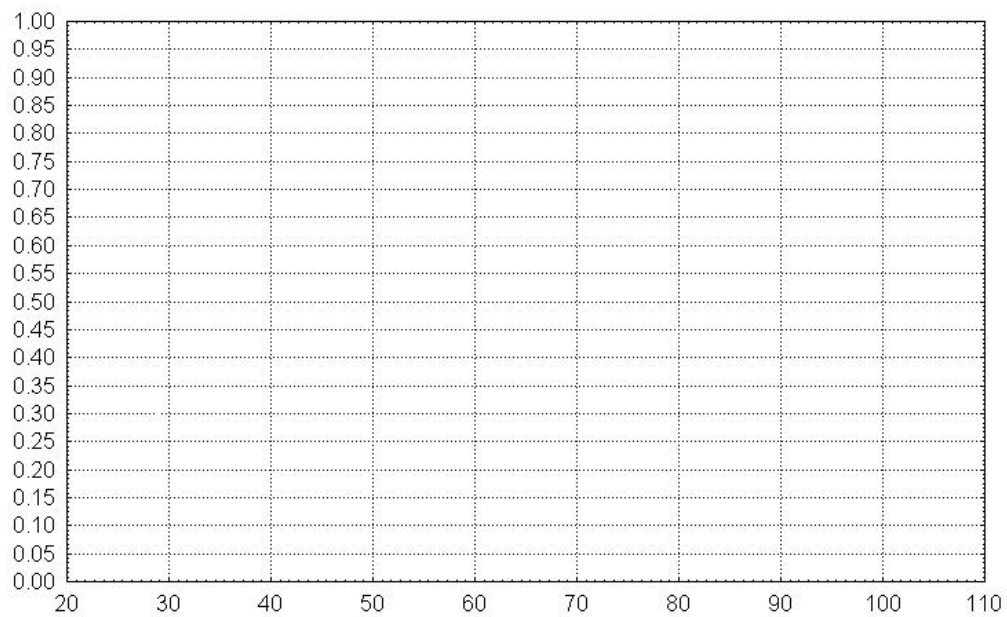
Source : données fournies par Alexandre-Nicolas Menchikoff, Laboratoire de Psychologie Expérimentale & équipe enseignante Metus-Medi, Université René Descartes, Paris 5.

Lors d'une expérience portant sur la mémorisation d'une liste de 112 mots appartenant à diverses catégories sémantiques : instruments, minéraux, animaux etc., on demande à 8 sujets adultes de langue maternelle française, échantillonnés au hasard, de rappeler dans l'ordre qu'ils veulent (rappel libre) le plus de mots possible. On retient 2 variables dépendantes (VD) : le nombre de mots rappelés en 10 minutes (*Variable X*) et un coefficient d'organisation subjective (*Variable Y*) susceptible de varier entre 0 et 1 selon la qualité des regroupements catégoriels effectués. Ce coefficient vaut zéro si les catégories dont relèvent les mots sont restituées dans un ordre aléatoire ; il vaut 1 si les mots sont tous regroupés par catégorie sémantique, quel que soit le nombre d'items rappelés pour chacune d'entre elles. Le tableau ci-dessous donne les résultats obtenus par les 8 sujets.

	s1	s2	s3	s4	s5	s6	s7	s8
Nombre de mots rappelés (<i>Variable X</i>)	38	29	45	56	65	76	96	75
Coefficient d'organisation (<i>Variable Y</i>)	0.14	0.30	0.40	0.72	0.67	0.74	0.88	0.91

Question initiale : existe-t-il une relation entre le nombre de mots rappelés et la qualité des regroupements catégoriels effectués ?

1) Construire le nuage bivarié ci-après (ne pas oublier les légendes) :



2) Commenter ce nuage bivarié :

3) Le coefficient de corrélation linéaire (r_{bp}) est égal à 0.89. Donner la valeur de ce coefficient arrondie à 4 décimales :

$r_{bp} =$

4) Dédire de la valeur calculée en 3) la part de variance de la variable "*Coefficient d'organisation*" prise en compte par la régression linéaire de "*Coefficient d'organisation*" sur "*Nombre de mots rappelés*". Indiquer le nom du coefficient utilisé ainsi que sa formule. Donner la valeur arrondie à 2 décimales.

5) A partir de ces indices (cf. calculs effectués en 3) et 4)), que peut-on dire de la liaison entre "*Nombre de mots rappelés*" et "*Coefficient d'organisation*" pour ces 8 sujets ?

6) Calculer et reporter ci-après l'équation de la droite de régression de "*Coefficient d'organisation*" sur "*Nombre de mots rappelés*" (2 décimales) :

7) Représenter cette droite sur le nuage de points ci-dessus.

8) Interpréter la valeur du coefficient de régression a :

9) Analyse inférentielle. Que peut-on inférer pour l'ensemble des adultes de langue maternelle française ?

a) Indiquer le nom du test que l'on va utiliser, ainsi que sa formule, pour répondre à cette question et développer le calcul ci-après :

b) Indiquer le résultat du test :

c) Rédiger une conclusion inférentielle :

Dossier Maternelle (13 points)

Source : D'après Hays, W.L. (1994). *Statistics*. Fort Worth : Harcourt Brace College Publishers (Fifth edition).

Quinze enfants âgés de 7 ans, échantillonnés au hasard dans une ville anglaise, ont été répartis de manière équilibrée en trois groupes de 5 enfants, selon le nombre d'années de préscolarisation effectuées (ou non) en maternelle :

$g1$: pas de préscolarisation en maternelle, $g2$: 1 an de préscolarisation, $g3$: plus de 1 an de préscolarisation. Chaque enfant a passé une épreuve de "maturité sociale" (notée sur 20).

On aimerait savoir s'il existe une différence de performances à l'échelle de maturité entre ces trois groupes. Les données sont les suivantes :

$g1$	4	8	11	5	3
$g2$	9	14	15	10	12
$g3$	15	18	20	10	14

$$m^{g1} = 6.20 \quad V^{g1} = 8.56$$

$$m^{g2} = 12.00 \quad V^{g2} = 5.20$$

$$m^{g3} = \quad V^{g3} =$$

Les résultats seront donnés arrondis à 2 décimales

1) Calculer la moyenne et la variance pour le groupe $g3$ et reporter les valeurs ci-dessus (il n'est pas nécessaire de donner le détail des calculs).

2) Calculer la moyenne générale des trois groupes (donner le détail des calculs) :

$$Moy =$$

3) Calculer la variance intra (donner la formule utilisée et le détail des calculs) :

$$V_{intra} =$$

4) On trouve $V_{inter} = 14.43$. Calculer η^2 (donner la formule utilisée et le détail des calculs) :

5) Rédiger une conclusion descriptive :

6) Analyse inférentielle. Que peut-on inférer quant à la population parente ?

a) Quelle est la population parente ?

b) Pour répondre à la question inférentielle, on utilise le test F de Fisher-Snedecor. On trouve : $F_{obs} = 10.14$. Donner le nombre de degrés de liberté (ddl) associé à ce test :

c) On trouve dans la table des valeurs critiques du F de Fisher-Snedecor, pour le nombre de degrés de liberté (ddl) associé à cette statistique :

.05	3.89
.01	6.93
.001	13.00

Indiquer le résultat du test :

d) Rédiger une conclusion inférentielle :

Dossier Régime (7 points)

Source : D'après Hays, W.L. (1994). *Statistics*. Fort Worth : Harcourt Brace College Publishers (Fifth edition).

Lors d'une étude sur les régimes effectués dans le but de perdre du poids, 15 couples (maris et femmes) ont été testés. On a relevé leur poids (en kg) avant régime, puis 2 semaines après. On aimerait savoir si les femmes perdent plus de poids que les hommes. Ci-dessous, les résultats obtenus après 2 semaines de régime (perte de poids en kg) :

$$\begin{array}{ll} m^h = 3.7 & V^h = 2.1 \\ m^f = 4.9 & V^f = 0.8 \end{array}$$

Critère sémantique : on jugera un effet important s'il est supérieur à 1kg, faible s'il est inférieur à 0.5 kg.

1) Calculer l'effet moyen (d_{obs}) du facteur "Sexe" :

2) Rédiger une conclusion descriptive :

3) On trouve $V_{intra} = 1.468$ et $V_{totale} = 1.824$. En déduire V_{inter} (donner le détail des calculs) :

$V_{inter} =$

4) Analyse inférentielle. On se demande si on peut généraliser ces résultats à l'ensemble des couples ayant les mêmes particularités que ceux testés dans cette étude ?

a) Indiquer le nom et la formule du test que l'on pourrait mettre en œuvre ainsi que le nombre de degrés de liberté qui lui est associé :

b) Est-il pertinent ici de généraliser ces résultats ? Pourquoi ? (justifier).

Dossier Décomposition de la variance (10 points)

1) Donner une définition de la variance intra :

2) Donner une définition de la variance inter :

3) Proposer des modifications des données (fictives) ci-dessous de manière à modifier les valeurs de certaines statistiques.

Il est inutile de recalculer les nouvelles valeurs des statistiques après les modifications. Par contre, expliquer pour chaque cas pourquoi vous avez procédé ainsi.

a) Modifier une, ou s'il le faut plusieurs valeur(s) de $g1$ de manière à diminuer la variance intra (sans se préoccuper de la variance inter):

$g1$

12	8	11	9
----	---	----	---

 $g2$

19	15	12	14	12	18
----	----	----	----	----	----

b) Modifier une, ou s'il le faut plusieurs valeur(s) de $g1$ de manière à augmenter la variance inter (sans se préoccuper de la variance intra) :

$g1$

12	8	11	9
----	---	----	---

 $g2$

19	15	12	14	12	18
----	----	----	----	----	----

c) Modifier une, ou s'il le faut plusieurs valeur(s) de $g1$ de manière à augmenter la variance inter sans modifier la variance intra.

$g1$

12	8	11	9
----	---	----	---

 $g2$

19	15	12	14	12	18
----	----	----	----	----	----

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

a	.05	.01	.001
ddl			
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
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.12
9	16.92	21.67	27.88
10	18.31	23.21	29.59

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

a/2	.025	.005	.0005
a	.05	.01	.001
ddl			
1	12.706	63.657	636.619
2	4.303	9.925	31.599
3	3.182	5.841	12.924
4	2.776	4.604	8.610
5	2.571	4.032	6.869
6	2.447	3.707	5.959
7	2.365	3.499	5.408
8	2.306	3.355	5.041
9	2.262	3.250	4.781
10	2.228	3.169	4.587

- Quelques formules...

$$t_{obs} = \sqrt{n-2} \times \sqrt{\frac{V_{inter}}{V_{intra}}} \qquad t_{obs} = \sqrt{n-2} \times \frac{r}{\sqrt{1-r^2}}$$

$$F = \frac{(n-K)}{(K-1)} \times \frac{V_{inter}}{V_{intra}} \qquad ddl = K - 1 \qquad ddl2 = n - K$$

$$r = \frac{Cov(x, y)}{\sqrt{Ety x} \sqrt{Ety y}} \qquad Cov(x, y) = \frac{\sum_i (x^i y^i)}{n} - \bar{x} \bar{y} \qquad a = \frac{Cov(x, y)}{Var x} = r \times \frac{Ety y}{Ety x} \qquad b = \bar{y} - a \bar{x}$$

$$t^{jk} = \frac{n_{jk} - \hat{n}_{jk}}{\hat{n}_{jk}} \qquad F^2 = \sum_{J,K} Cta_{jk} \qquad \mathbf{c}_{obs}^2 = \frac{(|n_{+-} - n_{-+}| - 1)^2}{n_{+-} + n_{-+}}$$

Les différents calculs ont été réalisés à l'aide des logiciels Statistica et DS3Win