

Chapitre 5a – Analyse de la variance

STT-1920 Méthodes statistiques

Département de mathématiques et de statistique



UNIVERSITÉ
LAVAL

Introduction

Objectifs de l'analyse de la variance (ANOVA)

Analyse de la variance à un facteur

- Comparer les **moyennes** d'une variable dans **plusieurs populations** (normales, de même variance);
- Déterminer si un **facteur** (variable discrète ou qualitative) a une influence sur une **variable réponse** (variable aléatoire continue).

Exemples

- Un ingénieur aux transports veut comparer la durabilité moyenne de 7 mélanges de peinture pour asphalte.

Exemples

- Un ingénieur aux transports veut comparer la durabilité moyenne de 7 mélanges de peinture pour asphalte.
- Un chercheur développe 3 traitements contre l'hypertension. Il désire savoir si l'un d'entre eux a un délai d'action plus court que les autres.

Exemples

- Un ingénieur aux transports veut comparer la durabilité moyenne de 7 mélanges de peinture pour asphalte.
- Un chercheur développe 3 traitements contre l'hypertension. Il désire savoir si l'un d'entre eux a un délai d'action plus court que les autres.
- Un agronome veut savoir si les 4 variétés de blé qui résistent le mieux à la moisissure donnent des rendements moyens égaux.

Exemples

- Un ingénieur aux transports veut comparer la durabilité moyenne de 7 mélanges de peinture pour asphalte.
- Un chercheur développe 3 traitements contre l'hypertension. Il désire savoir si l'un d'entre eux a un délai d'action plus court que les autres.
- Un agronome veut savoir si les 4 variétés de blé qui résistent le mieux à la moisissure donnent des rendements moyens égaux.
- Un directeur de magasin d'appareils électroniques veut comparer 5 méthodes de publicité : sont-elles équivalentes ou l'une d'elles amène-t-elle plus de clients que les autres ?

Principes d'expérimentation

Pourquoi planifier une expérience ?

Pour qu'une expérience mène à une conclusion fiable,

1. elle doit être **bien planifiée** ;
2. ses résultats doivent être **adéquatement analysés** (c'est l'ANOVA).

On planifie pour se protéger contre les **biais** et pour **contrôler la variabilité** due à d'autres facteurs que celui qui nous intéresse.

Exemple 1 : régimes alimentaires chez la souris

Une chercheuse veut savoir si trois régimes (standard, riche en gras, riche en fibres) mènent au même **gain de poids** moyen chez la souris après 8 semaines. Elle dispose de 24 souris, chacune logée dans sa propre cage.

Pourquoi planifier une expérience ?

Pour qu'une expérience mène à une conclusion fiable,

1. elle doit être **bien planifiée** ;
2. ses résultats doivent être **adéquatement analysés** (c'est l'ANOVA).

On planifie pour se protéger contre les **biais** et pour **contrôler la variabilité** due à d'autres facteurs que celui qui nous intéresse.

Exemple 1 : régimes alimentaires chez la souris

Une chercheuse veut savoir si trois régimes (standard, riche en gras, riche en fibres) mènent au même **gain de poids** moyen chez la souris après 8 semaines. Elle dispose de 24 souris, chacune logée dans sa propre cage.

Or, le gain de poids varie aussi selon le poids initial, l'âge, le sexe, la portée d'origine, l'emplacement de la cage (température, lumière), la personne qui manipule les souris...

Comment isoler l'effet du régime ?

Trois principes essentiels

Principes d'expérimentation

1. **Contrôle** : identifier les sources de variabilité, puis les garder constantes ou les équilibrer entre les groupes;
2. **Répétition** : appliquer chaque traitement à plusieurs unités, pour pouvoir estimer la variabilité naturelle (σ^2);
3. **Randomisation** : attribuer les traitements au hasard, pour éliminer les biais systématiques.

Trois principes essentiels

Principes d'expérimentation

1. **Contrôle** : identifier les sources de variabilité, puis les garder constantes ou les équilibrer entre les groupes;
2. **Répétition** : appliquer chaque traitement à plusieurs unités, pour pouvoir estimer la variabilité naturelle (σ^2);
3. **Randomisation** : attribuer les traitements au hasard, pour éliminer les biais systématiques.

Dans l'exemple 1 :

1. souris de même sexe, même âge et même lignée; mêmes conditions d'élevage;
2. 8 souris par régime (et non une seule) : on pourra mesurer à quel point deux souris d'un même régime diffèrent;
3. on tire au hasard les souris qui reçoivent chaque régime. Si on donnait le régime gras aux plus grosses souris, on ne pourrait plus séparer l'effet du régime de celui du poids initial!

Expérience ou étude observationnelle ?

- **Expérience** : le chercheur **impose** les traitements aux unités, idéalement au hasard. Ex. : les régimes sont attribués aux souris (exemple 1).
- **Étude observationnelle** : on compare des groupes **déjà formés**, sans intervenir. Ex. : comparer la pression artérielle moyenne de non-fumeurs, d'ex-fumeurs et de fumeurs.

Expérience ou étude observationnelle ?

- **Expérience** : le chercheur **impose** les traitements aux unités, idéalement au hasard. Ex. : les régimes sont attribués aux souris (exemple 1).
- **Étude observationnelle** : on compare des groupes **déjà formés**, sans intervenir. Ex. : comparer la pression artérielle moyenne de non-fumeurs, d'ex-fumeurs et de fumeurs.

Attention !

Les calculs de l'ANOVA sont les mêmes dans les deux cas, mais seule une **expérience randomisée** permet de conclure à un lien de **cause à effet**. Dans une étude observationnelle, une différence entre les groupes peut être due à une **variable confondante** (âge, alimentation, activité physique...).

Vocabulaire

Vocabulaire de l'expérimentation

- **Variable réponse** (Y) : la variable dont on compare les moyennes;
- **Facteur** : la caractéristique (qualitative) qui distingue les groupes;
- **Niveaux** (ou modalités) : les valeurs possibles du facteur. Dans une expérience, chaque niveau est un **traitement**;
- **Unité expérimentale** : l'individu ou l'objet auquel on applique **un** traitement, et sur lequel on mesure Y.

Vocabulaire

Vocabulaire de l'expérimentation

- **Variable réponse** (Y) : la variable dont on compare les moyennes;
- **Facteur** : la caractéristique (qualitative) qui distingue les groupes;
- **Niveaux** (ou modalités) : les valeurs possibles du facteur. Dans une expérience, chaque niveau est un **traitement**;
- **Unité expérimentale** : l'individu ou l'objet auquel on applique **un** traitement, et sur lequel on mesure Y.

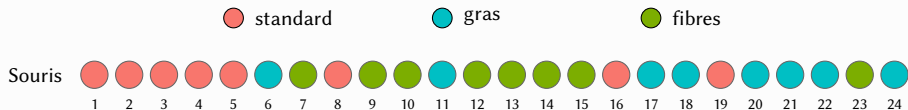
Exemple 1 : variable réponse = gain de poids (g); facteur = régime, à 3 niveaux (standard, gras, fibres); unité expérimentale = une souris.

En ANOVA à un facteur, les échantillons sont **indépendants** : une seule mesure de Y par unité expérimentale.

Devis complètement aléatoire

Devis complètement aléatoire

Les N unités expérimentales sont attribuées aux I traitements **entièrement au hasard** (n_i unités reçoivent le traitement i).



Tirage au hasard en R (8 souris par régime) :

```
> set.seed(1920)
> sample(rep(c("standard", "gras", "fibres"), each = 8))
[1] "standard" "standard" "standard" "standard" "standard" "gras"
[7] "fibres"   "standard" "fibres"   "fibres"   "gras"     "fibres"
[13] "fibres"   "fibres"   "fibres"   "standard" "gras"     "gras"
[19] "standard" "gras"     "gras"     "gras"     "fibres"   "gras"
```

Dans une étude observationnelle, on obtient plutôt I échantillons indépendants en tirant au hasard n_i unités dans chacune des I populations.

QUIZ!

Une agronome compare 4 engrais (A, B, C, D) sur la croissance de plants de tomate. Elle dispose de 24 plants, chacun dans son pot. Elle tire au hasard 6 pots pour chaque engrais et mesure la hauteur des plants (cm) après 6 semaines.

Identifiez :

- la variable réponse;
- le facteur et ses niveaux;
- l'unité expérimentale;
- s'agit-il d'une expérience ou d'une étude observationnelle?

QUIZ!

1. Un biologiste compare la croissance de truites élevées à 3 températures de l'eau. Il dispose de 6 bassins (2 par température, tirés au hasard) contenant chacun 15 truites. Il mesure la croissance de chaque truite.
 - Quelle est l'unité expérimentale ?
 - Combien de répétitions y a-t-il par traitement ?
2. Une écologiste compare la masse moyenne de souris sylvestres capturées dans 3 habitats (forêt, champ, milieu périurbain).
 - Expérience ou étude observationnelle ?
 - Peut-elle conclure que l'habitat **cause** les différences de masse ?

L'idée de l'ANOVA

ANOVA : décomposition de la variance

L'analyse de la variance consiste à identifier et à comparer les **sources de variation** d'une variable :

ANOVA : décomposition de la variance

L'analyse de la variance consiste à identifier et à comparer les **sources de variation** d'une variable :

- la variabilité **inter-traitements**, entre les moyennes des échantillons, que l'on associe au **facteur** (la caractéristique qui distingue les groupes à l'étude);

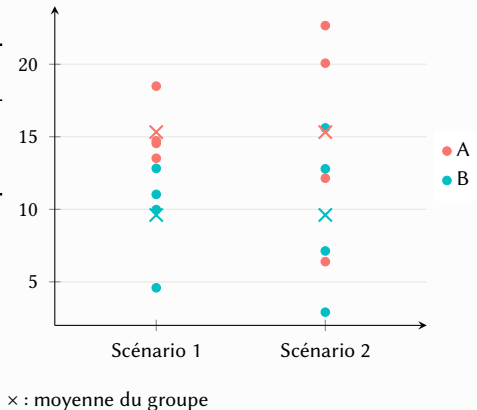
ANOVA : décomposition de la variance

L'analyse de la variance consiste à identifier et à comparer les **sources de variation** d'une variable :

- la variabilité **inter-traitements**, entre les moyennes des échantillons, que l'on associe au **facteur** (la caractéristique qui distingue les groupes à l'étude);
- la variabilité **intra-traitement**, entre les observations d'un même échantillon, que l'on associe à l'**erreur échantillonnale** (non expliquée par le modèle).

Comparons deux scénarios

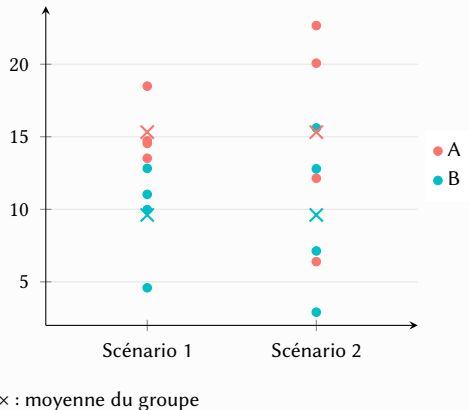
Moyennes	Scénario 1	Scénario 2
Traitement A	15,32	15,32
Traitement B	9,61	9,61



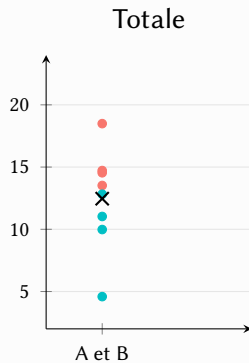
Comparons deux scénarios

Moyennes	Scénario 1	Scénario 2
Traitement A	15,32	15,32
Traitement B	9,61	9,61

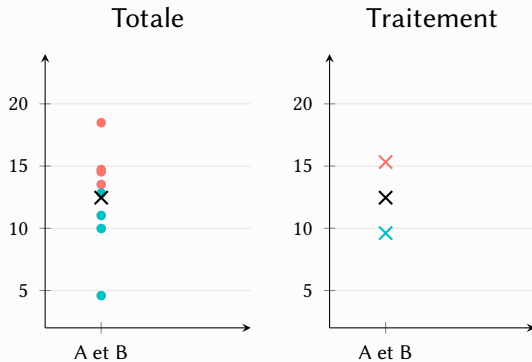
Les moyennes sont **identiques** dans les deux scénarios. Mais la dispersion à l'intérieur des groupes, elle, ne l'est pas...



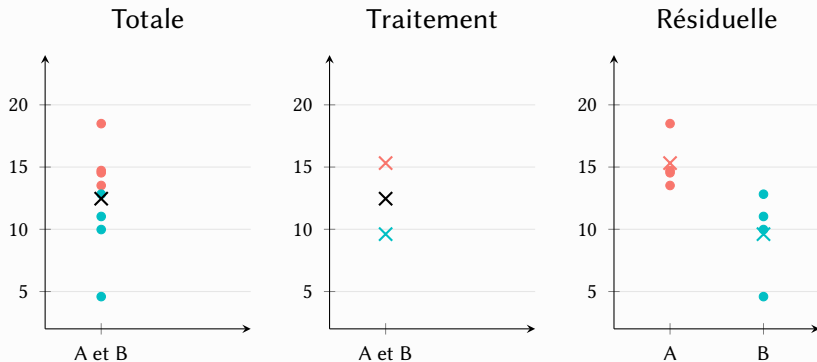
Scénario 1 : décomposition de la variance



Scénario 1 : décomposition de la variance



Scénario 1 : décomposition de la variance



Scénario 1 : table d'analyse de la variance

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Traitement	1	65.27	65.27	7.543	0.0334	*
Residuals	6	51.91	8.65			

Signif. codes: 0 '***' 0.001 '**' 0.01 '.' 0.1 ' ' 1

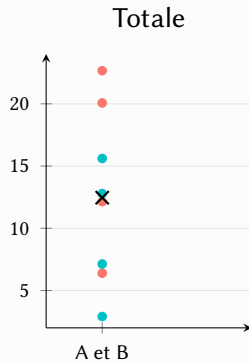
Scénario 1 : table d'analyse de la variance

	Df	Sum Sq	Mean Sq	F	value	Pr(>F)
Traitement	1	65.27	65.27	7.543	0.0334	*
Residuals	6	51.91	8.65			

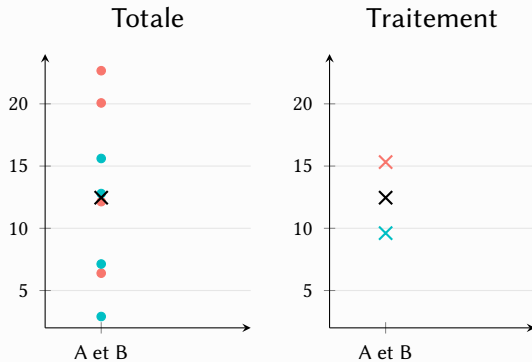
Signif. codes: 0 '***' 0.001 '**' 0.01 '.' 0.1 ' ' 1

La variabilité **entre** les traitements (65,27) est grande par rapport à la variabilité **à l'intérieur** des traitements (8,65).

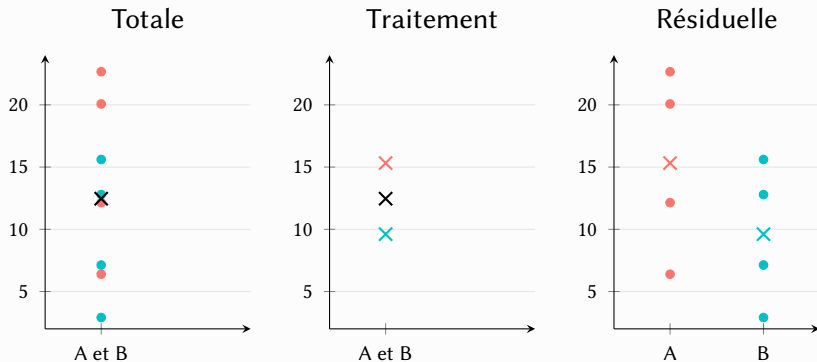
Scénario 2 : décomposition de la variance



Scénario 2 : décomposition de la variance



Scénario 2 : décomposition de la variance



Scénario 2 : table d'analyse de la variance

	Df	Sum Sq	Mean Sq	F	value	Pr(>F)
Traitement	1	65.15	65.15	1.484		0.269
Residuals	6	263.45	43.91			

Scénario 2 : table d'analyse de la variance

	Df	Sum Sq	Mean Sq	F	value	Pr(>F)
Traitement	1	65.15	65.15	1.484	0.269	
Residuals	6	263.45	43.91			

Même écart entre les moyennes, mais la variabilité à l'intérieur des traitements est maintenant beaucoup plus grande : on ne peut plus conclure à une différence entre A et B.

Mathématique de l'ANOVA

Cadre

But : comparer les moyennes $\mu_1, \dots, \mu_I$ d'une variable Y dans I populations.

- $Y_{11}, Y_{12}, \dots, Y_{1n_1}$ sont issues d'une population $N(\mu_1, \sigma^2)$;
- $Y_{21}, Y_{22}, \dots, Y_{2n_2}$ sont issues d'une population $N(\mu_2, \sigma^2)$;
- $\vdots$
- $Y_{I1}, Y_{I2}, \dots, Y_{In_I}$ sont issues d'une population $N(\mu_I, \sigma^2)$.

Cadre

But : comparer les moyennes $\mu_1, \dots, \mu_I$ d'une variable Y dans I populations.

- $Y_{11}, Y_{12}, \dots, Y_{1n_1}$ sont issues d'une population $N(\mu_1, \sigma^2)$;
- $Y_{21}, Y_{22}, \dots, Y_{2n_2}$ sont issues d'une population $N(\mu_2, \sigma^2)$;
- $\vdots$
- $Y_{I1}, Y_{I2}, \dots, Y_{In_I}$ sont issues d'une population $N(\mu_I, \sigma^2)$.

On note $N = n_1 + n_2 + \dots + n_I$ le nombre total d'observations.

Notation pour les sommes et les moyennes

y_{ij} = j^{e} observation du groupe (traitement) i . Le symbole $\bullet$ remplace un indice sur lequel on a **sommé**; la barre indique une **moyenne**.

Notation

$$y_{i\bullet} = \sum_{j=1}^{n_i} y_{ij} \text{ (somme du groupe } i),$$

$$\bar{y}_{i\bullet} = \frac{y_{i\bullet}}{n_i} \text{ (moyenne du groupe } i),$$

$$y_{\bullet\bullet} = \sum_{i=1}^I \sum_{j=1}^{n_i} y_{ij} \text{ (somme totale),}$$

$$\bar{y}_{\bullet\bullet} = \frac{y_{\bullet\bullet}}{N} \text{ (moyenne globale).}$$

Notation pour les sommes et les moyennes

y_{ij} = j^{e} observation du groupe (traitement) i . Le symbole $\bullet$ remplace un indice sur lequel on a **sommé**; la barre indique une **moyenne**.

Notation

$$\begin{aligned} y_{i\bullet} &= \sum_{j=1}^{n_i} y_{ij} \text{ (somme du groupe } i), & \bar{y}_{i\bullet} &= \frac{y_{i\bullet}}{n_i} \text{ (moyenne du groupe } i), \\ y_{\bullet\bullet} &= \sum_{i=1}^I \sum_{j=1}^{n_i} y_{ij} \text{ (somme totale),} & \bar{y}_{\bullet\bullet} &= \frac{y_{\bullet\bullet}}{N} \text{ (moyenne globale).} \end{aligned}$$

Scénario 1 (A = groupe 1, B = groupe 2; $n_1 = n_2 = 4$, $N = 8$) :

$$y_{1\bullet} = 14,72 + 13,52 + 14,54 + 18,49 = 61,27 \text{ et } \bar{y}_{1\bullet} = 61,27/4 = 15,32;$$

$$y_{2\bullet} = 38,42 \text{ et } \bar{y}_{2\bullet} = 9,61; \quad y_{\bullet\bullet} = 99,69 \text{ et } \bar{y}_{\bullet\bullet} = 99,69/8 = 12,46.$$

Hypothèses à confronter

Hypothèses de l'ANOVA à un facteur

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_I$$

H_1 : au moins 2 moyennes différent

Le modèle à un facteur

Modèle

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \quad Y_{ij} \sim N(\mu_i, \sigma^2), \quad i = 1, \dots, I, \quad j = 1, \dots, n_i$$

- $\mu_i =$
- $\varepsilon_{ij} =$

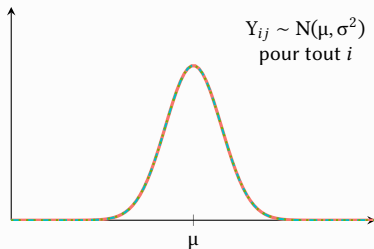
Estimation des paramètres :

- μ_i estimé par
- μ estimé par
- σ^2 estimé par

Loi de Y_{ij} sous H_0 et sous H_1

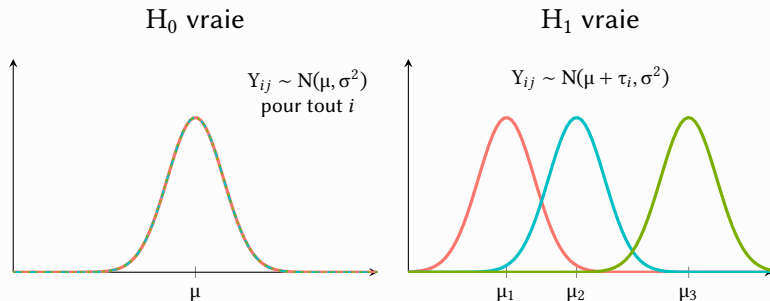
On peut écrire $\mu_i = \mu + \tau_i$, où μ est la moyenne globale et $\tau_i = \mu_i - \mu$ est l'effet du traitement i . Alors $H_0 : \mu_1 = \dots = \mu_I \iff H_0 : \tau_1 = \dots = \tau_I = 0$.

H_0 vraie



Loi de Y_{ij} sous H_0 et sous H_1

On peut écrire $\mu_i = \mu + \tau_i$, où μ est la moyenne globale et $\tau_i = \mu_i - \mu$ est l'effet du traitement i . Alors $H_0 : \mu_1 = \dots = \mu_I \iff H_0 : \tau_1 = \dots = \tau_I = 0$.



Sous H_0 , les I groupes proviennent de la même loi normale. Sous H_1 , les lois ont la même forme (même σ^2), mais au moins deux sont décalées.

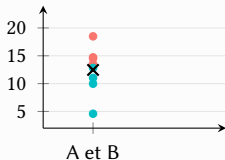
Décomposition de la variance

Décomposition des sommes de carrés

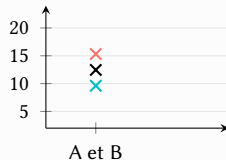
$$SST = SST_R + SSE$$

$$\sum_{i=1}^I \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^I n_i (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^I \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$$

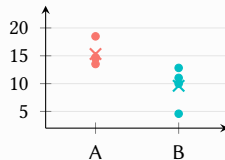
Totale



Traitement



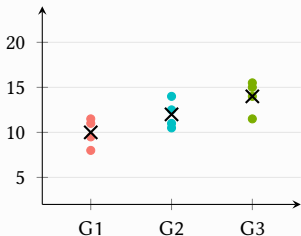
Résiduelle



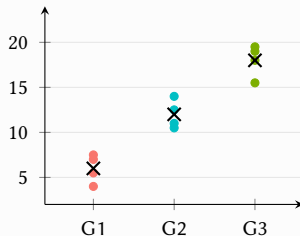
QUIZ!

Deux expériences ont obtenu les mêmes **variances** échantillonnales (même dispersion dans chaque groupe) :

Expérience 1



Expérience 2



- Quelle expérience a le plus grand SST_R ? le plus grand SSE ? le plus grand SST ?
- Et pour les scénarios 1 et 2 vus plus tôt (mêmes **moyennes**) ?

Table d'ANOVA

Source de variation	Somme de carrés	d.l.	Carrés moyens	F_{obs}
Traitements (inter-trt)	SST_R	$I - 1$	$MST_R = \frac{SST_R}{I - 1}$	$\frac{MST_R}{MSE}$
Erreur (intra-trt)	SSE	$N - I$	$MSE = \frac{SSE}{N - I}$	
Total	SST	$N - 1$		

Comparer MST_R et MSE

- S'il n'y a pas de différence entre les traitements (H_0 vraie), MST_R et MSE sont tous les deux de bons estimateurs de σ^2 .

Comparer MST_R et MSE

- S'il n'y a pas de différence entre les traitements (H_0 vraie), MST_R et MSE sont tous les deux de bons estimateurs de σ^2 .
- Si l'effet des traitements est important, on aura

$$MST_R \gg MSE.$$

Comparer MST_R et MSE

- S'il n'y a pas de différence entre les traitements (H_0 vraie), MST_R et MSE sont tous les deux de bons estimateurs de σ^2 .
- Si l'effet des traitements est important, on aura

$$MST_R \gg MSE.$$

Question : à partir de quand considère-t-on que MST_R est vraiment plus grand que MSE ?

Pourquoi $F \approx 1$ sous H_0 ? Espérance des carrés moyens

Espérance des carrés moyens

On peut montrer que

$$E(\text{MSE}) = \sigma^2 \quad \text{et} \quad E(\text{MST}_R) = \sigma^2 + \frac{\sum_{i=1}^I n_i \tau_i^2}{I - 1}, \quad \text{où } \tau_i = \mu_i - \mu.$$

- MSE estime toujours σ^2 , que H_0 soit vraie ou non.

Pourquoi $F \approx 1$ sous H_0 ? Espérance des carrés moyens

Espérance des carrés moyens

On peut montrer que

$$E(\text{MSE}) = \sigma^2 \quad \text{et} \quad E(\text{MST}_R) = \sigma^2 + \frac{\sum_{i=1}^I n_i \tau_i^2}{I - 1}, \quad \text{où } \tau_i = \mu_i - \mu.$$

- MSE estime toujours σ^2 , que H_0 soit vraie ou non.
- Sous H_0 : tous les $\tau_i = 0 \Rightarrow E(\text{MST}_R) = \sigma^2 \Rightarrow F = \text{MST}_R / \text{MSE}$ devrait être **proche de 1**.
- Sous H_1 : au moins un $\tau_i \neq 0 \Rightarrow E(\text{MST}_R) > \sigma^2 \Rightarrow F$ tend à être **plus grand que 1**.

Pourquoi $F \approx 1$ sous H_0 ? Espérance des carrés moyens

Espérance des carrés moyens

On peut montrer que

$$E(\text{MSE}) = \sigma^2 \quad \text{et} \quad E(\text{MST}_R) = \sigma^2 + \frac{\sum_{i=1}^I n_i \tau_i^2}{I - 1}, \quad \text{où } \tau_i = \mu_i - \mu.$$

- MSE estime toujours σ^2 , que H_0 soit vraie ou non.
- Sous H_0 : tous les $\tau_i = 0 \Rightarrow E(\text{MST}_R) = \sigma^2 \Rightarrow F = \text{MST}_R/\text{MSE}$ devrait être **proche de 1**.
- Sous H_1 : au moins un $\tau_i \neq 0 \Rightarrow E(\text{MST}_R) > \sigma^2 \Rightarrow F$ tend à être **plus grand que 1**.
- On rejette donc H_0 seulement pour les **grandes** valeurs de F .

Statistique de Fisher

Théorème.Si

- les observations sont **indépendantes** (dans les groupes et entre les groupes),
- les données proviennent d'une distribution **normale**,
- la **variance est égale** dans chaque groupe,

et que H_0 est vraie (μ_i toutes égales), alors

$$F = \frac{SST_R/(I - 1)}{SSE/(N - I)} = \frac{MST_R}{MSE} \sim F_{I-1, N-I}.$$

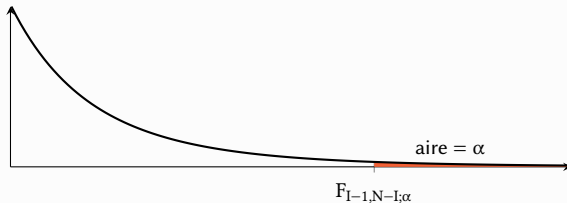
Règle de décision

Test F de l'ANOVA

On rejette H_0 au seuil α si

$$F_{obs} > F_{I-1, N-I; \alpha}.$$

Dans ce cas, on a une valeur-p inférieure à α .



Retour au scénario 1

	y1	Traitement
1	14.72	A
2	13.52	A
3	14.54	A
4	18.49	A
5	9.98	B
6	4.59	B
7	12.82	B
8	11.03	B

	Moyenne	Écart-type
A	15,32	2,18
B	9,61	3,54

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Traitement	1	65.27	65.27	7.543	0.0334 *
Residuals	6	51.91	8.65		

Retour au scénario 1

	y1	Traitement
1	14.72	A
2	13.52	A
3	14.54	A
4	18.49	A
5	9.98	B
6	4.59	B
7	12.82	B
8	11.03	B

	Moyenne	Écart-type
A	15,32	2,18
B	9,61	3,54

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Traitement	1	65.27	65.27	7.543	0.0334 *
Residuals	6	51.91	8.65		

Vérifier : $SST_R = 4(15,32 - 12,46)^2 + 4(9,61 - 12,46)^2 \approx 65,27$ et
 $SSE = 3(2,18^2) + 3(3,54^2) \approx 51,9$.

QUIZ!

Retour aux 4 engrais appliqués à 24 plants de tomate (6 par engrais). On obtient $SST_R = 84,0$ et $SST = 224,0$. Complétez la table d'ANOVA.

Source	Somme de carrés	d.l.	Carré moyen	F_{obs}
Engrais	84,0			
Erreur				
Total	224,0			

- Sachant que $F_{3, 20; 0,05} = 3,10$, que concluez-vous au seuil de 5 % ?
- Combien vaut l'estimation de σ^2 ? Et celle de σ ?

Aurait-on pu faire un test de Student ?

ANOVA :

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Traitement	1	65.27	65.27	7.543	0.0334 *
Residuals	6	51.91	8.65		

Test de Student :

Two Sample t-test
data: y1 by Traitement
t = 2.7464, df = 6, p-value = 0.03345
alternative hypothesis: true difference in means between group A and group B
is not equal to 0
95 percent confidence interval:
0.6230128 10.8019872
sample estimates:
mean in group A mean in group B
15.3175 9.6050

Student vs ANOVA

- Est-ce un hasard si les valeurs-p sont les mêmes ?
- Pourquoi faire l'ANOVA ?

Exemple : agent de conservation naturel pour la viande

Contexte : agent de conservation à partir de sang de porc ([ULaval Nouvelles](#)).

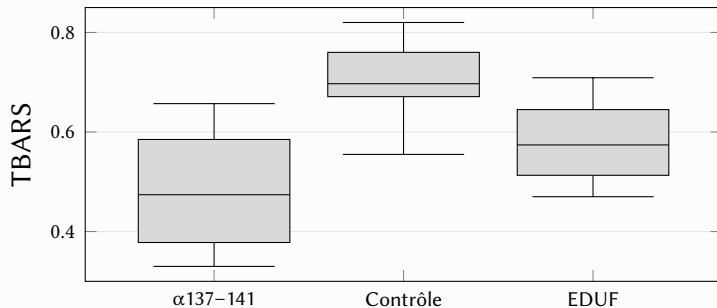
- En 2016, environ 30 millions de litres de sang bovin et porcin ont été produits au Québec.
- Une nouvelle méthode d'extraction (EDUF) permet une plus grande concentration de peptides bioactifs.
- Selon les chercheurs, ces extraits de sang se sont montrés aussi efficaces que le BHT, un additif chimique couramment utilisé par l'industrie alimentaire, lors de tests sur du bœuf haché.
- Nous avons **simulé** des données à partir des moyennes et écarts-types de la section 3.6 de l'article.

Contexte

- On veut comparer l'effet antioxydant de l'**EDUF** (la nouvelle méthode) au peptide original nommé **α 137-141** et à un **contrôle**.
- On crée un échantillon de taille 6 pour chaque traitement.
- On mesure le **TBARS** (en mg MDA/kg de viande hachée) après 4 jours (une mesure de rancissement).

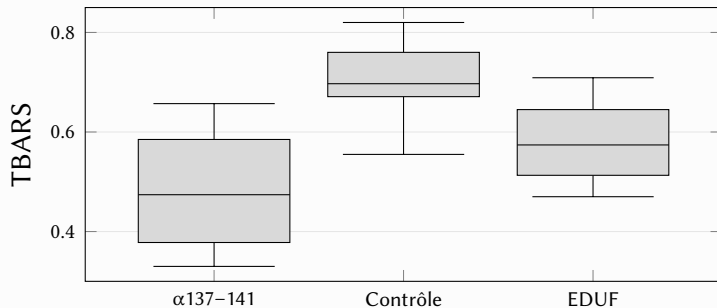
Premier réflexe : regarder les données

Avant toute analyse, on **visualise** les données de chaque groupe (diagrammes en boîte ou diagrammes en points).



Premier réflexe : regarder les données

Avant toute analyse, on **visualise** les données de chaque groupe (diagrammes en boîte ou diagrammes en points).



- Les TBARS moyens semblent-ils différer d'un traitement à l'autre ?
- La dispersion est-elle semblable dans les trois groupes ? Y a-t-il des valeurs extrêmes ?

Test ANOVA

1. Hypothèses

H_0 : les trois moyennes sont égales

H_1 : au moins une moyenne est différente d'une autre

Test ANOVA

1. Hypothèses

H_0 : les trois moyennes sont égales

H_1 : au moins une moyenne est différente d'une autre

2. Choix du test et définition de la zone de rejet

Si les données sont distribuées normalement et que les variances sont égales, on fait une analyse de variance à un facteur.

On choisit α . On rejette H_0 si la valeur-p est inférieure à α . Choisissons arbitrairement $\alpha = 0,05$.

Table d'ANOVA à compléter

3. On fait le test avec un logiciel

On obtient (en partie) la table suivante. Complétez-la.

Source de variation	Somme de carrés	d.l.	Carrés moyens	F_{obs}
Traitements (inter-trt)				
Erreur (intra-trt)			0,0104	
Total	0,2971			

Sortie R

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Traitement	2	0.1412	0.0706	6.791	0.00794	**
Residuals	15	0.1559	0.0104			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

4. Décision

5. Conclusion

QUIZ!

Lorsque H_0 est rejetée dans une ANOVA à un facteur, lesquelles des conclusions suivantes sont correctes ?

- Toutes les moyennes sont égales.
- Toutes les moyennes sont différentes.
- Au moins deux moyennes sont différentes.
- La variable réponse a un effet sur le facteur.
- Le facteur a un effet sur la variable réponse.
- Le contrôle a un TBARS moyen plus élevé que l'EDUF.

Contrastes

- Comment savoir **quelles** moyennes sont différentes? (section 5.8)
- Peut-on simplement faire plusieurs tests de Student 2 à 2?

Le problème des comparaisons multiples

Chaque test de Student au seuil $\alpha = 5\%$ a 5 % de chances de conclure à tort à une différence. Plus on fait de tests, plus le risque qu'**au moins un** d'entre eux donne une fausse alerte augmente (comme acheter plusieurs billets de loterie!).

Groupes (I)	Comparaisons 2 à 2 (m)	P(au moins une fausse alerte) $\approx 1 - 0,95^m$
3	3	0,14
4	6	0,26
5	10	0,40
10	45	0,90

(Approximation qui suppose les tests indépendants.)

Le problème des comparaisons multiples

Chaque test de Student au seuil $\alpha = 5\%$ a 5 % de chances de conclure à tort à une différence. Plus on fait de tests, plus le risque qu'**au moins un** d'entre eux donne une fausse alerte augmente (comme acheter plusieurs billets de loterie!).

Groupes (I)	Comparaisons 2 à 2 (m)	P(au moins une fausse alerte) $\approx 1 - 0,95^m$
3	3	0,14
4	6	0,26
5	10	0,40
10	45	0,90

(Approximation qui suppose les tests indépendants.)

Méthode de Tukey

Compare toutes les paires de moyennes en ajustant les valeurs-p (et les intervalles de confiance) pour que le **taux d'erreur global** (probabilité d'au moins une fausse alerte) soit égal à α .

Comparaisons multiples de Tukey

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

```
Fit: aov(formula = TBARS ~ Traitement, data = conservation)
```

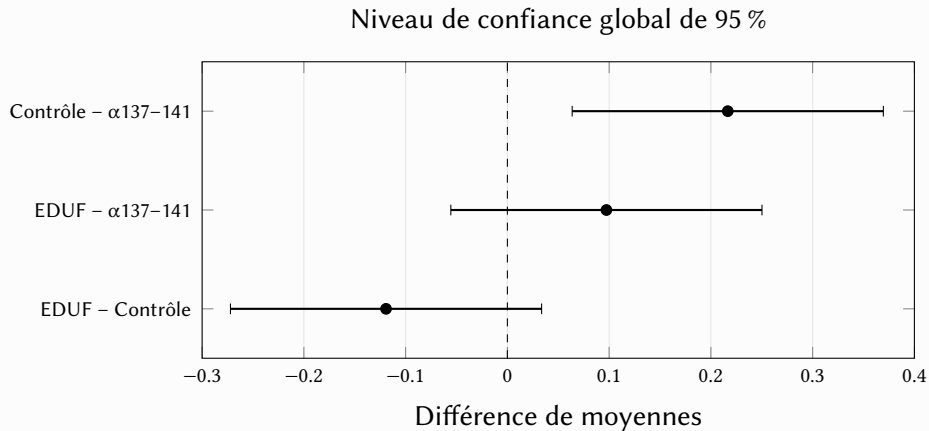
Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)	
Controle - a137_141 == 0	0.21658	0.05887	3.679	0.00599	**
EDUF - a137_141 == 0	0.09732	0.05887	1.653	0.25496	
EDUF - Controle == 0	-0.11926	0.05887	-2.026	0.14031	

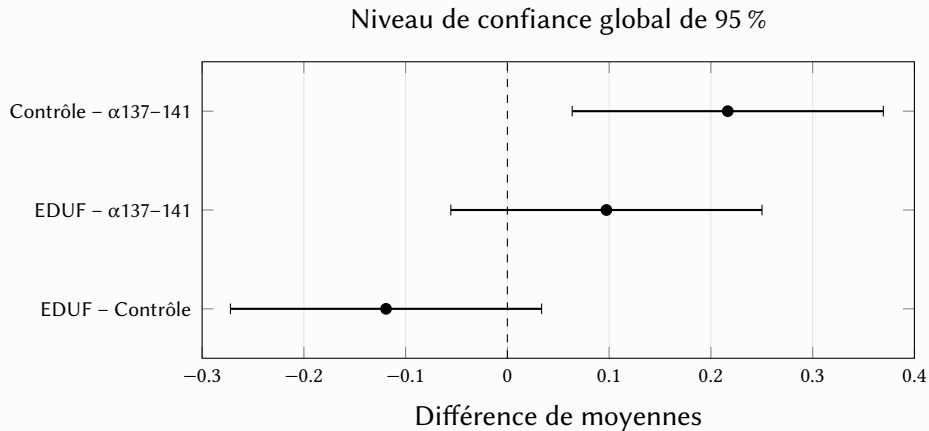
Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Adjusted p values reported -- single-step method)

Intervalles de confiance simultanés (Tukey)



Intervalles de confiance simultanés (Tukey)



Seul l'intervalle Contrôle - α 137-141 ne contient pas 0.

Vérification des postulats

Indépendance des observations : structure de l'expérience

- Y a-t-il des liens entre les observations d'un même traitement ? De traitements différents ?
- A-t-on pris plusieurs mesures sur une même unité expérimentale ?

Homogénéité des variances : test de Bartlett

On veut tester les hypothèses

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2$$

H_1 : au moins 2 de ces variances sont inégales

On utilise une statistique B (appelée K-squared dans R) :

Homogénéité des variances : test de Bartlett

On veut tester les hypothèses

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2$$

H_1 : au moins 2 de ces variances sont inégales

On utilise une statistique B (appelée K-squared dans R) :

- si H_0 est vraie, B suit une loi χ^2_{I-1} ;

Homogénéité des variances : test de Bartlett

On veut tester les hypothèses

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2$$

H_1 : au moins 2 de ces variances sont inégales

On utilise une statistique B (appelée K-squared dans R) :

- si H_0 est vraie, B suit une loi χ_{I-1}^2 ;
- si H_1 est vraie, B prendra des valeurs plus élevées que celles attendues sous la χ_{I-1}^2 ;

Homogénéité des variances : test de Bartlett

On veut tester les hypothèses

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_I^2$$

H_1 : au moins 2 de ces variances sont inégales

On utilise une statistique B (appelée K-squared dans R) :

- si H_0 est vraie, B suit une loi χ_{I-1}^2 ;
- si H_1 est vraie, B prendra des valeurs plus élevées que celles attendues sous la χ_{I-1}^2 ;
- alors, au seuil α , on rejette l'égalité des variances si...

Example

Bartlett test of homogeneity of variances

data: TBARS by Traitement

Bartlett's K-squared = 0.74383, df = 2, p-value = 0.6894

Exemple

Bartlett test of homogeneity of variances

data: TBARS by Traitement

Bartlett's K-squared = 0.74383, df = 2, p-value = 0.6894

Conclusion?

Normalité des observations : analyse des résidus

Rappel du modèle :

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

Puisque $Y_{ij} \sim N(\mu_i, \sigma^2)$, il suit que $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Résidus

Les erreurs échantillonnales ε_{ij} sont estimées par les **résidus** du modèle :

$$\hat{\varepsilon}_{ij} = y_{ij} - \bar{y}_{i\cdot}$$

Au lieu d'analyser directement les observations, nous analyserons les résidus.

Test de normalité sur les résidus

Shapiro-Wilk normality test

```
data:  AnovaModel.3$residuals  
W = 0.97151, p-value = 0.8255
```

Test de normalité sur les résidus

Shapiro-Wilk normality test

```
data:  AnovaModel.3$residuals  
W = 0.97151, p-value = 0.8255
```

Conclusion?

Réponses aux exemples

- **Quiz (tomates)** : variable réponse = hauteur des plants (cm) après 6 semaines; facteur = engrais, à 4 niveaux (A, B, C, D); unité expérimentale = un plant (pot); c'est une **expérience** (engrais attribués au hasard) : devis complètement aléatoire avec $n_i = 6$.
- **Quiz (truites)** : le traitement (température) est appliqué au **bassin** : l'unité expérimentale est le bassin, et il n'y a que **2 répétitions** par température (et non 30). Les truites d'un même bassin ne sont pas indépendantes.
- **Quiz (souris sylvestres)** : étude observationnelle; on ne peut pas conclure à un lien de cause à effet (l'habitat est confondu avec la nourriture disponible, les prédateurs, la densité de population...).
- **Modèle** : μ_i est la moyenne de la population i ; $\varepsilon_{ij} = Y_{ij} - \mu_i$ est l'erreur aléatoire, $\varepsilon_{ij} \sim N(0, \sigma^2)$. On estime μ_i par $\bar{y}_{i\cdot}$, μ par $\bar{y}_{\cdot\cdot}$ et σ^2 par MSE.
- **Quiz (même variance)** : l'expérience 2 a le plus grand SST_R (288 contre 32) et le plus grand SST (312,5 contre 56,5); les SSE sont égales (24,5). Scénarios 1 et 2 : SST_R presque égales (65,27 et 65,15), mais SSE (51,91 contre 263,45) et SST (117,18 contre 328,60) plus grandes dans le scénario 2.
- **Quiz (table des engrais)** : $SSE = 224 - 84 = 140$; d.l. 3, 20 et 23; $MST_R = 28$, $MSE = 7$; $F_{obs} = 4,00 > 3,10$: on rejette H_0 (valeur-p = 0,022); au moins deux engrais mènent à des hauteurs moyennes différentes. $\hat{\sigma}^2 = MSE = 7$, $\hat{\sigma} = \sqrt{7} = 2,65$ cm.

Réponses aux exemples (suite)

- **Student vs ANOVA** : avec deux groupes, $F_{obs} = t_{obs}^2$ ($2,7464^2 \approx 7,543$) et $F_{1,v} = t_v^2$: les deux tests sont équivalents. L'ANOVA permet de comparer plus de deux groupes en un seul test.
- **Regarder les données (TBARS)** : le contrôle semble avoir un TBARS plus élevé que les deux autres traitements ; les dispersions sont semblables et on ne voit pas de valeur extrême.
- **Table à compléter** : $SST_R = 0,1412$, $SSE = 0,1559$, d.l. 2, 15, 17 ; $MST_R = 0,0706$; $F_{obs} = 6,79$.
- **Décision** : valeur-p = 0,008 < 0,05, on rejette H_0 . **Conclusion** : au seuil de 5 %, au moins un traitement a un TBARS moyen différent des autres.
- **Quiz (H_0 rejetée)** : seules « Au moins deux moyennes sont différentes » et « Le facteur a un effet sur la variable réponse » sont correctes. Le test F ne dit pas **quelles** moyennes diffèrent (d'ailleurs, Tukey ne détecte pas de différence significative entre le contrôle et l'EDUF, $p = 0,14$).
- **Plusieurs tests 2 à 2** : le risque global d'erreur de type I augmente avec le nombre de comparaisons ; d'où les comparaisons multiples (Tukey).
- **Bartlett** : valeur-p = 0,69, on ne rejette pas l'égalité des variances. On rejette si $B > \chi_{1-1;\alpha}^2$.
- **Shapiro-Wilk** : valeur-p = 0,83, pas d'évidence contre la normalité.