

Analyse de la variance

Frédéric Sur
Mines Nancy
16 mai 2026

1 Introduction

L'**analyse de la variance** (*analysis of variance*, **ANOVA**), développée par le statisticien britannique Ronald A. Fisher dans les années 1920, est une méthode statistique permettant de comparer simultanément les moyennes de plusieurs populations (ou groupes). Cette procédure peut être vue comme une extension à $k > 2$ groupes de la comparaison de moyennes entre deux groupes à l'aide d'un test de Student (voir section 1.1 de la séance 2 sur les tests : comparaison de deux échantillons gaussiens indépendants).

Exemple 1

Connaissant les notes d'examen des dix groupes de TD du cours *Inférence statistique*, peut-on dire si les notes des groupes sont équivalentes? Plus précisément, les différences observées entre les moyennes des différents groupes est-elle compatible avec les fluctuations aléatoires d'échantillonnage?

Dans cet exemple 1, on cherche à mettre en évidence l'influence d'un **facteur de variabilité** (le groupe de TD) sur la note obtenue. C'est ce qui est appelé **analyse de variance à un facteur**. D'autres facteurs peuvent être envisagés : homme / femme, langue vivante étudiée, etc. Ce sera l'objet de l'**analyse de variance multifactorielle**. Les facteurs sont des variables qualitatives qui peuvent prendre un nombre fini k de **modalités** (on parle aussi de **niveaux**). Ici, le « nom » du groupe de TD (A1, A2...) est une modalité, et le facteur de variabilité a dix modalités.

Pour que les conclusions tirées d'une ANOVA soient statistiquement valides, les données doivent respecter trois hypothèses fondamentales :

1. **indépendance des observations** : les individus de chaque modalité doivent être choisis de manière aléatoire et indépendante.
2. **normalité des distributions** : dans chacune des k modalités, la variable étudiée (ou plus précisément les résidus du modèle) doit suivre la loi normale $\mathcal{N}(\mu_i, \sigma_i^2)$, pour $1 \leq i \leq k$.
3. **homoscédasticité (homogénéité des variances)** : la variance de la variable étudiée doit être la même pour chacune des k modalités étudiées ($\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$).

Dans la suite, ces hypothèses seront supposées vérifiées.

Dans l'exemple 1, cela veut dire que les élèves sont répartis aléatoirement dans les groupes, indépendamment l'un de l'autre (on ne fait pas des groupes de niveau, ou des groupes par filière d'origine). Cela semble raisonnable de supposer une répartition normale des notes. Supposer les variances toutes égales est une hypothèse plus forte qui mériterait d'être testée. Cela veut dire qu'on suppose que l'effet du facteur « groupe de TD » ne se traduit que par un décalage des moyennes.

2 ANOVA à un facteur

On considère un facteur explicatif qualitatif (appelée **variable indépendante**) présentant k modalités. On cherche à savoir si ce facteur influence une variable quantitative (appelée **variable dépendante**).

L'objectif est de tester l'hypothèse $\mathcal{H}_0 : \mu_1 = \mu_2 = \dots = \mu_k$.

Il serait possible d'effectuer des tests de Student d'égalité deux-à-deux : l'inconvénient est que cela représente un grand nombre de tests et que les risques d'erreur de type 1 s'« ajoutent ». Par exemple, si $k = 10$, on est amené à faire $k \times (k - 1)/2 = 45$ tests de Student. Si le risque est $\alpha = 0,05$, dans 5% des cas le test de Student rejette l'hypothèse d'égalité de deux moyennes « par hasard ». Dans une proportion $1 - (1 - \alpha)^{k(k-1)/2}$, au moins une des égalités est rejetée, conduisant à rejeter l'hypothèse \mathcal{H}_0 ¹. Lorsque $k = 10$, cette proportion atteint 90%! La stratégie des tests multiples n'est donc pas envisageable. Il faut un test « global ».

2.1 Modèle statistique

Soit y_{ij} la j -ème observation de la i -ème modalité, considérée comme une variable aléatoire. On suppose que la i -ème modalité comprend n_i observations (donc $1 \leq j \leq n_i$ pour $1 \leq i \leq k$), de manière à ce que le nombre total d'observations soit $N = \sum_{i=1}^k n_i$. Le modèle s'écrit de la manière suivante :

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

où :

- μ représente la moyenne générale de la population globale;
- α_i représente l'effet spécifique de la modalité i (avec $\sum_{i=1}^k n_i \alpha_i = 0$: l'effet moyen est nul);
- ε_{ij} représente l'erreur aléatoire (résidu), supposée suivre la loi normale de moyenne nulle et d'écart-type constant σ . Les ε_{ij} sont supposés indépendants.

Ainsi, les observations y_{ij} respectent bien les trois hypothèses fondamentales de l'ANOVA.

2.2 Décomposition de la variance

La somme des carrés totale T est décomposée en deux parties : la variation due au facteur étudié A et la variation résiduelle R . Notons

$$T = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2, \quad A = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y}_{..})^2, \quad R = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

où $\bar{y}_{..}$ est la moyenne générale et \bar{y}_i est la moyenne du groupe i .

On a :

$$T = A + R$$

D'où vient cette relation? On calcule :

$$\begin{aligned} T &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i + \bar{y}_i - \bar{y}_{..})^2 \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y}_{..})^2 + 2 \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(\bar{y}_i - \bar{y}_{..}) \end{aligned}$$

Par ailleurs :

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y}_{..})^2 = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y}_{..})^2$$

1. En effet, la probabilité que toutes les hypothèses $\mu_i = \mu_j$ ($1 \leq i < j \leq k$) soient rejetées par hasard est $(1 - \alpha)^{k(k-1)/2}$. Notons qu'à l'ordre 1 pour α petit, $1 - (1 - \alpha)^{k(k-1)/2} = k(k-1)/2 \alpha + o(\alpha)$: en ce sens, les risques d'erreur de type 1 s'ajoutent.

et :

$$\begin{aligned}
 \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(\bar{y}_i - \bar{y}_{..}) &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} \bar{y}_i - y_{ij} \bar{y}_{..} - \bar{y}_i^2 + \bar{y}_i \bar{y}_{..}) \\
 &= \sum_{i=1}^k n_i \bar{y}_i^2 - n \bar{y}_{..}^2 - \sum_{i=1}^k n_i \bar{y}_i^2 + \sum_{i=1}^k n_i \bar{y}_i \bar{y}_{..} \\
 &= 0
 \end{aligned}$$

D'où la relation cherchée.

En notant $S_T^2 = T/N$, $S_A^2 = A/N$, et $S_R^2 = R/N$, S_T^2 est la variance totale, S_A^2 est la variance expliquée (par le facteur étudié), et S_R^2 est la variance résiduelle.

On a donc la relation :

$$S_T^2 = S_A^2 + S_R^2$$

2.3 Analyse des variances

Remarque. La justification des lois peut être omise.

2.3.1 Loi de la variance résiduelle S_R^2

Comme les observations de la modalité i suivent la loi normale $\mathcal{N}(\mu_i, \sigma^2)$, la statistique

$$S_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

est telle que $n_i S_i^2 / \sigma^2$ suit la loi du χ^2 à $n_i - 1$ degrés de liberté².

Les observations étant indépendantes et la variance résiduelle vérifiant :

$$S_R^2 = \frac{1}{N} \sum_{i=1}^k n_i S_i^2$$

S_R^2 est telle que $N S_R^2 / \sigma^2$ suit la loi du χ^2 à $\sum_{i=1}^k (n_i - 1) = N - k$ degrés de liberté³.

2.3.2 Loi de la variance expliquée S_A^2

Sous l'hypothèse \mathcal{H}_0 , \bar{y}_i suit la loi $\mathcal{N}(\mu, \sigma^2/n_i)$, dont $\sqrt{n_i}(\bar{y}_i - \mu)/\sigma$ suit la loi $\mathcal{N}(0, 1)$. Ainsi, $\sum_{i=1}^k n_i (\bar{y}_i - \mu)^2 / \sigma^2$ suit la loi du χ^2 à k d.d.l.

On ne connaît pas la moyenne μ , mais :

$$\sum_{i=1}^k n_i (\bar{y}_i - \mu)^2 = \sum_{i=1}^k n_i ((\bar{y}_i - \bar{y}_{..}) - (\mu - \bar{y}_{..}))^2 = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y}_{..})^2 + N(\bar{y}_{..} - \mu)^2$$

Comme $N(\bar{y}_{..} - \mu)^2 / \sigma^2$ suit la loi du χ^2 à un degré de liberté, le théorème de Cochran (pas vu dans ce cours) permet de conclure que $N S_A^2 / \sigma^2$ suit la loi du χ^2 à $k - 1$ degrés de liberté.

2.3.3 Loi de la variance totale S_T^2

Sous l'hypothèse \mathcal{H}_0 , $N S_T^2 / \sigma^2$ suit la loi du χ^2 à $N - 1$ degrés de liberté².

Notons que les lois de $N S_A^2 / \sigma^2$ et $N S_T^2 / \sigma^2$ données ici ne sont valables que sous hypothèse \mathcal{H}_0 , contrairement à celle de $N S_R^2 / \sigma^2$.

2. Rappelons la propriété (conséquence du théorème de Cochran) : si (X_1, \dots, X_n) est un n -échantillon de loi $\mathcal{N}(\mu, \sigma^2)$ et S^2 est la variance empirique (non corrigée), alors $n S^2 / \sigma^2$ suit la loi du χ^2 à $n - 1$ d.d.l. Voir l'exercice 3 de la séance 2 de la première partie du cours.

3. La somme de deux variables aléatoires indépendantes suivant les lois du χ^2 à a et b d.d.l. suit la loi du χ^2 à $a + b$ d.d.l.

2.4 Test de Fisher

L'hypothèse nulle (\mathcal{H}_0) postule qu'il n'y a aucune différence entre les moyennes des groupes :

$$\mathcal{H}_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

Sous l'hypothèse alternative (\mathcal{H}_1), au moins deux moyennes sont différentes.

On admet que les statistiques NS_A^2/σ^2 et NS_R^2/σ^2 sont indépendantes (c'est encore une conséquence du théorème de Cochran).

La statistique de test est donnée par le rapport :

$$F = \frac{S_A^2/(k-1)}{S_R^2/(N-k)} = \frac{A/(k-1)}{R/(N-k)}$$

Sous \mathcal{H}_0 , cette statistique suit la loi de Fisher-Snedecor à $(k-1)$ et $(N-k)$ degrés de liberté. Pour un risque α donné, la règle de décision est donc : si la valeur F calculée est supérieure à la valeur critique F_α , on rejette \mathcal{H}_0 .

Bien que le nom *Analysis of variance* suggère une analyse des variances, l'objectif principal de l'ANOVA est d'inférer une propriété sur les moyennes. Le principe central consiste à comparer la variance estimée *entre* les différents groupes (variabilité expliquée par le facteur étudié) à la variance estimée *à l'intérieur* des groupes (variabilité résiduelle ou bruit).

2.5 Le tableau récapitulatif de l'ANOVA

Les résultats d'une ANOVA sont classiquement présentés sous la forme du tableau synthétique suivant.

Source de variation	Somme des carrés	ddl	Carré moyen	F_{obs}
Facteur (Inter-groupes)	A	$k-1$	$\frac{A}{k-1}$	$\frac{A/(k-1)}{R/(N-k)}$
Résiduelle (Intra-groupe)	R	$N-k$	$\frac{R}{N-k}$	
Total	T	$N-1$		

Les logiciels ou bibliothèques Python donnent une colonne supplémentaire avec la p -valeur de F_{obs} qui permet de prendre une décision selon le risque α assumé.

2.6 Comparaisons multiples (tests *post-hoc*)

Le test de l'ANOVA est dit *omnibus* (« pour tous » en latin) : il teste les moyennes dans leur ensemble. Lorsqu'il est significatif (rejet de \mathcal{H}_0), il indique qu'au moins une moyenne diffère des autres, mais il ne précise pas laquelle. Pour identifier où se situent ces différences, on réalise des tests de comparaisons multiples des moyennes (c'est-à-dire, d'après la première séance sur les tests, des tests de Student). On parle de test *post-hoc* (« après cela » en latin) : ce sont des tests réalisés après analyse de la variance. On peut être amené à réaliser un grand nombre de tests deux-à-deux : si le facteur a 10 modalités, il faudra procéder à $m = 10 \times 9/2 = 45$ tests.

On a déjà discuté du problème méthodologique posé par un grand nombre de tests sur le même ensemble d'observations (cf. première séance sur les tests et remarque en section 2) : multiplier les tests augmente le risque de première espèce (conclure à tort à une différence). On peut être amené à utiliser la correction de Bonferroni, qui consiste à diviser le risque α par le nombre de comparaisons m . D'autres méthodologies de tests *post-hoc* existent.

2.7 Cas pratique

Imaginons une étude où nous testons l'efficacité de 3 méthodes d'apprentissage (A, B, C) sur les scores d'un test (sur 20) obtenu par 15 étudiants (5 par méthode).

Données :

— Méthode A : 10, 12, 11, 14, 13 $\rightarrow \bar{y}_A = 12$

— Méthode B : 16, 15, 18, 17, 19 $\rightarrow \bar{y}_B = 17$

— Méthode C : 8, 9, 11, 10, 7 $\rightarrow \bar{y}_C = 9$

On calcule alors :

Moyenne générale $\bar{y}_{..} = 12.67$

Calcul des sommes des carrés :

$$A = 5 \times [(12 - 12.67)^2 + (17 - 12.67)^2 + (9 - 12.67)^2] = 163.33$$

$$R = \sum_{i,j} (\bar{y}_{ij} - \bar{y}_{i.})^2 = (10 - 12)^2 + \dots + (7 - 9)^2 = 30.00$$

Tableau de l'ANOVA :

Source	SS	ddl	Carrés moyens	F_{obs}
Méthode	163.33	$3 - 1 = 2$	$163.33/2 = 81.67$	$81.67/2.5 = 32.67$
Erreur	30.00	$15 - 3 = 12$	$30.00/12 = 2.50$	
Total	193.33	14		

Conclusion. La valeur critique de F pour $\alpha = 0.05$ avec 2 et 12 degrés de liberté est d'environ 3.89 (voir la table). Puisque $F_{\text{obs}} = 32.67 > 3.89$, on rejette formellement l'hypothèse nulle. La méthode d'apprentissage a un effet statistiquement très significatif sur le score des étudiants.

3 ANOVA à deux facteurs

Lorsqu'on étudie l'influence simultanée de deux variables qualitatives (facteurs A et B) sur une variable quantitative X , on utilise une ANOVA à deux facteurs.

Exemple 2

On souhaite tester l'influence du groupe de TD et du choix de Langue Vivante 2 sur la note au cours *Inférence statistique*.

Notons p et q le nombre de modalités respectives des facteurs A et B , et n_{ij} ($1 \leq i \leq p$ et $1 \leq j \leq q$) le nombre d'observations de X pour le couple (i, j) de modalités de A et B . On suppose disposer d'un modèle **équilibré** : $n_{ij} = r \geq 1$. À tout couple de modalités correspond le même nombre d'observations. L'intérêt de cette hypothèse est de simplifier les calculs dans ce cours. De cette manière, le nombre total d'observations est $N = pqr$.

Dans l'exemple 2, cela signifie qu'il y a le même nombre r d'élèves de chaque langue vivante dans chaque groupe de TD.

3.1 Effet principal et interaction

Ce modèle permet de tester trois hypothèses :

1. L'effet principal du facteur A est-il significatif?
2. L'effet principal du facteur B est-il significatif?
3. Y a-t-il une interaction entre les facteurs A et B?

Exemple 3

Nous voulons mesurer la croissance d'une plante (variable dépendante) en fonction de deux fac-

teurs prenant deux modalités chacun :

- ensoleillement (facteur A : Faible ou Fort)
- arrosage (facteur B : Hebdomadaire ou Quotidien)

L'ANOVA va tester trois hypothèses distinctes :

1. l'effet principal du facteur A : l'ensoleillement a-t-il un impact significatif sur la croissance, indépendamment de l'arrosage?
2. l'effet principal du facteur B : l'arrosage a-t-il un impact significatif sur la croissance, indépendamment de l'ensoleillement?
3. l'effet d'interaction : l'effet de l'ensoleillement sur la croissance dépend-il de la fréquence d'arrosage? C'est souvent l'effet le plus intéressant dans le cas étudié.

3.2 Le modèle statistique avec interaction

Soit un facteur A à p niveaux et un facteur B à q niveaux, avec r répétitions par combinaison. On suppose $r > 1$. Les données recueillies sont les y_{ijk} pour ($1 \leq i \leq p$, $1 \leq j \leq q$, et $1 \leq k \leq r$).

Explicitons cette notation dans le cas de l'exemple 2 : pour tout groupe de TD i il y a $r = 4$ étudiants qui suivent la LV2 j . Dans le groupe de TD 2, les notes des 4 étudiants qui suivent la LV2 3 sont : $y_{2,3,1}, y_{2,3,2}, y_{2,3,3}, y_{2,3,4}$.

On suppose que y_{ijk} suit la loi $\mathcal{N}(\mu_{ij}, \sigma^2)$, pour $1 \leq k \leq r$.

On peut alors décomposer :

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

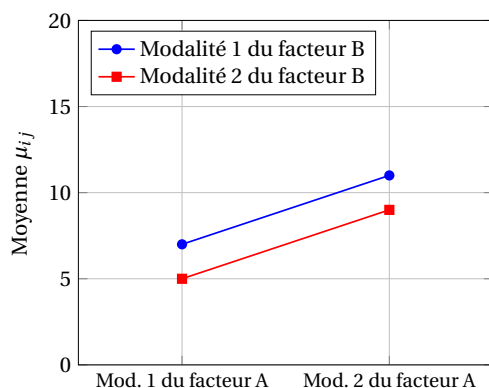
où μ est l'effet moyen, α_i l'effet principal de la modalité i du facteur A ($\sum_i \alpha_i = 0$), β_j l'effet principal de la modalité j du facteur B ($\sum_j \beta_j = 0$), γ_{ij} l'effet d'interaction causé par la combinaison du niveau i de A et j de B ($\sum_i \gamma_{ij} = 0$ et $\sum_j \gamma_{ij} = 0$), de manière à ce que les effets représentent un écart à la moyenne μ . Par ailleurs, ε_{ijk} est l'effet résiduel : les ε_{ijk} sont i.i.d. selon $\mathcal{N}(0, \sigma^2)$.

Ainsi :

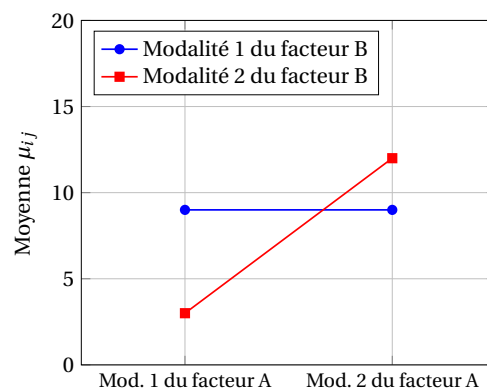
$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

Dans le cas de deux facteurs à deux modalités chacun, le **graphique d'interaction** peut prendre les allures de l'exemple suivant.

Exemple 4

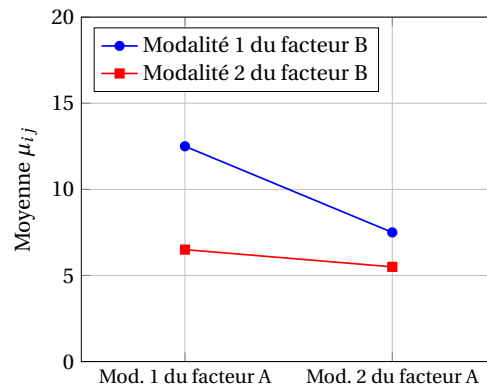
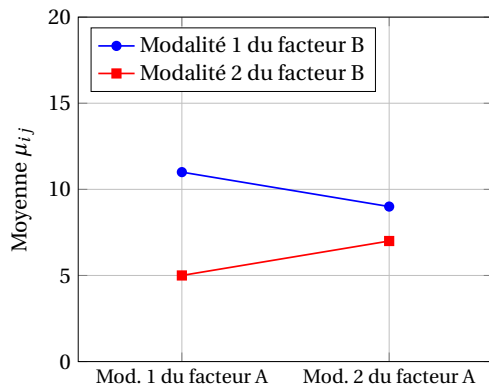


$$\mu = 8, \alpha = (-2; 2), \beta = (1, -1), \gamma = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$



$$\mu = 8, \alpha = (-2; 2), \beta = (1, -1), \gamma = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

En l'absence d'interaction (ci-dessus) les modalités de B sont translattées l'une par rapport à l'autre.



$$\mu = 8, \alpha = (0; 0), \beta = (2, -2), \gamma = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad \mu = 8, \alpha = (1.5; -1.5), \beta = (2, -2), \gamma = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

En présence d'interaction (ci-dessus), difficile de juger l'effet de A et B à partir des graphiques.

Remarque. Il est tout à fait possible d'avoir interaction sans effet principal des deux facteurs. Par exemple, imaginons une étude sur l'efficacité de deux types de médicaments (facteur A) sur deux groupes de patients (facteur B : Hommes et Femmes).

- Médicament 1 : très efficace sur les hommes (+10 points de santé), mais réduit la santé des femmes (-10 points).
- Médicament 2 : réduit la santé des hommes (-10 points), mais très efficace sur la santé des femmes (+10 points).

Effet principal du facteur A : si nous faisons la moyenne globale, les +10 et -10 s'annulent. La moyenne du Médicament 1 est 0, celle du Médicament 2 est 0, l'effet principal est non significatif.

Effet principal du facteur B : la moyenne des hommes est 0, celle des femmes est 0, l'effet principal est non significatif.

L'interaction est, pour sa part, hautement significative. L'efficacité du médicament dépend totalement de si vous êtes un homme ou une femme.

Il est souvent erroné (ou du moins incomplet) de conclure qu'un facteur n'a « pas d'effet » simplement parce que son effet principal n'est pas significatif. Dans le cas d'une interaction significative, le facteur a bien un impact, mais cet impact est conditionnel à la modalité.

Bien entendu, la difficulté est qu'en pratique, la situation dans laquelle on se trouve est difficile à identifier à cause des fluctuations aléatoires sur les observations.

Selon des calculs semblables à ceux de la section 2.2 pour l'ANOVA à un facteur, la décomposition de la variance devient dans le cas de deux facteurs :

$$S_T^2 = S_A^2 + S_B^2 + S_I^2 + S_R^2$$

avec $NS_A^2 = qr \sum_i (\bar{y}_{i..} - \bar{y}_{...})^2$, $NS_B^2 = pr \sum_i (\bar{y}_{.j.} - \bar{y}_{...})^2$, $NS_I^2 = r \sum_i \sum_j (\bar{y}_{ij.} - \bar{y}_{.j.} - \bar{y}_{i..} + \bar{y}_{...})^2$, et enfin $NS_R^2 = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij.})^2$.

Si on suppose que les moyennes « annulent » les termes aléatoires ε_{ijk} dans l'expression des y_{ijk} , on obtient, grâce aux hypothèses $\sum_i \alpha_i = \sum_j \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$, les approximations suivantes : $\sum_i (\bar{y}_{i..} - \bar{y}_{...})^2 \simeq \sum_i \alpha_i^2$, $\sum_i (\bar{y}_{.j.} - \bar{y}_{...})^2 \simeq \sum_j \beta_j^2$, $\sum_i \sum_j (\bar{y}_{ij.} - \bar{y}_{.j.} - \bar{y}_{i..} + \bar{y}_{...})^2 \simeq \sum_i \sum_j \gamma_{ij}^2$, et enfin $\sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij.})^2 \simeq N\sigma^2$. Ceci explique (en divisant ces quantités par $N = pqr$), les dénominations suivantes. Les quantités S_T^2 , S_A^2 , S_B^2 , S_I^2 , et S_R^2 sont les variances, respectivement, totale, expliquée par le facteur A, expliquée par le facteur B, expliquée par l'interaction entre A et B, et résiduelle.

On calcule les sommes de carrés NS_A^2 , NS_B^2 , NS_I^2 à l'aide d'un tableau des moyennes et des moyennes marginales analogue au tableau suivant (pour $p = 2$ et $q = 3$). La valeur de NS_R^2 ne peut, pour sa part, pas se calculer juste avec ce tableau.

Facteur A \ Facteur B	B_1	B_2	B_3	Moy. factorielle A
A_1	\bar{y}_{11}	\bar{y}_{12}	\bar{y}_{13}	$\bar{y}_{1..}$
A_2	\bar{y}_{21}	\bar{y}_{22}	\bar{y}_{23}	$\bar{y}_{2..}$
Moy. factorielle B	$\bar{y}_{..1}$	$\bar{y}_{..2}$	$\bar{y}_{..3}$	$\bar{y}_{...}$

De la même manière que pour l'ANOVA à un facteur, sous l'hypothèse \mathcal{H}_0 selon laquelle tous les μ_{ij} sont égaux, on forme des statistiques qui suivent des lois du χ^2 indépendantes, dont les rapports suivent des lois de Fisher-Snedecor. On remplit alors le tableau suivant.

Source de variation	Somme des carrés	ddl	Carré moyen	F_{obs}
Facteur A	$A = NS_A^2$	$p - 1$	$\frac{A}{p-1}$	$\frac{A/(p-1)}{R/(pq(r-1))}$
Facteur B	$B = NS_B^2$	$q - 1$	$\frac{B}{q-1}$	$\frac{B/(q-1)}{R/(pq(r-1))}$
Interaction A/B	$I = NS_I^2$	$(p-1)(q-1)$	$\frac{I}{(p-1)(q-1)}$	$\frac{I/((p-1)(q-1))}{R/(pq(r-1))}$
Résiduelle	$R = NS_R^2$	$pq(r-1)$	$\frac{R}{pq(r-1)}$	
Total	$T = NS_T^2$	$pqr - 1$		

Les logiciels ou les bibliothèques Python donnent une colonne supplémentaire avec les p -valeurs des statistiques de test calculées sur les observations.

Il n'est pas facile de comprendre intuitivement ce que signifie « rejeter » ou « accepter » l'hypothèse \mathcal{H}_0 (tous les μ_{ij} sont égaux à μ) selon un des trois tests du tableau précédent. Notons que sous \mathcal{H}_0 , $\forall i, j, \alpha_i + \beta_j + \gamma_{ij} = 0$ par définition. Ainsi (en sommant sur j) $\forall i, \alpha_i = 0$, (en sommant sur i) $\forall j, \beta_j = 0$, et donc $\forall i, j, \gamma_{ij} = 0$.

C'est la raison pour laquelle certains auteurs parlent de trois hypothèses \mathcal{H}_0 distinctes pour F_A (qui serait $\mathcal{H}_0 : \forall i, \alpha_i = 0$), F_B ($\mathcal{H}_0 : \forall j, \beta_j = 0$), et F_{AB} ($\mathcal{H}_0 : \forall i, j, \gamma_{ij} = 0$).

L'implication importante de cette remarque est que rejeter \mathcal{H}_0 avec F_{AB} par exemple veut dire que les γ_{ij} ne sont pas tous nuls (ce qui est plus facile à interpréter que « il existe une différence entre deux moyennes μ_{ij} et $\mu_{i'j'}$ »). Certes, accepter \mathcal{H}_0 avec F_{AB} permet d'accepter la proposition équivalente $\forall i, j, \alpha_i = \beta_j = \gamma_{ij} = 0$, mais cela ne veut pas dire que \mathcal{H}_0 est vraie. Dans la statistique de test de la troisième ligne, la seule information mesurée par $I = NS_I^2$ porte sur les interactions, donc il est raisonnable de se contenter d'accepter $\forall i, j, \gamma_{ij} = 0$ et de conclure seulement à l'absence d'interaction.

À un niveau de risque α fixé, les tests à effectuer sont les suivants :

- l'interaction n'est pas significative si $\frac{I/((p-1)(q-1))}{R/(pq(r-1))}$ est inférieure au quantile $1 - \alpha$ de la loi de Fisher-Snedecor à $(p-1)(q-1)$ et $pq(r-1)$ degrés de liberté (on accepte $\forall i, j, \gamma_{ij} = 0$);
- le facteur A n'est pas significatif si $\frac{A/(p-1)}{R/(pq(r-1))}$ est inférieure au quantile $1 - \alpha$ de la loi de Fisher-Snedecor à $p-1$ et $pq(r-1)$ degrés de liberté (on accepte $\forall i, \alpha_i = 0$);
- le facteur B n'est pas significatif si $\frac{B/(q-1)}{R/(pq(r-1))}$ est inférieure au quantile $1 - \alpha$ de la loi de Fisher-Snedecor à $q-1$ et $pq(r-1)$ degrés de liberté (on accepte $\forall i, \beta_i = 0$);

Autrement dit, l'interaction, le facteur A, ou le facteur B ne sont pas significatifs si l'hypothèse H_0 reste plausible face à des valeurs, respectivement, des sommes des carrés expliqués par l'interaction (I), expliqués par A (A), ou expliqués par B (B) qui ne sont pas assez grandes par rapport aux carrés expliqués par les résidus (R).

D'après la discussion à la fin de l'exemple 4, ces tests sont à faire dans cet ordre. Si l'interaction est significative, l'effet des facteurs A et B devient difficile à interpréter. Il faudrait alors comparer l'effet

de A pour chaque niveau de B par exemple (tests d'effet simple). Si l'interaction n'est pas significative, on teste l'effet des facteurs A et B.

Si les trois effets ne sont pas significatifs (aucun des trois tests ne permet de rejeter \mathcal{H}_0), alors la population est homogène et distribuée selon la loi $\mathcal{N}(\mu, \sigma^2)$.

Remarque. Si $r = 1$ (une seule observation par couple de modalités), on a un problème car dire que R suit la loi du χ^2 à $pq(r-1) = 0$ d.d.l. n'a pas de sens. Il faut alors adapter l'ANOVA (voir un livre de statistique).

3.3 Cas pratique

On souhaite étudier l'influence de deux facteurs sur les notes obtenues par des étudiants à un examen de statistiques :

1. *le Facteur A (Méthode)* : Classique (A_1) vs Interactive (A_2);
2. *le Facteur B (Moment)* : Matin (B_1) vs Après-midi (B_2).

Pour chaque combinaison de facteurs, on a mesuré les notes de 3 étudiants (répétitions). Les résultats sont consignés dans le tableau ci-dessous :

Méthode \ Moment	Matin (B_1)	Après-midi (B_2)
Classique (A_1)	11; 12; 13	8; 9; 10
Interactive (A_2)	13; 14; 15	15; 16; 17

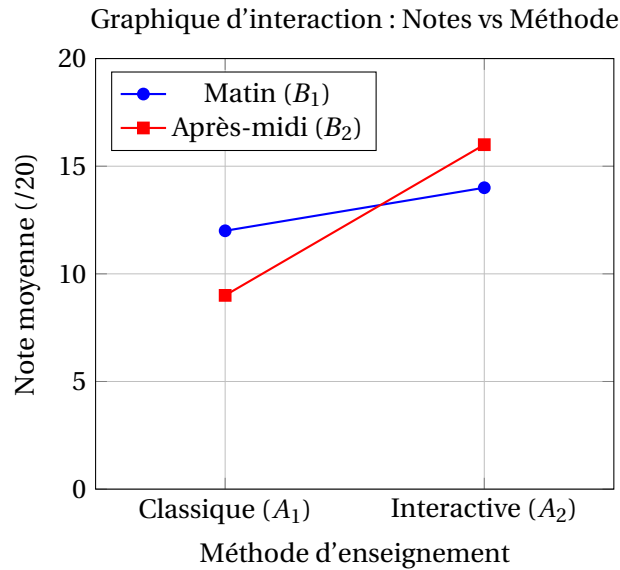
Pour l'ANOVA à deux facteurs, la première étape consiste à calculer les moyennes par cellule (combinaisons), les moyennes marginales (par facteur) et la moyenne générale. Notons $Y_{ij} = \bar{y}_{ij}$ la moyenne des observations de modalités A_i et B_j .

Méthode \ Moment	Matin (B_1)	Après-midi (B_2)	Moy. Factorielle A
Classique (A_1)	12,0	9,0	$\bar{Y}_{1.} = \mathbf{10,5}$
Interactive (A_2)	14,0	16,0	$\bar{Y}_{2.} = \mathbf{15,0}$
Moy. Factorielle B	$\bar{Y}_{.1} = \mathbf{13,0}$	$\bar{Y}_{.2} = \mathbf{12,5}$	$\bar{Y}_{..} = 12,75$

On trace le graphique d'interaction suivant, et on observe que :

- Pour la méthode *Classique*, la note diminue de 3 points l'après-midi.
- Pour la méthode *Interactive*, la note augmente de 2 points l'après-midi.

Les segments ne sont pas parallèles et se croisent. Cela suggère visuellement une forte interaction entre la méthode d'enseignement et le moment de la journée.



Voici le tableau de l'ANOVA rempli avec les calculs des carrés moyens et des rapports de Fisher-Snedecor F_{obs} .

Source	Somme des carrés	ddl	Carré moyen	F_{obs}
Facteur A (Méthode)	80,67	$2 - 1 = 1$	80,67	40,34
Facteur B (Moment)	0,67	$2 - 1 = 1$	0,67	0,34
Interaction (A \times B)	24,00	$(2 - 1)(2 - 1) = 1$	24,00	12,00
Résiduelle	16,00	$12 - 4 = 8$	2,00	
Total	121,34	11		

Interprétation statistique ($\alpha = 5\%$). La valeur critique donnée est $F_{\text{crit}}(1, 8) = 5,32$. Il se trouve que, dans cet exemple, la valeur critique est la même pour les trois tests car les degrés de liberté sont les mêmes. Cela n'a bien entendu aucune raison d'être le cas en général.

1. *Interaction* : $F_{\text{obs}} = 12,00 > 5,32$. L'interaction est significative.
2. *Effet Méthode* : $F_{\text{obs}} = 40,34 > 5,32$. L'effet est très significatif. La méthode interactive est globalement supérieure à la méthode classique (moyenne de 15,0 pour la première, et 10,5 pour la seconde).
3. *Effet Moment* : $F_{\text{obs}} = 0,34 < 5,32$. L'effet n'est pas significatif. Pris isolément, le moment de la journée n'influence pas la note moyenne (moyenne de 13,0 le matin, et 12,5 l'après-midi).

Même si le facteur Moment n'a pas une influence globale significative, comme l'effet de l'interaction entre Moment et Méthode est significative, l'étude montre que l'efficacité d'une méthode dépend sans doute du moment où elle est appliquée.

- Pour les cours du *matin*, les deux méthodes sont relativement proches, bien que la méthode interactive soit légèrement devant.
- Pour les cours de l'*après-midi*, la méthode classique provoque une chute des résultats (fatigue/passivité), tandis que la méthode interactive stimule les étudiants et améliore leurs performances.

La préconisation à la suite de l'étude pourrait être de privilégier la méthode interactive pour les créneaux de fin de journée. Néanmoins, la puissance du test est sans doute assez faible, les échantillons étant de taille limitée.

4 Exercices

4.1 Rendement agricole

Un agronome souhaite comparer l'effet de trois types d'engrais (A , B et C) sur le rendement du blé. Il a sélectionné 9 parcelles de terrain identiques et a appliqué chaque engrais sur 3 parcelles choisies au hasard. Les rendements obtenus (en quintaux par hectare) sont les suivants :

- **Engrais A** : 20; 22; 24
- **Engrais B** : 28; 30; 32
- **Engrais C** : 18; 20; 16

1) Analyse préliminaire

Calculer la moyenne de chaque groupe ($\bar{x}_A, \bar{x}_B, \bar{x}_C$) ainsi que la moyenne générale de l'échantillon (\bar{x}_{total}).

2) Sommes des Carrés

Calculer la Somme des Carrés Inter-groupes (SSB) et la Somme des Carrés Intra-groupe (SSW).
Rappel : $SSB = \sum n_i (\bar{x}_i - \bar{x}_{\text{total}})^2$ et $SSW = \sum \sum (y_{ij} - \bar{x}_i)^2$

3) Tableau de l'ANOVA

Compléter le tableau d'analyse de la variance ci-dessous :

Source de variation	Somme des carrés	ddl	Carré moyen	F observé
Inter-groupes (Facteur)
Intra-groupe (Résidus)	
Total		

4) Prise de décision

On donne la valeur critique de Fisher-Snedecor au seuil $\alpha = 5\%$ pour les degrés de liberté correspondants : $F_{\text{critique}}(2, 6) = 5,14$.

1. Formuler l'hypothèse nulle \mathcal{H}_0 et l'hypothèse alternative \mathcal{H}_1 .
2. Comparer F_{obs} et F_{critique} .
3. Peut-on conclure que le type d'engrais a un impact significatif sur le rendement?

4.2 Café et sommeil

On étudie l'influence de deux facteurs sur le score de concentration (sur 100) d'un groupe d'étudiants :

- **Facteur A (Sommeil)** : Court (A_1) ou Long (A_2).
- **Facteur B (Café)** : Sans café (B_1) ou Avec café (B_2).

On dispose de 2 répétitions par condition (8 observations au total). Les scores sont les suivants :

Sommeil \ Café	Sans Café (B_1)	Avec Café (B_2)
Court (A_1)	40; 50	70; 80
Long (A_2)	80; 90	85; 95

1) Moyennes des cellules

Calculer la moyenne de score pour chacune des 4 combinaisons (cellules), ainsi que les moyennes marginales pour les facteurs A et B, et enfin la moyenne générale G.

2) Analyse de la Variance

On donne les sommes des carrés suivantes :

- $SS_{\text{Sommeil}} = 1512,5$
- $SS_{\text{Café}} = 612,5$
- $SS_{\text{Résiduelle}} = 200,0$
- $SS_{\text{Total}} = 2637,5$

Question : par déduction, calculez la somme des carrés de l'interaction ($SS_{A \times B}$).

3. Tableau de l'ANOVA

Compléter le tableau d'analyse de la variance ci-dessous :

Source	Somme des carrés	ddl	Carré moyen	F observé
Sommeil (A)	1512,5
Café (B)	612,5
Interaction (AxB)
Résiduelle	200,0	
Total	2637,5	7		

4) Interprétation

Au seuil $\alpha = 5\%$, la valeur critique de Fisher-Snedecor $F(1, 4)$ est de **7,71**. Quelles sont vos conclusions concernant l'effet du sommeil, l'effet du café et l'existence d'une interaction ?

4.3 ANOVA en Python

Voir sur Arche.