CORRIGÉ - MEDI -

Durée de l'épreuve : 2 heures.
Aucun document n'est autorisé. Seule la calculette est autorisée, sans sa documentation.
Attention : les exercices (encadrés) sont indépendants. Le barème donné à titre indicatif est sur 60 ; la note finale sera donnée sur 20.
La page 9 (extraits des tables des valeurs critiques de la variable c^2, de la variable T de Student et formulaire) peut être détachée et conservée.
Indiquer les réponses exclusivement sur ce document.

Dossier Recherches (14 points)

Source : Herman, J. (1986). Analyse de données qualitatives. Paris : Masson.

Lors d'une étude sur l'efficacité de 1691 unités de recherche choisies au hasard (au niveau international), on a obtenu les résultats suivants en mettant en relation les Sciences étudiées (*Formelles*, *Naturelles*, *Appliquées ou Sociales*) et les méthodes qu'elles utilisent pour l'analyse de leurs données (*Test d'hypothèse, Déduction, Statistique Descriptive ou Modélisation*)

Tableau 1 : tableau de contingence Sciences étudiées / Méthodes utilisées

				Sciences étudiées		
		Formelles	*Naturelles*	*Appliquées*	*Sociales*	
Méthodes utilisées	*Test d'Hypo.*	35	457	178	216	*886*
	Déduction	19	83	49	42	*193*
	Stat. Descrip.	16	101	50	114	*281*
	Modélisation	13	114	72	132	*331*
		83	*755*	*349*	*504*	*1691*

***Question initiale** : les Sciences étudiées ont-elles toutes le même profil quant à l'utilisation des méthodes d'analyse de données ?*

Analyse descriptive

1) On se demande quelle est la Méthode d'analyse de données la plus fréquemment utilisée, indépendamment des Sciences étudiées (toutes Sciences confondues).

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

*Hypo : (886/1691) * 100 = 52%*
*Dédu. : (193/1691) * 100 = 11%*
*St.Desc. : (281/1691) * 100 = 17%*
*Modélis. : (331/1691) * 100 = 20%*

b) Commenter les résultats obtenus :

Par rapport aux autres méthodes, la méthode la plus fréquemment utilisée est le test d'hypothèse (52%), la moins fréquemment utilisée : la déduction (11%). Les statistiques decriptives et la modélisation sont utilisées respectivement à 17% et 20%.

2) Compléter le tableau 2 ci-dessous (case vide) :

Tableau 2 : tableau des effectifs théoriques

		Sciences étudiées			
		Formelles	*Naturelles*	*Appliquées*	*Sociales*
Méthodes utilisées	*Test d'Hypo.*	43.49	**395.58**	182.86	264.07
	Déduction	9.47	86.17	39.83	57.52
	Stat. Descrip.	13.79	125.46	57.99	83.75
	Modélisation	16.25	147.79	68.31	98.65

b) Indiquer précisément comment a été calculée la valeur correspondant à "Test d'hypothèse" pour les "Sciences Naturelles" :

*(755 * 886) / 1691 = 395.58*

c) Expliquer ce qu'indique la valeur 43.49 (Test d'Hypothèse/Sciences Formelles) :

Si la méthode utilisée et la science étudiée étaient indépendantes, on aurait observé 43 (43.49) unités de recherche qui étudient les Sciences Formelles et qui utilisent le test d'hypothèse comme méthode d'analyse de données.

3) Écart à l'indépendance

a) Construire le tableau des attractions/répulsions ci-après :

Tableau 3 : tableau des attractions/répulsions

		Sciences étudiées			
		Formelles	*Naturelles*	*Appliquées*	*Sociales*
Méthodes utilisées	*Test d'Hypo.*	-	+	-	-
	Déduction	+	-	+	-
	Stat. Descrip.	+	-	-	+
	Modélisation	-	-	+	+

b) Indiquer précisément comment les méthodes d'analyse de données diffèrent selon le type de recherche pratiquée ?

Le test d'hypothèse est plus fréquemment utilisé par les Sciences naturelles. La déduction est plus fréquemment utilisée par les Sciences Formelles et Appliquées. Les Statistiques descriptives sont plus fréquemment utilisées par les Sciences Formelles et Sociales et la modélisation par les Sciences Appliquées et les Sciences Sociales.

4) Écart global à l'indépendance

a) On trouve Phi² = 0.0432. Interpréter cette valeur :

On calcule le V² de Cramer. $\mathbf{V^2} = \frac{0.0432}{3} = 0.0144$ *. Avec Phi² max = plus petite dimension du tableau moins 1. V² < 0.04 ; ce qui indique une liaison faible entre Sciences étudiées et Méthodes utilisées.*

b) Rédiger une conclusion descriptive :

Pour ces 1691 unités étudiées, bien que l'on constate localement quelques préférences (par ex : Sciences formelles/Déduction et Description ; Naturelles / Test d'hypothèse,...), la liaison globale entre les 2 variables peut être considérée comme faible (V² = 0.0144 < 0.04)

5) <u>Analyse inférentielle</u> : on désire généraliser les résultats obtenus sur cet échantillon à une population plus vaste.

a) Donner le nom du test à utiliser, indiquer la formule et développer le calcul :

Test $c^2 = n \times f^2 = 1691 \times 0.0432 = 73.05$

(valeur exacte : $c^2 = 73.0607$*)*

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

ddl = (J - 1) (K -1) = (4 - 1) (4 -1) = 9

c) Indiquer le résultat du test :

$c^2_{obs} = 73.05 > c^2_{[9].001} = 27.88$

Résultat (très) significatif au seuil .001 (on peut rejeter H_0 *au seuil .001)*

d) Quelle est la condition préalable pour pouvoir donner une interprétation fréquentiste, en terme de test d'hypothèse, de ce résultat ? Est-elle vérifiée ?

<u>L'échantillonnage au hasard dans la population parente</u>. <u>Cette condition est vérifiée (1691 unités de recherche choisies au hasard)</u>.

6) Les résultats trouvés en ***4-a*** et ***5-c*** paraissent contradictoires. Indiquer pourquoi ils sont compatibles :

On pourrait s'étonner de constater à la fois <u>un effet faible descriptivement et un test significatif</u>. Le constat d'un test significatif permet <u>seulement de se prononcer sur l'existence d'un effet</u> dans la population parente (et non sur l'importance). Cet effet peut à la fois exister (différent de 0) et être faible. Ici le test est significatif du fait de <u>l'effectif important</u> : <u>Khi² est sensible aux effectifs et Phi² ne l'est pas</u>.

Dossier Formules (4 points)

La formule générale du coefficient de corrélation de Bravais-Pearson est :

$r = \dfrac{Cov(x,y)}{Ety(x) \times Ety(y)}$. Sachant que $Cov(x,y) = \dfrac{\sum (x - \bar{x})(y - \bar{y})}{n}$, montrer que dans le cas où X et Y sont des variables centrées-réduites, r peut s'écrire $\dfrac{\sum xy}{n}$:

Comme Ety(x) = Ety(y) = 1, alors r = Cov(x,y)

D'autre part, on a : $\bar{x} = 0$ *et* $\bar{y} = 0$ *,*

donc $r = Cov(x,y) = \dfrac{\sum (x - \bar{x})(y - \bar{y})}{n} = \dfrac{\sum xy}{n}$

Dossier UE (12 points)

On s'intéresse à la liaison entre les résultats obtenus à deux UE du Deug Sciences Humaines et Sociales de l'Université René Descartes : l'UE de Psychologie Expérimentale (EXPÉ) et l'UE de Psychologie Différentielle (DIFF) pour un échantillon de 9 étudiants de Premier Cycle, choisis au hasard dans la population des étudiants en Premier Cycle du Centre Universitaire de Boulogne. Les résultats sont les suivants (notes sur 20) :

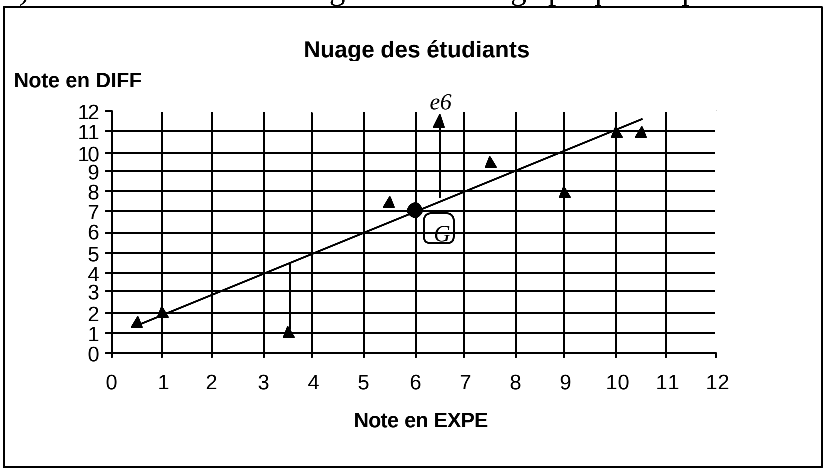
	e1	*e2*	*e3*	*e4*	*e5*	*e6*	*e7*	*e8*	*e9*
EXPÉ	3.5	9.0	10.0	0.5	5.5	6.5	7.5	1.0	10.5
DIFF	1.0	8.0	11.0	1.5	7.5	11.5	9.5	2.0	11.0

1) Le coefficient de corrélation de Bravais-Pearson entre ces deux variables est égal, arrondi à 2 décimales, à +.88. A l'aide de la calculette, recalculer cette valeur avec une précision de *6 décimales* :
$r_{(EXPÉ,\ DIFF)} = 0.878786$

2) Que deviendrait ce coefficient si la note à l'UE d'EXPÉ était donnée sur 40 et non sur 20, alors que la note en DIFF resterait sur 20 ? Justifier.
Le coefficient de corrélation reste <u>inchangé</u>, il est <u>invariant aux transformations linéaires</u>, comme c'est le cas ici où on a multiplié les notes d'EXPÉ par 2.

3) On cherche à estimer la note en DIFF à partir de la note en EXPÉ. Calculer et indiquer ci-dessous l'équation de la droite de régression (utiliser la calculette ou les formules indiquées page 9 ; *présentation des résultats : 2 décimales*) :
$\widetilde{Diff} = 1.03\, Expé + 0.81$ ou : $\tilde{y} = 1.03\, x + 0.81$

4) Tracer cette droite de régression sur le graphique ci-après :



5) Représenter le centre de gravité de ce nuage sur le graphique ci-dessus et indiquer ses coordonnées ci-après : *G (6.00 ; 7.00)*

6) Si l'on représentait l'autre droite de régression, qui exprime EXPÉ en fonction de DIFF, sur le graphique ci-dessus, quel serait le point d'intersection des deux droites de régression ?
G (6.00 ; 7.00) : le centre de gravité du nuage

7) On cherche à apprécier, par un indice global, la qualité de l'ajustement linéaire.

a) Indiquer le nom de cet indice et donner sa valeur :

<u>Le coefficient de détermination</u> : <u>R² = 0.77</u>

b) Conclure descriptivement :

<u>Pour ces 9 étudiants en DSH</u>, il existe une <u>corrélation forte et positive</u> entre la note obtenue en Expé et la note obtenue en Diff : <u>si un étudiant obtient une note élevée en Expé, il aura tendance à obtenir une note élevée en Diff</u>. Cette corrélation peut être considérée comme <u>importante</u> (0.88 > 0.40) - ou la qualité de l'ajustement linéaire est de 77% (>16%) -

8) Représenter sur le graphique précédent les écarts résiduels des étudiants *e1* et *e6*.

9) On cherche maintenant à généraliser les résultats obtenus par cet échantillon à l'ensemble des étudiants de premier cycle. Indiquer le nom et la formule du test à utiliser ainsi que le nombre de degrés de liberté associé à ce test (sans effectuer de calcul) :

Test T de Student pour l'inférence sur une corrélation :

$t_{obs} = \sqrt{n-2} \times \sqrt{\frac{r^2}{1-r^2}}$ *ou formule équivalente (donnée en 2002) : tobs =* $\sqrt{n-2} \times \frac{r}{\sqrt{1-r^2}}$

ddl = n - 2

10) La valeur obtenue pour ce test est égale à : 4.87. Indiquer le résultat du test :

$t_{obs} = 4.87 > t_{[7].01} = 3.499$ *; résultat significatif au seuil .005 unilatéral*

Dossier MAC (5 points) HORS PROGRAMME 2003-2004

D'après : Frugier, G. (1992). *Exercices ordinaires de probabilités*. Paris : Marketting, Ellipses.

Olivier est étudiant. Ses revenus l'obligent à se nourrir exclusivement de hamburgers. Il va chez Mac A dans 70% des cas et chez Mac B le reste du temps. Il attrape une gastrite dans 10% des cas chez Mac A et dans 20% des cas chez Mac B.
Aujourd'hui il a une gastrite. Où a-t-il pris, le plus probablement, son dernier hamburger ?

1) Appliquer l'une ou l'autre des deux formules du théorème de Bayes (cf. ci-dessous) pour répondre à la question posée :

$$P(j/k) = \frac{P(j) \times P(k/j)}{\sum_j P(j) \times P(k/j)} = \frac{P(j) \times P(k/j)}{P(k)}$$

	Gastrite	*Pas de Gastrite*
Mac A : 70%	*10%*	*90%*
Mac B : 30%	*20%*	*80%*

$$P(A/G) = \frac{P(A) \times P(G/A)}{P(A) \times P(G/A) + P(B) \times P(G/B)} = \frac{0.70 \times 0.10}{(0.70 \times 0.10) + (0.30 \times 0.20)} = \frac{0.07}{0.13} = 0.54$$

ou : $P(A/G) = \frac{P(A) \times P(G/A)}{P(G)} = \frac{0.70 \times 0.10}{0.13} = 0.54$

Dans 54% il a attrapé sa gastrite chez Mac A (et dans 46% des cas chez Mac B). Il a donc pris son dernier hamburger plus probablement chez Mac A.

2) En supposant que le nombre total de "repas" mangés est de 100, construire le tableau de contingence qui répartit les hamburgers selon ces deux variables : Mac et Gastrite.

	Gastrite	Pas de Gastrite	Total
Mac A	*7*	*63*	*70*
Mac B	*6*	*24*	*30*
Total	*13*	*87*	*100*

3) Vérifier à partir du tableau construit ci-dessus le résultat trouvé en 1) et indiquer le détail des calculs :

P (A/G) = 7/13 = 0.54 *ou P(B/G) = 6/13 = 0.46*

Dossier Stratégies (10 points)

On s'intéresse aux stratégies de résolution d'une tâche par 60 enfants, choisis au hasard et testés deux fois à un an d'intervalle. Deux stratégies sont observées : une stratégie A performante et une stratégie B peu performante. On obtient le nombre de réussites suivant pour chaque stratégie utilisée et chaque tranche d'âge testée :

		13 ans		
		Stratégie A	***Stratégie B***	
12 ans	***Stratégie A***	30	15	*45*
	Stratégie B	12	3	*15*
		42	*18*	*60*

On jugera un effet faible s'il est inférieur ou égal à 5 points de pourcentage, important s'il est supérieur ou égal à 10 points de pourcentage.

On fait l'hypothèse que la fréquence d'utilisation de la stratégie A (la plus performante) évolue entre 12 ans et 13 ans.

1) La structure de ce tableau de contingence est-elle de type : S <G> ou S * T ? Justifier.

*S * T car ce sont les mêmes sujets qui passent 2 tests à un an d'intervalle*

2) Calculer l'effet de l'âge sur la fréquence d'utilisation de la stratégie A dans cet échantillon :

$f_{12ans} = \frac{30+15}{60} = 0.75$ $f_{13ans} = \frac{30+12}{60} = 0.70$ [ou dobs = $\frac{15-12}{60} = 0.05$]

dobs = 0.75 - 0.70 = 0.05 Þ 5%

3) Élaborer une conclusion descriptive :

Pour ces 60 enfants, contrairement à ce que l'on pouvait attendre, le pourcentage de réussite à la stratégie A est légèrement plus élevé à 12 ans (75%) qu'à 13 ans (70%). L'effet de l'âge (dobs) est égal à 5 points de pourcentage. Cette différence peut être considérée comme faible selon la valeur-repère énoncée (effet £ 5 points de %)

4) On désire généraliser les résultats à la population des enfants âgés de 12 ans et 13 ans.

a) Indiquer le nom du test à utiliser, donner sa formule et développer le calcul :

Test c² de Mac Nemar corrigé : $c^2_{obs} = \frac{(|n_{+-} - n_{-+}| - 1)^2}{n_{+-} + n_{-+}}$

$c^2_{obs} = \frac{(|15-12|-1)^2}{15+12} = \frac{4}{27} = 0.148$ *(ou 0.15)*

b) Indiquer le résultat du test :

$c^2_{obs} = 0.148 < c^2_{[1]\ .05} = 3.84$;

résultat non significatif au seuil .05 ; on ne peut pas rejeter l'hypothèse nulle (d'absence d'évolution de la stratégie A entre 12 ans et 13 ans)

c) Pour l'ensemble des enfants âgés de 12 ans et 13 ans, peut-on penser qu'il existe une évolution de la fréquence d'utilisation de la stratégie A ?

On ne peut pas conclure (constat d'ignorance). En effet, un test non significatif signifie qu'il n'est pas possible de généraliser l'effet observé à une population plus vaste. Cela ne signifie pas qu'il n'existe pas d'effet dans cette population.

Dossier Mémoire (15 points)

D'après : Richard, J.F., Lecoutre, M.P. (1994). Inférence statistique. In R. Ghiglione & J.F. Richard (Eds.), *Cours de Psychologie (Tome 4 : Mesures et analyse).* Paris : Dunod, pp. 387-482.

Deux groupes de sujets : 7 adultes étudiants et 6 enfants de Cours Moyen, choisis au hasard, ont à résoudre une suite de problèmes. Dans chaque problème, ils doivent exécuter mentalement une suite d'additions ; pour le dernier calcul du problème, ils doivent utiliser certains des résultats de calculs précédents et donc rechercher en mémoire les résultats de ces calculs (il peut y avoir selon les cas 2, 3 ou 4 chiffres à récupérer en mémoire). On note l'exactitude de la dernière addition et le temps mis pour faire le dernier calcul. Les données ci-après concernent uniquement la condition "4 chiffres à récupérer en mémoire" pour la variable dépendante "Temps de résolution" (exprimé en demi-secondes) :

Adultes							*Enfants*					
s1	*s2*	*s3*	*s4*	*s5*	*s6*	*s7*	*s8*	*s9*	*s10*	*s11*	*s12*	*s13*
28	23	42	33	22	26	15	74	112	61	134	32	70

Moyenne Générale : 51.692
Variance Totale : 1264.213

***Question initiale** : les enfants mettent-ils plus de temps que les adultes pour faire le calcul dans le cas de 4 chiffres à récupérer en mémoire ?*

***Critère sémantique** : on jugera un effet faible s'il est inférieur ou égal à 15 demi-secondes, important s'il est supérieur ou égal à 30 demi-secondes.*
***Présentation des résultats** : 2 décimales*

A - On cherche à répondre à la question initiale en comparant ces 2 groupes.

1) <u>Analyse descriptive</u>

a) Calculer l'effet moyen du facteur "Âge" :

dobs = 80.500 - 27.00 = 53.500

b) Élaborer une conclusion :

Chez ces <u>13 sujets</u>, <u>le temps</u> (en demi-secondes) mis pour résoudre la suite de problèmes est <u>plus court</u> pour les 7 adultes (27 demi-secondes) que pour les 6 enfants (80.5 demi-secondes). <u>La différence de 53.5</u> demi-secondes peut être considérée comme <u>importante</u> en regard du critère sémantique (> 30 demi-secondes).

2) Calculer la variance inter de ce protocole (indiquer la procédure de calcul) :

$$V_{inter} = \frac{7 \times (27.00 - 51.692)^2 + 6 \times (80.50 - 51.692)^2}{13} = 711.33$$

ou : 1264.213 - 552.88 = 711.33

3) Calculer la variance intra de ce protocole (indiquer la procédure de calcul) :

$$V_{intra} = \frac{(7 \times 64.00) + (6 \times 1123.25)}{13} = 552.88$$ *ou : 1264.213 - 711.33 = 552.88*

5) Analyse inférentielle : on désire pouvoir généraliser les résultats obtenus sur l'échantillon à une population plus vaste.

a) Quel test va-t-on utiliser (indiquer seulement la formule du test et le nombre de degrés de liberté associé à ce test, sans calculer) ?

Test T de Student : $t_{obs} = \sqrt{n-2} \times \sqrt{\frac{V_{inter}}{V_{intra}}}$ *ddl = n - 2*

b) La valeur obtenue pour ce test est égale à : 3.76. Indiquer le résultat du test :

$t_{obs} = 3.76 > t_{[11].01} = 3.106$ *; résultat significatif au seuil .005 unilatéral*

(ou on peut rejeter H_0 au seuil unilatéral .005)

B - Un troisième groupe constitué de 10 sujets (adultes âgés sélectionnés dans une maison de retraite parce qu'ils ont obtenu les meilleures performances à un pré-test) vient participer à l'expérience "Mémoire".

Voici ses résultats : $m^{g3} = 100.100$ $V^{g3} = 409.490$

Par ailleurs, on a calculé pour les 3 groupes de sujets (Adultes/Enfants/Adultes âgés) :

$V_{inter} = 977.915$ $V_{intra} = 490.539$

1) Comment pourrait-on trouver la moyenne générale des 3 groupes (indiquer seulement la procédure, sans effectuer de calcul) ?

En effectuant la moyenne pondérée des 3 moyennes

(*ou :* $\frac{7}{23} \times 27.00 + \frac{6}{23} \times 80.500 + \frac{10}{23} \times 100.100$)

2) Analyse descriptive

a) On aimerait savoir si le facteur "Âge" (sur les 3 groupes) peut expliquer descriptivement les différences individuelles. Calculer la statistique appropriée pour répondre à cette question (donner le nom de la statistique, sa formule, le détail des calculs et le résultat) :

On calcule la statistique Eta-Deux : $h^2 = \frac{V\ inter}{Vtot} = \frac{977.915}{977.915 + 490.539} = \frac{977.915}{1468.454} = 0.67$

b) Interpréter ce résultat :

$h^2 = 0.67$ Þ 67% de la dispersion des temps est prise en compte par le facteur "Âge". Ce qui peut être considéré comme important en regard des valeurs-repères (67% > 16%).

3) Analyse inférentielle

a) Quelle procédure pourrait-on envisager pour généraliser les résultats obtenus à une population plus vaste (indiquer seulement le nom du test) ?

F de Snedecor (ou Anova)

b) Quelles sont les conditions de validité de ce test ?

- Echantillonnage au hasard
- Egalité des variances parentes
- Distributions approximativement normales

c) Répondre à la question inférentielle : peut-on généraliser les résultats à une population plus vaste ?

On ne peut pas répondre à la question inférentielle puisque la première condition de validité n'est pas observée (échantillonnage au hasard) : les personnes âgées ont en effet été choisies et non tirées au hasard pour l'expérience "mémoire".

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

a **ddl**	**.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
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** a **ddl**	**.025** **.05**	**.005** **.01**	**.0005** **.001**
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
11	2.201	3.106	4.437
12	2.179	3.055	4.318
13	2.160	3.012	4.221
14	2.145	2.977	4.140
15	2.131	2.947	4.073

Extrait de la table des valeurs critiques de la variable *F* de Snedecor :

Ddl1	!	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ddl2	*a*															
19	.05	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23
	.01	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30	3.24	3.19	3.15
	.001	15.1	10.2	8.28	7.27	6.62	6.18	5.85	5.59	5.39	5.22	5.08	4.97	4.87	4.78	4.70
20	.05	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20
	.01	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23	3.18	3.13	3.09
	.001	14.8	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24	5.08	4.94	4.82	4.72	4.64	4.56
21	.05	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.22	2.20	2.18
	.01	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.24	3.17	3.12	3.07	3.03
	.001	14.6	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11	4.95	4.81	4.70	4.60	4.51	4.44
22	.05	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15
	.01	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12	3.07	3.02	2.98
	.001	14.4	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99	4.83	4.70	4.58	4.49	4.40	4.33

- Quelques formules...

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

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

$$F = \frac{(n-G)}{(G-1)} \times \frac{V_{inter}}{V_{intra}}$$

$ddl1 = G - 1$ $ddl2 = n - G$

$$Cov(x,y) = \frac{\sum_i (x^i\ y^i)}{n} - \bar{x}\ \bar{y}$$

$$a = \frac{Cov(x,y)}{Var\ x} = r \times \frac{Ety\ y}{Ety\ x}$$

$$b = \bar{y} - a\ \bar{x}$$

$$t^{jk} = \frac{n_{jk} - \hat{n}_{jk}}{\hat{n}_{jk}}$$

$$c^2_{obs} = \frac{(|n_{+-} - n_{-+}| - 1)^2}{n_{+-} + n_{-+}}$$