Révisions et tests de connaissances

Exercice 1 *Statistiques descriptives* `Données ages`
Dans un groupe de travaux dirigés de licence, on a relevé les dates de naissance de l'ensemble des étudiants :

04/07/74	05/09/75	21/05/74	20/01/76
01/10/76	18/05/75	29/07/76	21/07/77
18/11/76	17/03/76	02/02/77	11/05/65
10/09/77	19/02/76	11/03/76	01/06/77
07/05/77	06/10/74	30/05/77	16/12/73
23/02/77	16/12/75	25/10/74	23/01/73
28/12/76	21/12/75	28/01/72	06/09/75
21/01/77	24/07/61	19/01/76	16/02/74
14/10/75	22/08/64	09/06/75	04/01/77
27/01/74	05/09/63	24/10/76	14/11/74
11/07/76	18/12/77	25/03/75	12/01/73
26/03/72	09/03/74	11/10/69	14/06/60

1) Calculer l'âge en nombre entier d'années de chacun des sujets, au 1/10/1997. Pourquoi s'intéresser à l'âge plutôt qu'à la date de naissance ?

2) Calculer la moyenne, la médiane, le mode, la variance et l'écart-type de l'ensemble des valeurs observées de la variable "âge".

3) On s'intéresse à différents codages de cette variable :
- codage numérique 1 : âge en années *(date de référence : 1/10/1997)*
- codage numérique 2 : âge en mois *(compté au 1/10/97)*
- codage en rangs *(rappel : rang 1 pour le plus jeune, convention habituelle pour les ex-æquos : attribution du rang moyen)*
- codage ordinal ; *par exemple, utiliser le codage suivant :*

Age	Code
$19 \leq A < 21$	1
$21 \leq A < 24$	2
$24 \leq A < 27$	3
$27 \leq A < 30$	4
$30 \leq A < 38$	5

- codage binaire : par exemple, coder 0 pour les individus en dessous de la moyenne, 1 pour ceux au dessus.
- codage centré : écarts à la moyenne
- codage centré réduit : écarts réduits.

a) Compléter un tableau du type de celui figurant sur la page suivante.
b) Calculer la moyenne, la variance et l'écart type de la série regroupée en classes (cf. codage ordinal).
4) Construire un histogramme de la distribution en années.
Réponses : 2) $m = 22.90$; $s = 4.28$; $s_c = 4.33$, $Md = 21.5$, $Mode = 20$.
3-b) $m' = 23.34$; $s'^2 = 16.39$; $s' = 4.05$

Suj.	Date-Nais	Age-1	Age-2	Rang	Ord.	Bin.	Centre	Cent-Red
i42	18/12/77	19	237	1	1	0	-43,9375	-0,86543
i29	21/10/77	19	239	2	1	0	-41,9375	-0,826036
i13	10/09/77	20	240	3	1	0	-40,9375	-0,806339
i08	21/07/77	20	242	4	1	0	-38,9375	-0,766946
i16	01/06/77	20	244	6	1	0	-36,9375	-0,727552
i17	07/05/77	20	244	6	1	0	-36,9375	-0,727552
i19	30/05/77	20	244	6	1	0	-36,9375	-0,727552
i11	02/02/77	20	247	8,5	1	0	-33,9375	-0,668461
i21	23/02/77	20	247	8,5	1	0	-33,9375	-0,668461
i36	04/01/77	20	248	10	1	0	-32,9375	-0,648765
i25	28/12/76	20	249	11	1	0	-31,9375	-0,629068
i09	18/11/76	20	250	12	1	0	-30,9375	-0,609371
i05	01/10/76	20	251	13,5	1	0	-29,9375	-0,589674
i39	24/10/76	20	251	13,5	1	0	-29,9375	-0,589674
i07	29/07/76	21	254	15,5	2	0	-26,9375	-0,530584
i41	11/07/76	21	254	15,5	2	0	-26,9375	-0,530584
i10	17/03/76	21	258	17,5	2	0	-22,9375	-0,451796
i15	11/03/76	21	258	17,5	2	0	-22,9375	-0,451796
i14	19/02/76	21	259	19	2	0	-21,9375	-0,432099
i04	20/01/76	21	260	20,5	2	0	-20,9375	-0,412403
i31	19/01/76	21	260	20,5	2	0	-20,9375	-0,412403
i22	16/12/75	21	261	22,5	2	0	-19,9375	-0,392706
i26	21/12/75	21	261	22,5	2	0	-19,9375	-0,392706
i33	14/10/75	21	263	24	2	0	-17,9375	-0,353312
i02	05/09/75	22	264	25,5	2	0	-16,9375	-0,333615
i28	06/09/75	22	264	25,5	2	0	-16,9375	-0,333615
i35	09/06/75	22	267	27	2	0	-13,9375	-0,274525
i06	18/05/75	22	268	28	2	0	-12,9375	-0,254828
i43	25/03/75	22	270	29	2	0	-10,9375	-0,215434
i40	14/11/74	22	274	30	2	0	-6,9375	-0,136647
i18	06/10/74	22	275	31,5	2	0	-5,9375	-0,11695
i23	25/10/74	22	275	31,5	2	0	-5,9375	-0,11695
i01	04/07/74	23	278	33	2	0	-2,9375	-0,0578595
i03	21/05/74	23	280	34	2	0	-0,9375	-0,0184658
i46	09/03/74	23	282	35	2	1	1,0625	0,0209279
i32	16/02/74	23	283	36	2	1	2,0625	0,0406247
i37	27/01/74	23	284	37	2	1	3,0625	0,0603216
i20	16/12/73	23	285	38	2	1	4,0625	0,0800184
i24	23/01/73	24	296	39,5	3	1	15,0625	0,296684
i44	12/01/73	24	296	39,5	3	1	15,0625	0,296684
i45	26/03/72	25	306	41	3	1	25,0625	0,493652
i27	28/01/72	25	308	42	3	1	27,0625	0,533046
i47	11/10/69	27	335	43	4	1	54,0625	1,06486
i12	11/05/65	32	388	44	5	1	107,063	2,10879
i34	22/08/64	33	397	45	5	1	116,063	2,28606
i38	05/09/63	34	408	46	5	1	127,063	2,50273
i30	27/07/61	36	434	47	5	1	153,063	3,01485
i48	14/06/60	37	447	48	5	1	166,063	3,27091

Exercice 2 *Distributions théoriques - loi binomiale*
Dans le cadre d'une étude sur les familles nombreuses, on veut constituer un échantillon de 100 familles représentatif des familles de 5 enfants.
1) Combien de familles comportant 1 garçon et 4 filles doit-on retenir pour constituer l'échantillon ?
2) Même question pour les familles comportant 0, 2, 3, 4, 5 garçons.
N.B. On fera l'hypothèse que la probabilité d'avoir un garçon ou une fille est égale à 0.5
3) Représenter la distribution obtenue à l'aide d'un diagramme en bâtons.
Réponse : Les effectifs respectifs des familles comportant 0, 1,..., 5 garçons sont donnés par : 3, 16, 31, 31, 16, 3

Exercice 3 *Distributions théoriques - loi normale*
Dans une maternité, on a relevé le poids de 200 enfants à la naissance. Après répartition en 9 classes, les valeurs observées sont les suivantes :

Poids (en kg)	Effectif
[2.0, 2.6[	17
[2.6, 2.8[	13
[2.8, 3.0[	16
[3.0, 3.2[	30
[3.2, 3.4[	41
[3.4, 3.6[	36
[3.6, 3.8[	17
[3.8, 4.0[	14
[4.0, 4.6]	16

1) Calculer la moyenne de la variable "poids", puis la variance d'échantillon et la variance corrigée (c'est-à-dire l'estimation ponctuelle de la variance sur la population).
2) On souhaite tester la normalité de la distribution des poids dans la population parente. A cet effet, on calcule les fréquences théoriques des classes précédentes lorsque la variable suit une loi normale de paramètres $\mu = 3.306$ et $\sigma = 0.504$. On obtient les résultats suivants :

Classe	Fréquence
$]-\infty,\ 2.6[$	0.0808
[2.6, 2.8[	0.0771
[2.8, 3.0[	0.1142
[3.0, 3.2[	0.1447
[3.2, 3.4[	0.1571
[3.4, 3.6[	0.1461
[3.6, 3.8[	0.1163
[3.8, 4.0[	0.0793
$[4.0,\ +\infty[$	0.0844

Refaire les calculs concernant les deux premières lignes et la dernière ligne de ce tableau. Construire le tableau des effectifs théoriques des classes, lorsque l'effectif total est de 200 unités statistiques.

Echantillonnage

Exercice 4
The mean of a large sample is K and σ_K is 2.50([1]). What are the chances that the sample mean misses the true mean by more than : (a) $\pm$1.00, (b) $\pm$3.00, (c) $\pm$10.00.
Réponses : (a) 69%, (b) 23%, (c) moins de 1%.

Exercice 5
Un test comportemental a été étalonné auprès d'une population de 3500 personnes. On sait que la moyenne au test est de 54.24 et que l'écart type est de 6.82. On sait en outre que la variable étudiée est distribuée selon une loi normale.
1) Quelle est la proportion de sujets qui ont une note supérieure à 62 ? Quel est leur nombre ?
2) Quelle est la proportion de sujets qui ont une note comprise entre 45 et 70 ? Quel est leur nombre ?
3) Quelle est la valeur de la variable correspondant au quartile inférieur (Q_1) de la distribution ?
Même question pour le quartile supérieur (Q_3).
4) On tire au hasard, avec remise, un échantillon de 100 sujets de la population étudiée. Soit $\overline{X}$ la variable "moyenne observée sur l'échantillon tiré".
a) Quelle est la loi suivie par cette variable ? Quels sont ses paramètres ?
b) Donner un intervalle rassemblant 95% des valeurs de $\overline{X}$ observées sur de tels échantillons.
Réponses : 1) $P(X > 62) = 0.1291$*, d'où* $N_1 = 452$.
2) $P(45 \leq X < 70) = 0.9011$ *d'où* $N_2 = 3153$.
3) $P(X \leq a) = 0.25$ *pour* $a = 49.63$. $P(X \leq a) = 0.75$ *pour* $a = 58.84$.
4) La distribution d'échantillonnage a pour paramètres : $\mu = 54.24$ *et* $S = 0.682$. *L'intervalle, centré autour de la moyenne, rassemblant 95% des échantillons est :* $[52.90; 55.58]$.

Exercice 6
Opinion upon an issue seems about equally divided. How large a sample (N) would you need to be sure (at .01 level) that a deviation of 3% in a sample is not accidental (due to chance) ?
Réponse : 1850

Intervalle de confiance

Exercice 7
For a given group of 500 soldiers the mean AGCT([2]) score is 95.00 and the SD is 25.
a) Determine the .99 confidence intervall for the true mean.
b) It is unlikely that the true mean is larger then what value ?
Réponses : a) $[92.11, 97.89]$*, b)* 97.89

[1] σ_K désigne ici l'écart type de la distribution d'échantillonnage
[2] AGCT : Army General Classification Test

Exercice 8

Quelle est la capacité de la mémoire à court terme ? Pour répondre à cette question, deux chercheurs[3] ont présenté à 210 élèves de lycée une liste de 16 mots communs sur un écran de télévision à raison de 1 mot toutes les deux secondes. La moyenne du rappel est de 6.91 tandis que l'écart type est égal à 2.08. On souhaite estimer, avec un degré de confiance de 95%, la moyenne de la population dont sont issus les sujets composant l'échantillon.
Réponse : Intervalle de confiance : $[6.63; 7.19]$ *avec un degré de confiance de 95%.*

Introduction aux tests statistiques

Exercice 9

In which of the following experimental problems would it be more important to avoid Type I errors of inference than Type II errors in determining the significance of a difference ?
a) Sex differences in reading rate and comprehension in the fifth grade.
b) Effects of a new drug upon reaction time – especially when the drugs are potent and probably dangerous.
c) Comparison of two methods of learning a new skill.
d) Acceptance of a program which involves much time and money and rejection of a less expensive program.
e) Comparative efficiency of a speed-up and a normal rate of work in a factory.
Réponses : a, c, d, e

Exercice 10

"Le problème lié à l'interprétation d'une hypothèse nulle qui n'a pas été rejetée tourmente depuis plus de 50 ans les étudiants inscrits à un cours de statistiques. Tous les statisticiens s'accordent cependant sur un point : on ne peut jamais prétendre avoir "prouvé" l'hypothèse nulle.
Pour Fisher, un résultat non significatif n'est pas un résultat concluant. D'après lui, on a le choix entre rejeter une hypothèse nulle et suspendre son jugement.
Neymann et Pearson adoptent une approche légèrement différente. Soit on rejette l'hypothèse nulle, soit on l'*accepte*. Toutefois, quand on dit que l'on accepte une hypothèse nulle, cela ne veut pas dire que l'on estime en avoir prouvé l'exactitude. Cela veut dire que l'on fera *comme si* elle était vraie, au moins jusqu'à ce que l'on obtienne des données plus adéquates."

Quels commentaires ces deux opinions suggèrent-elles ? Donner deux exemples de situations dans lesquelles on sera conduit à adopter l'une ou l'autre des deux démarches.

Réponse : Une hypothèse nulle peut ne pas être rejetée à cause d'une trop faible taille d'échantillon, ou de l'existence de variables non contrôlées qui ont artificiellement augmenté la dispersion. Il y a des cas où l'on est conduit à accepter l'hypothèse nulle, jusqu'à ce que des informations contraires viennent la contredire. Exemple : absence d'effet indésirable d'un médicament.

[3] Expérience rapportée par A. Lieury, Liège, Mardaga, 1992

Inférence sur une moyenne. Comparaison à une norme

Exercice 11

Dans ce qui suit, on se donne une population de référence dans laquelle la distribution du QI est une loi normale de paramètres $\mu = 100$ et $\sigma = 15$.
a) Dans une classe de rattrapage, on examine un groupe de 5 élèves et on trouve un QI moyen $\overline{QI} = 89.2$. Peut-on dire que le groupe de 5 élèves est, pour le QI, atypique de la population de référence aux seuils traditionnels ?
b) La classe de rattrapage compte 25 élèves ; on examine tous les élèves et on trouve un QI moyen $\overline{QI} = 89.2$. Mêmes questions que précédemment.
Réponses : a) $z_{obs} = -1.61$. *On ne peut pas retenir l'hypothèse d'un groupe atypique au seuil de 5% unilatéral. b)* $z_{obs} = -3.6$. *Groupe atypique au seuil de 1% unilatéral*

Exercice 12

Compas et ses collègues (étude non publiée) ont constaté avec surprise que les jeunes enfants soumis au stress présentent en fait moins de symptômes d'angoisse et de dépression que ce à quoi l'on pourrait s'attendre. Toutefois, ils ont aussi remarqué que les scores obtenus par ces enfants sur une échelle de désirabilité sociale sont étonnament élevés. On sait que la moyenne de population de l'échelle de désirabilité sociale est égale à 3.87. Pour un échantillon de 36 enfants soumis au stress, Compas *et al.* ont relevé une moyenne d'échantillon de 4.39, avec un écart type de 2.61.
1) Par quel test pourrait-on savoir si ce groupe présente une tendance accrue à donner des réponses socialement acceptables ?
2) Quelles seraient l'hypothèse nulle et l'hypothèse alternative ?
3) Mettre en œuvre le test choisi et conclure.
Réponse : Il s'agit ici de comparer la moyenne observée sur l'échantillon à la valeur 3.87, considérée comme une norme. Compte tenu de la taille de l'échantillon, on peut utiliser la statistique :

$$Z = \frac{\overline{x} - 3.87}{E} \quad \textit{avec} \quad E^2 = \frac{2.61^2}{36}$$

qui suit une loi normale centrée réduite.
On obtient alors $Z_{obs} = 1.20$, *ce qui n'est pas significatif d'une différence.*

Exercice 13

Dans une grande entreprise (américaine), le salaire moyen des hommes possédant entre 3 et 5 ans d'expérience est de 28 000 \$. Les salaires (en milliers de dollars) d'un échantillon aléatoire de 10 femmes possédant entre 3 et 5 ans d'expérience sont les suivants :

24 27 31 21 19 26 30 22 15 36

1) Au niveau descriptif, quelle estimation peut-on faire du salaire moyen féminin ?
2) Y a-t-il des preuves attestant de niveaux de salaires différents pour les hommes et les femmes ?
Réponses : 1) La moyenne des salaires sur l'échantillon est $\overline{x} = 25.1$. *On peut donc estimer le salaire féminin moyen à 25 100 \$.*
2) On a par ailleurs : $s_c = 6.23$. *La statistique de test (comparaison d'une moyenne à une norme sur un petit échantillon) vaut :* $t = \frac{25.1-28}{E}$ *avec* $E^2 = \frac{6.23^2}{10}$, *d'où* $t_{obs} = -1.47$. *Cette statistique suit une loi de Student à 9 ddl, et donc, pour un test unilatéral au seuil*

de 5%, $t_{crit} = 1.83$. *On conclut donc sur* H_0 *: on n'a pas mis en évidence de différence entre les hommes et les femmes.*

Exercice 14

1) En 2003-2004, la population des étudiants inscrits dans le secteur Lettres à l'UBO était constituée de 3635 femmes et 1338 hommes.

L'échantillon constitué par les étudiants inscrits au centre littéraire de Quimper comprenait 405 femmes pour 174 hommes.

Du point de vue de la variable étudiée, peut-on considérer cet échantillon comme tiré au hasard dans la population des étudiants littéraires de l'UBO (seuil utilisé : 1%) ?

2) De même, la population des étudiants inscrits dans le secteur Sciences était constituée de 1361 femmes pour 2232 hommes.

a) L'échantillon constitué par les étudiants inscrits à l'IUP Génie Mécanique et Productique comprenait 11 femmes pour 138 hommes.

Du point de vue de la variable étudiée, peut-on considérer cet échantillon comme tiré au hasard dans la population des étudiants scientifiques de l'UBO (seuil utilisé : 1%) ?

b) Même question pour les étudiants de l'IUP Génie des Systèmes Indistriels (69 femmes pour 25 hommes).

Réponses : 1) On fait ici un test de comparaison d'une proportion à une norme. La proportion de femmes dans la population est $p_0 = 0.7309$, *tandis que la proportion observée sur l'échantillon est* $f = 0.6995$, *avec une taille d'échantillon* $n = 579$. *On obtient* $Z_{obs} = -1.70$. *On retient donc* H_0, *ce qui revient à répondre positivement à la question posée.*

2) a) On a ici : $p_0 = 0.3788$ *et* $f = 0.0738$ *et donc* $Z_{obs} = -7.67$. *Du point de vue de la répartition par sexe, l'échantillon formé par les étudiants de l'IUP GMP est donc tout à fait atypique de la population des étudiants scientifiques.*

b) On a : $p_0 = 0.3788$ *et* $f = 0.7340$ *d'où* $Z_{obs} = 7.10$ *et une conclusion similaire à la précédente.*

Tests d'égalité de moyennes, de fréquences

Exercice 15 *Dossier "Plpc", "Plpc-cor"*

Dans une recherche de psychologie, on a présenté à chaque sujet des photographies de chiens, de diverses races, pour moitié à poils courts et pour moitié à poils longs. Chaque sujet évalue le chien présenté en cochant l'une des 6 cases d'une échelle d'attirance. Les évaluations ont été codées numériquement de 0 (faible attirance) à 5 (forte attirance). Dans le tableau ci-dessous, on donne pour chacun des 16 sujets les deux valeurs moyennes des évaluations concernant les deux types de chiens.

	s1	s2	s3	s4	s5	s6	s7	s8
Poils longs	3,50	2,75	3,50	2,75	1,50	3,00	4,00	3,00
Poils courts	2,50	1,25	2,75	0,50	2,00	3,00	2,00	2,00

	s9	s10	s11	s12	s13	s14	s15	s16
Poils longs	2,75	3,25	4,00	3,00	3,75	4,00	4,00	3,25
Poils courts	2,50	0,75	0,50	1,00	1,50	1,75	2,00	2,00

Un *effet individuel* est défini comme la différence entre les deux notes concernant un même sujet. Construire le protocole des effets individuels.

Peut-on affirmer, au seuil de 5%, que les chiens à poils longs sont préférés à ceux à poils courts.
Remarque. La variable statistique de départ (classement sur une échelle d'attirance) est en fait ordinale et non numérique. Nous reprendrons cet exemple avec d'autres outils dans un chapitre ultérieur.
Réponses : Le protocole des effets individuels a pour moyenne $\overline{d} = 1.5$ *et pour écart type corrigé* $s_c = 1.0448$. *On obtient* $t_{obs} = 5,74$. *La différence est significative au seuil de 5%.*

Exercice 16 *Dossier "Internat"*
Pour une recherche internationale sur l'évaluation des savoir-faire en situation de vie quotidienne, on a construit un test comportant une version en français et une version en anglais. Pour s'assurer de l'équivalence des deux instruments, on administre les deux versions du test à un groupe de 24 paires de sujets, chaque paire comprenant un sujet de langue maternelle anglaise et un sujet de langue maternelle française. Pour constituer les paires, les sujets sont appariés sur un ensemble de variables permettant de s'assurer de l'équivalence de leur efficience. Le plan du protocole est P24*T2.
On cherche à répondre à la question "les sujets obtiennent-ils le même résultat au test en anglais (t_1) et au test en français (t_2)".

Sujet	t_1	t_2
S1	27.0	25.5
S2	23.0	23.5
S3	21.5	18.5
S4	24.0	22.0
S5	22.5	28.0
S6	24.0	19.0
S7	21.5	19.0
S8	22.0	25.5

Sujet	t_1	t_2
S9	27.0	28.5
S10	28.5	26.0
S11	26.5	23.0
S12	24.0	24.5
S13	26.0	23.5
S14	21.5	22.0
S15	26.5	22.0
S16	20.0	16.0

Sujet	t_1	t_2
S17	21.5	20.5
S18	19.5	23.5
S19	23.5	23.0
S20	21.0	24.0
S21	23.0	24.0
S22	20.0	20.5
S23	20.0	17.5
S24	21.0	16.0

Réponses : On introduit le protocole des différences individuelles D défini par : $d_i = t_{i2} - t_{i1}$. *On a :* $\overline{d} = -0.8125$, $s_c = 2.9185$. $t_{obs} = -1.36386$. *Ici, ddl = 23 et donc, au seuil de 5% bilatéral,* $t_{crit} = 2.07$. *On ne peut donc pas rejeter l'hypothèse nulle*

Exercice 17 *Dossier "Manage"*
Dans une étude de psychologie du travail, on cherche à déterminer l'influence de deux conceptions managériales, l'une "hiérarchique" (H), l'autre "démocratique" (D) sur la sécurité dans l'entreprise. A cet effet, on prélève au hasard, dans un vaste ensemble de données, quinze intervalles trimestriels pour lesquels on relève ensuite le nombre d'incidents qui se sont produits sur chacun des sites, par ailleurs comparables à tous égards. Les données résultant des tirages aléatoires sont les suivantes :

Trim.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Conc. H	11	11	14	21	12	10	15	15	17	9	12	12	15	12	11
Conc. D	8	13	12	17	14	9	10	12	13	10	8	13	12	9	9

1) Calculer la moyenne et l'écart type de chacun des deux échantillons proposés. A partir de ces éléments, formuler une conclusion descriptive concernant le problème posé.
2) Les données observées définissent-elles des échantillons indépendants ou appariés ?

3) Réaliser un test statistique approprié en précisant avec soin l'hypothèse nulle et l'hypothèse alternative, la statistique de test utilisée, la loi suivie par cette statistique et le seuil choisi.

Réponses : 1) $\overline{x}_1 = 13.33$, $\overline{x}_2 = 11.27$, $s_{1c} = 3.09$, $s_{2c} = 2.55$. *3)* $\overline{d} = 1.87$. $s_c = 2.33$. $t_{obs} = 3.10$. *La différence est significative au seuil de 1% (test bilatéral).*

Exercice 18

Douze sujets ont eu à résoudre une série de problèmes particulièrement difficiles dont les réponses reposent toutes sur le même ensemble de règles très difficilement explicitables. Une première série de 80 problèmes servait de phase d'apprentissage. Dans la seconde phase (test), ils avaient à résoudre une seconde série de 80 problèmes. Pour chaque phase, on a mesuré le pourcentage de réponses correctes à la question posée. Les observations sont les suivantes :

Apprentissage	31	16	26	24	23	32	22	20	14	25	9	13
Test	37	22	27	25	26	25	23	22	20	31	9	15

1) Un test bilatéral de comparaison de moyennes sur des groupes appariés fait à l'aide d'un logiciel de traitements statistiques fournit le résultat suivant :

Test t pour des Echantillons Appariés (Feuille de données dans Classeur2) Différences significatives marquées à p < ,05000								
	Moyenne	Ec-Type	N	Differ.	Ec-Type Differ	t	dl	p
Appr.	21,25	7,12						
Test	23,50	7,14	12	-2,25	3,72	-2,10	11	0,0601

Quelle interprétation peut-on donner des résultats fournis par le logiciel ?

2) Les chercheurs souhaitent en fait montrer que les résultats en phase de test sont meilleurs que ceux obtenus en phase d'apprentissage. Leur hypothèse de recherche les conduit donc à prendre l'hypothèse alternative H_1 sous la forme unilatérale suivante : $\mu_{\text{Appr}} < \mu_{\text{Test}}$.

a) Examiner la consistance des observations avec l'hypothèse H_1 choisie.

b) Quel est le niveau de significativité (ou p-value) du test unilatéral ?

c) Quelle conclusion les chercheurs peuvent-ils énoncer ?

Réponses. 1) La p-value calculée par le logiciel est 6%. *Le test bilatéral ne permet donc pas de conclure à une différence significative entre la phase d'apprentissage et celle de test. 2) a) On constate que les moyennes observées sur l'échantillon étudié sont ordonnées dans le même sens que les moyennes théoriques figurant dans l'hypothèse unilatérale. Les observations sont donc consistantes avec l'hypothèse* H_1. *b) Le niveau de significativité du test unilatéral est alors la moitié de la p-value précédente, c'est-à-dire* 3%. *c) L'expérience semble montrer, au seuil de* 5%, *que les résultats en phase de test sont meilleurs que les résultats en phase d'apprentissage.*

Exercice 19 *Dossier "Usine"*

Dans une usine, on cherche à voir si un changement de l'environnement (musique dans les ateliers en particulier) peut modifier le rendement. Ce dernier est mesuré ici par le nombre moyen de pièces produites à l'heure par chaque ouvrier.

1) On note pour chacun des 33 ouvriers observés, son rendement avant et après l'introduction de ces changements. Les résultats sont les suivants.

Avant	Après	Avant	Après	Avant	Après
45	48	30	35	40	45
36	40	45	50	40	35
47	53	30	40	38	35
40	40	45	50	35	40
45	46	40	45	40	45
35	30	50	50	35	37
36	40	40	40	38	35
50	60	50	45	50	50
50	60	40	35	45	50
40	40	55	50	30	33
40	40	30	35	38	38

Effectuer les calculs permettant de répondre à la question posée et conclure.
2) On suppose maintenant que les données ci-dessus concernent deux groupes d'ouvriers distincts, les premiers (colonne "Avant") placés dans un environnement ordinaire, les seconds (colonne "Après") placés dans un environnement modifié. Quels sont alors les calculs à effectuer et quelle est alors la conclusion ?
Réponses : 1) $\overline{d} = 2.03$, $s_c = 4.41$. $z_{obs} = 2.64$. *La différence est significative pour un test unilatéral au seuil de 1%.*
2) $\overline{x}_1 = 40.85$, $\overline{x}_2 = 42.88$, $s_{1c} = 6.60$, $s_{2c} = 7.52$, $z_{obs} = 1.17$. *Dans le cas d'un test unilatéral au seuil de 5%, on ne peut refuser* H_0.

Exercice 20 *Dossier "pedago"*
Lors d'une expérience pédagogique, on s'intéresse à l'effet comparé de deux pédagogies des mathématiques chez deux groupes de 10 sujets :

– pédagogie traditionnelle (p_1)
– pédagogie moderne (p_2)

On note la performance à une épreuve de combinatoire.

p_1 traditionnelle	
s1	5.0
s2	4.0
s3	1.5
s4	6.0
s5	3.0
s6	3.5
s7	3.0
s8	2.5
s9	1.5
s10	2.5

p_2 moderne	
s11	4.0
s12	5.5
s13	4.5
s14	6.5
s15	4.5
s16	5.5
s17	1.0
s18	2.0
s19	4.5
s20	4.5

1) Vérifier que les paramètres des deux échantillons sont donnés par :

	p_1	p_2
Moyenne	3.250	4.250
Ecart-type	1.365	1.553
Variance	1.863	2.413
Ecart-type corrigé	1.439	1.637
Variance corrigée	2.069	2.681

2) Ces données expérimentales permettent-elles d'affirmer que la pédagogie a un effet sur les résultats à l'épreuve de combinatoire ?
Réponse : $t_{obs} = -1.45$. *Dans le cas d'un test unilatéral au seuil de 5%, on ne peut refuser* H_0.

Exercice 21
On veut savoir si une maladie M modifie le taux de certaines protéines dans le sang. On a mesuré leur concentration dans un échantillon de sujets atteints par M et dans un autre échantillon formé de sujets en bonne santé (sujets témoins). Les résultats (dans une unité convenable) sont les suivants :

	Effectifs	Moyenne échantillon	Variance échantillon
Malades	77	141	40
Témoins	33	131	32

Tester l'hypothèse "taux identiques chez les malades et les témoins" contre l'hypothèse :
a) "taux différent chez les malades et les témoins",
b) "taux supérieur chez les malades".
Réponses : $z_{obs} = 8.094$. *Dans les deux cas,* H_0 *est refusée.*

Exercice 22
Dans une étude publiée en 1996, les auteurs se sont intéressés au rôle de la valeur affective d'un texte dans la récupération du souvenir chez les personnes âgées. L'expérience a porté sur 20 sujets ; 10 d'entre eux présentent un déficit mnésique, les 10 autres n'en présentent pas.
Dans un premier temps, on veut s'assurer que les deux groupes ainsi constitués :
- d'une part, sont homogènes du point de vue de l'âge
- d'autre part obtiennent des scores significativement différents au test "Wescher Mémoire".
Les données relatives aux sujets des deux groupes figurent dans les tableaux ci-dessous.

Sujets déficitaires		
	Age	Wescher Mémoire
Sujet 1	80	66
Sujet 2	91	59
Sujet 3	82	84
Sujet 4	87	68
Sujet 5	82	80
Sujet 6	85	75
Sujet 7	84	72
Sujet 8	85	82
Sujet 9	88	78
Sujet 10	87	76

Sujets non déficitaires		
	Age	Wescher Mémoire
Sujet 11	80	113
Sujet 12	81	94
Sujet 13	82	87
Sujet 14	84	98
Sujet 15	85	103
Sujet 16	85	110
Sujet 17	86	97
Sujet 18	89	119
Sujet 19	91	88
Sujet 20	92	91

1) Calculer la moyenne et l'écart type des variables âge et Weschler mémoire pour chacun des deux groupes.
2) Comparer les deux groupes du point de vue de l'âge à l'aide d'un test de comparaison de moyennes.
3) De même, comparer les deux groupes du point de vue de la seconde variable.
Réponses : 1) Pour la variable âge : $\overline{x_1} = 85.1$, $\overline{x_2} = 85.5$, $s_1 = 3.11$, $s_2 = 3.88$. *Pour la seconde variable,* $\overline{x'_1} = 74$, $\overline{x'_2} = 100$, $s'_1 = 7.42$, $s'_2 = 10.40$.
2) On réalise un test de comparaison de moyennes sur deux groupes indépendants. Compte tenu de la taille des échantillons, c'est la loi de Student qui s'applique. Ici, $n_1 = n_2 = 10$ *(groupes équilibrés) ; ddl* $= 18$. *Au seuil de 5% unilatéral,* $t_{crit} = 1.7341$. *On obtient ici* $t_{obs} = 0.24$. *On retient donc* H_0 *et on conclut à l'homogénéïté des âges dans les deux groupes.*
3) Test analogue au précédent. On obtient ici $t_{obs} = 6.10$*, ce qui significatif d'une différence aux seuils traditionnels.*

Exercice 23

Le modèle de la mémorisation proposé par Craik et Lockhart (1972) stipule que le degré auquel un sujet se rappelle un matériel verbal est fonction du degré auquel ce matériel a été traité lors de sa présentation initiale. Eysenck (1974) voulait tester ce modèle et examiner s'il pouvait contribuer à expliquer certaines différences relevées entre des sujets jeunes et âgés concernant leur aptitude à se rappeler du matériel verbal. L'étude qu'il a menée incluait 50 sujets dont l'âge se situait entre 18 et 30 ans et 50 sujets compris dans la tranche d'âge 55–65 ans. Dans chacune des tranches d'âge, Eysenck a réparti les 50 sujets dans cinq groupes. Le premier devait lire une liste de mots et se contenter de compter le nombre de lettres de chacun d'eux. Le deuxième groupe devait lire chaque mot et lui trouver une rime. Le troisième groupe devait donner un adjectif qui aurait pu être utilisé pour modifier chaque mot de la liste. Le quatrième devait essayer de se former une image précise de chaque mot. Aucun de ces quatre groupes ne savait qu'il faudrait se rappeler les mots ultérieurement. Enfin, le cinquième groupe, ou groupe d'apprentissage intentionnel, devait lire la liste et mémoriser tous les mots. Après avoir passé trois fois en revue la liste de 27 mots, les sujets devaient retranscrire tous les mots dont ils se souvenaient. Le nombre de mots rappelés par chacun des 100 sujets est indiqué par le tableau 1 page 13.
1) On veut étudier chez les sujets âgés s'il existe une différence de performance entre le groupe 2 (traitement syntaxique) et le groupe 3 (traitement sémantique), en faveur de ce dernier. Le calcul permet d'obtenir les résultats de statistiques descriptives suivants :

	Groupe 2	Groupe 3
Moyenne	6.9	11.0
Effectif	10	10
Ecart-type	2.02	2.37
Ecart-type corrigé	2.13	2.49

2) Etudier de même s'il existe une différence de performance due à l'âge parmi les sujets du groupe 2.
Réponses : 1) Il s'agit d'une comparaison de moyennes sur deux groupes indépendants. Avec les données fournies, on obtient $t_{obs} = -3.80$ *alors que, pour un test bilatéral au seuil de 5%, on obtient* $t_{crit} = 2.10$ *(ddl* $= 18$*). Il existe donc une différence significative entre le traitement syntaxique et le traitement sémantique.*

	Gr. 1	Gr. 2	Gr. 3	Gr. 4	Gr. 5
Sujets âgés	9	7	11	12	10
	8	9	13	11	19
	6	6	8	16	14
	8	6	6	11	5
	10	6	14	9	10
	4	11	11	23	11
	6	6	13	12	14
	5	3	13	10	15
	7	8	10	19	11
	7	7	11	11	11
Sujets jeunes	8	10	14	20	21
	6	7	11	16	19
	4	8	18	16	17
	6	10	14	15	15
	7	4	13	18	22
	6	7	22	16	16
	5	10	17	20	22
	7	6	16	22	22
	9	7	12	14	18
	7	7	11	19	21

TAB. 1 – Données Eysenck

2) Les sujets du groupe 2 forment deux sous-groupes indépendants du point de vue de l'âge. On obtient $t_{obs} = -0.766$. La différence de performance n'est pas significative dans ce cas.

Exercice 24 Un débat télévisé est organisé entre deux candidats à une élection. Un sondage fait auprès d'un échantillon de 200 électeurs a eu lieu avant le débat ; 95 électeurs déclaraient alors vouloir voter pour le candidat A.
Un autre sondage est effectué après le débat. Sur 150 électeurs interrogés, 84 déclarent alors vouloir voter pour le candidat A.
La proportion d'intentions de vote pour A a-t-elle été influencée par le débat ? Répondre à cette question à l'aide d'un test de comparaison de proportions au seuil de 5%.
Réponse. Il s'agit ici d'une comparaison de fréquences sur deux groupes indépendants. On trouve : $f_1 = 0.475$, $f_2 = 0.56$, $p = 0.5114$, d'où $E = 0.054$ et $z_{obs} = -1.57$. Au seuil de 5% pour un test unilatéral, on a : $z_{crit} = -1.645$. On n'a donc pas mis en évidence d'influence du débat sur les intentions de vote pour A.

Exercice 25 χ^2 de MacNemar
Soit deux questions d'un test, A et B, auxquelles un groupe de 184 sujets a répondu. On voudrait savoir si la fréquence des réussites et des échecs est la même pour A et B. On a :

A : 99 échecs et 85 réussites

B : 64 échecs et 120 réussites.

On sait que 62 sujets ont répondu avec succès aux deux questions.
Justifier le tableau d'effectifs observés suivant :

	B : réussite	B : échec	Total
A : réussite	62	23	85
A : échec	58	41	99
Total	120	64	184

Pour répondre à cette question, le statisticien forme le tableau d'effectifs attendus suivant :

	B : réussite	B : échec
A : réussite	62	40.5
A : échec	40.5	41

Justifier la construction de ce tableau et comparer les deux distributions à l'aide de la distance du χ^2 et d'un test (le nombre de degrés de liberté est ici $ddl = 1$. Pourquoi ?). Vérifier que le χ^2 de MacNemar peut aussi être obtenu par la formule :

$$\chi^2 = \frac{(|\text{effectif d'une case} - \text{effectif de l'autre case}| - 1)^2}{\text{effectif d'une case} + \text{effectif de l'autre case}}$$

Réponses : Sans correction de Yates, on obtient : $\chi^2_{obs} = 15.12$. *La différence est donc significative au seuil de 1%.*
Avec la correction de Yates, on obtient : $\chi^2_{obs} = 14.27$. *En utilisant la formule indiquée ci-dessus, on obtient de même :* $\chi^2_{obs} = 14.27$.

Exercice 26
Lors d'une enquête par questionnaire relative au traitement de la violence dans les établissements scolaires, un groupe de parents est interrogé avant et après la projection d'un film illustrant ce problème. Au total, 45 personnes ont fourni des réponses. On remarque que, avant la projection, 29 personnes pensaient que ces difficultés devaient être traitées en rendant plus sévère le règlement de l'établissement alors que les autres préconisaient des aides individualisées aux élèves concernés. Après la projection, on trouve 18 personnes qui préfèrent une solution de type règlementaire dont 14 avaient déjà cette opinion avant le film.
1) Dresser le tableau de contigence obtenu en croisant la condition (avant ou après la projection) et le type de solution préconisé.
2) L'opinion des sujets est-elle influencée par la projection du film ? Répondre à cette question à l'aide d'un test de comparaison de proportions appariées, puis à l'aide d'un test du χ^2 de Mac Nemar, au seuil de 5% (on suppose l'échantillon de taille suffisante pour pouvoir appliquer les méthodes envisagées).
Réponses. 1)

		Après Projection		
		Règlement	*Aide*	*Total*
Avant projection	*Règlement*	*14*	*15*	*29*
	Aide	*4*	*12*	*16*
	Total	*18*	*27*	*45*

Méthode des proportions appariées : $Z = \frac{15-4}{\sqrt{15+4}} = 2.52$. *Pour un test bilatéral au seuil de 5%,* $z_{crit} = 1.96$. *L'opinion des sujets est donc influencée par la projection du film.*
Méthode du χ^2 *de Mac Nemar (avec correction de Yates) :* $\chi^2_{obs} = \frac{(15-4-1)^2}{15+4} = 5.26$, $ddl = 1$, $\chi^2_{crit} = 3.84$, *et la conclusion reste la même.*

Exercice 27

Une compagnie d'assurances désire évaluer le comportement des automobilistes qu'elle assure. Les statisticiens de la compagnie choisissent 18 conducteurs qu'ils répartissent en deux groupes en fonction de leurs antécédents d'assurés : un groupe "à fort risque d'accident" et un groupe "à faible risque d'accident". Ils étudient les façons de conduire de ces automobilistes sur autoroute, en ville et sur un circuit spécialement aménagé. Ils enregistrent, pour chaque conducteur, le nombre d'erreurs de conduite. Les données observées sont les suivantes :

Sujet	Risque	Autoroute	Ville	Circuit
S1	Faible	4	5	12
S2	Faible	8	7	15
S3	Faible	9	15	17
S4	Faible	5	4	11
S5	Faible	5	12	13
S6	Faible	9	10	18
S7	Faible	1	8	10
S8	Faible	7	9	11
S9	Faible	5	9	13
S10	Faible	2	2	8
S11	Fort	4	6	5
S12	Fort	3	4	3
S13	Fort	9	10	16
S14	Fort	6	9	11
S15	Fort	6	10	8
S16	Fort	3	5	8
S17	Fort	5	7	6
S18	Fort	4	3	2

1) A l'aide d'un test de comparaison de moyennes, comparer le nombre d'erreurs commises par les assurés du groupe "à fort risque d'accident" en ville et sur circuit. Réaliser un test bilatéral au seuil de 5%.

2) Pour la conduite sur autoroute, y a-t-il une différence significative entre les deux groupes ? Pour répondre à cette question, on a réalisé un test de comparaison de moyennes sur deux groupes indépendants à l'aide de R . On a obtenu le résultat suivant :

```
Two Sample t-test
data:  Autoroute by Risque
t = 0.4292, df = 16, p-value = 0.6735
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.969513  2.969513
sample estimates:
mean in group Faible   mean in group Fort
                 5.5                  5.0
```

Interpréter les résultats fournis par R.

Réponses : 2) Le traitement réalisé par R est un test de Student sur 2 groupes indépendants. La phrase "alternative hypothesis ..." indique qu'il s'agit d'un test bilatéral. On obtient $t_{obs} = 0.4292$*, ce qui correspond à un niveau de significativité (bilatéral) de 67%. C'est donc ici l'hypothèse* H_0 *qui doit être retenue : aucune différence entre le groupe "à faible risque" et le groupe "à fort risque" n'a été mise en évidence du point de vue de la conduite sur autoroute.*

Exercice 28

Dans un article publié en 1995, R. Champagnol étudie l'assemblage sémantique de l'adjectif épithète et du nom en un syntagme. L'expérience décrite ici s'intéresse à un éventuel effet du placement de l'adjectif, avant ou après le nom.
Dans certaines langues, comme l'anglais, l'adjectif épithète est toujours préposé ; dans d'autres, comme l'arabe écrit, il est postposé. En langue française, la mobilité positionnelle montre que l'intégration sémantique du couple adjectif–nom (ou nom–adjectif) est relativement indifférente à leur ordre d'énonciation, mais on peut se demander si les traitements sont les mêmes dans les deux cas d'ordre, étant donné que dans le discours parlé, mais aussi écrit, les mots sont perçus ou produits dans leur succession linéaire.

Principe de l'expérience : Le principe de l'expérience est le suivant : on présente aux sujets un indice (amorce) suivi d'un texte dans lequel ils doivent détecter un mot cible "amorcé" par l'indice. La cible est un nom dit *spécifique*, sémantiquement analysable en deux ensembles sémantiques dont l'un correspond à un nom *général* et l'autre à un adjectif. Par exemple le nom "une berceuse" peut être traduit par "une chanson douce", le nom "une fissure" par "une petite fente", etc.

L'hypothèse est que si l'intégration sémantique du couple est effectivement plus facile avec l'amorce nom-adjectif, la détection de la cible sera meilleure et plus rapide qu'avec l'amorce adjectif-nom.

Le matériel verbal est composé de 16 textes contenant chacun un nom spécifique défini comme cible. Les textes sont affichés sur l'écran d'un ordinateur par la méthode d'autoprésentation segmentée, qui permet de mesurer le temps de lecture de chaque segment présenté.

On a trois conditions expérimentales définies par l'amorce utilisée : un groupe travaille avec des amorces NA (ex. un homme pauvre), un autre avec des amorces AN (ex. un précieux bijou) et un troisième dit mixte avec moitié NA et moitié AN. Dans ce dernier cas, on a retenu les séquences NA ou AN allant "le mieux" selon les estimations d'un groupe de juges.

Les sujets sont 30 élèves de CM2 (âge moyen 10 ans 9 mois), soit 10 sujets (5 filles et 5 garçons) par condition.

Le dispositif expérimental permet de relever le temps de lecture de l'amorce, le nombre de détections de la cible et le temps de détection de la cible. Les données relatives au temps de détection de la cible, observées lors d'une réplique de l'expérience, sont reproduites ci-dessous : (cf tableau 2)

1) On compare tout d'abord les temps de détection observés dans la situation NA et ceux observés dans la condition AN. A l'aide d'un test bilatéral de comparaison de moyennes au seuil de 5%, étudier si les temps de détection de la cible sont significativement différents pour ces deux types d'amorces.

	NA	AN	Mixte
	123	137	96
	134	143	102
	160	184	136
	169	192	144
	177	211	157
	195	214	161
	219	241	172
	246	280	195
	331	327	222
	345	342	246
Moyenne	209.9	227.1	163.1
Ecart type	72.85	67.00	45.62
Ecart type corrigé	76.80	70.62	48.09

TAB. 2 – Temps de détection de la cible (en centisecondes)

2) On utilise R Commander pour comparer la condition "amorces mixtes" à chacune des deux autres conditions. Interpréter les résultats fournis par le logiciel :

```
Two Sample t-test
data:  Temps.Detection by Amorce
t = -1.6332, df = 18, p-value = 0.1198
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -107.00110   13.40110
sample estimates:
mean in group Mixte   mean in group N.A
              163.1               209.9

Two Sample t-test
data:  Temps.Detection by Amorce
t = 2.3687, df = 18, p-value = 0.02924
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
   7.236299 120.763701
sample estimates:
   mean in group AN mean in group Mixte
              227.1               163.1
```

Réponses : 1) Dans cette question, on compare les variables X_{NA} et X_{AN}. Il s'agit ici d'une comparaison de moyennes sur deux groupes indépendants. Désignons par μ_{NA} et μ_{AN} les moyennes de X_{NA} et X_{AN} dans les populations parentes. Les hypothèses du test s'écrivent :

– $H_0 : \mu_{NA} = \mu_{AN}$

– H_1 : $\mu_{NA} \neq \mu_{AN}$

La statistique de test est $t = \dfrac{\overline{x_{AN}} - \overline{x_{NA}}}{E}$ *avec* $E^2 = \dfrac{s^2_{c,AN} + s^2_{c,NA}}{n}$. *Elle suit une loi de Student à* $2n - 2$*, c'est-à-dire 18 ddl. Pour un seuil de 5% bilatéral, on lit dans la table :* $t_{crit} = 2.10$. *La règle de décision du test est :*

– *Si* $|t_{obs}| \leq 2.10$*, on retient* H_0 *;*
– *Si* $|t_{obs}| > 2.10$*, on retient* H_1.

Le calcul donne ici : $E^2 = \dfrac{76.80^2 + 70.62^2}{10} = 1088.54$ *et* $E = 32.99$. *D'où* $t_{obs} = \dfrac{227.1 - 209.9}{32.99} = 0.52$. *Par conséquent, on retient* H_0 *: les temps de détection de la cible pour les amorces de type NA et les amorces de type AN ne sont pas significativement différents.*

2) Pour la comparaison entre les amorces de type NA et les amorces mixtes, R obtient $t_{obs} = -1.63$*, et un niveau de significativité (bilatéral) de 0.1198, c'est-à-dire 11.98%. Comme ce niveau de significativité est supérieur au seuil de 5% (utilisé dans tout le reste de l'exercice), on retient l'hypothèse* H_0 *: on n'a pas mis en évidence de différence significative entre les conditions définies par ces deux types d'amorces.*
Pour la comparaison entre les amorces de type AN et les amorces mixtes, R obtient $t_{obs} = 2.3687$*, et un niveau de significativité (bilatéral) de 0.0292, c'est-à-dire 2.92%. Comme ce niveau de significativité est inférieur au seuil de 5%, on conclut à une différence significative entre les deux groupes. Autrement dit, il existe une différence significative entre les conditions définies par ces deux types d'amorces*

Tests non paramétriques

Exercice 29 *Test du* χ^2 *sur un tableau de contingence.*
Lors d'une enquête, on a interrogé 150 personnes prises au hasard sur leurs connaissances en langues étrangères. Les résultats obtenus sont les suivants :

	Hommes	Femmes
Anglais	37	24
Allemand	9	19
Espagnol	16	15
Aucune	20	10

Les connaissances en langues étrangères dépendent-elles du sexe dans la population dont est issu l'échantillon étudié ? On répondra à cette question en effectuant un test au seuil de 5%.
Réponses numériques : $\chi^2_{obs} = 8.48$, $ddl = 3$, $\chi^2_c = 7.815$. La différence entre les deux sexes est donc significative.

Exercice 30
Deux groupes de séropositifs asymptomatiques ont été constitués sur la base d'une affectation au hasard des patients aux traitements. Le premier groupe, traité à l'aide d'un placebo, comprenait 428 personnes ; le second comprenait 453 personnes traitées par 500 mg/jour d'AZT. Après cinquante-cinq semaines de traitement, 33 personnes sous placebo sont tombées malades, contre seulement 11 dans le groupe AZT. *D'après* Le Monde *du 18 avril 1990.*

L'analyse statistique visera à répondre à la question : la prise de 500 mg/jour d'AZT retarde-t-elle l'apparition du sida ?
1) Calculer, pour chacun des deux groupes, la fréquence d'apparition de la maladie. Formuler une conclusion descriptive.
2) Comparer les fréquences d'apparition de la maladie dans les deux groupes à l'aide d'un test de comparaison de proportions au seuil de 1% unilatéral.
3) a) On considère les deux variables statistiques *groupe* (modalités : *Placebo* et *AZT*) et *apparition de la maladie* (modalités : *oui* et *non*). Dresser un tableau de contingence résumant les observations précédentes.
b) Comparer les groupes à l'aide d'un test du χ^2, au seuil de 1%.
Réponses :
1) Les fréquences sont respectivement de $f_1 = \frac{33}{428} = 7.7\%$ *et* $f_2 = \frac{11}{453} = 2.4\%$.
2) On obtient $z_{obs} = 3.60$. *Les fréquences sont donc significativement différentes au seuil de 5% unilatéral (*$z_{crit} = 2.33$*).*
2) Le tableau de contingence, le tableau des effectifs théoriques et celui des contributions au χ^2 *sont donnés par :*

	Placébo	*AZT*	*Total*
oui	*33*	*11*	*44*
non	*395*	*442*	*837*
Total	*428*	*453*	*881*

	Plac.	*AZT*
oui	*21.4*	*22.6*
non	*406.6*	*430.4*

	Plac.	*AZT*
oui	*6.28*	*5.95*
non	*0.33*	*0.31*

Ainsi, $\chi^2_{obs} = 12.89$. *On peut affirmer, au seuil de 1%, l'existence d'une liaison entre le type de traitement et l'apparition de la maladie.*

Exercice 31

Dans une étude devenue classique (1939), des chercheurs ont montré à des enfants noirs une poupée noire et une poupée blanche en leur demandant de choisir celle avec laquelle ils voudraient jouer. Sur 252 enfants, 169 ont choisi la poupée blanche tandis que 83 préféraient la poupée noire.
Un deuxième groupe de recherche a reproduit l'expérience précédente en 1970. Les études n'étaient pas exactement identiques, mais les résultats se sont avérés intéressants : sur 89 enfants noirs, 28 ont choisi la poupée blanche et 61 ont préféré la poupée noire.
Pour les deux études, un test adéquat montre que les enfants ne choisissent pas la poupée au hasard.
Une troisième équipe de chercheurs réunit les résultats précédents dans un tableau de contingence et procède à un test du χ^2 sur ce tableau :

	Exp. 1	Exp. 2
Poupée noire	83	61
Poupée blanche	169	28

Formuler avec précision l'hypothèse nulle et l'hypothèse alternative correspondant à ce test. Réaliser le test et conclure.
Réponses : Les variables mises en jeu dans le tableau de contingence proposé sont d'une part la couleur de la poupée (noire ou blanche), d'autre part le numéro d'ordre de l'expérience (expérience 1 ou expérience 2). L'hypothèse correspondant au test du χ^2 *est donc ici : le choix pour l'une ou l'autre des poupées dépend-il de l'expérience considérée ? ou encore : les deux expériences fournissent-elles des résultats identiques ? On trouve :*

$\chi^2_{obs} = 34.15$; *ddl* $= 1$ *; pour* $\alpha = 5\%, \chi^2_{crit} = 3.84$ *et donc les deux expériences fournissent des résultats différents.*

Exercice 32 *Dossier "Psymot"*
On dispose des résultats de deux groupes à une épreuve de psycho-motricité, l'un des groupes ayant reçu une éducation particulière.

G1 : 38 34 18 45 5 1 21 22 11 7 52 24 93 17 56
G2 : 79 96 76 89 53 39 78 81 99 77 29 27.

1) a) Etudier, à l'aide d'un test de la médiane au seuil de 5%, si l'éducation a une influence sur les scores des sujets.
b) Que devient le résultat du test si l'on choisit un seuil de 1% ?
2) Reprendre la question précédente à l'aide d'un test de Wilcoxon Mann Whitney au seuil de 1% unilatéral.
Réponses : 1) La médiane de l'ensemble des observations est $M = 39$*. Le tableau de contingence obtenu en croisant "position par rapport à la médiane" et "groupe" est le suivant :*

	G1	*G2*	*Tot.*
$\leq M$	*11*	*3*	*14*
$> M$	*4*	*9*	*13*
Tot.	*15*	*12*	*27*

On obtient : $\chi^2_{obs} = 6.24$*. Ce résultat est significatif d'une différence entre les groupes au seuil de 5%, mais n'est pas significatif au seuil de 1%.*
2) $\overline{R_1} = 9.6$, $\overline{R_2} = 19.5$ *et* $z_{obs} = -3.22$*. Au seuil de 1%, on refuse l'hypothèse* H_0*. C'est ici le groupe G2 qui a l'effectif le plus faible. Avec les notations de la table du test de Wilcoxon Mann Whitney, on a :* $W = 234$ *et, au seuil de 1 %,* $W'_s = 216$*. On refuse donc* H_0*.*

Exercice 33
Deux enseignants se sont partagé un lot de 20 copies. L'enseignant A a corrigé $n_1 = 8$ copies, l'enseignant B a corrigé $n_2 = 12$ copies, d'où deux groupes de copies. On range les 20 copies par ordre de mérite. Les résultats sont les suivants :

Correcteur A : 2 4 5 7 9 10 12 17
Correcteur B : 1 3 6 8 11 13 14 15 16 18 19 20.

1) Calculer la moyenne des rangs des copies notées par chacun des correcteurs. Quelle conclusion peut-on formuler au niveau descriptif ?
2) A l'aide d'un test de Wilcoxon-Mann-Whitney, étudier si l'on peut conclure à l'hétérogénéïté des deux groupes aux seuils traditionnels.
Réponses : 1) Noter que l'énoncé donne le protocole des rangs. On obtient : $\overline{R_1} = 8.25$, $\overline{R_2} = 12$ *d'où, au niveau descriptif, une différence au bénéfice du correcteur B.*
2) Avec les notations de la table du test de Wilcoxon Mann Whitney, on a : $N_1 = 8$, $N_2 = 12$, $W = 66$ *et, au seuil de 5 %,* $W_s = 62$, $W'_s = 106$*. On accepte donc* H_0*.*

Exercice 34
On veut comparer les résultats de deux classes de même niveau scolaire en utilisant une épreuve commune de performances linguistiques. On obtient, selon le barème adopté, les résultats suivants :

Classe A : 41, 49, 14, 81, 18, 54, 95, 13, 38, 15, 5, 17, 25, 96, 8.
Classe B : 80, 96, 86, 99, 89, 32, 7, 49, 98, 12, 67, 46, 26, 18, 80.

1) Effectuer un test de la médiane afin de déterminer si les deux classes obiennent des résultats comparables (seuil 5%).
2) Reprendre le problème à l'aide d'un test de Wilcoxon Mann Whitney.
Réponses : 1) $\chi^2_{obs} = 3.33$*, au seuil de 5%,* $\chi^2_{crit} = 3.84$*. On accepte donc* H_0*.*
2) On a ici : $\sum R_A = 192.5$*,* $\sum R_B = 272.5$*,* $N_1 = N_2 = 15$ *d'où* $Z_{obs} = 1.6591$*. Pour un seuil de 5% unilatéral, on a :* $Z_{crit} = 1.645$*. On accepte donc l'hypothèse* H_1 *d'une différence entre les deux groupes, au bénéfice de la classe B. Mais, la différence entre* Z_{obs} *et* Z_{crit} *est ici très faible, et il serait sans doute utile de faire une correction tenant compte des ex aequo.*

Exercice 35

On reprend l'énoncé 15 (chiens à poils longs, chiens à poils courts).
1) On recode les effets en ne retenant que leur signe (+ ou −). A partir du protocole obtenu, procéder au test du signe (N.B. Compte-tenu de la faible taille de l'échantillon, il faut utiliser ici une loi binomiale de paramètres $p = 0.5$ et $n = 15$).
2) Construire le protocole des rangs appliqué au protocole des effets individuels (différence des notes attribuées par un même sujet). L'observation correspondant à un effet nul ne sera pas prise en compte.
3) Procéder au test d'homogénéïté des rangs de Wilcoxon.
Réponses : 1) $D_+ = 14$*,* $D_- = 1$*,* $N = 15$*. Pour une variable* X *distribuée selon une loi binomiale de paramètres* $p = 0.5$ *et* $N = 15$*, on a :* $P(X \leq 1) = 0.00048$*. Sous l'hypothèse* H_0*, on a moins de 5 chances sur 10000 d'observer un échantillon aussi extrême. On accepte donc* H_1*.*
Test de Wilcoxon : $N = 15$*,* $T_+ = 118$*,* $T_- = 2$*. D'après la table, on a, au seuil de 1%,* $T_{m,crit} = 19$*. On accepte donc* H_1*.*

Exercice 36

Nurcombe et *al.* ont mené en 1984 une étude sur les enfants présentant un poids réduit à la naissance (PRN). Ces enfants posent des problèmes particuliers à leurs parents parce qu'ils sont, en apparence, apathiques et imprévisibles ; en outre, ils risquent de connaître des problèmes physiques et comportementaux. L'étude a porté sur trois groupes d'enfants ;

- Un groupe expérimental de 25 enfants PRN dont les mères bénéficiaient d'un apprentissage particulier : elles étaient sensibilisées aux signaux émis par ces enfants, afin de leur permettre de mieux répondre à leurs besoins ;
- Un groupe témoin de 31 enfants PRN dont les mères ne bénéficiaient d'aucun programme particulier ;
- Un groupe d'enfants dont le poids à la naissance était normal.

Le tableau 3 page 22 représente une partie des données observées. Il indique l'indice de développement mental (IDM) à 6 mois et à 24 mois pour le groupe témoin PRN, ainsi que la différence entre les deux indices.
1) Des travaux antérieurs sur de tels enfants laissent penser que les scores IDM des enfants PRN du groupe témoin pouvaient diminuer de façon sensible entre 6 et 24 mois. Réaliser un test de comparaison de moyennes afin de déterminer si les données observées confirment cette hypothèse (choisir un seuil de 5%).
2) Peu satisfait du résultat obtenu à la question 1, un chercheur souhaite comparer les données relatives aux enfants du groupe témoin PRN à l'aide d'un test du signe. Réaliser le test et conclure.

PRN Témoin			
	IDM-6	IDM-24	Diff
s1	124	114	-10
s2	94	88	-6
s3	115	102	-13
s4	110	127	17
s5	116	104	-12
s6	139	104	-35
s7	116	91	-25
s8	110	96	-14
s9	129	104	-25
s10	120	106	-14
s11	105	91	-14
s12	88	102	14
s13	120	104	-16
s14	120	100	-20
s15	116	114	-2
s16	105	109	4
s17	100	109	9
s18	91	119	28
s19	129	91	-38
s20	84	81	-3
s21	91	114	23
s22	116	119	3
s23	100	102	2
s24	113	111	-2
s25	89	80	-9
s26	102	119	17
s27	110	119	9
s28	116	123	7
s29	124	119	-5
s30	126	114	-12
s31	123	132	9
Moyenne	111.00	106.71	-4.29
Var. corrigée	191.8667	167.8129	257.2129

TAB. 3 – Données PRN-Témoin

Réponses : 1) Soit μ_6 et μ_{24} les moyennes de l'IDM à 6 mois et à 24 mois dans la population dont est issu l'échantillon. L'hypothèse de recherche se traduit par les hypothèses statistiques suivantes :
H_0 : $\mu_6 = \mu_{24}$
H_1 : $\mu_6 > \mu_{24}$
Il s'agit d'un test de comparaison de moyennes sur des groupes appariés. La statistique de test est $t = \frac{\overline{d}}{E}$ *avec* $E^2 = \frac{s_c^2}{n}$. *Compte tenu de l'effectif* ($n = 31$)*, on peut utiliser ici une loi de Student avec 30 ddl ou une loi normale. Pour un seuil de 5% (et un test unilatéral avec la zone de rejet "à gauche"), on obtient* $t_{crit} = -1.70$ *(ou* $z_{crit} = -1.65$*).*
On obtient : $E^2 = \frac{257.21}{31} = 8.30$; $E = \sqrt{8.30} = 2.88$; $t_{obs} = \frac{-4.29}{2.88} = -1.49$. *L'hypothèse* H_0 *est donc retenue : l'hypothèse d'une baisse de l'IDM n'a pas été démontrée.*
2) Les hypothèses statistiques s'expriment ici par :
H_0 *: Autant de différences positives que de différences négatives dans la population.*
H_1 *: Dans la population parente, la fréquence des différences négatives est supéieure à 50%.*
Comme $N > 30$*, on peut utiliser l'approximation de la loi binomiale par une loi normale et prendre comme statistique de test :* $Z = \frac{2D - 1 - N}{\sqrt{N}}$ *où* $D = \max(D_+, D_-)$*. Pour un seuil de 5%, on obtient* $z_{crit} = 1.65$*. On a ici 12 différences positives pour 19 différences négatives, d'où :* $z_{obs} = \frac{38 - 1 - 31}{\sqrt{31}} = 1.07$*. On accepte donc l'hypothèse* H_0*, ce qui confirme le résultat de la question 1.*

Exercice 37

Dans une recherche médicale, on a examiné 15 patients atteints de la maladie de Parkinson, dans deux conditions : à jeun et sous médicament L-Dopa. Les résultats ont permis d'établir, pour la variable *Vitesse de la marche*, le protocole suivant des effets individuels.

Sujet	Effet	Sujet	Effet	Sujet	Effet
s1	+0,23	s6	+0,12	s11	+0,30
s2	+0,32	s7	+0,33	s12	+0,10
s3	+0,04	s8	+0,28	s13	-0,12
s4	+0,19	s9	0,00	s14	+0,15
s5	+0,22	s10	+0,12	s15	+0,09

Ordonner le protocole des effets individuels et construire le protocole des rangs. L'observation correspondant à un effet nul ne sera pas prise en compte.
Procéder à un test d'homogénéïté des rangs de Wilcoxon et formuler une conclusion en termes d'effet du médicament (seuil : 5% unilatéral).
Réponse : $T_- = 5$, $T_+ = 100$*. Avec les notations de la table du test de Wilcoxon, on a :* $N = 14$, $T_m = 5$*. Or, au seuil de 5% unilatéral,* $T_{m,crit} = 25$*. On accepte l'hypothèse alternative au seuil de 5% unilatéral.*

Exercice 38

On étudie les comportements agressifs d'un groupe d'enfants ayant des difficultés de comportement, pendant une demi-journée, avant et après la projection d'un film d'aventures. Peut-on dire que ce film a eu une influence sur le comportement agressif des enfants ? (on trouvera dans le tableau 4 page 24 trois colonnes indiquant le numéro du sujet, le nombre de comportements agressifs avant la projection et le nombre de comportements agressifs après la projection). Apporter une réponse à la question posée en utilisant un

Sujet	Avant	Après
1	46	71
2	31	79
3	65	27
4	61	39
5	47	75
6	32	28
7	58	32
8	14	36
9	7	61
10	48	83
11	43	80
12	33	28
13	14	72

Sujet	Avant	Après
14	84	35
15	2	28
16	5	39
17	99	53
18	88	66
19	75	40
20	71	84
21	23	14
22	84	15
23	2	99
24	34	36
25	92	13
26	61	87

Tab. 4 – Comportements agressifs

test de Wilcoxon.
Réponse : $T = T_{-} = 194.5$, $z_{obs} = 0.47$. *On ne peut pas refuser l'hypothèse nulle.*
En utilisant la table du test de Wilcoxon, on a : $N = 26$, $T_{min} = T_{+} = 156.5$ *et, au seuil de 5% unilatéral,* $T_{m,crit} = 110$. *On ne peut donc pas refuser* H_0.

Exercice 39

Dans une crèche, on a observé des enfants en présence de deux puéricultrices différentes ; on a compté, suivant une grille préalablement établie, le nombre de comportements verbaux de ces enfants, pendant une période déterminée. Les résultats sont donnés dans les tableaux suivants. Peut-on dire qu'il y a une différence de comportements verbaux entre les deux observations ?

Séance 1	Séance 2	Diff.	Rang-	Rang+
64	88	24		17
62	42	-20	11	
10	65	55		24
61	87	26		19
17	9	-8	5,5	
20	28	8		5,5
65	87	22		13
10	78	68		27
67	9	-58	25	
25	58	33		22
1	92	91		30
42	36	-6	4	
90	66	-24	17	
87	22	-65	26	
93	4	-89	29	

Séance 1	Séance 2	Diff.	Rang-	Rang+
16	39	23		14,5
16	66	50		23
83	70	-13	8	
45	13	-32	20,5	
33	10	-23	14,5	
21	10	-11	7	
30	54	24		17
67	50	-17	10	
34	2	-32	20,5	
93	10	-83	28	
78	73	-5	2,5	
63	42	-21	12	
49	65	16		9
71	73	2		1
93	98	5		2,5

Répondre à la question posée à l'aide d'un test des signes puis d'un test de Wilcoxon.
Réponses : Test des signes : $D = D_- = 16$, $z_{obs} = 0.18$. *On ne peut pas refuser l'hypothèse nulle.*
Test de Wilcoxon : $T_- = 240.5$, $z_{obs} = 0.15$. *Même conclusion.*
Avec les notations de la table du test de Wilcoxon : $N = 30$, $T_{min} = T_+ = 224.5$. *Au seuil de 5% unilatéral,* $T_{m,crit} = 151$. *On accepte donc* H_0.

Exercice 40

On dit que dans une famille, les aînés ont tendance à être plus indépendants que leurs cadets. Un chercheur élabore une échelle d'indépendance en 25 points et procède à l'évaluation de 20 aînés et du frère ou de la sœur qui suit directement chacun des aînés. Imaginons qu'il obtienne les résultats suivants :

Paire	Aîné	Cadet
1	8	9
2	13	15
3	8	10
4	5	7
5	12	10
6	15	13
7	5	8
8	15	12
9	16	13
10	5	9

Paire	Aîné	Cadet
11	17	13
12	12	8
13	2	7
14	13	8
15	19	14
16	18	12
17	14	8
18	17	11
19	18	12
20	20	10

Quels sont les individus statistiques et les variables étudiés ?
A partir des données ci-dessus, construire le protocole des rangs signés.
N.B. Malgré leur désordre apparent, les données ont été classées de manière à faciliter ce travail. . .
Tester l'hypothèse du chercheur à l'aide d'un test de Wilcoxon, au seuil de 5%.
Réponse : $T_+ = 164$; $T_- = 46$. *En utilisant la table de Wilcoxon, on trouve, au seuil de 5%,* $T_{m,c} = 60$. *L'hypothèse "les ainés sont plus indépendants que leurs cadets" est donc confirmée. On peut aussi utiliser l'approximation par une loi normale. On obtient* $Z_{obs} = 2.18$ *et la conclusion est la même.*