Université Abderrahmane Mira de Bejaia

Faculté des Sciences Humaines et Sociales

Département Des Sciences et Techniques des Activités Physiques Et Sportives (STAPS)

Support de cours : Statistiques déductives

Niveau : licence

Semestre 04 (licence 02 semestre 02)

Filière : Entrainement Sportif

Unité méthodologique

Crédit : 05

Coefficient :03

Volume horaire semestriel: 42h (21h cours + 21h TD)

Elaboré par Dr Hadji Abderrahmen

Année universitaire 2020/2021

Table des matières

1. INTRODUCTION A LA STATISTIQUE DEDUCTIVE 1

2. LES TESTS D'HYPOTHESES (DIFFERENCES STATISTIQUES) 2

2.1. TESTS D'HYPOTHESES POUR 2 ECHANTILLONS 3
2.1.1. Tests paramétriques 3
2.1.1.1. Test Student pour échantillons indépendants 3
2.1.1.2. Test Student pour échantillons appariés 10
2.1.2. Tests non-paramétriques 14
2.1.2.1. Test de « Mann et Whitney » pour 2 échantillons indépendants 14
2.1.2.2. Test de Wilcoxon pour 2 échantillons appariés 17
2.2. TESTS D'HYPOTHESES POUR PLUS DE 2 ECHANTILLONS 20
2.2.1. Tests paramétriques 21
2.2.1.1. Analyse de la variance (ANOVA) pour des échantillons indépendants 21
2.2.1.2. ANOVA à mesures répétées (ANOVA MR) 26
2.2.2. Tests non-paramétriques 29
2.2.2.1. Kruskal-Wallis pour des échantillons indépendants 29
2.2.2.2. Friedmann pour des échantillons appariés 31

3. LE Z-SCORE ET LE T-SCORE : 34

3.1. LE Z-SCORE 34
3.2. LE T-SCORE 34
3.3. APPLICATION 34
3.4. CLASSER LES INDIVIDUS DANS DES CATEGORIES DE PERFORMANCE 36
3.5. ÉTABLIR DES VALEURS NORMATIVES 37

4. LE COEFFICIENT DE CORRELATION DE PEARSON 39

4.1. LA CORRELATION : CONCEPT 39
4.2. COEFFICIENT DE CORRELATION DE PEARSON : 39
4.3. LA FORMULE : 40
Décision (conclusion du test) : 41
4.4. TAILLE DE L'ECHANTILLON ET SIGNIFICATION 41
4.5. COEFFICIENT DE CORRELATION SOUS EXCEL 42
4.5.1. Coefficient de corrélation avec « Utilitaire d'analyse » 42
4.5.2. Coefficient de corrélation avec « Formule Excel » 43

5. LE KHI2 45

5.1. TEST D'INDEPENDANCE (DEUX VARIABLES QUALITATIVES) 45
5.2. TEST D'HOMOGENEITE (ECHANTILLON ET POPULATION) 48

6. ANALYSE FACTORIELLE 51

6.1. L'ANALYSE EN COMPOSANTES PRINCIPALES (ACP) 51
6.1.1. Procédure ACP sous SPSS 52
6.1.2. Interprétation des résultats de l'ACP 53

Annexe du programme des enseignements de la 2ème année de la licence

domaine "Sciences et techniques des activités physiques et sportives" Filière "Entrainement sportif"

Semestre 4

Unités d'enseignement	Matières – Intitulé	Crédits	Coefficient	Volume horaire hebdomadaire – Cours	TD	TP	VHS (14-16 semaines)	Autre*	Mode d'évaluation – Contrôle Continu	Examen
UE Fondamentale Code : UEF 2.2 Crédits : 18 Coefficients : 12	Didactique des jeux	3	2	1h30		02h00	49h00		X	X
	Karaté	3	2	1h30		02h00	49h00		X	X
	Volleyball	3	2	1h30		02h00	49h00		X	X
	Physiologie de l'effort physique	3	2	1h30	1h30		42h00		X	X
	Planification de l'entrainement	3	2	1h30			21h00		~~X~~	X
	Pédagogie de l'entrainment 2	3	2			04h00	56h00		X	
UE Méthodologique Code : UEM 2.2 Crédits : 9 Coefficients : 5	Statistiques déductives	5	3	1h30			21h00			X
	Mesure et évaluation en sport	4	2	1h30	1h30		42h00		X	X
UE Découverte Code : UED 2.2 Crédits : 1 Coefficients : 1	Déontologie et corruption	1	1	1h30			21h00			X
UE Transversale Code : UET 2.2 Crédits : 2 Coefficients : 2	Médecine du sport et premiers secours	1	1	1h30			21h00			X
	Psychologie du sport	1	1	1h30			21h00			X
	Total semestre 4	30	20	15h00	3h00	10h00	392h00			

السداسي: الرابع
عنوان الوحدة : وحدة التعليم المنهجية
المادة: الإحصاء الاستدلالي

أهداف التعليم:

- القواعد الأساسية والمعارف النظرية والتطبيقية المرتبطة بعلم الإحصاء.
- جعل الطالب قادر على استخدام الوسائل الإحصائية في مجال بحثه والاستدلال.

المعارف المسبقة المطلوبة :

- معرفة القواعد الأساسية للحساب والرياضيات والإحصاء التطبيقي.
- الاطلاع على أهم النظريات المتخلفة في الرياضيات وقواعدها.

محتوى المادة:

1. تعريف الإحصاء الاستدلالي
2. الإحصاء التطبيقي والبحوث العلمية
3. الفروق الإحصائية
4. اختبار ستودانت
5. معامل كاف التربيعي
6. المعامل الزادي
7. تحليل التباين
8. تحليل التباين لأكثر من متوسطين حسابيين
9. معامل بيرسون
10. التحليل العاملي
11. تمارين حول الفروق الإحصائية
12. تمارين حول التحليل العاملي

طريقة التقييم: المتابعة الدائمة والامتحانات.

المراجع باللغة العربية:

1. فؤاد البهي السيد: الإحصاء الوصفي في التربية البدنية والرياضة، 1987.
2. حسن أحمد الشافعي: الإحصاء في البحوث النفسية، 1999.
3. محمد خيري مامسر: التحيل الإحصائي في التربية البدنية والرياضة، 2004.
4. عزام صبري: الإحصاء الوصفي ونظام **SPSS** ، 2006.

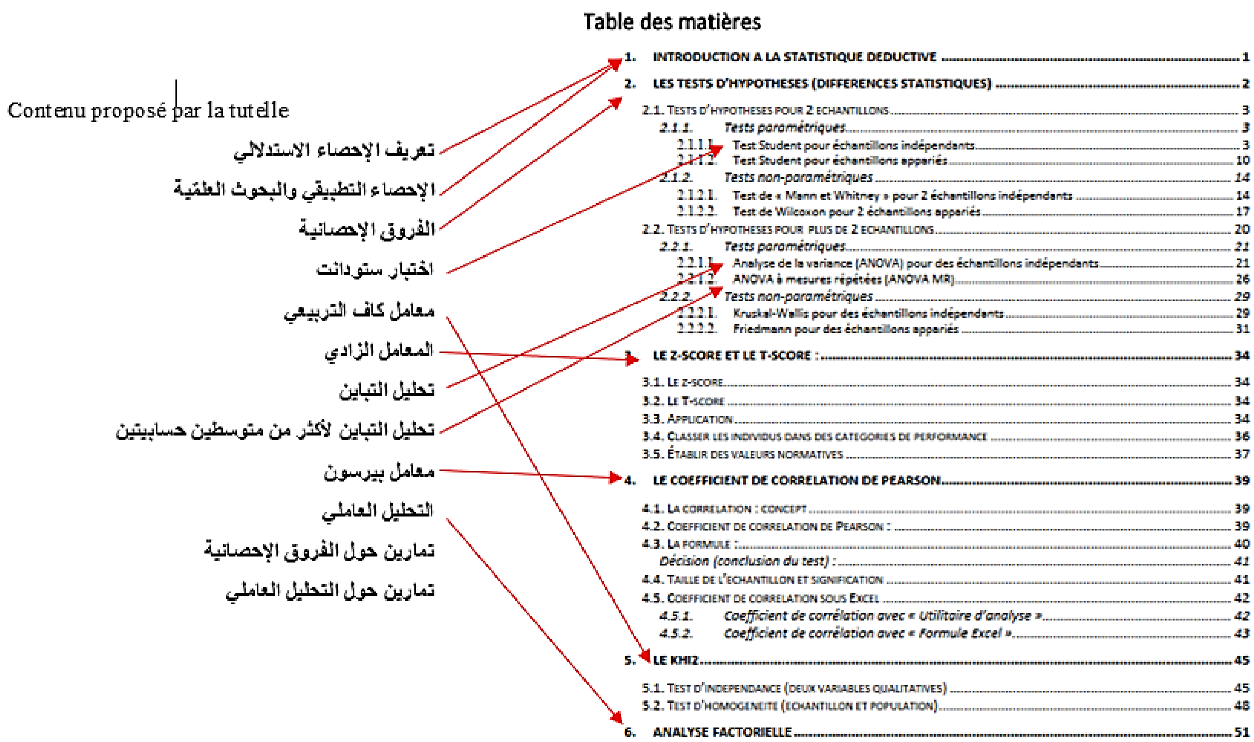
المراجع باللغات الأجنبية:

Contenu de la matière (Traduction)

Traduction	**Contenu proposé par la tutelle**
Définition de la statistique déductives	تعريف الإحصاء الاستدلالي
Statistique appliquée et recherche scientifique	الإحصاء التطبيقي والبحوث العلمّية
Différences statistiques	الفروق الإحصائية
Tests de Student	اختبار ستودانت
Test de Khi2	معامل كاف التربيعي
Z-scores	المعامل الزادي
Analyse de la variance	تحليل التباين
Analyse de la variance pour plus de 2 moyennes	تحليل التباين لأكثر من متوسطين حسابيتين
Coefficient de Pearson	معامل بيرسون
Analyse factorielle	التحليل العاملي
Exercices sur la différence statistique	تمارين حول الفروق الإحصائية
Exercices sur l'analyse factorielle	تمارين حول التحليل العاملي

Concordance du support avec le contenu proposé par la tutelle

Afin d'éviter la redondance et de donner au cours une certaine continuité et progressivité, nous avons apporté quelques ajustements au classement des cours. Les exercices prévus en bas du programme ont été insérés avec leurs techniques respectives. Ce schéma représente la concordance entre ce support et le contenu du programme proposé par la tutelle :



Objectifs de l’enseignement

les règles et connaissances de base théoriques et appliquées liées à la statistique

Rendre l’étudiant capable d’utiliser les techniques statistiques dans son travail de recherche

Prérequis

Connaissances de base en calcul mathématique et statistique appliquée

Connaissances des théories mathématiques basiques et leurs applications

Présentation du cours

Le cours de statistique appliquée est prévu pour le semestre 04 (licence). Il est la suite du cours statistique descriptive (semestre 03). Dans ce cours l’étudiant va découvrir la statistique appliquée à la recherche en STAPS . La conception du cours suit la logique de la recherche scientifique : poser le problème et trouver la technique adéquate pour infirmer ou confirmer les hypothèses. Le cours comporte les techniques les plus utilisées dans la recherche en STAPS : (1) qualitative : Khi2 et analyse factorielle et (2) quantitative : Tests d’hypothèses paramétriques et non paramétriques et la corrélation.

Outils didactiques

En plus des calculs manuels et afin de passer en revue les logiciels statistiques les plus populaires (gratuits et payants), nous avons utilisé une multitude de logiciels (Microsoft Excel, SPSS, Jamovi, Jasp et JMP) pour la démonstration des étapes de la conduite des techniques statistiques ainsi que l’interprétation des résultats obtenus par ces logiciels. Les étapes ont été présentées en détail avec des captures d’écran explicatives.

La représentation graphique, partie importante dans la présentation des résultats, a été utilisée pour l’ensemble des techniques afin d’aider le chercheur à rendre visibles ses résultats. Pour cela, nous avons utilisé Microsoft Excel et Prism GraphPad.

Pour les tests d’hypothèses, nous avons conçu des diagramme expliquant les conditions d’application et de choix pour chaque test.

1. Introduction à la statistique déductive

La collecte , le traitement et l'analyse de l'information et donnée sont au cœur de tous les processus de gestion et de décision. Les méthodes de description, de prévision et de décision se sont considérablement enrichies et développées, ce qui place la statistique appliquée (déductive) au carrefour de l'observation et de la modélisation (Goldfarb & Pardoux, 2011, p. IX).

La statistique déductive consiste à utiliser les données disponibles (récoltés) pour tirer des conclusions ou des décisions permettant une généralisation d'une relation, dépendance, ou différence dans une population donnée. Diverses techniques sont proposées par les chercheurs afin de vérifier des hypothèses scientifiques de la plus simple (Différence) à la plus complexe (Modélisation).

En se basant sur des lois de probabilité (Student, gaussienne, Khi2.etc.), les chercheurs ont proposé des techniques standardisées pour réduire au maximum l'erreur lors de la prise de décision. L'utilisation des indices tels que le seuil de signification ou la taille d'effet permet au chercheur d'affirmer ou infirmer objectivement une hypothèse scientifique.

Par exemple, deux groupes avec des moyennes de 10 et 9.5 peuvent être jugés comme très proches par un individu alors qu'un autre peut estimer qu'ils sont différents. Un test statistique standardisé peut résoudre cette problématique et nous aider à prendre la décision la plus adéquate de l'existence ou non d'une différence. Cette décision est validée et approuvée par la communauté scientifique en tant que « décision valide ».

En fonction du type de variables « Cibles= dépendantes » et « indépendantes=Prédicteurs » et en rapport avec l'hypothèse scientifique, le chercheur opte pour la technique adéquate. Il faut rappeler que nous comptons deux types de variables à savoir qualitative et quantitative. Il est important de signaler que le nombre de modalités de la variable cible lorsqu'elle est qualitative (2 ou plus de 2), oriente le choix de la technique utilisée.

Les études en STAPS s'étalent sur un vaste champ scientifique, elles touchent les sciences biologiques et médicales, sciences humaines et sociales, sciences de gestion, biomécaniques et sa propre discipline (sports). Avec cette multidisciplinarité vient une large gamme de techniques statistiques adaptées à chaque domaine de recherche. Dans ce cours nous essayons de passer sur les plus utilisés et les plus accessibles en prenant en compte le profil des étudiants et des recherches récurrentes.

Nous avons observé que les recherches en STAPS tournent autour des : enquêtes qui sollicitent le Khi2 ou l'analyse factorielle des correspondances multiples ; différences entre populations qui sollicitent le T de Student, ANOVA et leurs homologues de tests non-paramétriques ; relations qui sollicitent le coefficient de corrélation, ACP (analyse en composante principale) et AFD (analyse factorielle discriminante) et tous types de régressions linéaires ou logistiques .

2. Les tests d'hypothèses (différences statistiques)

Dans les mémoires de fin d'études en STAPS, plusieurs études tentent de vérifier la présence d'une différence significative entre des échantillons ou entre sessions.

Pour vérifier ce genre d'hypothèse, nous devons appliquer ce qu'on appelle « les tests d'hypothèse » dans le but de prendre une décision. On entend par décision l'existence ou non d'une différence significative.

Exemple 01 :

Dans notre hypothèse de recherche, nous supposant l'exitance d'une différence dans la variable « Endurance (VMA) » entrent les milieux de terrain et les attaquants en football.

En calculant la moyenne et l'écart-type des deux groupes, nous trouvons ce qui suit :

	Milieux de terrain	Attaquants
VMA (km/h)	18±2.5	16.5±3.8

Selon ces résultats, nous pouvons observer que les milieux de terrain sont plus endurants que les attaquants, mais dans la recherche scientifique nous sommes dans l'obligation de prouver et d'argumenter nos résultats. C'est ici qu'entre les tests d'hypothèse pour nous aider à avancer l'existence ou l'absence d'une différence dite : statistiquement significative.

2.1.Tests d'hypothèses pour 2 échantillons

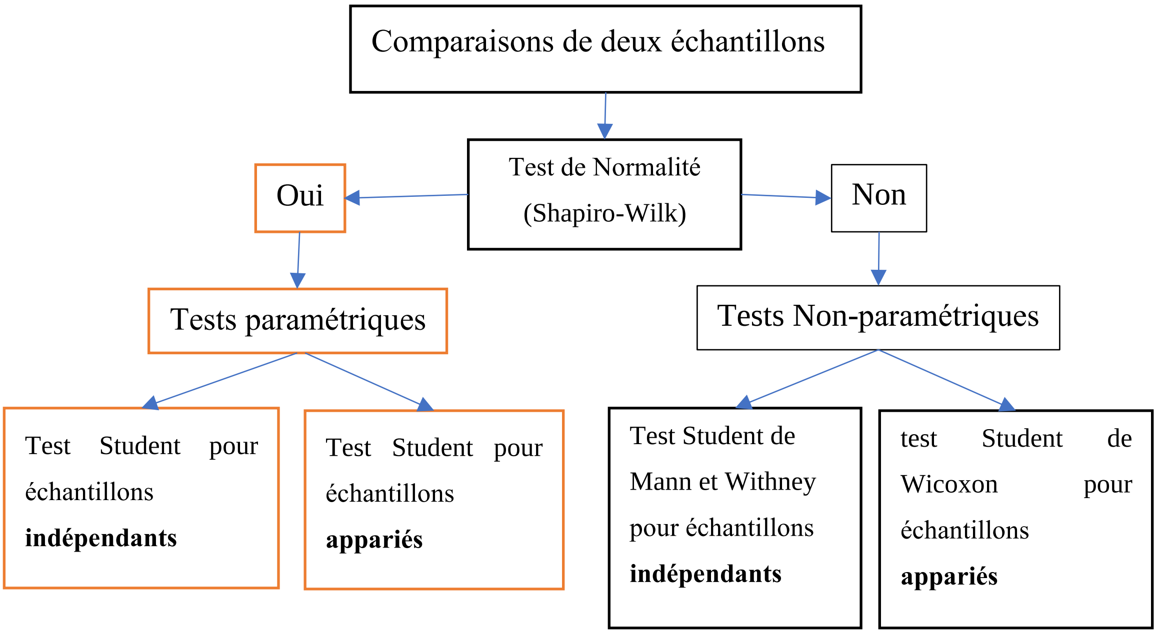


Figure 1. Comparaisons de deux échantillons

2.1.1. Tests paramétriques

2.1.1.1. Test Student pour échantillons indépendants

Échantillons indépendants : dans les études en STAPS, cela signifie deux groupes différents (voir exemple 01).

Pour vérifier l'existence de différence entre les deux groupes, il faut tout d'abord calculer la valeur de t_{cal}, pour la comparer avec celle de t_{tab} .

$$\boxed{t_{Cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}}$$

Où :

$\mathbf{t_{cal}}$: valeur de t calculé

$\bar{x}_1$: moyenne du groupe 01

$\bar{x}_2$: moyenne du groupe 02

s^2 : variance commune

n_1 : nombre d'individus dans le groupe 01

n_2 : nombre d'individus dans le groupe 02

$$s^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

Où :

s_1^2 : variance du groupe 01

s_2^2 : variance du groupe 02

n_1 : nombre d'individus dans le groupe 01

n_2 : nombre d'individus dans le groupe 02

Décision (interprétation) :

Si $t_{cal} > t_{tab}$: il existe une différence statistiquement significative au seuil $\alpha<0.05$.

Où : t_{tab} est la valeur de T Student dans la table, qui correspond à :

DDL : n1 +n2 – 2 et **α : 0.05**

Exemple 02 :

Les résultats d'un test de vitesse sur 100 m pour deux groupes (élèves au CEM) sont présentés dans ce tableau. Le groupe 01 est constitué d'élèves pratiquant une activité sportive extrascolaire, alors que le groupe 02 est constitué d'élèves sédentaires.

groupe 01	groupe 02
12	15
14	16
13	17
12	15
15	16
13	17
13	16
14	15
12	16
10	14
13	12
12	12
14	15
18	13
	14
	13

Problématique de la recherche : l'activité physique extrascolaire peut-elle améliorer la vitesse des élèves ?

Pour répondre à cette problématique, nous supposant l'existence d'une différence significative entre les deux groupes.

Solution :

Pour calculer le t_{cal}, il faut tout d'abord calculer :

1. la moyenne et la variance des deux groupes,
2. la variance commune

Rappel :

La moyenne : $$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$$

la variance : $$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

groupe 01	variance 01	groupe 02	variance 02
12	$(12-13.21)^2$=1.47	15	$(15-14.75)^2$=0.06
14	0.62	16	1.56
13	0.05	17	5.06
12	1.47	15	0.06
15	3.19	16	1.56
13	0.05	17	5.06
13	0.05	16	1.56
14	0.62	15	0.06
12	1.47	16	1.56
10	10.33	14	0.56
13	0.05	12	7.56
12	1.47	12	7.56
14	0.62	15	0.06
18	22.90	13	3.06
		14	0.56
		13	3.06
moyenne01= 13.21	Σ=44.36 s_1^2=44.36/(14-1)=**3.41**	moyenne02 = 14.75	Σ=39.00 s_2^2=39/(16-1)=**2.6**
n_1=14		n_2=16	

La variance commune :

$$s^2 = \frac{s_1^2(n_1-1) + s_2^2(n_2-1)}{n_1 + n_2 - 2}$$

$$s^2 = \frac{3.41(14-1)+2.6(16-1)}{14+16-2} = 2.98$$

La valeur de t calculé :

$$t_{Cal} = \frac{13.21 - 14.75}{\sqrt{\frac{2.98}{14} + \frac{2.98}{16}}} = 2.43$$

Lire la valeur de T_{tab} dans la table (voir table Student)

DDL= 14+16-2=28

DDL (*v*)	0.05	0.01
1	12.7062	63.6567
⋮	⋮	⋮
28	**2,0484**	2.7633

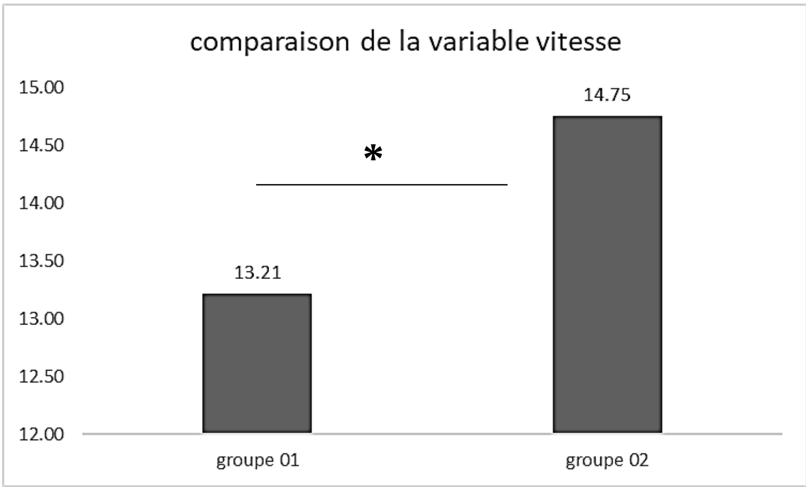
T_{tab}= 2.0484

Tcal (2.43)>T_{tab} (2.0484) : il existe une différence significative entre les deux groupes.

Interprétation :

Puisque la performance en vitesse est inversement liée au temps réalisé, nous pouvons conclure que :

Les élèves qui pratiquent une activité physique extrascolaire (groupe 01 : 13.21) sont plus rapides que les élevés sédentaires (groupe 02 : 14.75) au seuil α<0.05.



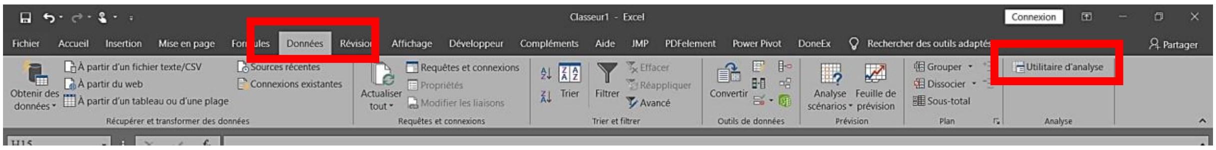
*** : différence significative au seuil α<0.05**

Figure 2.Comparaison du test de vitesse sur 100 m entre le groupe 01 et le groupe 02

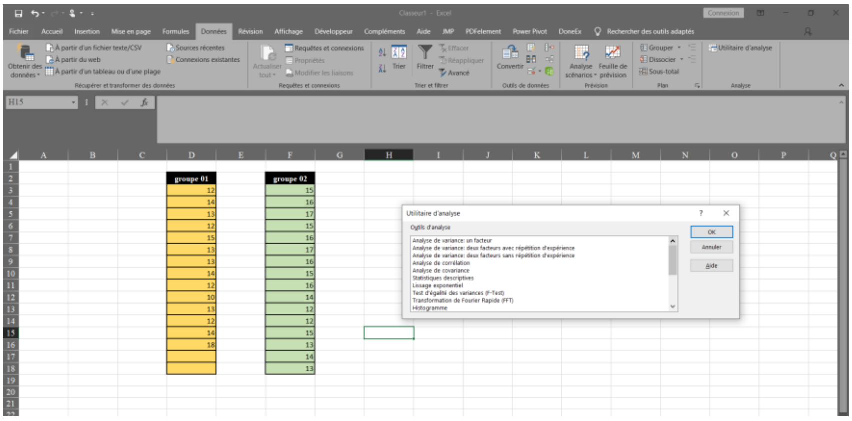
Test de Student pour échantillons indépendants sous Microsoft Excel

Il faut tout d'abord activer l' **Utilitaire d'analyse .** Selon la version de Microsoft Office que vous disposez, cherchez sur Google comment l'activé (un lien sur YouTube : https://www.youtube.com/watch?v=hy-lTUZ4whk).

1. Après l'activation, allez dans l'onglet « **Données** » pour trouver l' **Utilitaire d'analyse.**

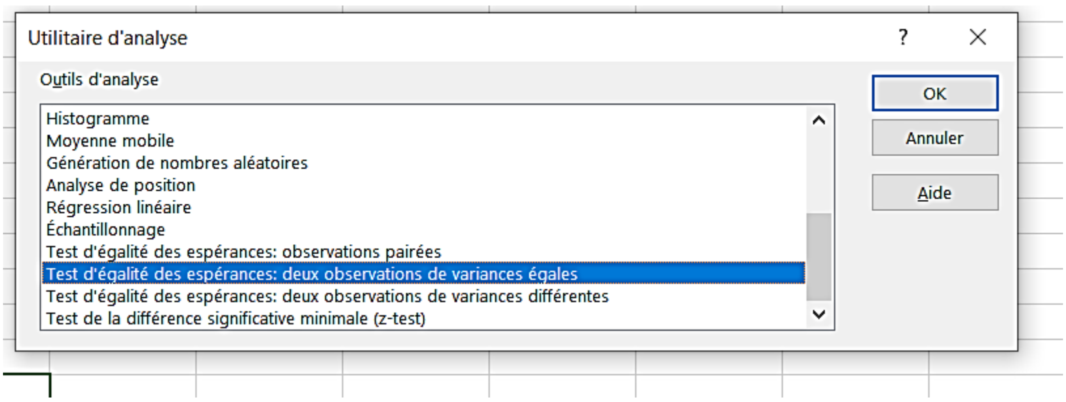


2. Lorsque vous cliquez sur **Utilitaire d'analyse,** une boite de dialogue s'ouvre

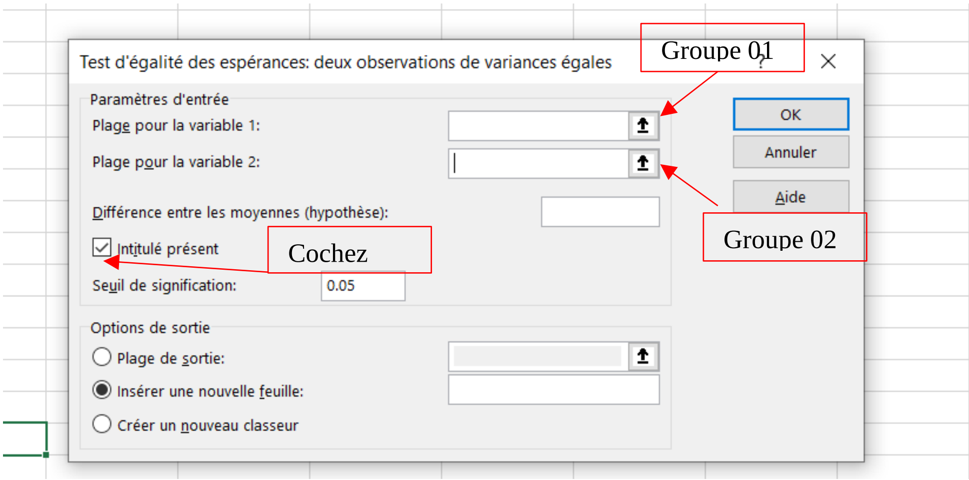


3. Dans cette boite de dialogue, allez tout en bas et choisissez :

   Test d'égalité des espérances : deux observations de variance égale



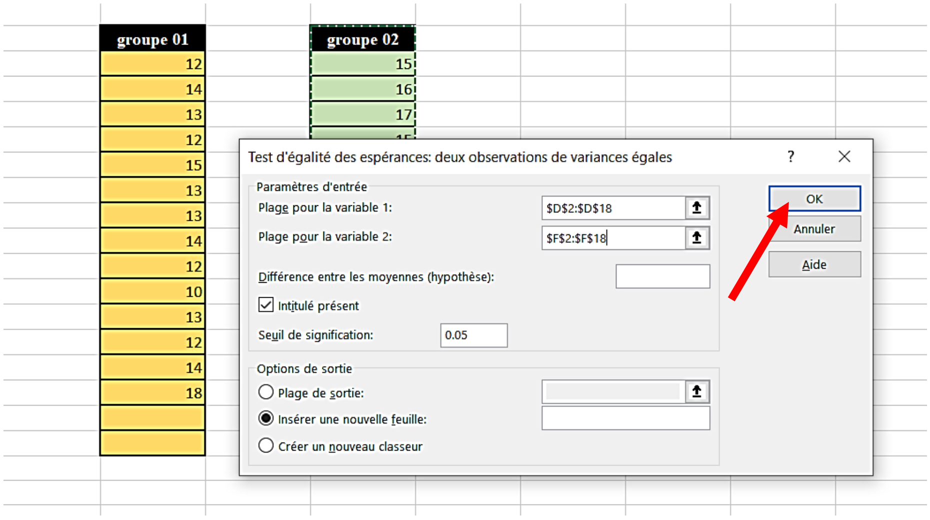
4. En cliquant sur « OK », une autre boite de dialogue s'ouvre



5. **En cliquant sur la petite flèche du groupe 01, choisissez la plage du premier groupe, ensuit cliquez sur le clavier « Entré »**

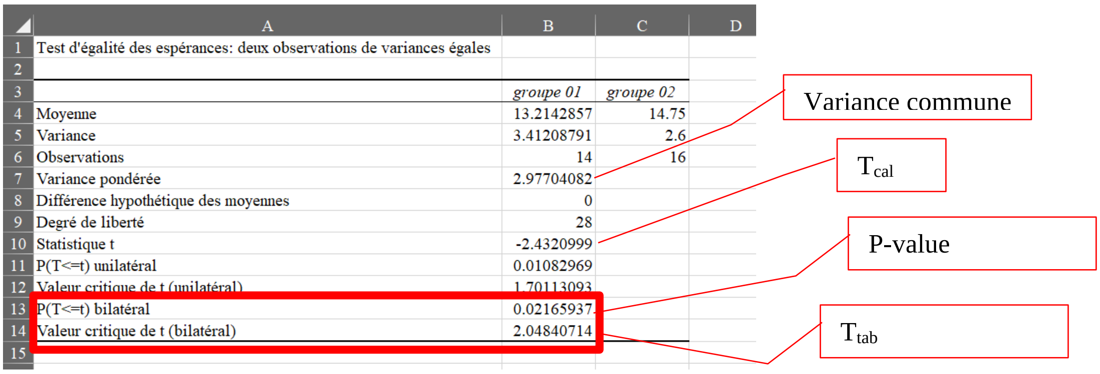
Remarque : l'entête (groupe 01) doit être choisi avec les valeurs

6. **Faire la même chose pour le groupe 02 et cliquez « OK »**



7. Une nouvelle « Feuille » est ajoutée

8. Lecture des résultats :



	A	B	C
1	Test d'égalité des espérances: deux observations de variances égales		
2			
3		*groupe 01*	*groupe 02*
4	Moyenne	13.2142857	14.75
5	Variance	3.41208791	2.6
6	Observations	14	16
7	Variance pondérée	2.97704082	
8	Différence hypothétique des moyennes	0	
9	Degré de liberté	28	
10	Statistique t	-2.4320999	
11	P(T<=t) unilatéral	0.01082969	
12	Valeur critique de t (unilatéral)	1.70113093	
13	P(T<=t) bilatéral	0.02165937	
14	Valeur critique de t (bilatéral)	2.04840714	
15			

Comme vous pouvez le remarquer, les résultats sont identiques à ceux trouvez avec le calcul manuel.

La valeur P-value = 0.021 : cet indice ne peut être calculé manuellement. Il donne la valeur exacte du type d'erreur 01. En termes plus simples, cette valeur nous permet simplement de dire que la différence est significative à P= 0.021 qui se trouve <0.05.

2.1.1.2. Test Student pour échantillons appariés

Échantillons appariés : le même groupe avec un pré-test et post-test

En STAPS, plusieurs études proposent des programmes et des contenus pour impacter une variable donnée. Pour cela, l'étude suit les étapes suivantes :

1. le chercheur réalise un test (pré-test)
2. intervention : application du programme (contenu, exercice, méthodes etc.)
3. le chercheur refait le même test (post-test)

En comparant le pré-test avec le post-test, le chercheur étudie l'évolution de la variable en question. Par conséquent, il peut conclure si cette variable a progressé ou régresser.

Pour vérifier cette évolution, il faut tout d'abord calculer la valeur de t_{cal}, pour la comparer avec celle de t_{tab} .

$$t_{cal} = \frac{\sum D}{\sqrt{\frac{n\sum D^2 - (\sum D)^2}{n-1}}}$$

Où :

D= la différence entre la valeur du pré-test et du post-test **(post-test – pré-test)**

T_{tab} : la valeur qui correspond au DDL : n-1 et $\alpha=0.05$ dans la table de Student (voir la table)

Si valeur absolue de t_{cal} est supérieure à la valeur T_{tab} → il existe une évolution dans le sens de la différence ($T_{cal} >0$= progression ; $T_{cal}<0$= régression)

Exemple01

Dans l'objectif d'évaluer l'efficacité de la méthode d'entrainement en « Intermittent » pour le développement de puissance aérobie, le chercheur a suivi les étapes suivantes :

1. **Réalisé un premier test de VMA au début de la saison (pré-test);**
2. **Appliquer un programme d'entrainement basé sur la méthode d'entrainement en « intermittent » pendant 2 mois ;**
3. **Refaire le même test (post-test)**

Problématique de la recherche : est-ce que la méthode d'entrainement en « intermittent » améliore la qualité d'endurance chez les footballeurs.

Hypothèse : nous supposons que l'application de la méthode d'entrainement en intermittent améliore la qualité d'endurance .

Les résultats de l'expérimentation sont présentés dans le tableau suivant.

individus	pré-test	post-test

1	11	14
2	13	14
3	11	12
4	14	13
5	15	11
6	13	14
7	15	15
8	12	17
9	13	13
10	12	14
11	11	14
12	13	13
13	11	15
14	12	15
15	10	14
16	15	16

Solution :

individus	**pré-test**	**post-test**	**D**	**D²**
1	11	14	14-11=3	3^2=9
2	13	14	1	1
3	11	12	1	1
4	14	13	-1	1
5	15	11	-4	16
6	13	14	1	1
7	15	15	0	0
8	12	17	5	25
9	13	13	0	0
10	12	14	2	4
11	11	14	3	9
12	13	13	0	0
13	11	15	4	16
14	12	15	3	9
15	10	14	4	16
16	15	16	1	1
N=16			ΣD=23 $(\Sigma D)^2$=529	ΣD^2=109

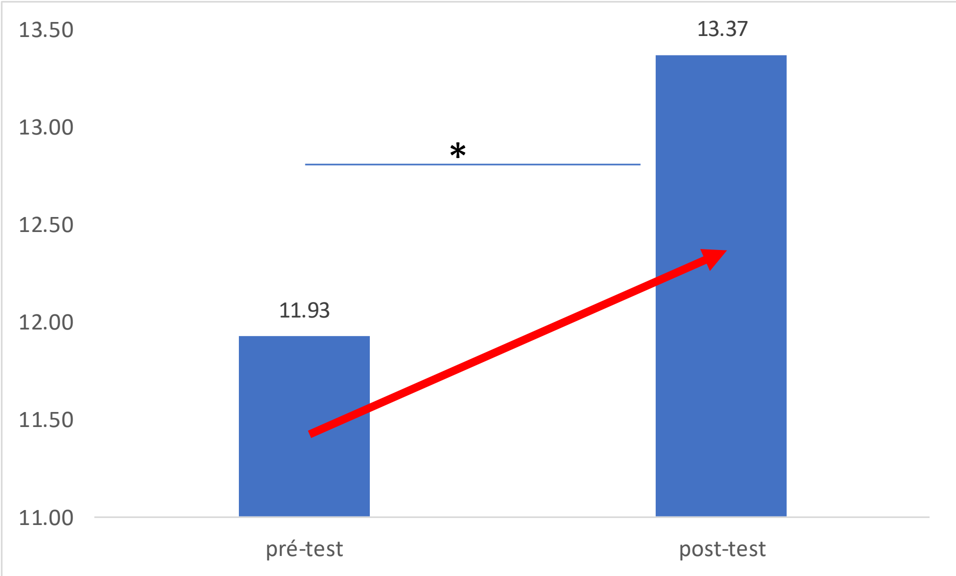
$$t_{cal} = \frac{\sum D}{\sqrt{\frac{n\sum D^2-(\sum D)^2}{n-1}}} \qquad t_{cal} = \frac{23}{\sqrt{\frac{16*190-229}{16-1}}} = +2.56$$

DDL = 16-1=15; α= 0.05 → ***T_{tab} = 2.13***

T_{cal} (2.56) > T_{tab} (2.13) → il existe une évolution significative au seuil α<0.05.

Puisque la valeur de T_{cal} >0, nous concluons qu'il a eu une progression

Conclusion :

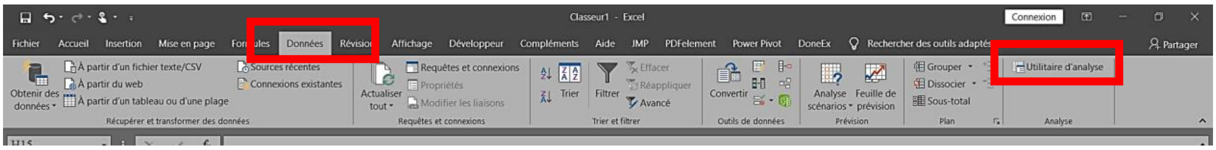


L'application de la méthode d'entrainement en intermittent pendant 2 mois améliore la qualité d'endurance chez les footballeurs.

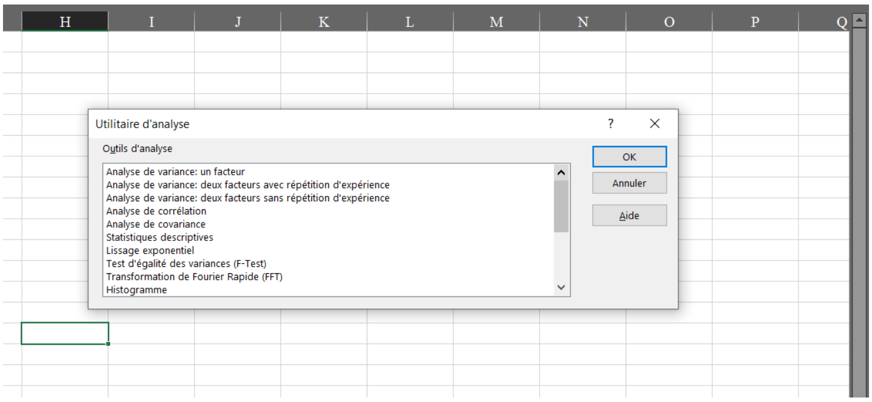
Test de Student pour échantillons appariés sous Microsoft Excel

Il faut tout d'abord activer l' **Utilitaire d'analyse .** Selon la version de Microsoft Office que vous disposez, cherchez sur Google comment l'activé (un lien sur YouTube : [https://www.youtube.com/watch?v=hy-lTUZ4whk).](https://www.youtube.com/watch?v=hy-lTUZ4whk)

1. Après l'activation, allez dans l'onglet « **Données** » pour trouver l' **Utilitaire d'analyse.**

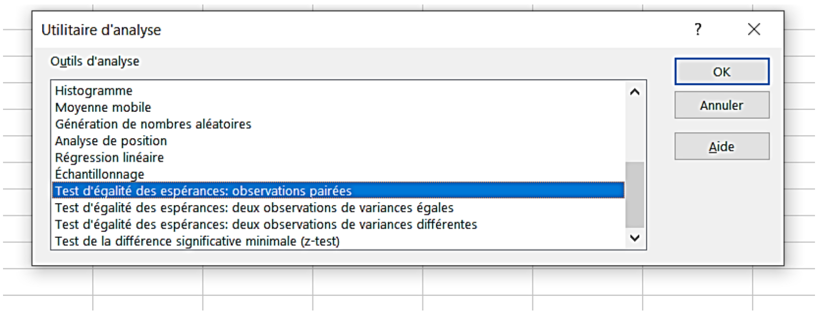


2. Lorsque vous cliquez sur **Utilitaire d'analyse,** une boite de dialogue s'ouvre

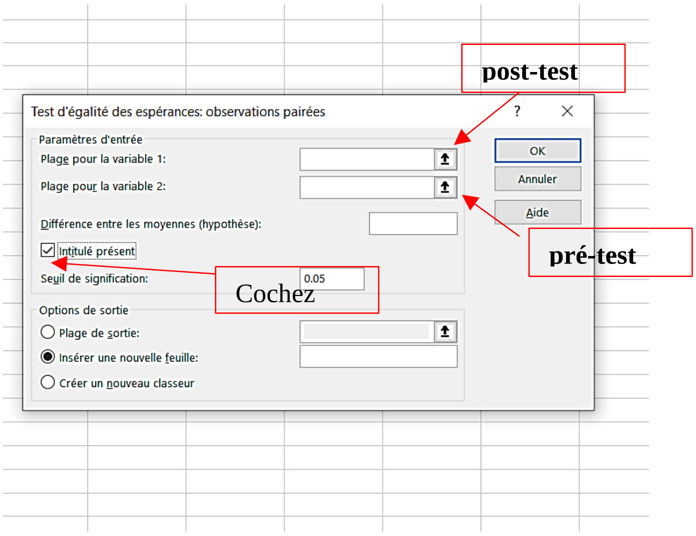


3. Dans cette boite de dialogue, défilez vers le bas et choisissez :

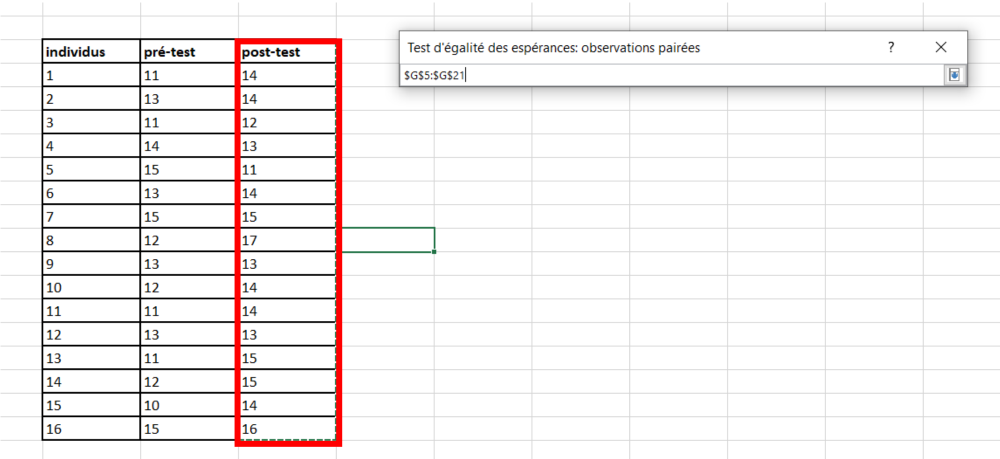
Test d'égalité des espérances : observations pariées



4. **En cliquant sur « OK », une autre boite de dialogue s'ouvre**



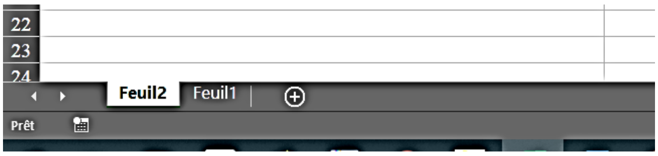
5. **En cliquant sur la petite flèche du post-test, choisissez la plage du deuxième test, ensuit cliquez sur le clavier « Entré »**



individus	pré-test	post-test
1	11	14
2	13	14
3	11	12
4	14	13
5	15	11
6	13	14
7	15	15
8	12	17
9	13	13
10	12	14
11	11	14
12	13	13
13	11	15
14	12	15
15	10	14
16	15	16

Remarque : l'entête (post-test) doit être introduit avec les valeurs

6. **Faire la même chose pour le pré-test et cliquer « OK »**
7. **Une nouvelle « Feuil » est ajoutée**



8. Lecture des résultats :

	A	B	C	D
1	Test d'égalité des espérances: observations pairées			
2				
3		*post-test*	*pré-test*	
4	Moyenne	14	12.5625	
5	Variance	2.133333333	2.52916667	
6	Observations	16	16	
7	Coefficient de corrélation de Pearson	-0.086101737		
8	Différence hypothétique des moyennes	0		
9	Degré de liberté	15		
10	Statistique t	2.555555556		
11	P(T<=t) unilatéral	0.010977384		
12	Valeur critique de t (unilatéral)	1.753050356		
13	P(T<=t) bilatéral	0.021954768		
14	Valeur critique de t (bilatéral)	2.131449546		
15				
16				

$T_{cal=2.56}$

$T_{tab=2.13}$

p-valut =0.022

Comme vous pouvez le remarquer, les résultats sont identiques à ceux trouvez avec le calcul manuel. L'interprétation du P-value est la même que le test de Student indépendant (voir cours 01).

2.1.2. Tests non-paramétriques

2.1.2.1. Test de « Mann et Whitney » pour 2 échantillons indépendants

Ce test est appliqué pour comparer entre deux échantillons indépendants. Ex : comparer entre des volleyeurs et des footballeurs.

Il est aussi recommandé pour la comparaison des échantillons avec un petit effectif (<30). Ce test est basé sur les Rangs, ce qui le rend insensible aux valeurs extrêmes.

Tout d'abord il est nécessaire de comprendre la notion des Rangs. Le rang c'est l'emplacement de la valeur dans une série ordonnée.

Exemple :

Série : 5 ; 2 ; 6 ; 3 ; 4 ; 9 ; 10

Quel est le rang de la valeur « 4 » et « 6 » ??

Série ordonnée : 2 ; 3 ; 4 ; 5 ; 6 ; 9 ; 10 (ordonner la série du plus petit au plus grand : croissant)

Donc, le rang de la valeur « 4 » est **3** et le rang de la valeur « 6 » est **5.**

U de Mann et Whitney :

$$\boldsymbol{U_1 = \sum Rangs1 - \frac{n_1(n_1+1)}{n_1}}$$ **Équation 1**

$$\boldsymbol{U_2 = \sum Rangs2 - \frac{n_2(n_2+1)}{n_2}}$$ **Équation 2**

U_{cal}= Min [U_1,U_2] (U_{cal}= la valeur la plus petite des deux U)

Exemple :

Groupe 01	Groupe 02
7	8
4	9
3	11
5	13
3	9
6	3
4	4
	3

Question : est-ce que le groupe 01 diffère du groupe 02 (statistiquement) ?

Solution :

Etape 01 :

Classer l'ensemble des valeurs (groupe 01 et 02) dans un ordre croissant, et donner à chaque valeur son rang dans cette série

groupe 01 et 02	rangs	rang moyen
3	1	(1+2+3+4)/4=2.5 2.5
3	2	
3	3	
3	4	
4	5	(5+6+7)/3=6 6
4	6	
4	7	
5	8	8
6	9	9
7	10	**10**
8	11	11
9	12	(12+13)/2=12.5 12.5
9	13	
11	14	14

13	15	15

Tableau 1. Rangs et rangs moyens des valeurs des deux échantillons

Lorsque la valeur est répétée, on calcule le rang moyen de cette valeur. Ex : la valeur « 3 » s'est répétée 4 fois, donc on divise la somme des rangs (1+2+3+4) sur le nombre de répétitions (ici 4 fois) :

Le rang moyenne de la valeur « 3 » est égal à (1+2+3+4)/4= 2.5

Etape 02 :

Rapporter chaque rang dans sa colonne

rangs du groupe 01	**Groupe 01**	**Groupe 02**	rangs du groupe 02
10	**7**	8	11
6	4	9	12.5
2.5	3	11	14
8	5	13	15
2.5	3	9	12.5
9	6	3	2.5
6	4	4	6
		3	2.5

Pour remplir la colonne « Rangs du groupe 01 », on cherche le rang de la valeur « 7 » dans le Tableau 1, il se trouve qu'elle correspond à 10. Pour chaque valeur du groupe 01 et 02 on rapporte les rangs correspondants.

Etape 03 :

Calculer U1 et U2 en appliquant l'équation 01 et 02

rangs du groupe 01	**Groupe 01**	**Groupe 02**	rangs du groupe 02
10	**7**	8	11
6	4	9	12.5
2.5	3	11	14
8	5	13	15
2.5	3	9	12.5
9	6	3	2.5
6	4	4	6
		3	2.5
	n1= 7	n2 =8	
Σ rangs1=44			Σ rangs 2=76

$$U_1 = \sum Rangs1 - \frac{n_1(n_1+1)}{n_1} = 44 - \frac{7(7+1)}{7} = 16$$

$$U_2 = \sum Rangs2 - \frac{n_2(n_2+1)}{n_2} = 76 - \frac{8(8+1)}{8} = 40$$

U_{cal}= min[16,40]= 16 (16 c'est la valeur la plus petite)

U_{cal} =16

Nous comparons ensuite le U_{cal} avec le U_{tab} . le U_{tab} (α=0.05 ; n1 et n2)

Dans notre exemple n1= 7 ; n2 = 8 donc U_{tab}= 10

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	--	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36

Interprétation

Contrairement au test de Student, la différence est significative lorsque le $U_{cal}<U_{tab}$.

dans notre exemple U_{cal} (16) >U_{tab} (10) donc :

il n'existe pas de différence significative entre les deux groupes.

2.1.2.2. Test de Wilcoxon pour 2 échantillons appariés

Il est appelé aussi test des **Rangs Signés**, en référence au signe des rangs positif et négatif. Ce test non-paramétrique est l'équivalant du test Student pour échantillons appariés. Souvent, il est utilisé pour comparer entre pré-test et post-test. Il est aussi recommandé pour la comparaison des échantillons avec un petit effectif (<30). Ce test est basé sur les Rangs, ce qui le rend insensible aux valeurs extrêmes.

Vu la complexité de la formule, nous procédons aux calculs par étapes et sans mentionner cette formule.

Exemple : voici les résultats de la VMA d'un groupe de sportif. Le pré-test est réalisé avant le début de la saison, alors que le deuxième s'est effectué après 30 jours d'entrainement.

Pré-Test	11	10	13	15	16	14	13	13	11
Post-test	15	12	14	17	18	17	11	10	12

Question : la VMA de ce groupe a-t-elle progressé ?

Solution

Etape 01 : calculer la différence (post-test)-(pré-test) et la valeur absolue de la différence

Pré-Test	Post-test	Différence	Absolu de différence
11	15	(15-11)=4	4
10	12	2	2
13	14	1	1
15	17	2	2
16	18	2	2
14	17	3	3
13	11	**-2**	2
13	10	**-3**	3
11	12	1	

Etape 02 : attribuer les rangs aux valeurs absolues de la « Différence ».

Pour ce faire, il faut rapporté les valeurs dans ordre croissant

Abs (Différence) ordonné	Rang	Rang moyen
1	1	(1+2)/2=1.5
1	2	
2	3	(3+4+5+6)/2=4.5
2	4	
2	5	
2	6	
3	7	(7+8)/2=7.5
3	8	
4	9	9

Donc :

- le rang de la valeur « 1 » est 1.5
- le rang de la valeur « 2 » est 4.5
- le rang de la valeur « 3 » est 7.5
- le rang de la valeur 4 est 9

Etape 03 : calculer les rangs positifs et les rangs négatifs :

Rapporter les rangs dans le tableau suivant :

Différence	Absolu de différence	rangs moyens
(15-11)=4	4	9
2	2	4.5
1	1	1.5
2	2	4.5
2	2	4.5
3	3	7.5
-2	2	4.5
-3	3	7.5

Somme des rangs positifs : 9+4.5+1.5+4.5+4.5+7.5=33

Somme des rangs négatifs : 4.5+7.5=12

W_{Cal} = Min [33,12]= 12 (W_{Cal} = la valeur la plus petite = 12)

Etape 04

Comparer W_{cal} avec le W_{tab} en sachant que :

W_{tab} = la valeur qui correspond à N et α=0.05 (Bilatéral dans le tableau)

Dans notre exemple ; N=9 donc W_{cal}= 6

N	Niveau de signification, test unilatéral		
	0,025	0,01	0,005
	Niveau de signification, test bilatéral		
	0,05	0,02	0,01
6	0		
7	2	0	
8	4	2	0
9	6	3	2
10	8	5	3
11	11	7	5
12	14	10	7
13	17	13	10
14	21	16	13
15	25	20	16

Interprétation

Comme pour le test de Mann et Whitney et contrairement au test de Student, la différence est significative lorsque le $W_{cal} < W_{tab}$.

Dans notre exemple Wcal (12) > Wtab(6) donc nous concluons qu'il n'existe pas de différence statistiquement significative.

Remarque :

Toutes les étapes peuvent être réalisées dans un seul tableau.

pré-test	post-test	Différence	ABS dif	Rang	Rang Moyen
11	15	4	4	9	9
10	12	2	2	3	4.5
13	14	1	1	1	1.5
15	17	2	2	2	4.5
16	18	2	2	2	4.5
14	17	3	3	3	7.5
13	11	-2	2	2	4.5
13	10	-3	3	2	7.5
11	12	1	1	1	1.5
	Somme Positive	33			
	Somme Négative	12			

2.2.Tests d'hypothèses pour plus de 2 échantillons

Comme nous l'avons déjà vu dans les cours précédents, la comparaison entre deux échantillons peut être réalisée par les tests paramétriques ou non paramétriques. Il en est de même pour la comparaison entre plusieurs groupes (échantillons) . Dans ce cours nous étudions la comparaison des échantillons indépendants.

Avant de commencer il est nécessaire de comprendre le schéma de la comparaison des échantillons de « plus de 2 groupes »

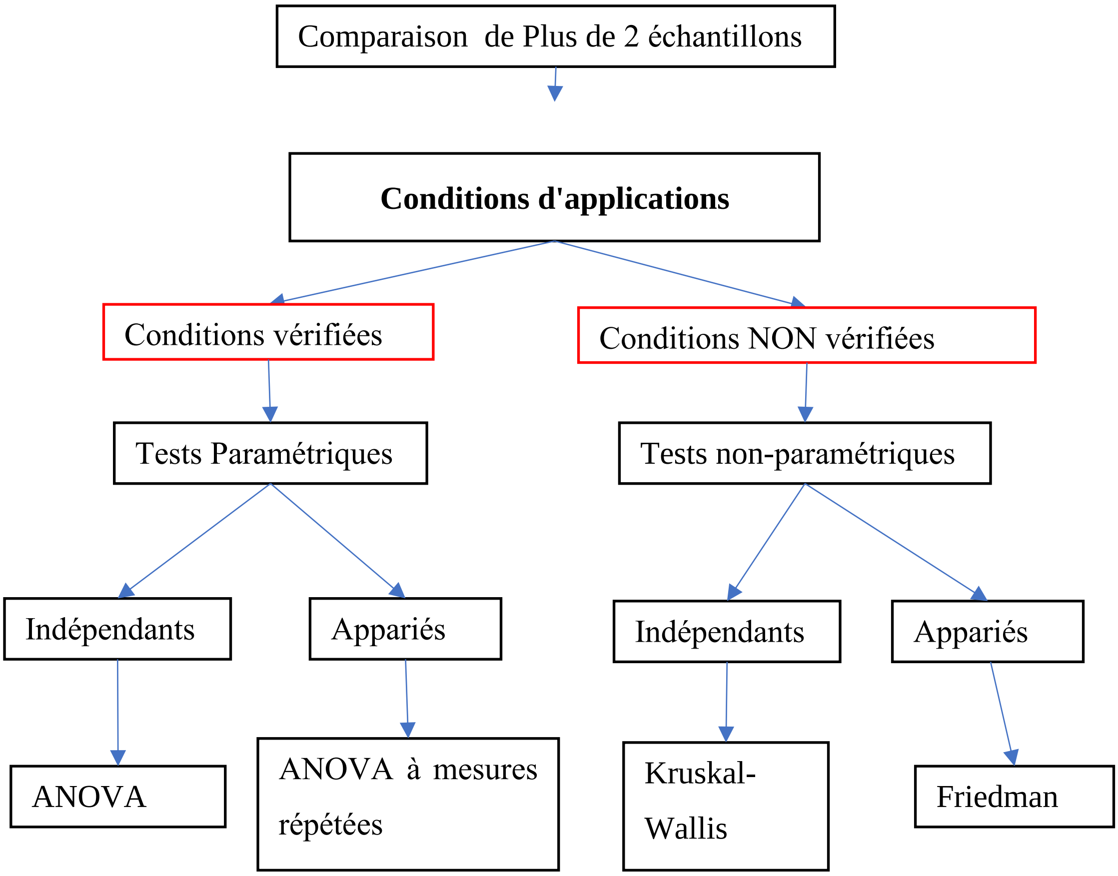


Figure 3. schéma des comparaisons de « plus de 2 échantillons »

<u>Tests post-hoc</u>

Avant d'expliquer les tests post-hoc, il important de savoir que le résultat d'un test -ANOVA par exemple- doit être comparé à une norme (tiré d'une table),exactement comme pour la comparaison de deux échantillons. En revache, avec 2 échantillons il est facile de comprendre le résultat du test. Ex : on compare groupe « A » avec « B », si la différence est significative, nous retournons vers les moyennes (ou médianes) pour décider lequel des deux est meilleur.

Alors que lorsqu'on compare « A », « B » et « C » et on trouve une différence significative, nous ne pouvons pas savoir si la différence est significative entre « A-B », « A-C » ou « B-C ». C'est pour cela qu'on réalise le test post-hoc afin de comparer par paires (2 groupes).

Une question importante : « pourquoi ne pas procéder directement à la comparaison par paires en utilisant les tests de deux échantillons ?? »

Pour vérifier la même hypothèse (l'existence ou pas d'une différence), le nombre de tests réaliser influent sur l'erreur de type I. dans la mesure où on réalise un seul test comme pour le cas du T de Student pour deux échantillons , la marge d'erreur acceptée est 0.05.

Maintenant si on a 4 groupes à comparer par paires, le nombre de comparaisons s'élève à 6 (3+2+1). Donc en multipliant l'erreur (0.05) par le nombre de comparaison on trouve : 6 x 0.05= 0.3. de ce fait, on accepte 30% d'erreur comparé à 5% ciblé. .

C'est pour cela que les statisticiens ont développé des tests spécifiques pour réaliser des comparaisons entre un ensemble d'échantillons (+2).

Aussi, il faut savoir que pour chaque test, il existe plusieurs tests post-hoc proposés dans les logiciels d'analyse. Pour notre cours, nous allons choisir un seul test post-hoc pour chaque test. Tous les tests Post-Hoc prennent en considération le nombre de comparaisons pour fixer le seuil de signification.. Il aussi possible d'utiliser les tests standards (après vérification de l'existence d'une différence) avec la correction de Bonferroni. Cette technique consiste à diviser le seuil ciblé (généralement 0.05) sur le nombre de comparaisons. Pour 6 comparaisons, le alpha (α) de Bonferroni = 0.05/6= 0,0083. Donc pour qu'une différence soit considérée comme étant significative, le P-Value doit être inferieur à 0.0083 et non pas à 0.05.

2.2.1. Tests paramétriques

2.2.1.1. Analyse de la variance (ANOVA) pour des échantillons indépendants

Dans cette partie et vu la difficulté de faire des calculs assez compliqués et très lourds, nous allons réaliser l'analyse statistique sous le logiciel Jamovi.

Les conditions d'application d'une ANOVA :

Pour pouvoir appliquée une ANOVA de Fisher nous devons suivre le schéma présenté dans la Figure 4.

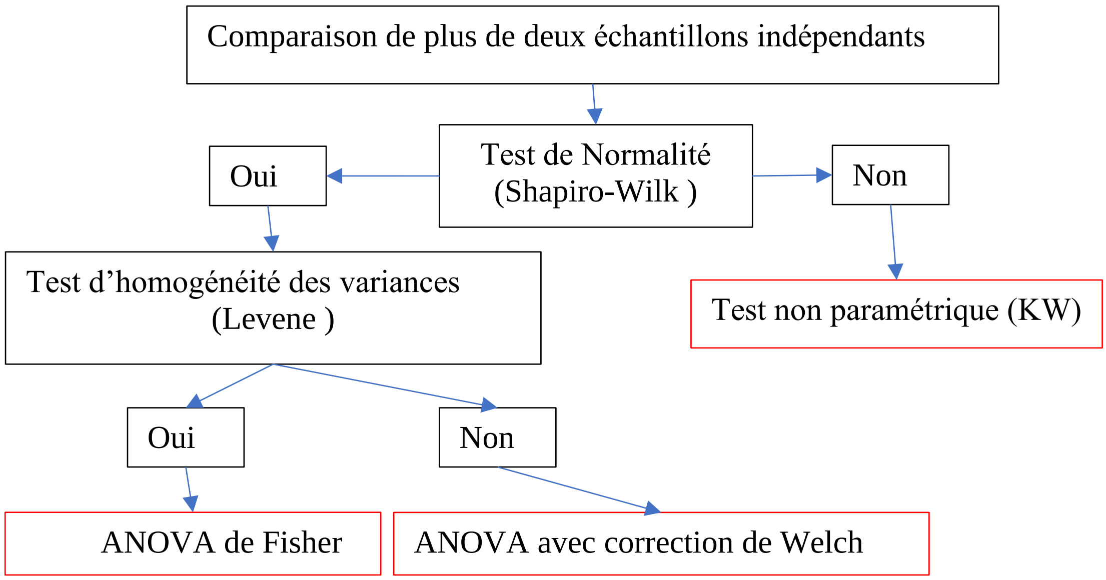


Figure 4. Conditions d'application d'une ANOVA pour échantillons indépendants

2.2.1.1.1. Préparation des données

Voici un exemple des notes d'un examen pour trois groupes

Dans la majorité du temps, les étudiants présentent leurs données de la façon suivante.

G01	14	15	14	16	15	16	16	10	10	17	10	19
G02	19	12	11	19	8	19	10	10	3	13	12	18
G03	12	2	10	13	7	6	11	9	12	5	10	7

Il faut savoir que l'ensemble des logiciels de statistique n'accepte pas cette forme. La forme correcte est la suivante :

Chaque colonne correspond à une variable.

Combien de variables dans ce tableau ?????

Il y a deux variables :

1. Quantitative = note d'examen
2. Qualitative = groupe

Cela signifie que chaque individu possède deux valeurs : sa note d'examen et le groupe d'appartenance.

Après manipulation sur Excel (copier-coller), le tableau devient comme suit.

Notes	Groupes
14	1
⋮	⋮
19	1
19	2
· ·	⋮
18	2

12	3
⋮	⋮
7	3

Nb. Nous avons supprimé quelque valeur (pointillés) pour la démonstration.

Avec ce tableau il est possible de faire des analyses statistiques avec n'importe quel logiciel.

2.2.1.1.2. ANOVA sous Jamovi

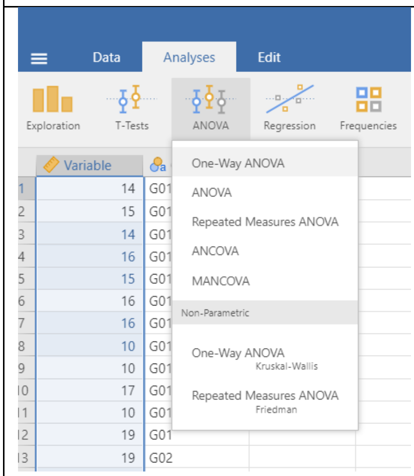
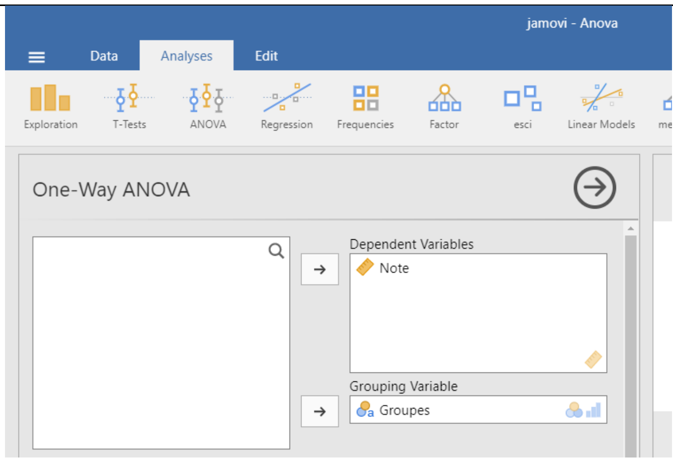
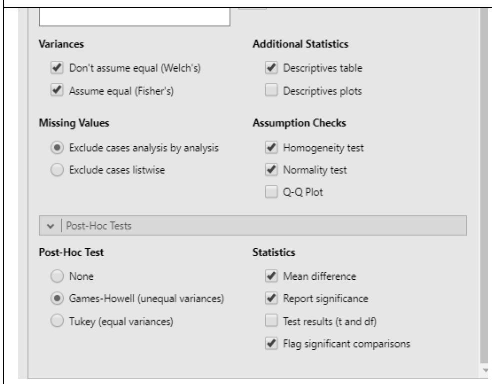
<table>
<tr><td>Dans l'onglet analyse → ANOVA → One-way ANOVA</td><td>Note → « dependent variable »
Groupes → « Grouping variable »</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>cocher les éléments suivants</td><td></td></tr>
<tr><td>
</td><td></td></tr>
</table>

Figure 5. Procédure d'une ANOVA sous Jamovi

2.2.1.1.3. Interprétation des résultats ANOVA sous Jamovi

Tableau 2. Résultat de l'ANOVA

		F	df1	df2	p
Note	Welch's	9.73	2	21.2	0.001
	Fisher's	6.87	2	33	0.003

Ce tableau représente les résultats de l'ANOVA. Selon la condition de l'homogénéité des variances, nous devons choisir de lire la première ligne (homogénéité non vérifiée) ou la deuxième ligne (homogénéité vérifiée).

La valeur le P (P-value) renseigne sur l'existence ou pas d'une différence entre les trois groupe. Dans cet exemple, les deux tests démontrent l'existence d'une différence entre les groupes (Welch=0.001 ; Fisher=0.003).

Tableau 3. Statistique descriptive

	Groupes	N	Mean	SD	SE
Note	G01	12	14.33	2.93	0.847
	G02	12	12.83	5.06	1.461
	G03	12	8.67	3.31	0.956

Mean = Moyenne ; SD= écart-type ; SE = erreur standard (à ignorer pour vous)

A partir de ce tableau, nous pouvons déjà avoir un aperçu sur les trois groupes. Nous aurons besoins des moyennes pour interpréter les résultats des comparaisons multiples.

Conditions d'application

Les deux tableaux suivants représentent les résultats des tests de normalité et d'homogénéité des variances. Il convient de rappeler que les conditions sont remplies lorsque la valeur du P est supérieure à 0.05.

Tableau 4. Test de normalité Shapiro-Wilk

	W	p
Note	0.978	0.662

Étant donné que la valeur de P= 0.662 est supérieure à 0.05, nous pouvons dire que la **condition de normalité est remplie** .

Tableau 5. Test d'homogénéité des variances (Levene)

	F	df1	df2	p
Note	1.95	2	33	0.158

Étant donné que la valeur de P= 0.158 est supérieure à 0.05, nous pouvons dire que la **condition d'homogénéité des variances est remplie. Ce qui implique que, pour le résultat de l'ANOVA, nous nous referons à la deuxième ligne du** Tableau 2**. (F de Fisher)**

Tableau 6. Résultat du test Post-Hoc de Tukey

		G01	G02	G03
G01	Mean difference	—	1.50	5.67 **
	p-value	—	0.615	0.003
G02	Mean difference		—	4.17 *
	p-value		—	0.034
G03	Mean difference			—
	p-value			—

Note. * p < .05, ** p < .01, *** p < .001

Ce tableau représente les comparaison par paires entre les trois groupes. Les résultats démontrent ce qui suit :

- Aucune différence significative entre le G01 et le G02 (P= 0.615 ; NS)
- Une différence significative entre le G01 et le G03 (P= 0.003 ; <0.01)
- Une différence significative entre le G02 et G03 (P=0.034 ; <0.05)

Conclusion de la lecture des résultats

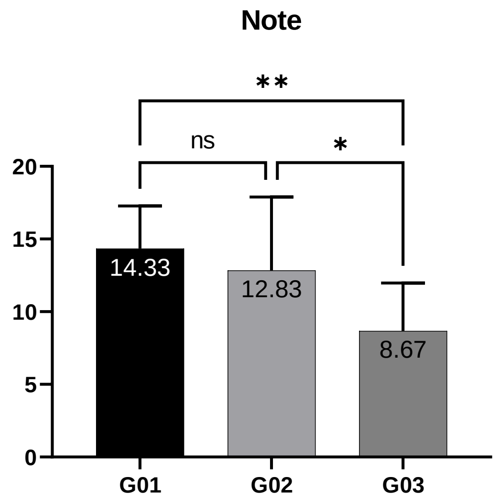


Figure 6. Comparaison de la variable « Note » entre les trois groupes

L'ANOVA a révélée l'existence d'une différence significative au seuil α< 0.01 (P=0.003). les résultats démontrent que le groupe 03 enregistrent des valeurs (8.67 ± 3.31) inférieures par rapport au groupe 01(14.33 ± 2.93 ; P<0.01) et groupe 02 (12.83 ± 5.06 ; P<0.01), alors les deux groupes 01 et 02 ne présentent pas de différences entre eux.

2.2.1.2. ANOVA à mesures répétées (ANOVA MR)

L'ANOVA à mesures répétées est utiliser pour comparer l'évolution d'un groupe sur plusieurs sessions (Plus de 2). C'est l'équivalent du test de Student pour deux échantillons appariés.

Exemples

Afin de suivre l'évolution d'un groupe de sportifs (N=15) ayant comme objectif la « diminution du poids corporel » avec un programme d'entrainement fractionné (Interval Training). Le chercheur a planifié 4 sessions de peser, dont deux au début et à la fin du programme et deux au milieu du programme . Chaque peser (session) est réalisé après 10 jours d'entrainement. Les résultats sont présentés dans le tableau suivant :

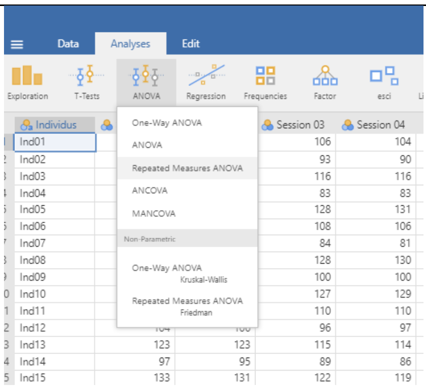
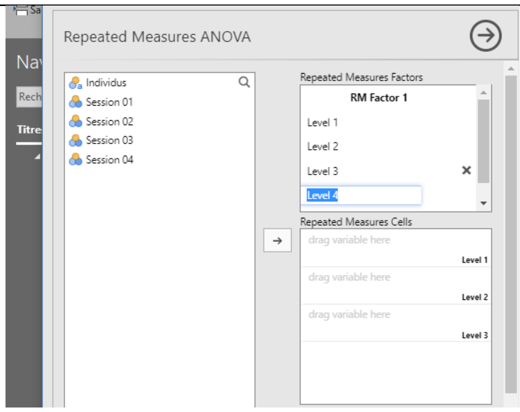
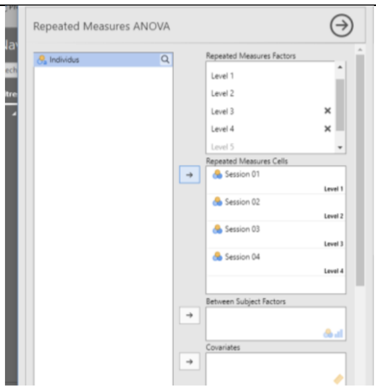
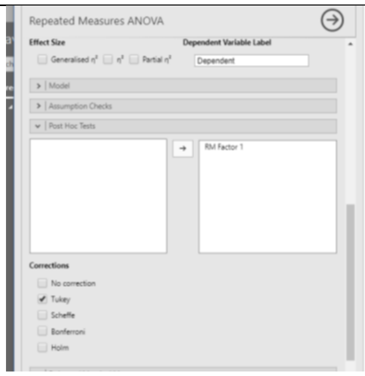
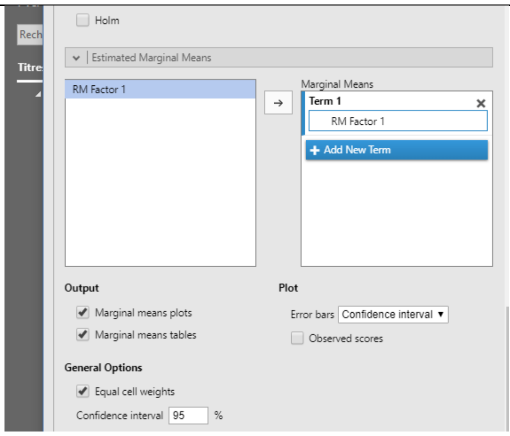
Individus	Session 01	Session 02	Session 03	Session 04
Ind01	110	111	106	104
Ind02	99	98	93	90
Ind03	124	125	116	116
Ind04	91	90	83	83
Ind05	134	133	128	131
Ind06	119	118	108	106
Ind07	96	94	84	81
Ind08	138	137	128	130
Ind09	107	109	100	100
Ind10	139	137	127	129
Ind11	120	118	110	110
Ind12	104	106	96	97
Ind13	123	123	115	114
Ind14	97	95	89	86
Ind15	133	131	122	119

Problématique :

Comment évolue le poids corporel avec un travail en « Interval Training » ?

2.2.1.2.1. ANOVA à mesures répétées sous Jamovi

Pour réaliser l'ANOVA à mesures répétées, il faut suivre les étapes suivantes :

<table>
<tr><td>Analyses → ANOVA → Repeated Measures ANOVA</td><td>Dans l’onglet “RM factor 1” → cliquer sur « Level 3 » et taper « Entrer » → cliquer sur « Level 4 » et taper « Entrer » (pour créer les 4 sessions)</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td></td><td></td></tr>
<tr><td>Dans l’onglet «Repeated Measures cells » Mettre chaque session dans une ligne, en respectant l’ordre (session 1 ;2 ;3 ;4)</td><td>Développer le titre « Post Hoc tests » et déplacer « Rm Factor 1 » à droite.
cocher Tukey dans « Correction »</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td></td><td></td></tr>
<tr><td colspan="2">Développer le titre « Estimated marginal Means » → Déplacer « RM Factor 1 » à droite → Cocher « Marginal Means Plots » et « Marginal Means Tables »</td></tr>
<tr><td colspan="2">
</td></tr>
</table>

Interprétation des résultats

Tableau 7. Résultats de l'ANOVA MR

	Sum of Squares	df	Mean Square	F	p
RM Factor 1	1115	3	371.60	138	**< .001**
Residual	113	42	2.68		

Les résultats de l'ANOVA MR révèlent l'existence d'une différence significative entre les sessions à α <0.001.

Tableau 8. Tests Post Hoc (comparaison par paires)

Comparison							
RM Factor 1		RM Factor 1	Mean Difference	SE	df	t	p_{tukey}
Level 1	-	Level 2	0.600	0.598	42.0	1.00	**0.748**
	-	Level 3	8.600	0.598	42.0	14.38	**< .001**
	-	Level 4	9.200	0.598	42.0	15.38	**< .001**
Level 2	-	Level 3	8.000	0.598	42.0	13.37	**< .001**
	-	Level 4	8.600	0.598	42.0	14.38	**< .001**
Level 3	-	Level 4	0.600	0.598	42.0	1.00	**0.748**

« Level » = session

Ce tableau regroupe toutes les comparaisons par paires. L'identification des différence se base sur les valeurs de la dernière colonne (P_{Tukey}). Les résultats démontrent que :

- Aucune différence n'est observée entre les deux premières sessions (01 et 02) et les deux dernières sessions (03 et 04) (P= 0.748)
- Les autres comparaisons démontrent une différence très significative à P< 0.001

Tableau 9. Statistique descriptive

			95% Confidence Interval	
RM Factor 1	Mean	SE	Lower	Upper
Level 1	116	4.19	106.6	125
Level 2	115	4.19	106.0	124
Level 3	107	4.19	98.0	116
Level 4	106	4.19	97.4	115

Pour être simple, ce tableau représente les statistiques descriptives

Conclusion de l'interprétation de l'ANOVA MR

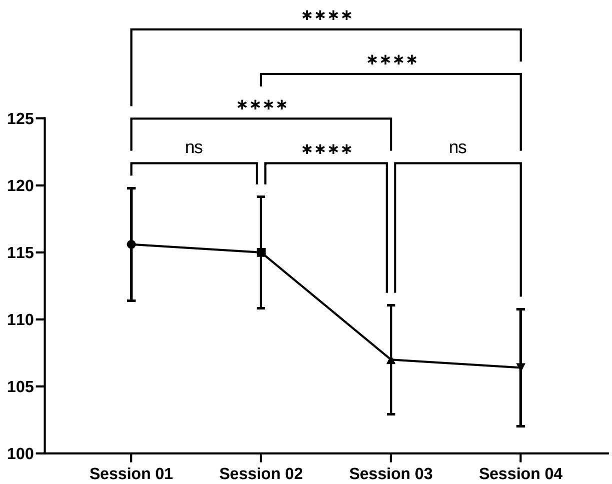


Figure 7. évolution du poids des sportifs sur les 4 sessions (Moyenne et erreur standard)

Ce graphique nous aide à bien visualiser l'évolution du poids des sportifs. Après 10 jours d'entrainement, aucune perte de poids n'est enregistrée (S01= 116 ; S02= 115 ; P= 0.748). Après 20 jours d'entrainement (S03=107), une perte considérable est enregistrée que ce soit par rapport à la session 01 ou à la session 02 (P<0.001). une stagnation du poids est ensuite enregistrée lors des derniers 10 jours d'entrainement (S03 = 107 ; S04=106 ; P=0.748).

2.2.2. Tests non-paramétriques

2.2.2.1. Kruskal-Wallis pour des échantillons indépendants

C'est l'équivalent non paramétrique de l'ANOVA de Fisher. Ce test ne nécessite aucune condition. Il souvent conseillé pour les études où la taille de l'échantillon est réduite .

Comme pour les tests non paramétriques pour deux échantillon, ce test est basé sur les rangs. Il fonctionne comme les autres tests d'hypothèse où l'on compare une valeur calculée - appelée H- avec une valeur tabulé. La valeur tabulée suit la distribution du Khi2. Donc on compare la valeur calculée (H) avec une valeur du Khi2.

$$H = \frac{12}{n(n+1)} \sum \frac{R_i^2}{ni} - 3(n+1)$$

Où

n : le nombre d'individus statistique de l'ensemble des groupes

R_i : somme des rangs pour un seul groupe

n_i : le nombre d'individus pour un seul groupe

Exemple

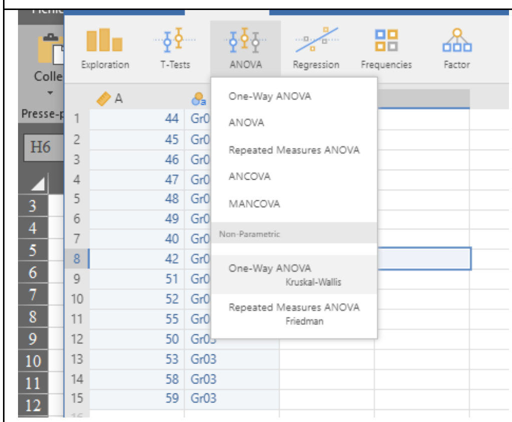
	Gr01	**Gr02**	**Gr03**		**Rang01**	**Rang02**	**Rang03**
	44	40	50		3	1	9
	45	42	53		4	2	12
	46	51	58		5	10	14
	47	52	59		6	11	15
	48	55			7	13	
	49				8		
ni	**6**	**5**	**4**	**Ri**	**33**	**37**	**50**
				R_i^2/ni	181.5	273.8	625

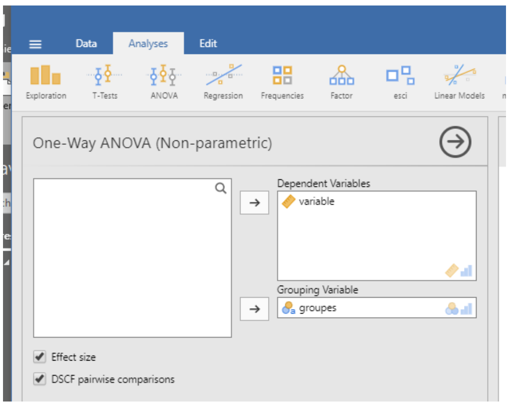
$$H = \frac{12}{15(15+1)}(181.5 + 273.8 + 625) - 3(15 + 1) = 6.02$$

Khi2 (α=0.05 ; ddl= k-1=2)=5.99

H(6.02)>khi2 (5.99)→ il existe une différence significative entre les groupes.

2.2.2.1.1. Kruskal-Wallis sous Jamovi





Résultat du test Kruskal-Wallis

	χ^2 (H)	df	p	ε^2
variable	6.01	2	0.049	0.430

Les résultats démontrent l'existence d'une différence significative entre les groupes au seuil α< 0.05 (P=0.049).

Comparaisons par paires

		W	p
Gr01	Gr02	0.775	0.848
Gr01	Gr03	3.618	**0.028**
Gr02	Gr03	2.078	0.306

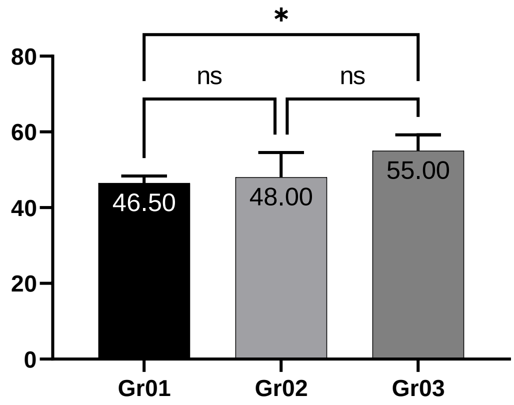


Figure 8. comparaison des trois groupe avec le test Post Hoc de KW (Moyenne ; écart-type)

Les comparaisons par paires démontrent que le Gr01 et différent du Gr03 à α<0.05 (P=0.028), mais aucune différence n'est observée pour les autres groupes (P>0.05) : Gr01 vs Gr02 ;P=0.848 et Gr02 vs Gr03 ; P=0.306.

2.2.2.2. Friedmann pour des échantillons appariés

C'est l'équivalent non paramétrique du test ANOVA à mesures répétées. Il basé sur les rangs. Il est utilisé pour vérifier l'évolution d'un échantillon sur plusieurs sessions.

$$F_r = \frac{12}{Nk(k+1)} \sum R_i^2 - 3N(k+1)$$

Où :

N : Nombre d'individus

K : Nombre de sessions

Ri : Somme des rangs d'une session

Nous comparons le Fr calculé à la valeur du Khi2 sur la table avec α=0.05 et DDL= K-1.

Exemple

Le chercheur a évalué (Note su 20) son échantillon sur trois sessions. Les résultats sont présentés ci-dessous :

Session 01	Session 02	Session 03
8	10	13
11	13	11
11	16	18
10	15	17
12	17	15
10	16	14

Problématique : est-ce qu'il existe un évolution dans le niveau des étudiants ?

Solution :

Le rang de chaque valeur est calculé par rapport à la ligne. Par exemple : pour la première ligne la valeur 8 est la plus petite donc son rang est « 1 » ; de même 10 (2) et 13 (3).

Session 01	Session 02	Session 03	R1	R2	R3
8	10	13	1	2	3
11	13	14	1	2	3
11	16	18	1	2	3
10	15	17	1	2	3
12	17	15	1	3	2
10	16	14	1	3	2
N	**6**	**Ri**	18	10	8
K	**3**	**R_i^2**	324	100	64
		ΣR_i^2	**488**		

$$F_r = \frac{12}{6 * 3 * (3+1)} * (18^2 + 10^2 + 8^2) - 3 * 6 * (3+1) = 9.33$$

Fr= 9.33

Le khi2 correspondant à α=0.05 et DDL= 2 (K-1=3-1=2) égal à : $\chi^2 = 5.99$

Fr >χ^2 → il existe une différence significative au seuil α=0.05

Figure 9. comparaison par paire de Dunn (Post-Hoc Friedmann)

Dunn's multiple comparisons test	Adjusted P Value
Session 01 vs. Session 02	0,0628
Session 01 vs. Session 03	**0,0117***
Session 02 vs. Session 03	>0,9999

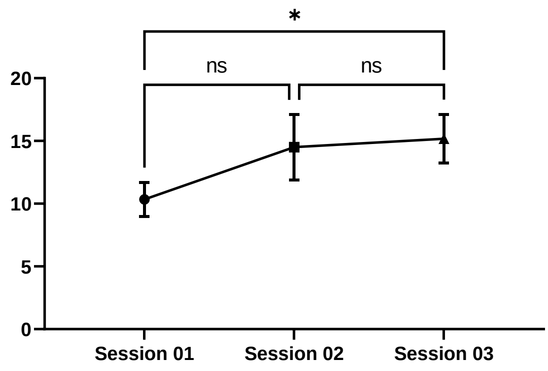


Figure 10. évolution des notes selon le test de Dunn (Moyenne ; écart-type)

Conclusion

Les résultats du test Friedmann démontrent l'existence d'une différence entre les trois sessions (Fr= 9.33 ; X^2 =5.99 ;P<0.05). la comparaison par paires a démontré que l'existence d'une progression significative entre la session 01 et la session 03 à 0.05. aucune différence n'est observée entre la session 01 et 02 et entre la session 02 et 03. Il convient de signaler une tendance vers une progression entre la session 01 et 02 (P=0.06 proche du seuil de signification) et c'est la progression la plus élevée lorsqu'on prend en compte que les sessions successives .

3. Le Z-score et le T-score :

Dans certaines situations, le chercheur travaille sur plusieurs variables dont l'unité diffère. Donc, une normalisation des données est très utile pour pouvoir visualiser, combiner et trier les données.

3.1.Le z-score

$$\boxed{z - score = \frac{\chi_i - \bar{x}}{s}}$$

Où :

$\bar{x}$: $Moyenne$

S : écart-type

3.2.Le T-score

Les valeurs du z-score sont toujours petites (entre -3 et 3), ce qui rend la visualisation des données relativement difficile. En revanche les valeurs de T-score sont plus élevées (entre 0 et 100) avec une moyenne de 50. Le classement de chaque individu devient très simple à localiser.

$$\boldsymbol{T - score = (Z \times 10) + 50}$$

3.3.Application

Nous voulons établir un classement de nos joueurs sur la base du profil physique (force, VMA).

Force (kg)	VMA (km/h)
121	13
90	16
89	13
80	13
154	17
81	17
89	15
149	13
99	13
154	16
92	17
112	13
122	14
97	16

Comme vous pouvez le constater, il est impossible de faire des calculs avec deux variables dont l'unité n'est pas la même (ici Kg et Km/h). nous procédons à la normalisation (standardisation) des valeurs.

individus	**Force (kg)**	**VMA(km/h)**	**Z-force**	**Z-VMA**	**T-Force**	**T-VMA**
1	**121**	**13**	0.44	-0.99	54.41	40.08
2	**90**	**16**	-0.72	0.74	42.81	57.44
3	**89**	**13**	-0.76	-0.99	42.44	40.08
4	**80**	**13**	-1.09	-0.99	39.07	40.08
5	**154**	**17**	1.68	1.32	66.75	63.22
6	**81**	**17**	-1.06	1.32	39.45	63.22
7	**89**	**15**	-0.76	0.17	42.44	51.65
8	**149**	**13**	1.49	-0.99	64.88	40.08
9	**99**	**13**	-0.38	-0.99	46.18	40.08
10	**154**	**16**	1.68	0.74	66.75	57.44
11	**92**	**17**	-0.64	1.32	43.56	63.22
12	**112**	**13**	0.10	-0.99	51.04	40.08
13	**122**	**14**	0.48	-0.41	54.78	45.87
14	**97**	**16**	-0.46	0.74	45.43	57.44
Moyenne	**109.21**	**14.71**				
Ecart-type	**26.73**	**1.73**				

Nous avons -en quelque sorte- situé le classement de chaque individu par rapport à la moyenne du groupe

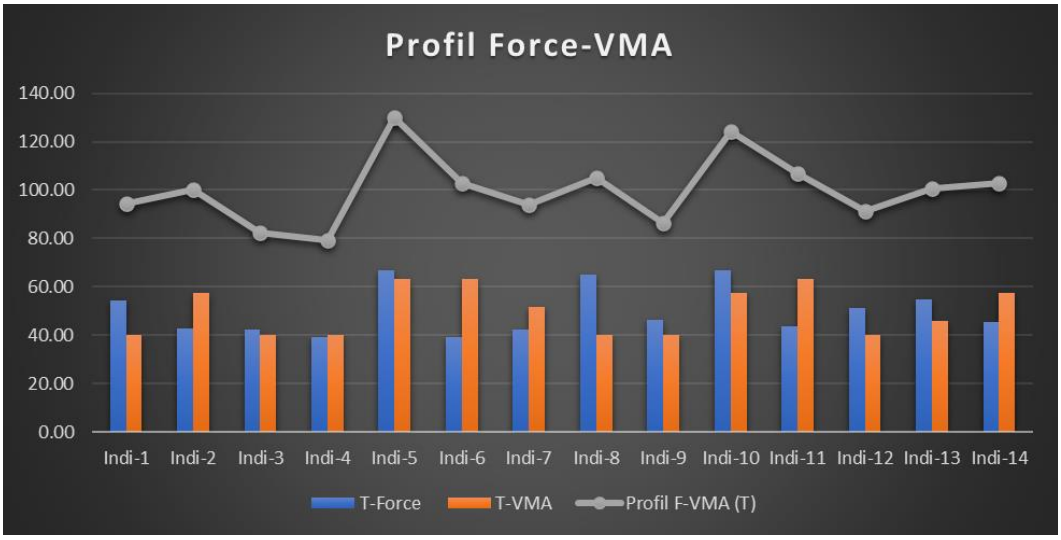
Ex :pour la variable « force », l'individu « 5 » se trouve à 1.68 au-dessus de la moyenne, alors que l'individu « 11 » se trouve à « -0.64 » en dessous de la moyenne. L'individu « 5 » présente un profil de force supérieur à la moyenne du groupe, contrairement à l'individu « 11 », on peut aussi dire que l'individu « 12 » est moyen par rapport au groupe

Indi	Force	VMA	Z-force	Z-VMA	T-Force	T-VMA	Profil F-VMA (T)	**Class**
1	121	13	0.44	-0.99	54.41	40.08	94.49	**9**
2	90	16	-0.72	0.74	42.81	57.44	100.25	**8**
3	89	13	-0.76	-0.99	42.44	40.08	82.52	**13**
4	80	13	-1.09	-0.99	39.07	40.08	79.16	**14**
5	154	17	1.68	1.32	66.75	63.22	129.97	**1**
6	81	17	-1.06	1.32	39.45	63.22	102.67	**6**
7	89	15	-0.76	0.17	42.44	51.65	94.09	**10**
8	149	13	1.49	-0.99	64.88	40.08	104.97	**4**
9	99	13	-0.38	-0.99	46.18	40.08	86.26	**12**
10	154	16	1.68	0.74	66.75	57.44	124.19	**2**
11	92	17	-0.64	1.32	43.56	63.22	106.78	**3**
12	112	13	0.10	-0.99	51.04	40.08	91.13	**11**
13	122	14	0.48	-0.41	54.78	45.87	100.65	**7**
14	97	16	-0.46	0.74	45.43	57.44	102.87	**5**

Après la standardisation des valeurs , il est possible de faire des calculs sur les deux variables. Nous avons choisi de faire une addition des T-scores pour dégager un profil physique. Ensuite, nous pouvons procéder au classement des joueurs selon ce profil combiné (Force-VMA).

Profil F-VMA (T)= T-score Force + T-score VMA.

Les résultats démontrent que l'individu « 5 » est le meilleur joueur en termes de force et VMA combinés, alors que l'individu « 4 » est le plus faible. L'individu « 11 » se trouve dans la 3ème position alors qu'il est au-dessous de la moyenne en termes de « Force », mais il a pu compenser avec une VMA très élevée (17 km/h).



3.4.Classer les individus dans des catégories de performance

En utilisant les valeurs de T-score, nous pouvons créer des catégories (classes) de niveau permettant de donner une appréciation catégorielle pour chaque performance. Cette appréciation est plus parlante que les chiffres. Il convient de préciser que le nombre et l'étendu des classes peuvent être choisis par le chercheur, mais généralement on utilise 5 classes d'un étendu de 20 :

Classes	T-scores
Très faible	Inférieur à 20
Faible	Entre 20 et 40
Moyenne	Entre 40 et 60
Bon	Entre 60 et 80
Très bon	Supérieur à 80

En utilisant ces classes, nous pouvons placer chaque individu dans une catégorie de performance en comparant son T-score aux valeurs fourchettes des catégories (classes) :

T-Force	Catégorie	T-VMA	Catégorie
54.41	Moyen	40.08	Moyen
42.81	Moyen	57.44	Moyen
42.44	Moyen	40.08	Moyen
39.07	Faible	40.08	Moyen
66.75	Bon	63.22	Bon
39.45	Faible	63.22	Bon
42.44	Moyen	51.65	Moyen
64.88	Bon	40.08	Moyen
46.18	Moyen	40.08	Moyen
66.75	Bon	57.44	Moyen
43.56	Moyen	63.22	Bon
51.04	Moyen	40.08	Moyen
54.78	Moyen	45.87	Moyen
45.43	Moyen	57.44	Moyen

Grâce à cette classification, nous pouvons détecter le nombre de joueurs par niveau de performance pour chaque variable.

Classes	Force	VMA
Très faible	0	0
Faible	2	0
Moyenne	9	11
Bon	3	3
Très bon	0	0

Il en ressort que deux joueurs ont un niveau force faible, trois avec un niveau bon et le reste des joueurs (9) se situe dans la moyenne du groupe . Pour la variable « VMA », trois sont bons et le reste des joueurs se trouve dans la catégorie « Moyen ».

3.5.Établir des valeurs normatives

Pour rendre les résultats de l'étude plus visible et claire, il est possible d'utiliser les T-scores pour établir des valeurs normatives permettant ainsi de classer n'importe quel nouvel individu dans une classe. Cette procédure est souvent utilisée lorsqu'on dispose d'un échantillon assez large pour représenter une population.

Pour ce faire, nous allons utiliser la formule inversement- de T-score et Z-score pour calculer les valeurs limites inférieures et supérieures réelles de chaque catégorie à partir des valeurs limites T-score :

Formule

$$\boxed{\text{Valeur limite réelle} = \left(\left(\frac{\text{Valeur limite Tscore} - 50}{10}\right) \times S\right) + \bar{x}}$$

Où :

S : l'écart-type

$\bar{x}$: La moyenne

Application numérique de cette formule nous donne les valeurs suivantes :

Valeurs limites Tscore	**Valeurs limites réelles**	
	Force	**VMA**
20	29.02	9.52
40	82.48	12.98
60	135.94	16.44
80	189.4	19.9

Avec ces données, nous pouvons désormais établir des valeurs normatives (avec unité) pour nos variables :

Classes	**Valeurs normatives**	
	Force (kg)	**VMA (km/h)**
Très faible	< 29.02	< 9.52
Faible	Entre 29.02 et 82.48	Entre 9.52 et 12.98
Moyenne	Entre 82.48 et 135.94	Entre 12.98 et 16.44
Bon	Entre 135.94 et 189.4	Entre 16.44 et 19.9
Très bon	> 189.4	> 19.9

Avec ces valeurs, nous pouvons estimer le niveau d'un individu qui théoriquement appartient à cette population. Par exemple, un joueur qui a réalisé 20km/h dans le test de VMA et a soulevé 170 kg est considéré comme étant « très bon » en endurance et « Bon » pour la force.

4. Le coefficient de corrélation de Pearson

4.1.La corrélation : concept

La corrélation est dite bi-variable, elle consiste à étudier de la relation entre deux variables quantitatives. Elle a comme objectif l'identification de la nature de la relation entre ces deux variables.

La nature de la corrélation peut prendre trois formes :

1. Positive : les deux variables évoluent dans le même sens ;
2. Négative : les deux variables évoluent dans un sens contraire ;
3. Absence de corrélation : les deux variables évoluent d'une façon indépendante l'une de l'autre.

4.2.Coefficient de corrélation de Pearson :

Pour évaluer cette relation, nous utilisons -souvent- le coefficient de corrélation de Pearson « r ».

Étant donné que « r » est une normalisation de la covariance, il varie entre -1 et 1

-1 : corrélation négative très forte

+1 : corrélation positive très forte

0 : absence de corrélation

Dans la réalité, il est rare, voire impossible, de trouver ces valeurs extrêmes (0, -1, 1). On est souvent face à des valeurs médianes (0.55 ; 0.60 ; -0.35…etc.). Plus la valeur de « r » se rapproche des deux extrémités (-1, 1), plus la corrélation est forte.

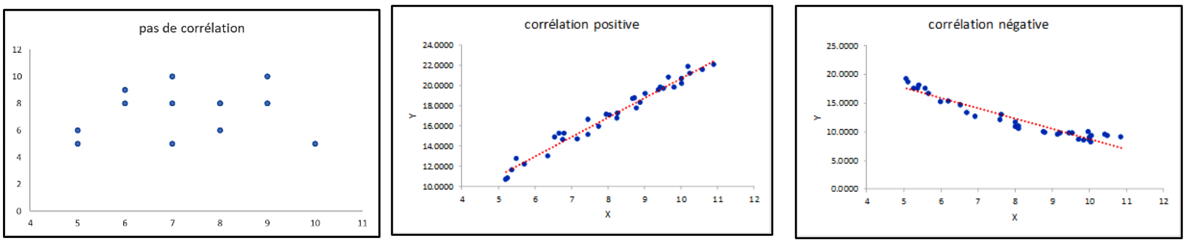


Figure 11. différents types de corrélation

Par contre, pour décider de la nature de cette corrélation nous sommes dans l'obligation de comparer ce « r » calculé avec la valeur de « r tab » :

Si |« r $_{cal}$ »| > « r $_{tab}$ » : la corrélation est significative.

Nb : | | renvoie la valeur absolue.

En fonction du signe de « r $_{cal}$ », la corrélation est positive ou négative

Exemple :

Dans une étude1 la valeur de la corrélation entre la variable « A » et « B » « r »= 0.90, alors que dans une autre étude2 « r » entre « C » et « D » = 0.70.

À première vue, nous concluons que la corrélation « A » « B » est plus forte que la corrélation « C » « D », car (0.90 >0.70). Cette conclusion est fausse, car la taille de l'échantillon joue un rôle crucial dans la décision. Si N1=10 et N2= 100, il est fort possible que la corrélation de l'étude2 soit plus forte. Nous illustrons cette idée un peu plus bas.

4.3.La formule :

$$r = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\sum x_i^2 - n\,\bar{y}^{\,2}} * \sqrt{\sum y_i^2 - n\,\bar{x}^{\,2}}}$$

$\bar{x}$: moyenne de la variable 01 (X)

$\bar{y}$ **:** moyenne de la variable 02 (Y)

n **:** la taille de l'échantillon

Pour décider de la puissance et la nature de la corrélation, nous comparons le coefficient de corrélation calculé « r » calc avec le coefficient de corrélation « r » tabulé.

Où la valeur de « r » tabulé est celle qui correspond dans la table :

- ddl : n-2
- α= 0.05

Exercice :

Identifiez la nature de la corrélation entre la variable « x » et la variable « y ».

x	9	2	14	4	2	5	7	9	5	11	22	30
y	3	4	4	2	8	11	5	8	5	25	11	18

Solution :

Comme d'habitude, nous fractionnons dans un tableau la formule où se présente la somme, tandis que le reste va être calculé séparément.

x	y	xi*yi	x_i^2	y_i^2
9	3	27	81	9
2	4	8	4	16
14	4	56	196	16
4	2	8	16	4
2	8	16	4	64
5	11	55	25	121
7	5	35	49	25
9	8	72	81	64
5	5	25	25	25
11	25	275	121	625
22	11	242	484	121

30	18	540	900	324
n= 12				
$\bar{x}$ =10.00	$\bar{y}$ = 8.67	Σ xi yi= 1359	Σx^2 = 1986	Σy^2 = 1414

Calculons toutes les composantes de la formule

$n\,\bar{x}^2 = 12 * 10^2 = 1200$

$n\,\bar{y}^2 = 12 * 8.67^2 = 902,03$

$\sum x_i y_i$ **= 1359**

$n\bar{x}\bar{y}$**= 12*10*8.67= 10 404**

$\Sigma x^2 = 1986$

$\Sigma y^2 = 1414$

$$r = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\sum x_i^2 - n\,\bar{y}^2} * \sqrt{\sum y_i^2 - n\,\bar{x}^2}}$$

Application numérique $r = \frac{1395-10404}{\sqrt{1968-902.03} * \sqrt{1414-1200}}$

$\boxed{r_{calc} = 0.503}$

$\boxed{r_{tab} = (0.05, ddl=10\ (12-2) = 0.576}$

Décision (conclusion du test) :

Étant donné que r_{cal} (0.503) < r_{tab} (0.756), nous pouvons conclure que :

la corrélation n'est pas statistiquement significative. Les deux variables sont indépendantes l'une de l'autre.

4.4.Taille de l'échantillon et signification

Comme nous l'avons déjà expliqué un peu plus haut, la taille de l'échantillon joue un rôle important dans la signification. À titre d'exemple, nous supposant que la taille de l'échantillon de l'exemple n = 1002 (et non pas 12) en maintenant toutes les valeurs calculées.

Donc le $\boxed{r_{calc} = 0.503}$

Alors que le $\boxed{r_{tab} = (0.05, ddl=1000\ (1002-2) = 0.062}$

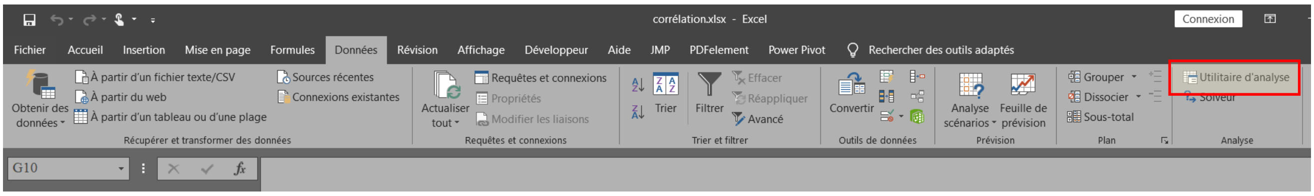
Dans ce cas de figure, la corrélation est positivement significative (même à 0.001 où r_{tab}= 0.104)

La corrélation comme pour les autres tests de comparaison est sensible à la taille de l'échantillon. C'est pour cela que dans la recherche scientifique il est fortement recommandé d'avoir une taille assez conséquente.

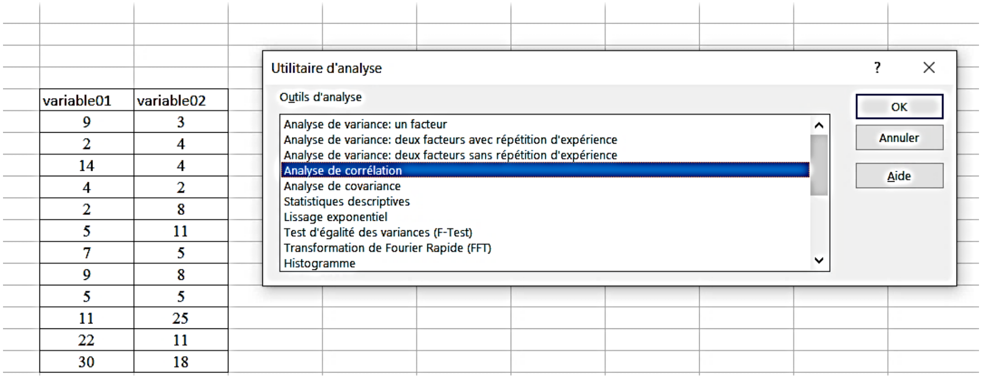
4.5.Coefficient de corrélation sous Excel

Il existe deux façons pour calculer le coefficient de corrélation sous Excel

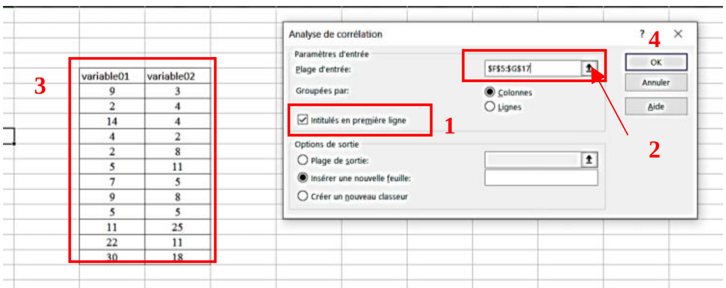
4.5.1. Coefficient de corrélation avec « Utilitaire d'analyse »



Étape 01 : aller dans l'onglet « Données ». cliquer sur « Utilitaire d'analyse »



Étape 02 : choisir « Analyse de corrélation » et cliquer sur « OK »



Étape 03 : définir la plage de donnée

1. Cocher « intitulé en première ligne »
2. Clique sur le bouton pour choisir
3. Sélectionner la plage de donnée en incluant les noms des variables (intitulés) comme démontré dans l'image.

	A	B	C
1		variable01	variable02
2	variable01	1	
3	variable02	0.50252986	1
4			
5			

Étape 04 : une nouvelle feuillue va apparaitre avec le tableau suivant

Ce tableau est une matrice de corrélation. Le coefficient entre la variable01 et la variable02 « r » ≈ 0.503 (comme nous l'avons trouvé avec le calcul manuel).

Cette méthode est plus intéressante dans la mesure où vous disposez de plusieurs variables et que vous voulez éviter de refaire l'analyse pour chaque paire.

Ex. : pour 7 variables nous avons 21 coefficients de corrélation à calculer.

4.5.2. Coefficient de corrélation avec « Formule Excel »

Dans la mesure où vous n'avez que deux variables, il est plus facile d'utiliser la formule « COEFFICIENT.CORRELATION » disponible dans Excel.

Pour le faire il suffit de :

- choisir une cellule vide
- taper « = »
- écrire « COEFFICIENT.CORRELATION »
- introduire la colonne « variable 01 » dans la « matrice 1 »
- taper « ; » point-virgule
- introduire la colonne « variable 02 » dans la « matrice 2 »
- taper « Entrée »

variable01	variable02
9	3
2	4
14	4
4	2
2	8
5	11
7	5
9	8
5	5
11	25
22	11
30	18

=COEFFICIENT.CORRELATION(F5:F17;G5:G17)

En suivant ces étapes, la valeur du coefficient de corrélation s'affiche dans la cellule : « 0.50252986 » (la même que les deux autres méthodes).

Conclusion

La corrélation est une technique très utilisée dans la recherche scientifique qui comporte des variables quantitatives. L'objectif de son utilisation consiste à étudier la relation (lien) entre ces variables.

Par exemple, il est intéressant de trouver une corrélation positive entre une variable « x » et la variable « performance sportive ». Ce qui signifie que la performance est plus élevée lorsque cette variable est élevée et *vice versa* . Cette relation peut éclaircir la nature et le profil de la réussite sportive.

Par contre, la corrélation n'établit pas un lien de « Cause à effet ». Ce qui veut dire qu'à aucun moment nous ne pouvons dire que la variable « x » influe sur la performance ou le contraire. Tout ce que nous pouvons dire est qu'elles sont liées positivement entre elles.

Tout de même, avec prudence et dans la présence des études scientifiques qui ont déjà établi ce lien de cause à effet entre ces deux variables, nous pouvons dire que la variable « x » influe sur la performance.

5. Le khi2

Ce test permet de vérifier l'absence de lien statistique entre deux variables qualitatives X et Y. Les deux sont dites indépendantes lorsqu'il n'existe aucun lien statistique entre elles.

Pour vérifier cette indépendance, nous calculons un $Khi2_{Calc}$ pour le comparer ensuite à une valeur tabulée *($Khi2_{tab}$)*.

Si $Khi2_{Calc} < Khi2_{tab}$: nous concluons qu'il n'existe pas de relation entre la variable X et la Variable Y.

Où :

$Khi2_{Calc}$: la valeur calculée du Khi2

$Khi2_{tab}$: la valeur tabulée qui correspond à **α=0.05** et **ddl = (Lignes -1)* (colonnes -1)**

Ce test est souvent utilisé pour le traitement des questionnaires en STAPS, où le chercheur émis l'hypothèse pour laquelle il y aurai une relation entre deux variables.

Formule :

$$\chi^2 = \sum \frac{\left(fo_i - f_{t_i}\right)^2}{f_{t_i}}$$

5.1.Test d'indépendance (deux variables qualitatives)

Dans ce cas, le chercheur étudie la relation entre deux variables qualitatives.

Les fréquences théoriques se calculent selon l'équation suivante :

$$f_{t_i} = \frac{\sum lignes\ X\ \sum colonnes}{Total}$$

Exemple :

Un chercheur se fixe l'objectif d'étudier la relation entre l'importance des facteurs de la performance ET le type de sport pratiqué. Les résultats de l'enquête sont présentés dans le tableau suivant.

	Physique	**Technique**	**Tactique**	**Psychologique**	**Σ**
Sports collectifs	35	40	70	35	**180**
Sport de combat	35	45	44	65	**189**
Sports cycliques	75	35	30	35	**175**
Gymnastique	30	60	10	30	**130**
Σ	**175**	**180**	**154**	**165**	**674**

1ère étape : calculer l'effectif théorique

Pour chaque case du tableau , nous appliquons la formule des fréquences théoriques

$$f_{t_i} = \frac{\sum lignes\ X\ \sum colonnes}{Total}$$

	Physique	Technique	Tactique	Psychologique
Sports collectifs	**(180*175)/674** =46.74	**(180*180)/674** = 48.07	41.13	44.07
Sport de combat	49.07	50.47	43.18	46.27
Sports cycliques	45.44	46.74	39.99	42.84
Gymnastique	33.75	34.72	29.70	31.82

2ème étape : calculer le Khi2

Pour chaque case nous appliquons la formule du Khi2. Le total du tableau représente le $Khi2_{Calc.}$

	Physique	Technique	Tactique	Psychologique	
Sports collectifs	**(35-46.74)2 /46.74=** 2.95	**$(40\text{-}48.07)^2/48.07=$** 1.36	20.27	1.86	
Sport de combat	4.04	0.59	0.02	7.58	
Sports cycliques	19.23	2.95	2.49	1.44	
Gymnastique	0.42	18.41	13.07	0.10	
					Σ=96.78

$$\boxed{\boldsymbol{Khi2_{cal} = 96.78}}$$

3ème étape : comparer la valeur du *Khi2 $_{cal}$* *avec celle du Khi2 $_{tab}$:*

ddl = (4-1)*(4-1)=9

dans la table nous lisons la valeur qui correspond à α= 0.05 ET ddl =9

α / n	0.8	0.7	0.5	0.3	0.2	0.1	0.05	0.(
1	0.06	0.15	0.46	1.07	1.64	2.71	3.84	5.
2	0.45	0.71	1.39	2.41	3.22	4.61	5.99	7.
3	1.01	1.42	2.37	3.66	4.64	6.25	7.81	9.
4	1.65	2.19	3.36	4.88	5.99	7.78	9.49	11
5	2.34	3.00	4.35	6.06	7.29	9.24	11.07	12
6	3.07	3.83	5.35	7.23	8.56	10.64	12.59	14
7	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16
8	4.59	5.53	7.34	9.52	11.03	13.36	15.51	17
9	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19
10	6.18	7.27	9.34	11.78	13.44	15.99	18.31	20
11	6.99	8.15	10.34	12.90	14.63	17.28	19.68	21
12	7.81	9.03	11.34	14.01	15.81	18.55	21.03	23
13	8.63	9.93	12.34	15.12	16.98	19.81	22.36	24
14	9.47	10.82	13.34	16.22	18.15	21.06	23.68	26
15	10.31	11.72	14.34	17.32	19.31	22.31	25.00	27

Le ***Khi2 $_{tab}$* = 16.92**

Interprétation :

Étant donné que la valeur du ***Khi2*** $_{cal}$ **(96.78)** > ***Khi2*** $_{tab}$ **= 16.92,** nous pouvons conclure qu'il existe une relation statistiquement significative (au seuil α=0.05) entre l'importance des facteurs de la performance ET le type de sport pratiqué.

Pour passer à l'interprétation détailler de cette relation , nous lisons tout d'abord les valeurs les élevées dans le tableau du Khi2 :

	Physique	Technique	Tactique	Psychologique	
Sports collectifs	2.95	1.36	**20.27**	1.86	
Sport de combat	4.04	0.59	0.02	7.58	
Sports cycliques	**19.23**	2.95	2.49	1.44	
Gymnastique	0.42	**18.41**	**13.07**	0.10	
					Σ=96.78

Après qu'on a détecté les valeurs les plus élevées, nous procédons à la comparaison de chaque valeur du tableau des fréquences observé qui correspond à ces cases avec les valeurs de la même ligne et la même colonne. Nous rapportons ensuite si cette valeur est la plus élevée ou la plus faible.

- **Khi2 = 20.27** : Les individus qui pratiquent un « **Sport collectif** » estiment que la **Tactique** est le facteur le plus important (70 est la fréquence la plus élevée pour ce type de sport)
- **Khi2 = 19.23 :** les individus qui pratiquent un « **Sport cyclique** » estiment que le facteur **physique** est le plus important **(**75 est la fréquence la plus élevée pour ce type de sport).
- **Khi2 = 18.41 :** les individus qui pratiquent la « **gymnastique** » estiment que la technique est **Technique** est le facteur le plus important (70 est la fréquence la plus élevée pour ce type de sport)
- **Khi2 = 13.07 : :** les individus qui pratiquent la « **gymnastique** » estiment que la technique est **Tactique** est le facteur le **moins** important (10 est la fréquence la plus faible pour ce type de sport).

5.2.Test d'homogénéité (échantillon et population)

Dans ce cas le chercheur compare une distribution observée (Echantillon) avec une distribution théorique (Population). Il cherche à vérifier si l'échantillon est tiré de cette population, autrement dit, l'échantillon peut-il représenter cette population.

Exemple :

Le chercheur veut vérifier si la distribution de son échantillon en termes de relation entre « le temps de pratique sportive par semaine » ET « l'âge » est différente de celle de la population.

	2h-4h	5h-7h	8h-10h	11h-13h	>14h
20ans-23ans	10	13	11	10	35
24ans-27ans	25	10	12	15	11
28ans-31ans	11	13	15	15	10
32ans-35ans	12	30	25	20	13
36ans-39ans	12	18	12	22	12
>40 ans	25	10	8	7	8
	Σ=450				

Tableau 10. Tableau des fréquences observées (échantillon)

	2h-4h	5h-7h	8h-10h	11h-13h	>14h
20ans-23ans	1043	1363	1179	1019	3588
24ans-27ans	2542	1039	1279	1599	1140
28ans-31ans	1127	1300	1537	1542	1070
32ans-35ans	1272	3036	2553	2048	1370
36ans-39ans	1282	1805	1293	2201	1164
>40 ans	2552	1049	879	798	844
	Σ=46513				

Tableau 11. Tableau de la distribution de la Population

1ère étape : Calculer le *Pi (Pourcentage)*

Pour chaque case du tableau de la population , nous calculons le Pi.

$$P_i = \frac{n_{pi}}{N_p}$$

Où :

n_{pi} : la fréquence de la case du tableau de la population

N_p : La taille de la population

	2h-4h	5h-7h	8h-10h	11h-13h	>14h
20ans-23ans	**1043/46513= 0.022**	**1363/46513= 0.029**	0.025	0.022	0.077
24ans-27ans	0.055	0.022	0.027	0.034	0.025
28ans-31ans	0.024	0.028	0.033	0.033	0.023

32ans-35ans	0.027	0.065	0.055	0.044	0.029
36ans-39ans	0.028	0.039	0.028	0.047	0.025
>40 ans	0.055	0.023	0.019	0.017	0.018

Tableau 12. Tableau des Pi

2ème étape : Calculer les fréquences théoriques

$$f_{oi} = N_E P_i$$

où :

N_E : la taille de l'échantillon

P_i : le pourcentage de la population

	2h-4h	**5h-7h**	**8h-10h**	**11h-13h**	**>14h**
20ans-23ans	10.09	13.19	11.41	9.86	34.71
24ans-27ans	24.59	10.05	12.37	15.47	11.03
28ans-31ans	10.90	12.58	14.87	14.92	10.35
32ans-35ans	12.31	29.37	24.70	19.81	13.25
36ans-39ans	12.40	17.46	12.51	21.29	11.26
>40 ans	24.69	10.15	8.50	7.72	8.17

Tableau 13. Tableau des fréquences théoriques

3ème étape : Calculer le Khi2

	2h-4h	**5h-7h**	**8h-10h**	**11h-13h**	**>14h**
20ans-23ans	**$(10-10.09)^2/10.09$=** 0.001	0.003	0.014	0.002	0.002
24ans-27ans	0.007	0.000	0.011	0.014	0.000
28ans-31ans	0.001	0.014	0.001	0.000	0.012
32ans-35ans	0.008	0.013	0.004	0.002	0.005
36ans-39ans	0.013	0.017	0.021	0.023	0.048
>40 ans	0.004	0.002	0.030	0.067	0.003
	Σ= 0.344				

Le Khi2 $_{cal}$ = 0.344

4ème étape : comparer la valeur du *Khi2 $_{cal}$* *avec celle du Khi2 $_{tab}$:*

Khi $_{tab}$ est la valeur qui correspond -dans la table du χ^2 – à α= 0.05 ET ddl= (lignes-1)*(colonnes -1)

ddl= (6-1)*(5-1)=20

α / n	0.8	0.7	0.5	0.3	0.2	0.1	0.05
1	0.06	0.15	0.46	1.07	1.64	2.71	3.84
2	0.45	0.71	1.39	2.41	3.22	4.61	5.99
3	1.01	1.42	2.37	3.66	4.64	6.25	7.81
4	1.65	2.19	3.36	4.88	5.99	7.78	9.49
5	2.34	3.00	4.35	6.06	7.29	9.24	11.07
6	3.07	3.83	5.35	7.23	8.56	10.64	12.59
7	3.82	4.67	6.35	8.38	9.80	12.02	14.07
8	4.59	5.53	7.34	9.52	11.03	13.36	15.51
9	5.38	6.39	8.34	10.66	12.24	14.68	16.92
10	6.18	7.27	9.34	11.78	13.44	15.99	18.31
11	6.99	8.15	10.34	12.90	14.63	17.28	19.68
12	7.81	9.03	11.34	14.01	15.81	18.55	21.03
13	8.63	9.93	12.34	15.12	16.98	19.81	22.36
14	9.47	10.82	13.34	16.22	18.15	21.06	23.68
15	10.31	11.72	14.34	17.32	19.31	22.31	25.00
16	11.15	12.62	15.34	18.42	20.47	23.54	26.30
17	12.00	13.53	16.34	19.51	21.61	24.77	27.59
18	12.86	14.44	17.34	20.60	22.76	25.99	28.87
19	13.72	15.35	18.34	21.69	23.90	27.20	30.14
20	14.58	16.27	19.34	22.77	25.04	28.41	31.41

$\boxed{Khi2_{tab} = 31.41}$

Interprétation :

$Khi2_{calc}$ (0.344) < $Khi2_{tab}$ (31.41): il n'existe pas de différence entre la distribution de l'échantillon et celle de la population. Nous pouvons conclure que l'échantillon peut représenter la population ou que l'échantillon est tiré cette population.

6. Analyse factorielle

Elle est utilisée pour simplifier la compréhension et le traitement d'ensembles de données complexes qui comptent de plusieurs variables. En effet, cette technique statistique permet de regrouper les variables similaires et de simplifier la compréhension des résultats de recherche.

L'analyse factorielle est un moyen de condenser les données de nombreuses variables en quelques variables seulement pour simplifier la compréhension des résultats. Pour cette raison, elle est aussi parfois appelée « réduction de dimension ». Vous pouvez réduire les « dimensions » de vos données en une ou plusieurs «variables latentes = Facteurs».

Elle est basée sur la relation entre les variables, en regroupons les variables les plus corrélées entre elles en un seul Facteur. Cette technique est utilisée pour :

- Explorer les données
- Confirmer l'existence d'une variable latente qui regroupe un ensemble de variables observables
- Représenter sur un graphique plusieurs variables à la fois

Dans ce cours, nous allons aborder l'Analyse en Composantes Principales (ACP) qui est la technique la plus utilisée dans la recherche en STAPS.

Il existe d'autres techniques qui suivent plus en moins la même procédure pour différent type de variables et différentes hypothèses. Nous citons par exemple l'Analyse Factorielle des Correspondances (AFC) pour traiter deux variables qualitatives issues des questionnaires par exemple ; l'analyse Factorielle Discriminante (AFD) qui est une forme de régression avec une variable dépendante qualitative (nominale). L'analyse des correspondances multiples (ACM) qui est une approche plus complexe de la AFC avec plus de 2 variables.

Les calculs de ces analyses sont très lourds à réaliser avec la main, de ce fait nous nous contentons de présenter la procédure à suivre avec les logiciels de statistique et mettre l'accent sur l'interprétation des résultats qui représente la finalité de l'analyse statistique.

6.1.L'analyse en composantes principales (ACP)

L'ACP consiste à transformer des variables corrélées statistiquement (c'est-à-dire liées entre elles) en nouvelles variables «Facteurs ». Ces nouvelles variables sont appelées « composantes principales » elle permet à un analyste de réduire le nombre de variables, de simplifier une analyse et de pouvoir identifier le facteur qui provoque le plus de variances.

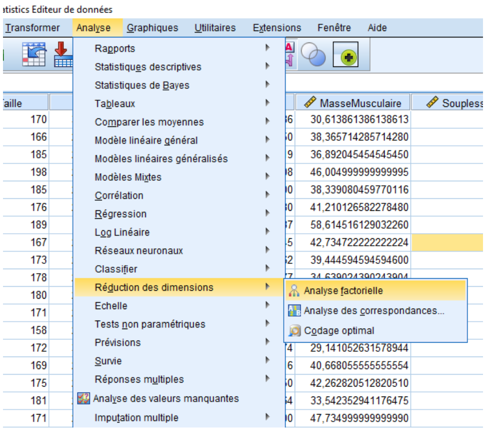
Exemple :

Un chercheur a réalisé des évaluations multiples sur un groupe de sportif (30). Il en a ressorti avec 8 variables. Dans le but de visualiser l'ensemble des relations entre ces différentes variables et la possibilité de trouver des variables latentes (caché) permettant de réduire le nombre élevé de variables en quelques facteurs représentatifs, il est nécessaire de conduire une ACP. Ce tableau représente un échantillon 3 individus de l'ensemble des données que vous trouvez à la fin de ce chapitre

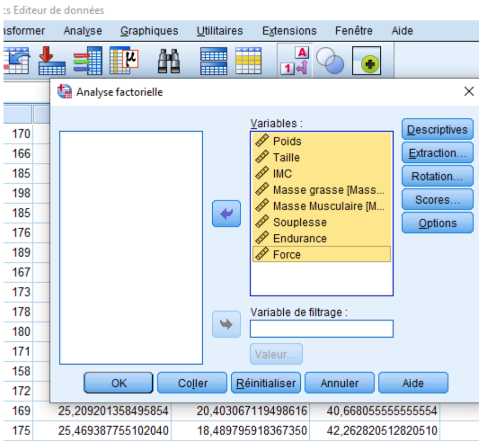
indi	Poids	Taille	IMC	Masse grasse	Masse Musculaire	Souplesse	Endurance	Force
1	101	170	34.95	19.65	30.6	28	19	102.26
2	70	166	25.40	17.47	38.4	33	15	126.33
3	88	185	25.71	12.57	36.9	15	13	154.85

6.1.1. Procédure ACP sous SPSS

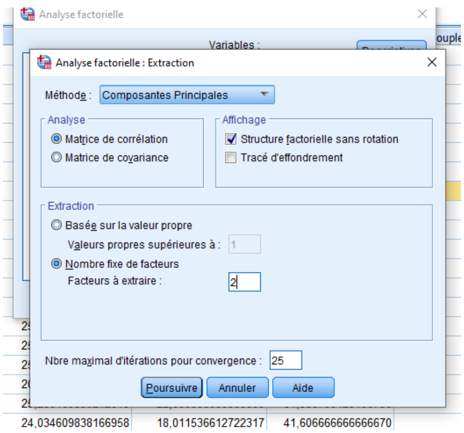
1. Aller sur « Analyse » →Réduction des dimensions → Analyse Factorielle



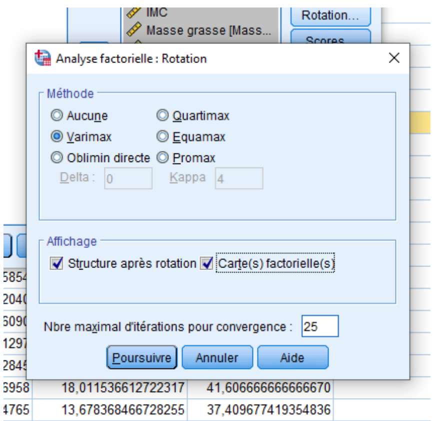
2. Dans la boite de dialogue, entrer toutes variable



3. Cliquer le bouton extraction et choisir « Nombre fixe de facteurs» et spécifier « 2 »



4. Cliquer sur le bouton rotation → cocher « Varimax »→cocher « cartes factorielles »



6.1.2. Interprétation des résultats de l'ACP

Tableau 14. Matrice des corrélations	**Poids**	**Taille**	**IMC**	**Masse grasse**	**Masse Musculaire**	**Souplesse**	**Endurance**	**Force**
Poids	1,0000	0,5358	0,6388	0,2659	-0,6646	-0,5462	-0,1386	0,0957
Taille	0,5358	1,0000	-0,2989	-0,2258	0,1221	-0,9905	-0,2306	0,1636
IMC	**0,6388**	-0,2989	1,0000	0,5147	-0,8237	0,2762	0,0844	-0,0830
Masse grasse	0,2659	-0,2258	0,5147	1,0000	-0,4694	0,2349	0,2400	-0,2299
Masse Musculaire	**-0,6646**	0,1221	**-0,8237**	-0,4694	1,0000	-0,1209	0,1622	-0,1527
Souplesse	-0,5462	**-0,9905**	0,2762	0,2349	-0,1209	1,0000	0,2278	-0,1605
Endurance	-0,1386	-0,2306	0,0844	0,2400	0,1622	0,2278	1,0000	-0,9377
Force	0,0957	0,1636	-0,0830	-0,2299	-0,1527	-0,1605	**-0,9377**	1,0000

Ce tableau représente les valeurs du coefficient de corrélation de Pearson. Il convient de rappeler que les valeurs de R varient entre -1 et 1, plus la valeur se rapproche de ces deux extrêmes, plus la corrélation est forte. Le signe (+ ou -) indique l'orientation de la corrélation positive ou négative. Les valeurs qui se rapprochent de 0 indiquent l'absence de relation entre les deux variables.

Une simple lecture permet parfois de détecter des relations importantes. Par exemple en choisissant seulement les coefficients dont la valeur absolue est supérieure à 0.6 (en gras), nous remarquons que :

- La taille est corrélée négativement avec la souplesse (R=-0,9905)
- L'endurance est corrélée négativement avec la force (R=-0,9377)
- L'IMC est corrélé négativement avec la masse musculaire (R=-0,8237)
- Le poids est corrélé positivement avec l'IMC (R=0,6388)
- Le poids est corrélé négativement avec la masse musculaire (R=-0,6646)

Déjà avec ces résultats, nous pouvons dresser une image de la nature de relation qu'existe entre ces variables.

Tableau 15. Variance totale expliquée

Variance totale expliquée

	Valeurs propres initiales			Sommes extraites du carré des chargements			Sommes de rotation du carré des chargements		
Composante	Total	% de la variance	% cumulé	Total	% de la variance	% cumulé	Total	% de la variance	% cumulé
1	2,803	35,037	35,037	2,803	35,037	35,037	2,749	34,359	34,359
2	2,686	33,576	68,613	2,686	33,576	68,613	2,740	34,254	68,613
3	1,722	21,529	90,142						
4	,564	7,044	97,186						
5	,152	1,899	99,085						
6	,060	,751	99,836						
7	,011	,133	99,969						
8	,002	,031	100,000						

Méthode d'extraction : Analyse en composantes principales.

Ce tableau représente la variance expliquée par les facteurs (Composantes principales) que nous avons retenus. Ces valeurs nous renseignent sur la variance expliquée par les deux facteurs, plus la valeur de la variance expliquée est élevée, plus les résultats sont proches de la réalité. Le tableau démontre que la Composante 1 explique ≈35% et la composante explique 2 ≈34%, ce qui nous donne une variance expliquée (cumulée) de 69%. Cela veut dire que seulement 31% de la variance n'a pas été expliquée ou représentée par ces deux composantes (facteurs).

En général, plus le nombre de facteurs retenu est élevé, plus la variance expliquée sera élevée. Dans notre exemple nous avons choisi de retenir que deux facteurs pour faciliter la lecture du graphique, mais il est possible de laisser le logiciel trouver le nombre adéquat de facteurs pour obtenir une variance expliquée plus élevé.

Tableau 16. Variance totale expliquée avec 3 composantes

Variance totale expliquée

	Valeurs propres initiales			Sommes extraites du carré des chargements			Sommes de rotation du carré des chargements		
Composante	Total	% de la variance	% cumulé	Total	% de la variance	% cumulé	Total	% de la variance	% cumulé
1	2,803	35,037	35,037	2,803	35,037	35,037	2,743	34,291	34,291
2	2,686	33,576	68,613	2,686	33,576	68,613	2,421	30,266	64,558
3	1,722	21,529	90,142	1,722	21,529	90,142	2,047	25,585	90,142
4	,564	7,044	97,186						
5	,152	1,899	99,085						
6	,060	,751	99,836						
7	,011	,133	99,969						
8	,002	,031	100,000						

Méthode d'extraction : Analyse en composantes principales.

Tableau 17. Valeurs des coefficients de corrélation entre les composantes de l'ACP et les variables après rotation (Varimax)

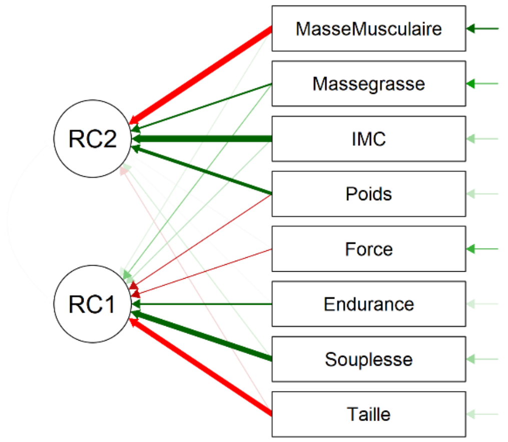
En laissant le logiciel choisir, il a retenu 3 composantes, ce qui a augmenté la valeur de la variance expliquée à 90%. En revanche, la présentation des variables sur des graphiques devient plus difficile à interpréter (3 dimensions XYZ).

Rotation de la matrice des composantes[a]

	Composante	
	1	2
Poids	-,564	,751
Taille	-,869	-,105
IMC	,170	,935
Masse grasse	,320	,657
Masse Musculaire	,071	-,922
Souplesse	,870	,096
Endurance	,652	-,001
Force	-,598	-,007

Ce tableau représente les valeurs des coefficients de corrélation entre les composantes de l'ACP et les variables après rotation (Varimax). À partir de ces coefficients, nous pouvons attribuer des libellés pour chaque facteur. En gardant que les coefficients les plus élevés pour chaque variable avec l'un des deux facteurs, nous pouvons remarquer que le premier facteur est corrélé avec la taille, et les trois qualités physiques (souplesse, endurance et force), ce qui nous permet de libeller la *première composante* par « Qualités physiques ». La deuxième composante est corrélée davantage avec le poids, l'IMC, la masse corporelle, et la masse grasse, ce qui nous permet de libeller la *deuxième composante* par « Composition Corporelle ».

	C1	C2
Poids		0.752
Taille	-0.869	
IMC		0.934
Masse grasse		0.657
Masse Musculaire		-0.922
Souplesse	0.870	
Endurance	0.652	
Force	-0.598	



RC1 & RC2 : Composante 1 & composante 2

Lignes vertes : corrélation positive ; Lignes rouges : corrélation négative ; épaisseur des traits : puissance de la corrélation

Figure 12. diagramme de corrélation des composantes principales avec les variables

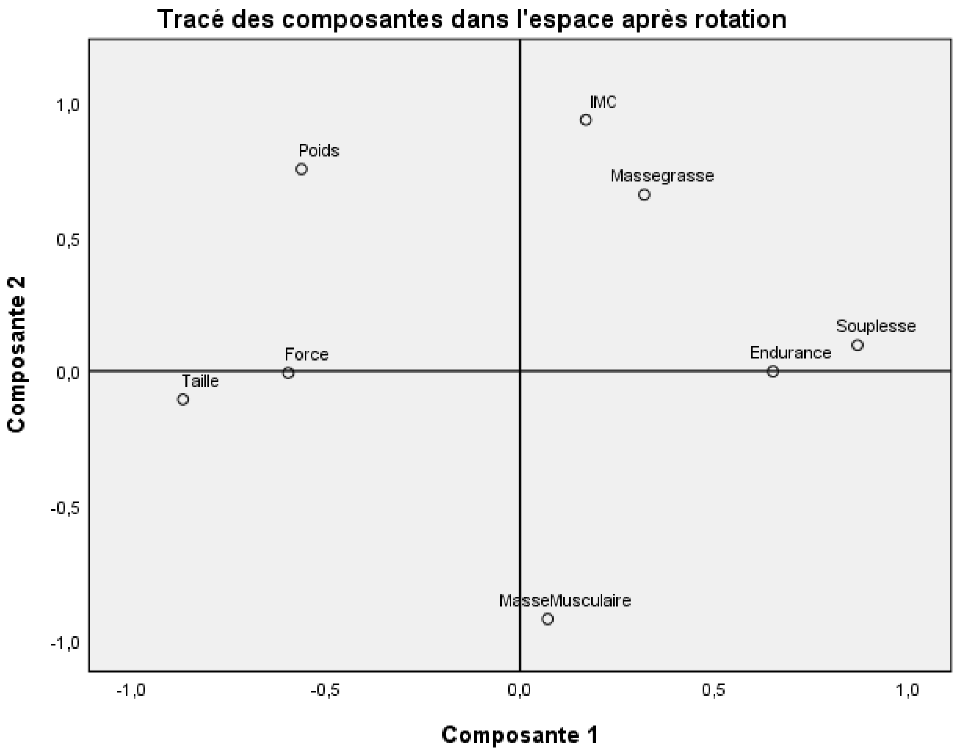


Figure 13. Illustration de la position des variables par rapport aux deux composantes principales

Cette figure représente la corrélation entre les variables étudiées et les deux composantes principales (facteurs). Pour interpréter ce graphique, il faut savoir les informations suivantes :

- La distance entre une variable et le centre du graphique (x=0 :y=0) représente l'importance de la variable dans l'analyse : plus elle est loin plus sa participation dans l'ACP est meilleure
- La distance entre la variable et l'axe (facteur) renseigne sur la puissance de corrélation entre cette variable et ce facteur
- Lorsque deux variables sont très proches l'une de l'autre, on considère qu'elles sont corrélées positivement entre elles
- Lorsque deux variables se trouvent à une distance similaire par rapport à un facteur, mais dans des côtés opposés, on considère que ces deux variables présentent une corrélation négative.

Lecture :

La souplesse est l'endurance sont corrélées positivement entre elle et avec le facteur 01. Mais elles sont corrélées négativement avec la taille et la force. Nous pouvons dire que les sportifs (dans cette étude) avec une grande taille tendent à être plus forts (force), mais moins endurants et moins souples.

D'autre part, ceux qui présentent un poids, une masse grasse et un IMC élevés possèdent une masse musculaire faible.

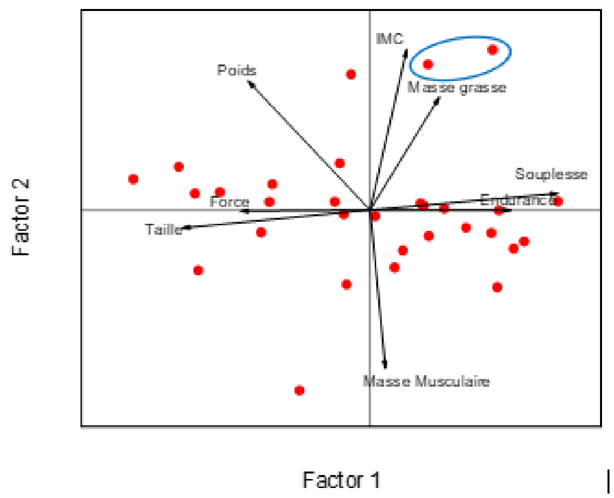


Figure 14. représentation des individus dans le graphique de l'ACP

En ajoutant les individus au graphique, nous pouvons détecter le profil de chaque individu par rapport à une variable donnée. Par exemple, les deux individus dans le cercle bleu présentent un IMC et une Masse grasse élevés, ils sont plus endurants et souples, mais ils présentent une qualité de force et une taille inférieure aux autres.

Tableau 18. Données complètes de l'exemple (ACP)

indi	Poids	Taille	IMC	Masse grasse	Masse Musculaire	Souplesse	Endurance	Force
1	101	170	34.95	19.65	30.6	28	19	102.26
2	70	166	25.40	17.47	38.4	33	15	126.33
3	88	185	25.71	12.57	36.9	15	13	154.85
4	80	198	20.41	14.80	46.0	4	15	125.33
5	87	185	25.42	13.00	38.3	13	16	122.00
6	79	176	25.50	9.50	41.2	25	13	159.85
7	62	189	17.36	6.79	58.6	10	17	114.65
8	72	167	25.82	14.61	42.7	31	19	115.26
9	74	173	24.73	16.24	39.4	25	14	132.86
10	82	178	25.88	9.63	34.6	21	17	116.65
11	75	180	23.15	6.72	46.2	19	18	113.11
12	76	171	25.99	14.66	41.5	28	16	135.00
13	63	158	25.24	9.41	40.6	40	13	155.85
14	95	172	32.11	22.70	29.1	28	15	132.33
15	72	169	25.21	20.40	40.7	30	16	127.00
16	78	175	25.47	18.49	42.3	23	20	100.00
17	85	181	25.95	9.65	33.5	20	13	154.85
18	60	171	20.52	18.00	47.7	30	20	109.00
19	81	179	25.28	22.00	34.6	23	16	134.00
20	60	158	24.03	18.01	41.6	42	13	162.85
21	93	189	26.04	13.68	37.4	13	18	120.11
22	65	160	25.39	11.00	42.4	42	17	107.65
23	97	195	25.51	13.50	36.2	4	14	146.86
24	65	161	25.08	19.36	36.9	41	19	101.26
25	70	165	25.71	15.00	43.9	37	18	111.11
26	94	189	26.32	14.77	35.0	13	13	147.85
27	69	164	25.65	14.55	37.0	34	19	106.26
28	65	162	24.77	12.26	40.4	40	19	108.26
29	94	160	36.72	22.24	30.2	39	18	114.11
30	62	162	23.62	7.87	40.3	37	14	149.86

Références

Références à consulter

Français

Albarello, L., Bourgeois, E., et Guyot, J.-L. (2007). *Statistique descriptive: Un outil pour les praticiens-chercheurs* : De Boeck Supérieur.

Baccini, A. (2010a). Statistique descriptive élémentaire. *Institut de Mathématiques de Toulouse*.

Baccini, A. (2010b). *Statistique Descriptive Multidimensionnelle (pour les nuls)*. Institut de Toulouse.

Beaufils, B. (2000). *Statistiques appliquées à la psychologie: statistiques descriptives* (Vol. 1): Editions Bréal.

Bressoud, E., et Kahané, J.-C. (2010). *Statistique descriptive* : Pearson Education France.

Champely, S. (2003). *Statistique vraiment appliquée au sport: cours et exercices* : De Boeck Supérieur.

Cheze, C. (2003). Statistique inférentielle: Estimation. *Techniques de l'ingénieur, AF168*.

Demotes-Mainard, M. (2003). La connaissance statistique de l'immatériel. *Groupe de Voorburg sur la statistique des services*.

Dodge, Y. (1999). *Premiers pas en statistique*. Paris: Springer-Verlag.

Dodge, Y. (2004). *Statistique: dictionnaire encyclopédique* : Springer Science & Business Media.

Dreyfus, G. (2008). *Apprentissage statistique* : Editions Eyrolles.

Goldfarb, B., et Pardoux, C. (2011). *Introduction à la méthode statistique-6e éd.: Économie, gestion* : Hachette.

Haccoun, R. R., et Cousineau, D. (2007). *Statistiques: Concepts et applications* : PUM.

Leboucher, L., et Voisin, M.-J. (2013). *Introduction à la statistique descriptive: cours et exercices avec tableur* : Cépaduès éd.

Lecoutre, J.-P. (2016). *Statistique et probabilités-6e éd.: Cours et exercices corrigés* : Dunod.

Lejeune, M. (2004). *Statistique: La théorie et ses applications* : Springer Science & Business Media.

Méot, A. (2003). *Introduction aux statistiques inférentielles: De la logique à la pratique* : De Boeck Supérieur.

Protassov, K. (2012). *Analyse statistique de données expérimentales* : EDP Sciences.

Rousson, V. (2013). *Statistique appliquée aux sciences de la vie* : Springer.

Vidal, A. (2010). *Statistique descriptive et inférentielle avec Excel: approche par l'exemple* : Presses universitaires de Rennes.

Anglais

Albert, J., Glickman, M. E., Swartz, T. B., et Koning, R. H. (2017). *Handbook of statistical methods and analyses in sports* : CRC Press.

Aron, A., et Aron, E. N. (1999). *Statistics for psychology* : Prentice-Hall, Inc.

McCall, R. B., et Kagan, J. (1975). *Fundamental statistics for psychology*. repéré à

Morgan, G. A., Leech, N. L., Gloeckner, G. W., et Barrett, K. C. (2004). *SPSS for introductory statistics: Use and interpretation* : Psychology Press.

Tabachnick, B. G., et Fidell, S. L. (2007). Using multivariate statistics. *Using Multivariate Statistics*.

Vogt, W. P. (1999). *Dictionary of statistics and methodology* (Vol. null).

Logiciels

JASP: https://jasp-stats.org/ (Gratuit)

Jamovi: https://www.jamovi.org/ (Gratuit)

SPSS: https://www.ibm.com/products/spss-statistics (Payant)

GraphPad: https://www.graphpad.com/scientific-software/prism/ (Payant)

Tables Statistiques

Table 7 : Coefficients de Shapiro-Wilk :

n = taille de l'échantillon, i = numéro de la différence d_i

n \ i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	0,707														
3	0,707	0													
4	0,687	0,168													
5	0,665	0,241	0												
6	0,643	0,281	0,088												
7	0,623	0,303	0,14	0											
8	0,605	0,316	0,174	0,056											
9	0,589	0,324	0,198	0,095	0										
10	0,574	0,329	0,214	0,122	0,04										
11	0,56	0,332	0,226	0,143	0,07	0									
12	0,548	0,333	0,235	0,159	0,092	0,03									
13	0,536	0,333	0,241	0,171	0,11	0,054	0								
14	0,525	0,332	0,246	0,18	0,124	0,073	0,024								
15	0,515	0,331	0,25	0,188	0,135	0,088	0,043	0							
16	0,506	0,329	0,252	0,194	0,145	0,101	0,059	0,02							
17	0,496	0,327	0,254	0,199	0,152	0,111	0,073	0,036	0						
18	0,489	0,325	0,255	0,203	0,159	0,12	0,084	0,05	0,016						
19	0,481	0,323	0,256	0,206	0,164	0,127	0,093	0,061	0,03	0					
20	0,473	0,321	0,257	0,209	0,169	0,133	0,101	0,071	0,042	0,014					
21	0,464	0,319	0,258	0,212	0,174	0,14	0,109	0,08	0,053	0,026	0				
22	0,459	0,316	0,257	0,213	0,176	0,144	0,115	0,088	0,062	0,037	0,012				
23	0,454	0,313	0,256	0,214	0,179	0,148	0,12	0,094	0,07	0,046	0,023	0			
24	0,449	0,31	0,255	0,215	0,181	0,151	0,125	0,1	0,076	0,054	0,032	0,011			
25	0,445	0,307	0,254	0,215	0,182	0,154	0,128	0,105	0,082	0,061	0,04	0,02	0		
26	0,441	0,304	0,253	0,215	0,184	0,156	0,132	0,109	0,088	0,067	0,048	0,028	0,009		
27	0,437	0,302	0,252	0,215	0,185	0,158	0,135	0,113	0,092	0,073	0,054	0,036	0,018	0	
28	0,433	0,299	0,251	0,215	0,186	0,16	0,137	0,116	0,097	0,078	0,06	0,042	0,025	0,008	
29	0,429	0,297	0,25	0,215	0,106	0,162	0,14	0,119	0,1	0,082	0,065	0,048	0,032	0,016	0
30	0,425	0,294	0,249	0,215	0,187	0,163	0,142	0,122	0,104	0,086	0,07	0,054	0,038	0,023	0,008

Table des valeurs limites de W

n	5%	1%
3	0.767	0.753
4	0.748	0.687
5	0.762	0.686
6	0.788	0.713
7	0.803	0.730
8	0.818	0.749
9	0.829	0.764
10	0.842	0.781
11	0.850	0.792
12	0.859	0.805
13	0.856	0.814
14	0.874	0.825
15	0.881	0.835
16	0.837	0.844
17	0.892	0.851
18	0.897	0.858
19	0.901	0.863
20	0.905	0.868
21	0.908	0.873
22	0.911	0.878
23	0.914	0.881
24	0.916	0.884
25	0.918	0.888
26	0.920	0.891
27	0.923	0.894
28	0.924	0.896
29	0.926	0.898
30	0.927	0.900

Table de Lilliefors

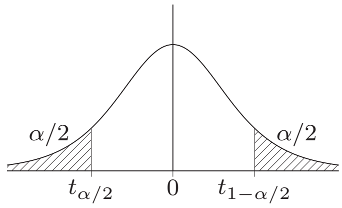
α \ n	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	25	30	n>30
0,1	0,352	0,15	0,294	0,276	0,261	0,249	0,239	0,23	0,223	0,214	0,207	0,201	0,195	0,189	0,184	0,179	0,174	0,158	0,144	$0,805/\sqrt{n}$
0,05	0,381	0,337	0,19	0,3	0,285	0,271	0,258	0,249	0,243	0,234	0,227	0,22	0,213	0,206	0,2	0,195	0,19	0,173	0,161	$0,886/\sqrt{n}$
0,01	0,417	0,405	0,364	0,348	0,311	0,311	0,294	0,284	0,242	0,268	0,261	0,257	0,25	0,245	0,239	0,235	0,231	0,2	0,187	$1,031/\sqrt{n}$

A.3. Lois de Student

Si T est une variable aléatoire suivant la loi de Student à ν degrés de liberté, la table donne, pour α fixé, la valeur $t_{1-\alpha/2}$ telle que

$$\mathbb{P}\{|T| \geqslant t_{1-\alpha/2}\} = \alpha.$$

Ainsi, $t_{1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ de la loi de Student à ν degrés de liberté.



ν \ α	0,900	0,500	0,300	0,200	0,100	0,050	0,020	0,010	0,001
1	0,1584	1,0000	1,9626	3,0777	6,3138	12,7062	31,8205	63,6567	636,6193
2	0,1421	0,8165	1,3862	1,8856	2,9200	4,3027	6,9646	9,9248	31,5991
3	0,1366	0,7649	1,2498	1,6377	2,3534	3,1824	4,5407	5,8409	12,9240
4	0,1338	0,7407	1,1896	1,5332	2,1318	2,7764	3,7469	4,6041	8,6103
5	0,1322	0,7267	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321	6,8688
6	0,1311	0,7176	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9588
7	0,1303	0,7111	1,1192	1,4149	1,8946	2,3646	2,9980	3,4995	5,4079
8	0,1297	0,7064	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	5,0413
9	0,1293	0,7027	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,1289	0,6998	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	4,5869
11	0,1286	0,6974	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	4,4370
12	0,1283	0,6955	1,0832	1,3562	1,7823	2,1788	2,6810	3,0545	4,3178
13	0,1281	0,6938	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	4,2208
14	0,1280	0,6924	1,0763	1,3450	1,7613	2,1448	2,6245	2,9768	4,1405
15	0,1278	0,6912	1,0735	1,3406	1,7531	2,1314	2,6025	2,9467	4,0728
16	0,1277	0,6901	1,0711	1,3368	1,7459	2,1199	2,5835	2,9208	4,0150
17	0,1276	0,6892	1,0690	1,3334	1,7396	2,1098	2,5669	2,8982	3,9651
18	0,1274	0,6884	1,0672	1,3304	1,7341	2,1009	2,5524	2,8784	3,9216
19	0,1274	0,6876	1,0655	1,3277	1,7291	2,0930	2,5395	2,8609	3,8834
20	0,1273	0,6870	1,0640	1,3253	1,7247	2,0860	2,5280	2,8453	3,8495
21	0,1272	0,6864	1,0627	1,3232	1,7207	2,0796	2,5176	2,8314	3,8193
22	0,1271	0,6858	1,0614	1,3212	1,7171	2,0739	2,5083	2,8188	3,7921
23	0,1271	0,6853	1,0603	1,3195	1,7139	2,0687	2,4999	2,8073	3,7676
24	0,1270	0,6848	1,0593	1,3178	1,7109	2,0639	2,4922	2,7969	3,7454
25	0,1269	0,6844	1,0584	1,3163	1,7081	2,0595	2,4851	2,7874	3,7251
26	0,1269	0,6840	1,0575	1,3150	1,7056	2,0555	2,4786	2,7787	3,7066
27	0,1268	0,6837	1,0567	1,3137	1,7033	2,0518	2,4727	2,7707	3,6896
28	0,1268	0,6834	1,0560	1,3125	1,7011	2,0484	2,4671	2,7633	3,6739
29	0,1268	0,6830	1,0553	1,3114	1,6991	2,0452	2,4620	2,7564	3,6594
30	0,1267	0,6828	1,0547	1,3104	1,6973	2,0423	2,4573	2,7500	3,6460
40	0,1265	0,6807	1,0500	1,3031	1,6839	2,0211	2,4233	2,7045	3,5510
60	0,1262	0,6786	1,0455	1,2958	1,6706	2,0003	2,3901	2,6603	3,4602
80	0,1261	0,6776	1,0432	1,2922	1,6641	1,9901	2,3739	2,6387	3,4163
120	0,1259	0,6765	1,0409	1,2886	1,6577	1,9799	2,3578	2,6174	3,3735
∞	0,1257	0,6745	1,0364	1,2816	1,6449	1,9600	2,3263	2,5758	3,2905

Lorsque $\nu = \infty$, $t_{1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ de la loi normale $\mathcal{N}(0,1)$.

TABLE DE MANN-WHITNEY

Valeurs critiques (U_{crit}) à comparer avec la valeur observée (U_{obs}) à partir de vos 2 échantillons pour un test **bilatéral** au seuil α = **0.05** ou **0.01**.

NB : n_1 et n_2 représentent le nombre d'observations dans chaque échantillon.

n_2	α	n_1																	
		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	--	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	.05	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	.01	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	.05	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
	.01	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	.05	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
	.01	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	56	60
14	.05	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
	.01	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	.05	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
	.01	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	.05	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
	.01	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	.05	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
	.01	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	.05	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
	.01	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	.05	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
	.01	3	7	12	17	22	28	33	39	45	51	56	63	69	74	81	87	93	99
20	.05	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
	.01	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

D.5 Valeurs critiques du test de rangs signés de Wilcoxon pour échantillons appariés

Source : http://www.cons-dev.org/elearning/stat/Tables/Tab5.html.

N	Niveau de signification, test unilatéral		
	0,025	0,01	0,005
	Niveau de signification, test bilatéral		
	0,05	0,02	0,01
6	0		
7	2	0	
8	4	2	0
9	6	3	2
10	8	5	3
11	11	7	5
12	14	10	7
13	17	13	10
14	21	16	13
15	25	20	16
16	30	24	20
17	35	28	23
18	40	33	28
19	46	38	32
20	52	43	38
21	59	49	43
22	66	56	49
23	73	62	55
24	81	69	61
25	89	77	68

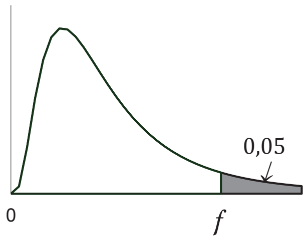
Fig. D.5. Table des valeurs critiques pour le test de Wilcoxon - Échantillons appariés

Table 5_1 : loi F de Fisher-Snedecor

ν_1 est le nombre de d.d.l. du numérateur,

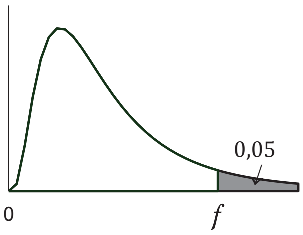
ν_2 est le nombre de d.d.l. du dénominateur.

Valeurs de f telles que $p(F \geq f) = 0,05$



ν_2 \ ν_1	1	2	3	4	6	8	10	12	15	18	21	25	29	33	37	41	45	49
1	161	200	216	225	234	239	242	244	246	247	248	249	250	250	251	251	251	252
2	18,5	19,0	19,2	19,2	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	8,94	8,85	8,79	8,74	8,70	8,67	8,65	8,63	8,62	8,61	8,60	8,59	8,59	8,58
4	7,71	6,94	6,59	6,39	6,16	6,04	5,96	5,91	5,86	5,82	5,79	5,77	5,75	5,74	5,72	5,71	5,71	5,70
5	6,61	5,79	5,41	5,19	4,95	4,82	4,74	4,68	4,62	4,58	4,55	4,52	4,50	4,48	4,47	4,46	4,45	4,45
6	5,99	5,14	4,76	4,53	4,28	4,15	4,06	4,00	3,94	3,90	3,86	3,83	3,81	3,80	3,78	3,77	3,76	3,76
7	5,59	4,74	4,35	4,12	3,87	3,73	3,64	3,57	3,51	3,47	3,43	3,40	3,38	3,36	3,35	3,34	3,33	3,32
8	5,32	4,46	4,07	3,84	3,58	3,44	3,35	3,28	3,22	3,17	3,14	3,11	3,08	3,07	3,05	3,04	3,03	3,02
9	5,12	4,26	3,86	3,63	3,37	3,23	3,14	3,07	3,01	2,96	2,93	2,89	2,87	2,85	2,84	2,82	2,81	2,80
10	4,96	4,10	3,71	3,48	3,22	3,07	2,98	2,91	2,85	2,80	2,76	2,73	2,70	2,69	2,67	2,66	2,65	2,64
11	4,84	3,98	3,59	3,36	3,09	2,95	2,85	2,79	2,72	2,67	2,64	2,60	2,58	2,56	2,54	2,53	2,52	2,51
12	4,75	3,89	3,49	3,26	3,00	2,85	2,75	2,69	2,62	2,57	2,53	2,50	2,47	2,45	2,44	2,42	2,41	2,40
13	4,67	3,81	3,41	3,18	2,92	2,77	2,67	2,60	2,53	2,48	2,45	2,41	2,39	2,37	2,35	2,34	2,33	2,32
14	4,60	3,74	3,34	3,11	2,85	2,70	2,60	2,53	2,46	2,41	2,38	2,34	2,31	2,29	2,28	2,26	2,25	2,24
15	4,54	3,68	3,29	3,06	2,79	2,64	2,54	2,48	2,40	2,35	2,32	2,28	2,25	2,23	2,21	2,20	2,19	2,18
16	4,49	3,63	3,24	3,01	2,74	2,59	2,49	2,42	2,35	2,30	2,26	2,23	2,20	2,18	2,16	2,15	2,14	2,13
17	4,45	3,59	3,20	2,96	2,70	2,55	2,45	2,38	2,31	2,26	2,22	2,18	2,15	2,13	2,11	2,10	2,09	2,08
18	4,41	3,55	3,16	2,93	2,66	2,51	2,41	2,34	2,27	2,22	2,18	2,14	2,11	2,09	2,07	2,06	2,05	2,04
19	4,38	3,52	3,13	2,90	2,63	2,48	2,38	2,31	2,23	2,18	2,14	2,11	2,08	2,06	2,04	2,02	2,01	2,00
20	4,35	3,49	3,10	2,87	2,60	2,45	2,35	2,28	2,20	2,15	2,11	2,07	2,05	2,02	2,01	1,99	1,98	1,97
21	4,32	3,47	3,07	2,84	2,57	2,42	2,32	2,25	2,18	2,12	2,08	2,05	2,02	1,99	1,98	1,96	1,95	1,94
22	4,30	3,44	3,05	2,82	2,55	2,40	2,30	2,23	2,15	2,10	2,06	2,02	1,99	1,97	1,95	1,93	1,92	1,91
23	4,28	3,42	3,03	2,80	2,53	2,37	2,27	2,20	2,13	2,08	2,04	2,00	1,97	1,94	1,93	1,91	1,90	1,89
24	4,26	3,40	3,01	2,78	2,51	2,36	2,25	2,18	2,11	2,05	2,01	1,97	1,95	1,92	1,90	1,89	1,88	1,86
25	4,24	3,39	2,99	2,76	2,49	2,34	2,24	2,16	2,09	2,04	2,00	1,96	1,93	1,90	1,88	1,87	1,86	1,84
26	4,23	3,37	2,98	2,74	2,47	2,32	2,22	2,15	2,07	2,02	1,98	1,94	1,91	1,88	1,87	1,85	1,84	1,83
27	4,21	3,35	2,96	2,73	2,46	2,31	2,20	2,13	2,06	2,00	1,96	1,92	1,89	1,87	1,85	1,83	1,82	1,81
28	4,20	3,34	2,95	2,71	2,45	2,29	2,19	2,12	2,04	1,99	1,95	1,91	1,88	1,85	1,83	1,82	1,80	1,79
29	4,18	3,33	2,93	2,70	2,43	2,28	2,18	2,10	2,03	1,97	1,93	1,89	1,86	1,84	1,82	1,80	1,79	1,78
30	4,17	3,32	2,92	2,69	2,42	2,27	2,16	2,09	2,01	1,96	1,92	1,88	1,85	1,82	1,80	1,79	1,77	1,76
31	4,16	3,30	2,91	2,68	2,41	2,25	2,15	2,08	2,00	1,95	1,91	1,87	1,83	1,81	1,79	1,78	1,76	1,75
32	4,15	3,29	2,90	2,67	2,40	2,24	2,14	2,07	1,99	1,94	1,90	1,85	1,82	1,80	1,78	1,76	1,75	1,74
33	4,14	3,28	2,89	2,66	2,39	2,23	2,13	2,06	1,98	1,93	1,89	1,84	1,81	1,79	1,77	1,75	1,74	1,73
34	4,13	3,28	2,88	2,65	2,38	2,23	2,12	2,05	1,97	1,92	1,88	1,83	1,80	1,78	1,76	1,74	1,73	1,72
35	4,12	3,27	2,87	2,64	2,37	2,22	2,11	2,04	1,96	1,91	1,87	1,82	1,79	1,77	1,75	1,73	1,72	1,71
36	4,11	3,26	2,87	2,63	2,36	2,21	2,11	2,03	1,95	1,90	1,86	1,81	1,78	1,76	1,74	1,72	1,71	1,70
37	4,11	3,25	2,86	2,63	2,36	2,20	2,10	2,02	1,95	1,89	1,85	1,81	1,77	1,75	1,73	1,71	1,70	1,69
38	4,10	3,24	2,85	2,62	2,35	2,19	2,09	2,02	1,94	1,88	1,84	1,80	1,77	1,74	1,72	1,70	1,69	1,68
39	4,09	3,24	2,85	2,61	2,34	2,19	2,08	2,01	1,93	1,88	1,83	1,79	1,76	1,73	1,71	1,70	1,68	1,67
40	4,08	3,23	2,84	2,61	2,34	2,18	2,08	2,00	1,92	1,87	1,83	1,78	1,75	1,73	1,71	1,69	1,67	1,66
41	4,08	3,23	2,83	2,60	2,33	2,17	2,07	2,00	1,92	1,86	1,82	1,78	1,74	1,72	1,70	1,68	1,67	1,66
42	4,07	3,22	2,83	2,59	2,32	2,17	2,06	1,99	1,91	1,86	1,81	1,77	1,74	1,71	1,69	1,67	1,66	1,65
43	4,07	3,21	2,82	2,59	2,32	2,16	2,06	1,99	1,91	1,85	1,81	1,76	1,73	1,71	1,69	1,67	1,65	1,64
44	4,06	3,21	2,82	2,58	2,31	2,16	2,05	1,98	1,90	1,84	1,80	1,76	1,73	1,70	1,68	1,66	1,65	1,64
45	4,06	3,20	2,81	2,58	2,31	2,15	2,05	1,97	1,89	1,84	1,80	1,75	1,72	1,69	1,67	1,66	1,64	1,63
46	4,05	3,20	2,81	2,57	2,30	2,15	2,04	1,97	1,89	1,83	1,79	1,75	1,71	1,69	1,67	1,65	1,64	1,62
47	4,05	3,20	2,80	2,57	2,30	2,14	2,04	1,96	1,88	1,83	1,78	1,74	1,71	1,68	1,66	1,64	1,63	1,62
48	4,04	3,19	2,80	2,57	2,29	2,14	2,03	1,96	1,88	1,82	1,78	1,74	1,70	1,68	1,66	1,64	1,62	1,61
49	4,04	3,19	2,79	2,56	2,29	2,13	2,03	1,96	1,88	1,82	1,78	1,73	1,70	1,67	1,65	1,63	1,62	1,61
50	4,03	3,18	2,79	2,56	2,29	2,13	2,03	1,95	1,87	1,81	1,77	1,73	1,69	1,67	1,65	1,63	1,61	1,60
55	4,02	3,16	2,77	2,54	2,27	2,11	2,01	1,93	1,85	1,79	1,75	1,71	1,67	1,65	1,63	1,61	1,59	1,58
60	4,00	3,15	2,76	2,53	2,25	2,10	1,99	1,92	1,84	1,78	1,73	1,69	1,66	1,63	1,61	1,59	1,57	1,56
65	3,99	3,14	2,75	2,51	2,24	2,08	1,98	1,90	1,82	1,76	1,72	1,68	1,64	1,61	1,59	1,57	1,56	1,55
70	3,98	3,13	2,74	2,50	2,23	2,07	1,97	1,89	1,81	1,75	1,71	1,66	1,63	1,60	1,58	1,56	1,55	1,53
75	3,97	3,12	2,73	2,49	2,22	2,06	1,96	1,88	1,80	1,74	1,70	1,65	1,62	1,59	1,57	1,55	1,53	1,52
80	3,96	3,11	2,72	2,49	2,21	2,06	1,95	1,88	1,79	1,73	1,69	1,64	1,61	1,58	1,56	1,54	1,52	1,51
85	3,95	3,10	2,71	2,48	2,21	2,05	1,94	1,87	1,79	1,73	1,68	1,64	1,60	1,57	1,55	1,53	1,52	1,50
90	3,95	3,10	2,71	2,47	2,20	2,04	1,94	1,86	1,78	1,72	1,67	1,63	1,59	1,57	1,54	1,52	1,51	1,49
95	3,94	3,09	2,70	2,47	2,20	2,04	1,93	1,86	1,77	1,71	1,67	1,62	1,59	1,56	1,54	1,52	1,50	1,49
100	3,94	3,09	2,70	2,46	2,19	2,03	1,93	1,85	1,77	1,71	1,66	1,62	1,58	1,55	1,53	1,51	1,49	1,48
105	3,93	3,08	2,69	2,46	2,19	2,03	1,92	1,85	1,76	1,70	1,66	1,61	1,58	1,55	1,52	1,50	1,49	1,47
110	3,93	3,08	2,69	2,45	2,18	2,02	1,92	1,84	1,76	1,70	1,65	1,61	1,57	1,54	1,52	1,50	1,48	1,47
115	3,92	3,08	2,68	2,45	2,18	2,02	1,91	1,84	1,75	1,69	1,65	1,60	1,57	1,54	1,51	1,49	1,48	1,46
120	3,92	3,07	2,68	2,45	2,18	2,02	1,91	1,83	1,75	1,69	1,64	1,60	1,56	1,53	1,51	1,49	1,47	1,46
125	3,92	3,07	2,68	2,44	2,17	2,01	1,91	1,83	1,75	1,69	1,64	1,59	1,56	1,53	1,51	1,49	1,47	1,46

Table 5_2 : loi F de Fisher-Snedecor (suite)



ν_1 est le nombre de d.d.l. du numérateur,

ν_2 est le nombre de d.d.l. du dénominateur.

Valeurs de f telles que $p(F \geq f) = 0,05$

ν_2 \ ν_1	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130	135
1	252	252	252	252	252	253	253	253	253	253	253	253	253	253	253	253	253	253
2	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5	19,5
3	8,6	8,58	8,57	8,57	8,57	8,56	8,56	8,56	8,56	8,56	8,55	8,55	8,55	8,55	8,55	8,55	8,55	8,55
4	5,70	5,69	5,69	5,68	5,68	5,68	5,67	5,67	5,67	5,67	5,66	5,66	5,66	5,66	5,66	5,66	5,66	5,65
5	4,44	4,44	4,43	4,43	4,42	4,42	4,41	4,41	4,41	4,41	4,41	4,40	4,40	4,40	4,40	4,40	4,40	4,39
6	3,75	3,75	3,74	3,73	3,73	3,73	3,72	3,72	3,72	3,71	3,71	3,71	3,71	3,71	3,70	3,70	3,70	3,70
7	3,32	3,31	3,30	3,30	3,29	3,29	3,29	3,28	3,28	3,28	3,27	3,27	3,27	3,27	3,27	3,27	3,26	3,26
8	3,02	3,01	3,01	3,00	2,99	2,99	2,99	2,98	2,98	2,98	2,97	2,97	2,97	2,97	2,97	2,97	2,96	2,96
9	2,80	2,79	2,79	2,78	2,78	2,77	2,77	2,76	2,76	2,76	2,76	2,75	2,75	2,75	2,75	2,75	2,74	2,74
10	2,64	2,63	2,62	2,61	2,61	2,60	2,60	2,60	2,59	2,59	2,59	2,59	2,58	2,58	2,58	2,58	2,58	2,58
11	2,51	2,50	2,49	2,48	2,48	2,47	2,47	2,47	2,46	2,46	2,46	2,45	2,45	2,45	2,45	2,45	2,44	2,44
12	2,40	2,39	2,38	2,38	2,37	2,37	2,36	2,36	2,36	2,35	2,35	2,35	2,34	2,34	2,34	2,34	2,34	2,34
13	2,31	2,30	2,30	2,29	2,28	2,28	2,27	2,27	2,27	2,26	2,26	2,26	2,26	2,25	2,25	2,25	2,25	2,25
14	2,24	2,23	2,22	2,22	2,21	2,21	2,20	2,20	2,19	2,19	2,19	2,18	2,18	2,18	2,18	2,18	2,17	2,17
15	2,18	2,17	2,16	2,15	2,15	2,14	2,14	2,13	2,13	2,13	2,12	2,12	2,12	2,12	2,11	2,11	2,11	2,11
16	2,12	2,11	2,11	2,10	2,09	2,09	2,08	2,08	2,07	2,07	2,07	2,07	2,06	2,06	2,06	2,06	2,06	2,05
17	2,08	2,07	2,06	2,05	2,05	2,04	2,03	2,03	2,03	2,02	2,02	2,02	2,02	2,01	2,01	2,01	2,01	2,01
18	2,04	2,03	2,02	2,01	2,00	2,00	1,99	1,99	1,98	1,98	1,98	1,98	1,97	1,97	1,97	1,97	1,96	1,96
19	2,00	1,99	1,98	1,97	1,97	1,96	1,96	1,95	1,95	1,94	1,94	1,94	1,93	1,93	1,93	1,93	1,93	1,92
20	1,97	1,96	1,95	1,94	1,93	1,93	1,92	1,92	1,91	1,91	1,91	1,90	1,90	1,90	1,90	1,89	1,89	1,89
21	1,94	1,93	1,92	1,91	1,90	1,90	1,89	1,89	1,88	1,88	1,88	1,87	1,87	1,87	1,87	1,86	1,86	1,86
22	1,91	1,90	1,89	1,88	1,88	1,87	1,86	1,86	1,86	1,85	1,85	1,85	1,84	1,84	1,84	1,84	1,83	1,83
23	1,88	1,87	1,86	1,86	1,85	1,84	1,84	1,83	1,83	1,83	1,82	1,82	1,82	1,82	1,81	1,81	1,81	1,81
24	1,86	1,85	1,84	1,83	1,83	1,82	1,82	1,81	1,81	1,80	1,80	1,80	1,79	1,79	1,79	1,79	1,79	1,78
25	1,84	1,83	1,82	1,81	1,81	1,80	1,80	1,79	1,79	1,78	1,78	1,78	1,77	1,77	1,77	1,77	1,76	1,76
26	1,82	1,81	1,80	1,79	1,79	1,78	1,78	1,77	1,77	1,76	1,76	1,76	1,75	1,75	1,75	1,75	1,74	1,74
27	1,81	1,79	1,79	1,78	1,77	1,76	1,76	1,75	1,75	1,75	1,74	1,74	1,74	1,73	1,73	1,73	1,73	1,72
28	1,79	1,78	1,77	1,76	1,75	1,75	1,74	1,74	1,73	1,73	1,73	1,72	1,72	1,72	1,71	1,71	1,71	1,71
29	1,77	1,76	1,75	1,75	1,74	1,73	1,73	1,72	1,72	1,71	1,71	1,71	1,70	1,70	1,70	1,70	1,69	1,69
30	1,76	1,75	1,74	1,73	1,72	1,72	1,71	1,71	1,70	1,70	1,70	1,69	1,69	1,69	1,68	1,68	1,68	1,68
31	1,75	1,74	1,73	1,72	1,71	1,70	1,70	1,69	1,69	1,69	1,68	1,68	1,68	1,67	1,67	1,67	1,67	1,66
32	1,74	1,72	1,71	1,71	1,70	1,69	1,69	1,68	1,68	1,67	1,67	1,67	1,66	1,66	1,66	1,65	1,65	1,65
33	1,72	1,71	1,70	1,69	1,69	1,68	1,67	1,67	1,66	1,66	1,66	1,65	1,65	1,65	1,64	1,64	1,64	1,64
34	1,71	1,70	1,69	1,68	1,68	1,67	1,66	1,66	1,65	1,65	1,65	1,64	1,64	1,64	1,63	1,63	1,63	1,63
35	1,70	1,69	1,68	1,67	1,66	1,66	1,65	1,65	1,64	1,64	1,63	1,63	1,63	1,63	1,62	1,62	1,62	1,62
36	1,69	1,68	1,67	1,66	1,66	1,65	1,64	1,64	1,63	1,63	1,62	1,62	1,62	1,62	1,61	1,61	1,61	1,61
37	1,68	1,67	1,66	1,65	1,65	1,64	1,63	1,63	1,62	1,62	1,62	1,61	1,61	1,61	1,60	1,60	1,60	1,60
38	1,68	1,66	1,65	1,64	1,64	1,63	1,62	1,62	1,61	1,61	1,61	1,60	1,60	1,60	1,59	1,59	1,59	1,59
39	1,67	1,66	1,65	1,64	1,63	1,62	1,62	1,61	1,61	1,60	1,60	1,59	1,59	1,59	1,58	1,58	1,58	1,58
40	1,66	1,65	1,64	1,63	1,62	1,61	1,61	1,60	1,60	1,59	1,59	1,59	1,58	1,58	1,58	1,57	1,57	1,57
41	1,65	1,64	1,63	1,62	1,61	1,61	1,60	1,59	1,59	1,59	1,58	1,58	1,57	1,57	1,57	1,57	1,56	1,56
42	1,65	1,63	1,62	1,61	1,61	1,60	1,59	1,59	1,58	1,58	1,57	1,57	1,57	1,56	1,56	1,56	1,56	1,55
43	1,64	1,63	1,62	1,61	1,60	1,59	1,59	1,58	1,58	1,57	1,57	1,56	1,56	1,56	1,55	1,55	1,55	1,55
44	1,63	1,62	1,61	1,60	1,59	1,59	1,58	1,57	1,57	1,56	1,56	1,56	1,55	1,55	1,55	1,54	1,54	1,54
45	1,63	1,61	1,60	1,59	1,59	1,58	1,57	1,57	1,56	1,56	1,55	1,55	1,55	1,54	1,54	1,54	1,54	1,53
46	1,62	1,61	1,60	1,59	1,58	1,57	1,57	1,56	1,56	1,55	1,55	1,54	1,54	1,54	1,53	1,53	1,53	1,53
47	1,61	1,60	1,59	1,58	1,57	1,57	1,56	1,56	1,55	1,55	1,54	1,54	1,53	1,53	1,53	1,53	1,52	1,52
48	1,61	1,60	1,59	1,58	1,57	1,56	1,56	1,55	1,54	1,54	1,54	1,53	1,53	1,53	1,52	1,52	1,52	1,51
49	1,60	1,59	1,58	1,57	1,56	1,56	1,55	1,54	1,54	1,53	1,53	1,53	1,52	1,52	1,52	1,51	1,51	1,51
50	1,60	1,59	1,58	1,57	1,56	1,55	1,54	1,54	1,53	1,53	1,52	1,52	1,52	1,51	1,51	1,51	1,51	1,50
55	1,58	1,56	1,55	1,54	1,54	1,53	1,52	1,52	1,51	1,51	1,50	1,50	1,49	1,49	1,49	1,48	1,48	1,48
60	1,56	1,55	1,53	1,52	1,52	1,51	1,50	1,50	1,49	1,49	1,48	1,48	1,47	1,47	1,47	1,46	1,46	1,46
65	1,54	1,53	1,52	1,51	1,50	1,49	1,49	1,48	1,47	1,47	1,46	1,46	1,46	1,45	1,45	1,45	1,44	1,44
70	1,53	1,52	1,50	1,49	1,49	1,48	1,47	1,46	1,46	1,45	1,45	1,45	1,44	1,44	1,44	1,43	1,43	1,43
75	1,52	1,50	1,49	1,48	1,47	1,47	1,46	1,45	1,45	1,44	1,44	1,43	1,43	1,43	1,42	1,42	1,42	1,41
80	1,51	1,49	1,48	1,47	1,46	1,45	1,45	1,44	1,44	1,43	1,43	1,42	1,42	1,41	1,41	1,41	1,40	1,40
85	1,50	1,48	1,47	1,46	1,45	1,45	1,44	1,43	1,43	1,42	1,42	1,41	1,41	1,40	1,40	1,40	1,39	1,39
90	1,49	1,48	1,46	1,45	1,44	1,44	1,43	1,42	1,42	1,41	1,41	1,40	1,40	1,39	1,39	1,39	1,39	1,38
95	1,48	1,47	1,46	1,45	1,44	1,43	1,42	1,42	1,41	1,40	1,40	1,39	1,39	1,39	1,38	1,38	1,38	1,37
100	1,48	1,46	1,45	1,44	1,43	1,42	1,41	1,41	1,40	1,40	1,39	1,39	1,38	1,38	1,38	1,37	1,37	1,37
105	1,47	1,46	1,44	1,43	1,42	1,42	1,41	1,40	1,40	1,39	1,39	1,38	1,38	1,37	1,37	1,37	1,36	1,36
110	1,47	1,45	1,44	1,43	1,42	1,41	1,40	1,40	1,39	1,38	1,38	1,37	1,37	1,37	1,36	1,36	1,36	1,35
115	1,46	1,45	1,43	1,42	1,41	1,40	1,40	1,39	1,38	1,38	1,37	1,37	1,36	1,36	1,36	1,35	1,35	1,35
120	1,46	1,44	1,43	1,42	1,41	1,40	1,39	1,39	1,38	1,37	1,37	1,36	1,36	1,36	1,35	1,35	1,35	1,34
125	1,45	1,44	1,42	1,41	1,40	1,40	1,39	1,38	1,37	1,37	1,36	1,36	1,35	1,35	1,35	1,34	1,34	1,34

TABLE DE LA LOI DU χ^2

α / n	0.8	0.7	0.5	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001
1	0.06	0.15	0.46	1.07	1.64	2.71	3.84	5.02	6.63	7.88	10.80
2	0.45	0.71	1.39	2.41	3.22	4.61	5.99	7.38	9.21	10.60	13.82
3	1.01	1.42	2.37	3.66	4.64	6.25	7.81	9.35	11.34	12.84	16.27
4	1.65	2.19	3.36	4.88	5.99	7.78	9.49	11.14	13.28	14.86	18.47
5	2.34	3.00	4.35	6.06	7.29	9.24	11.07	12.83	15.09	16.75	20.52
6	3.07	3.83	5.35	7.23	8.56	10.64	12.59	14.45	16.81	18.55	22.46
7	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16.01	18.48	20.28	24.32
8	4.59	5.53	7.34	9.52	11.03	13.36	15.51	17.53	20.09	21.95	26.12
9	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19.02	21.67	23.59	27.88
10	6.18	7.27	9.34	11.78	13.44	15.99	18.31	20.48	23.21	25.19	29.59
11	6.99	8.15	10.34	12.90	14.63	17.28	19.68	21.92	24.72	26.76	31.26
12	7.81	9.03	11.34	14.01	15.81	18.55	21.03	23.34	26.22	28.30	32.91
13	8.63	9.93	12.34	15.12	16.98	19.81	22.36	24.74	27.69	29.82	34.53
14	9.47	10.82	13.34	16.22	18.15	21.06	23.68	26.12	29.14	31.32	36.12
15	10.31	11.72	14.34	17.32	19.31	22.31	25.00	27.49	30.58	32.80	37.70
16	11.15	12.62	15.34	18.42	20.47	23.54	26.30	28.85	32.00	34.27	39.25
17	12.00	13.53	16.34	19.51	21.61	24.77	27.59	30.19	33.41	35.72	40.79
18	12.86	14.44	17.34	20.60	22.76	25.99	28.87	31.53	34.81	37.16	42.31
19	13.72	15.35	18.34	21.69	23.90	27.20	30.14	32.85	36.19	38.58	43.82
20	14.58	16.27	19.34	22.77	25.04	28.41	31.41	34.17	37.57	40.00	45.31
21	15.44	17.18	20.34	23.86	26.17	29.62	32.67	35.48	38.93	41.40	46.80
22	16.31	18.10	21.34	24.94	27.30	30.81	33.92	36.78	40.29	42.80	48.27
23	17.19	19.02	22.34	26.02	28.43	32.01	35.17	38.08	41.64	44.18	49.73
24	18.06	19.94	23.34	27.10	29.55	33.20	36.42	39.36	42.98	45.56	51.18
25	18.94	20.87	24.34	28.17	30.68	34.38	37.65	40.65	44.31	46.93	52.62
26	19.82	21.79	25.34	29.25	31.79	35.56	38.89	41.92	45.64	48.29	54.05
27	20.70	22.72	26.34	30.32	32.91	36.74	40.11	43.19	46.96	49.64	55.48
28	21.59	23.65	27.34	31.39	34.03	37.92	41.34	44.46	48.28	50.99	56.89
29	22.48	24.58	28.34	32.46	35.14	39.09	42.56	45.72	49.59	52.34	58.30
30	23.36	25.51	29.34	33.53	36.25	40.26	43.77	46.98	50.89	53.67	59.70

8 (Table 8) Valeurs tabulées et coefficient de corrélation linéaire

La table ci-dessous donne

$$r_\alpha(\nu) = \frac{t_\alpha(\nu)}{\sqrt{\nu + t_\alpha^2(\nu)}},$$

où $t_\alpha(\nu)$ est le réel vérifiant $\mathbb{P}(|T| \geq t_\alpha(\nu)) = \alpha$, $T \sim \mathcal{T}(\nu)$. Cette valeur est principalement utilisée dans le test de nullité du coefficient de corrélation linéaire.

ν \ α	0,05	0,01	0,001
1	0,9969	0,9999	1,0000
2	0,9500	0,9900	0,9990
3	0,8783	0,9587	0,9912
4	0,8114	0,9172	0,9741
5	0,7545	0,8745	0,9508
6	0,7067	0,8343	0,9249
7	0,6664	0,7976	0,8982
8	0,6319	0,7646	0,8721
9	0,6020	0,7348	0,8471
10	0,5760	0,7079	0,8233
11	0,5529	0,6836	0,8010
12	0,5324	0,6614	0,7800
13	0,5139	0,6411	0,7604
14	0,4973	0,6226	0,7419
15	0,4821	0,6055	0,7247
16	0,4683	0,5897	0,7084
17	0,4556	0,5750	0,6932
18	0,4438	0,5614	0,6788
19	0,4329	0,5487	0,6652
20	0,4227	0,5368	0,6524
21	0,4133	0,5256	0,6402
22	0,4044	0,5151	0,6287
23	0,3961	0,5051	0,8177
24	0,3883	0,4958	0,6073
25	0,3809	0,4869	0,5974
26	0,3740	0,4785	0,5880
27	0,3673	0,4706	0,5790
28	0,3609	0,4629	0,5703
29	0,3550	0,4556	0,5620
30	0,3490	0,4487	0,5541
∞	0,1218	0,1593	0,2018