

ECUE ASDP2

Durée de l'épreuve : 1 heure.

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 40 ; la note finale sera donnée sur 20.

La page 5 (formules et table) peut être détachée et conservée.

Indiquer les réponses exclusivement sur ce document.

Dossier Lecture (11 points)

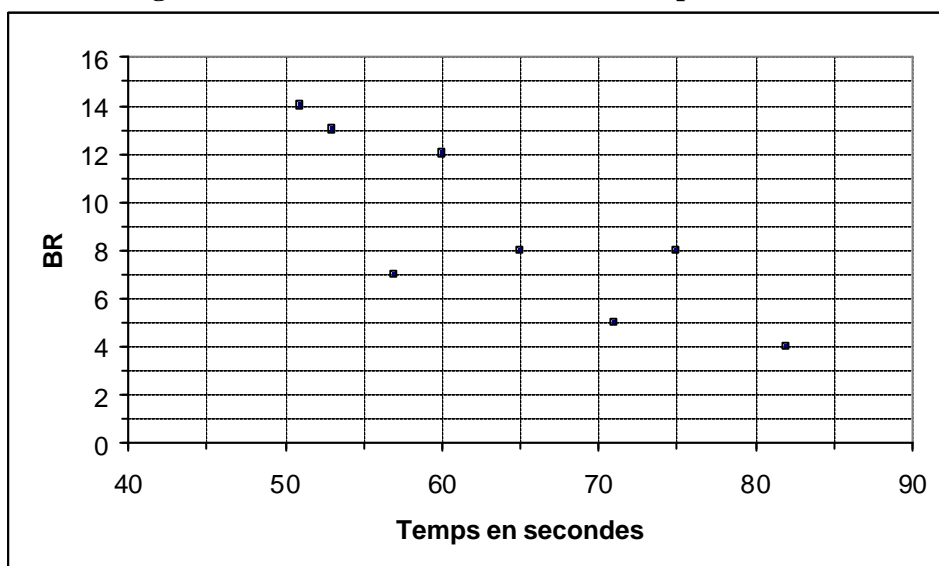
Source : données fournies par A.N. Menchikoff, LPE & équipe enseignante ASDP1-ASDP2.

Huit enfants sélectionnés dans une classe de CM1 sont invités à lire un texte à haute voix, à leur propre rythme. Après la lecture, ils subissent une épreuve de compréhension. On relève le temps mis en secondes pour lire ce texte, puis on compte le nombre de bonnes réponses (BR) données à cette épreuve de compréhension. Les résultats sont présentés dans le tableau ci-dessous et sur la figure 1.

Question initiale : on aimerait savoir d'une part, s'il existe une relation entre le temps de lecture du texte et sa compréhension par le lecteur, et d'autre part si on peut prédire le nombre de BR à l'épreuve de compréhension à partir du temps de lecture.

	s1	s2	s3	s4	s5	s6	s7	s8	Moy	Ety
Temps	57	53	51	60	65	71	75	82	64.250	10.305
BR	7	13	14	12	8	5	8	4	8.875	3.480

Figure 1 : nombre de BR en fonction du temps de lecture



1) Commenter le nuage de points représenté par la figure 1, et expliquer le sens de la relation :

2) Le coefficient de corrélation est égal à : $r_{bp} = -0.83$. Donner la valeur de ce coefficient arrondie à 3 décimales.

3) L'équation de la droite de régression est la suivante : $B\tilde{R} = -0.279temps + 26.781$. Interpréter la valeur "-0.279" :

- 4) Tracer cette droite de régression sur le graphique (figure 1).
5) Donner la qualité de l'ajustement (sans omettre le nom de l'indice) et interpréter :

6) Quel est le nombre prédit de BR pour un temps de lecture de 75 secondes ? Justifier.

7) Quelle est la valeur de l'écart résiduel ? Donner la procédure de calcul.

- Tracer cet écart sur le graphique (figure 1).

8) On aimerait généraliser les résultats obtenus sur cet échantillon à toute la population des enfants actuellement en classe de CM1. Que suggériez-vous pour procéder à cette généralisation ?

Dossier Allergie (11 points)

Source : d'après Prum, B. (1996). *Modèle linéaire : comparaison de groupes et régression*. Paris : Inserm.

On souhaite comparer trois traitements, notés A, B et C contre une allergie au pollen. On répartit par tirage au sort 43 patients habitant en Île de France et venant consulter dans un centre de soin parisien, puis on leur affecte l'un des trois traitements. Le groupe $g1$ ($n_{g1} = 11$) aura le traitement A, le groupe $g2$ ($n_{g2} = 17$), le traitement B et le groupe $g3$ ($n_{g3} = 15$), le traitement C. On mesure pour chaque patient le nombre de jours d'efficacité du traitement avant qu'une nouvelle crise d'allergie se déclenche. Ci-après les statistiques élémentaires de ces trois groupes :

$n_{g1} = 11$	$m^{g1} = 40.27$	$v^{g1} = 106.20$	
$n_{g2} = 17$	$m^{g2} = 52.71$	$v^{g2} = 94.68$	$Vinter = 33.182$
$n_{g3} = 15$	$m^{g3} = 41.47$	$v^{g3} = 162.78$	$Vintra = 121.382$

Question initiale : peut-on conclure que les traitements ont une efficacité différente pour les crises d'allergie au pollen ?

Tous les résultats seront présentés arrondis à 2 décimales

A. Analyse descriptive

1) Calculer la moyenne générale à partir des moyennes des trois groupes (donner la procédure de calcul) :

2) Donner la définition de la variance intra ($Vintra$) :

3) Calculer la variance totale (donner la procédure de calcul) :

4) On trouve Eta-Deux (η^2) = 0.21. Indiquer la formule qui permet de calculer cette valeur :

5) Rédiger une conclusion descriptive :

B. Analyse inférentielle

- 1) Indiquer quelle est la population parente :

- 2) On a utilisé un test pour cette analyse inférentielle. La valeur trouvée pour ce test est 5.47.
 - a) Indiquer le nom du test que l'on a utilisé :

 - b) Indiquer les nombres de degrés de liberté (ddl) associés à ce test (donner le détail des calculs) :

 - c) Indiquer le résultat du test :

 - d) Élaborer une conclusion inférentielle qui réponde à la question initiale :

Dossier Apprentissage (8 points)

Source : d'après Reuchlin, M. (1982). *Précis de statistique*. Paris : Puf.

Au cours d'une expérience étudiant l'effet des échanges interactifs sur l'apprentissage, on présente à des sujets échantillonnés au hasard une liste de syllabes sans signification et on note le nombre d'essais nécessaires à sa mémorisation. Le sujet travaille soit seul (condition *c1*), soit avec un autre sujet effectuant la même tâche que lui (condition d'échanges interactifs *c2*). Les résultats obtenus sur 30 sujets travaillant seuls et 30 sujets travaillant en condition d'échanges interactifs sont indiqués ci-après. On fait l'hypothèse que les sujets apprendront plus vite en condition d'échanges interactifs.

	Conditions	
	<i>c1: travail seul</i>	<i>c2 : travail avec échanges interactifs</i>
Moyenne	7.5	5.4
Écart-type	1.64	1.09

Critère sémantique : on jugera un effet important s'il est supérieur à 2 essais, faible s'il est inférieur à 1 essai.

Sens de calcul : modalité 1 - modalité 2

A. Analyse descriptive

- 1) Calculer l'effet du facteur "Condition" :

- 2) Rédiger une conclusion descriptive :

B. Analyse inférentielle

- a) On trouve : $t_{obs} = 5.20$. Indiquer la formule qui a permis de trouver ce résultat :

- b) Quel est le nombre de degrés de liberté (ddl) associé à ce test ?

c) Indiquer le résultat du test :

d) Rédiger une conclusion inférentielle :

Dossier Victimes (10 points)

Source : Carole Damiani, Psychologue. Données en cours d'analyse (2003).

On étudie le profil de 87 victimes (hommes et femmes ; *variable Sexe à 2 modalités*), échantillonnés au hasard lors de consultations ayant lieu dans un centre parisien, selon le type de traumatisme qu'elles ont subi (*variable Traumatisme à 6 modalités*) : agression sexuelle, viol, prise d'otage, hold-up, témoin d'homicide, accident du travail. Voir tableau 1, ci-après :

Tableau 1 : profils-lignes

Traumatisme Sexe	Agression sexuelle	Viol	Prise D'otage	Hold-up	Témoin homicide	Accident travail	Total
Hommes	8%	8%	12%	36%	28%	8%	100%
Femmes	31%	24%	11%	16%	10%	8%	100%
Profil moyen	24%	20%	11%	22%	15%	8%	100%

Question initiale: le type de traumatisme subi diffère-t-il selon le sexe ?

A. Analyse descriptive

1) Commenter à l'aide d'une phrase, la valeur 31% du tableau 1 ci-dessus (colonne de gauche, case grisée) :

2) Compléter le tableau 2 des attractions/répulsions ci-dessous :

Traumatisme Sexe	Agression sexuelle	Viol	Prise d'otage	Hold-up	Témoin homicide	Accident travail
Hommes						
Femmes						

3) Commenter les attractions de ce tableau 2 :

4) On trouve $\phi^2 = 0.1541$. Interpréter cette valeur (justifier) :

B. Analyse inférentielle

1) On procède à un test du χ^2 afin de généraliser les résultats obtenus sur cet échantillon.

a) Indiquer la formule de calcul, développer, et donner le résultat du calcul :

b) Donner le nombre de degrés de liberté (ddl) associé à ce test (rappeler la formule) :

2) Le résultat du test est significatif au seuil .05. Élaborer une conclusion inférentielle et répondre à la question initiale :

Extraits de table et formules. Attention ! Tout n'est pas utile

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

ddl1		1	2	3	4	5	6	7	8	9	10	11	12	13	14
ddl2	a														
30	.05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04
	.01	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84	2.79	2.74
	.001	13.3	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39	4.24	4.11	4.00	3.91	3.82
35	.05	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.07	2.04	2.01	1.99
	.01	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96	2.88	2.80	2.74	2.69	2.64
	.001	12.9	8.47	6.79	5.88	5.30	4.89	4.59	4.36	4.18	4.03	3.90	3.79	3.70	3.62
40	.05	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95
	.01	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66	2.61	2.56
	.001	12.6	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02	3.87	3.75	3.64	3.55	3.47
45	.05	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05	2.01	1.97	1.94	1.92
	.01	7.23	5.11	4.25	3.77	3.45	3.23	3.07	2.94	2.83	2.74	2.67	2.61	2.55	2.51
	.001	12.4	8.09	6.45	5.56	5.00	4.61	4.32	4.09	3.91	3.76	3.64	3.53	3.44	3.36

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

a/2 a	.025 .05	.005 .01	.0005 .001
20	2.086	2.845	3.850
21	2.080	2.831	3.819
22	2.074	2.819	3.792
23	2.069	2.807	3.768
24	2.064	2.797	3.745
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
40	2.021	2.704	3.551
58	2.002	2.663	3.466
60	2.000	2.660	3.460
120	1.980	2.617	3.373

- Extrait de la table des valeurs critiques de la variable c² :

a	.05	.01	.001
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

- Quelques formules...

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

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

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

$$t_{jk} = \frac{n_{jk} - \hat{n}_{jk}}{\hat{n}_{jk}} \quad Cta_{\bar{k}} = \frac{(n_{\bar{k}} - \hat{n}_{\bar{k}})^2}{\hat{n}_{\bar{k}}} / n \quad Ctr_{jk} = Cta_{jk} / F^2 \quad c_{obs}^2 = \frac{(|n_{+-} - n_{-+}| - I)^2}{n_{+-} + n_{-+}}$$