

## Stage Laboratoire E3S

Examen des conditions qui augmentent ou diminuent les différences entre les sexes dans l'engagement dans l'activité physique pendant les cours d'éducation physique, dans le cadre écologique.

Kossi ABOTSI

August 15, 2024

# Table des matières

- 1 Introduction
- 2 Modèle
- 3 Validation des hypothèses
- 4 Anova de type II
- 5 Résultats et discussion
- 6 Limites des résultats et Conclusion

# Table des matières

- 1 Introduction
- 2 Modèle
- 3 Validation des hypothèses
- 4 Anova de type II
- 5 Résultats et discussion
- 6 Limites des résultats et Conclusion

# Introduction


- Les filles obtiennent en moyenne des résultats inférieurs à ceux des garçons lors des évaluations en éducation physique et sportive (EPS).
- L'engagement des filles dans l'activité physique à l'adolescence est un enjeu crucial pour leur santé et bien-être, associé aux questions d'équité entre les sexes.

## Objectif du stage :

Analyser les différences d'engagement physique entre filles et garçons pendant une leçon d'EPS de 2 heures, en tenant compte de la nature de l'activité (CA), de la catégorie socioculturelle de l'établissement (IPS) et du milieu géographique.

# Questions à répondre

- Examen des écarts de niveau d'engagement en EPS entre filles et garçons selon les champs d'apprentissage (CA1, CA2, CA3, CA4).
- Examen des écarts de niveau d'engagement en EPS entre filles et garçons selon la catégorie d'IPS du collège (élevé, moyen, faible).
- Examen des écarts de niveau d'engagement en EPS entre filles et garçons selon le milieu géographique (urbain, rural).

 Dans cette présentation, nous n'aborderons que l'examen des écarts de niveaux d'engagement en EPS entre filles et garçons selon les champs d'apprentissage.

# Table des matières

- 1 Introduction
- 2 Modèle**
- 3 Validation des hypothèses
- 4 Anova de type II
- 5 Résultats et discussion
- 6 Limites des résultats et Conclusion

# Présentation des données

- Les participants âgés de 11 à 15 ans ont rendu les autorisations parentales et ont accepté de participer à l'étude.
- MVPA mesuré chez les filles et garçons au cours d'une leçon d'EPS de 2 heures dans 4 champs d'apprentissage différents (1, 2, 3 et 4), dans les classes de 5ème à 3ème, dans différentes écoles en Alsace et Île-de-France.

A tibble: 462 x 7

classe <chr>	mvpa <dbl>	ecart_MVPA <dbl>	gender <chr>	cat_IPS <chr>	CA <dbl>	age <int>
3.P	23.833333	-5.282051...	F	moyen	3	15
3.P	32.666667	3.55128205	M	moyen	3	14
3.P	27.333333	-1.782051...	M	moyen	3	15
3.P	21.500000	-7.615384...	F	moyen	3	15
3.P	30.500000	1.38461538	M	moyen	3	15
3.P	28.000000	-1.115384...	F	moyen	3	15
3.P	28.500000	-0.615384...	M	moyen	3	15
3.P	32.500000	3.38461538	M	moyen	3	15
3.P	50.000000	20.88461...	M	moyen	3	15
3.P	28.333333	-0.782051...	F	moyen	3	15

1-10 of 462 rows

Previous 1 2 3 4 5 6 ... 47 Next

Figure: Extraits des données

# ANOVA à deux facteurs (1/2)

**⚠** Nous voulons étudier l'effet de deux facteurs sur la variable réponse. Le modèle statistique utilisée est donc une ANOVA à deux facteurs avec interaction.

## ● Les facteurs

- ① Champs d'apprentissage (CA) à 4 modalités (1, 2, 3 et 4)
  - **CA1:** sports de performance (athlétisme, natation, cyclisme, ...)
  - **CA2:** sports de plein air et d'aventure (randonnée, alpinisme, surf, ...)
  - **CA3:** activité artistique (gymnastique rythmique, patinage artistique, ...)
  - **CA4:** sports d'opposition (sport de combat, tennis, ...)
- ② Genre à deux modalités (F ou M)

## ● Choix de la variable dépendante :

Pour analyser les différence d'engagement nous nous basons sur l'écart de MVPA à la moyenne de MVPA de chaque classe.

- ① Cela ajuste les différences de durée d'activité et d'enseignants entre les classes.



# ANOVA à deux facteurs (2/2)

Modèle d'ANOVA à deux facteurs (genre et champ d'apprentissage) avec comme variable dépendante l'écart de MVPA par rapport à la moyenne de MVPA de chaque classe.

Le modèle s'écrit :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}, \quad i \in \{1, 2\}; j \in \{1, 2, 3, 4\}, \quad \epsilon_{ijk} \stackrel{\text{i.i.d}}{\sim} \mathcal{N}(0, \sigma^2) \quad (1)$$

où

- $\mu$  est l'effet moyen général,
- $\alpha_i$  représente l'effet principal du genre,
- $\beta_j$  représente l'effet principal du champ d'apprentissage,
- $\gamma_{ij}$  est le terme d'interaction,
- $\epsilon_{ijk}$  sont les résidus.
- $k$  l'indice de répétition pour le couple  $(i, j)$

# Identifiabilité du modèle

- Le modèle n'est pas identifiable car pour tout  $(1 + I + J + IJ)$ -uplet  $(\mu, \alpha_1, \dots, \alpha_I, \beta_1, \dots, \beta_J, \gamma_{11}, \dots, \gamma_{ij}, \dots, \gamma_{IJ})^\top$  et pour tout  $a \in \mathbb{R}$ , le  $(1 + I + J + IJ)$ -uplet  $(\mu - a, \alpha_1 + \frac{a}{3}, \dots, \alpha_I + \frac{a}{3}, \beta_1 + \frac{a}{3}, \dots, \beta_J + \frac{a}{3}, \gamma_{11} + \frac{a}{3}, \dots, \gamma_{ij} + \frac{a}{3}, \dots, \gamma_{IJ} + \frac{a}{3})^\top$  correspond au même modèle.
- On utilise la contrainte de somme :

$$\sum_i \alpha_i = 0; \quad \sum_j \beta_j = 0; \quad \forall i, \quad \sum_j \gamma_{ij} = 0; \quad \forall j, \quad \sum_i \gamma_{ij} = 0$$

# Table des matières

- 1 Introduction
- 2 Modèle
- 3 Validation des hypothèses**
- 4 Anova de type II
- 5 Résultats et discussion
- 6 Limites des résultats et Conclusion

# Hypothèses à vérifier par le modèle

- Indépendance des observations (ou des résidus).
- Égalité des variances des résidus du modèle (homoscédasticité).
- Normalité des résidus du modèle.

# Indépendance des observations

L'hypothèse d'indépendance des observations est vérifiée. La liaison potentielle due à l'appartenance à une même classe ou école est éliminée en utilisant l'écart à la moyenne du MVPA pour chaque classe.

# Égalité des variances des résidus du modèle (homoscédasticité)

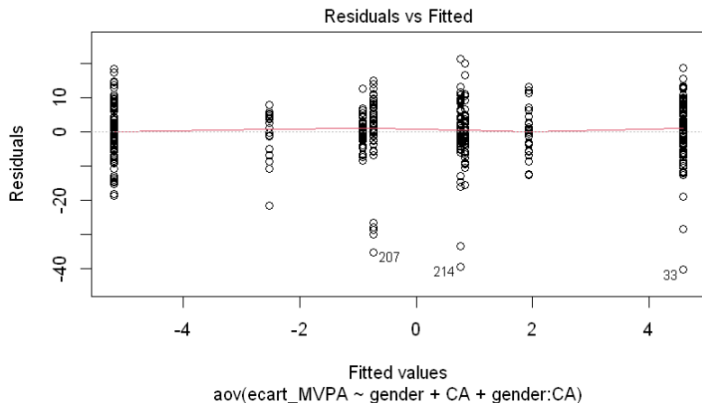


Figure: Graphique de diagnostic de l'homoscédasticité

# Normalité des résidus du modèle

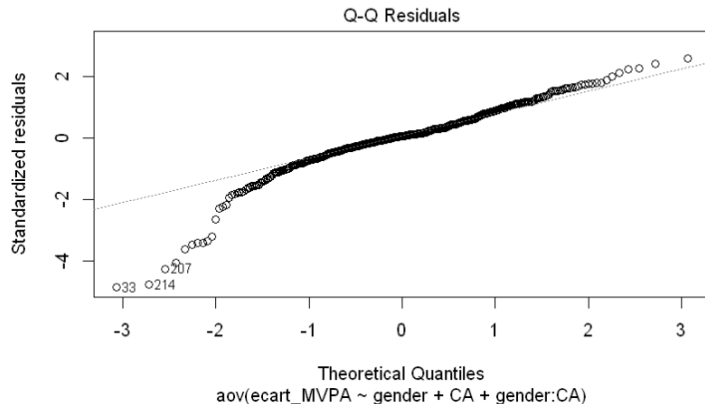


Figure: Q-Q plot

Petit commentaire...

# Table des matières

- 1 Introduction
- 2 Modèle
- 3 Validation des hypothèses
- 4 Anova de type II**
- 5 Résultats et discussion
- 6 Limites des résultats et Conclusion



# Type d'ANOVA (1/4)

- Il est courant de tester la nullité des paramètres pour chaque niveau d'un facteur pour évaluer son effet.
- Tester uniquement la nullité des paramètres n'est pas suffisant pour définir correctement l'hypothèse nulle ; il faut aussi considérer les hypothèses sur les autres facteurs et interactions.
- Dans un plan équilibré, le test de significativité d'un facteur reste le même, indépendamment des hypothèses sur les autres variables.
- Dans un plan déséquilibré, le test de significativité d'un facteur peut varier en fonction des hypothèses sur les autres facteurs et interactions.
- Pour spécifier correctement le rôle des autres facteurs et interactions dans un plan déséquilibré, on utilise la notion de réduction.

## Définition de la réduction

Soit un modèle contenant les effets  $(a_1, \dots, a_I)$  des facteurs  $(X_1, X_2, \dots, X_I)$ . On appelle réduction associée à l'introduction de  $a_{q_1}, \dots, a_{q_d}$  dans un modèle contenant les effets  $a_{i_1}, \dots, a_{i_m}$ , notée  $R(a_{q_1}, \dots, a_{q_d} | \mu, a_{i_1}, \dots, a_{i_m})$  la norme suivante :

$$R(a_{q_1}, \dots, a_{q_d} | \mu, a_{i_1}, \dots, a_{i_m}) = SCE_{i_1, i_2, \dots, i_m, q_1, q_2, \dots, q_d} - SCE_{i_1, i_2, \dots, i_m}, \quad (2)$$

avec  $SCE_{i_1, i_2}$  la somme des carrés expliquée par le modèle associée aux facteurs  $X_{i_1}, X_{i_2}$ .

Un test de l'effet d'un facteur est associé à une réduction donnée.

## Les hypothèses associés à la réduction $R(\gamma|\mu, \alpha, \beta)$ du modèle $M_1$

- $H_0$  : Il n'y a pas de différence significative entre le modèle ne contenant pas l'effet du facteur d'interaction entre genre et CA et le modèle complet.
- $H_1$  : Il y a une différence significative entre le modèle ne contenant pas l'effet du facteur d'interaction entre genre et CA et le modèle complet.

## ANOVA de type II

- La réduction de type II dans une ANOVA à deux facteurs évalue l'effet d'un facteur ou d'une interaction en tenant compte des autres facteurs principaux, sans dépendre de l'ordre d'introduction.
- Ce type de réduction considère que l'effet d'un facteur principal ne peut pas être supprimé si une interaction est présente dans le modèle.

# Type d'ANOVA (4/4)

**Table:** Table d'analyse de la variance des réductions de type II du modèle  $M_1$ .

Effet	Réduction type II	DDL	F	Loi de $F$ sous $H_0$	Question
$\alpha$	$R(\alpha \beta, \mu)$	$I - 1$	$\frac{\frac{R(\alpha \beta, \mu)}{I-1}}{\frac{SCR}{n-IJ}}$	$\mathcal{F}_{I-1, n-IJ}$	Est-il pertinent d'ajouter l'effet du facteur genre à un modèle contenant la constante et l'effet du facteur CA ?
$\beta$	$R(\beta \mu, \alpha)$	$J - 1$	$\frac{\frac{R(\beta \mu, \alpha)}{J-1}}{\frac{SCR}{n-IJ}}$	$\mathcal{F}_{J-1, n-IJ}$	Est-il pertinent d'ajouter l'effet du facteur CA à un modèle contenant la constante et l'effet du facteur genre ?
$\gamma$	$R(\gamma \mu, \alpha, \beta)$	$(I - 1) \times (J - 1)$	$\frac{\frac{R(\gamma \mu, \alpha, \beta)}{(I-1) \times (J-1)}}{\frac{SCR}{n-IJ}}$	$\mathcal{F}_{(I-1) \times (J-1), n-IJ}$	Est-il pertinent d'ajouter l'effet de l'interaction entre les deux facteurs à un modèle contenant la constante et les effets des deux facteurs ?

# Table des matières

- 1 Introduction
- 2 Modèle
- 3 Validation des hypothèses
- 4 Anova de type II
- 5 Résultats et discussion**
- 6 Limites des résultats et Conclusion

# Tableau de contingence

Table: Effectifs des CA par genre

	CA 1	CA 2	CA 3	CA 4
F	20	53	49	98
M	26	51	54	111

- Plan non équilibré
- ANOVA de type II

# Test de validité du modèle complet

On teste :

- $H_0 : \alpha_i = 0 \text{ et } \beta_j = 0 \text{ et } \gamma_{ij} = 0 \quad \forall i, j$
- $H_1 : \alpha_i \neq 0 \text{ ou } \beta_j \neq 0 \text{ ou } \gamma_{ij} \neq 0 \quad \forall i, j$

**Table:** Tableau d'analyse de variance pour le modèle  $M_1$

Res.Df	RSS	Df	Sum of Sq	F value	Pr(> F)
461	36916				
454	31593	7	5323.2	10.928	$8.944 \times 10^{-13}$

# Test de validité de sous modèle

**Table:** Effets des différents facteurs (type II) dans le modèle M1

Source	Sum Sq	Df	F value	Pr(> F)
genre	3585.2	1	51.5200	$2.911 \times 10^{-12}$
CA	6.4	3	0.0307	0.9927
genre:CA	1738.0	3	8.3252	$2.122 \times 10^{-5}$
Résidus	31592.8	454		

## Effet de l'interaction entre genre et CA :

- $H_0 : Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk} \quad \forall i, j, k$
- $H_1 : Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad \forall i, j, k$

Nous n'accorderons donc plus trop d'importance aux tests sur les facteurs principaux, car les deux facteurs principaux exercent leur influence par le biais de leur interaction



# Test de comparaison de la moyenne des écarts à la moyenne de MVPA de chaque classe des filles et garçons selon le CA (Test post hoc)

Nous voulons tester les hypothèses :

- $H_0 : \theta_{i_1j} = \theta_{i_2j}$
- $H_1 : \theta_{i_1j} \neq \theta_{i_2j}$

où  $\theta_{i_1j_1}$  et  $\theta_{i_2j_2}$  sont les moyennes dans chaque groupe différent et  $i_1, i_2 \in \{1, 2\}$  avec  $i_1 \neq i_2$  ;  $j \in \{1, \dots, 4\}$ .

Table: Test post hoc (Tukey-Cramer)

	CA1	CA2	CA3	CA4
Écart moyen entre filles et garçons des écarts à la moyenne de MVPA de chaque classe	-4.4548	-1.4898	-1.7470	-9.7647
Intervalle de confiance	-12.010 à 3.100	-6.4730 à 3.493	-6.759 à 3.265	-13.286 à -6.244
P-valeur	0.6235	0.9850	0.9600	0.0000

# Taille d'effet ( $\omega^2$ )

**Table:** Tailles d'effet ( $\omega^2$ ) et intervalles de confiance à 95% (unilatéraux)

Paramètre	$\omega^2$	95% CI
Genre	0.10	[0.06, 1.00]
CA	0.00	[0.00, 1.00]
Genre:CA	0.05	[0.02, 1.00]

- L'interaction entre le genre et le CA explique 5 % de la variance totale, ce qui représente une taille d'effet modérée
- En EPS cette taille d'effet indique que les différences entre les filles et les garçons dans le champs 4 doivent être prise en compte.

# Table des matières

- 1 Introduction
- 2 Modèle
- 3 Validation des hypothèses
- 4 Anova de type II
- 5 Résultats et discussion
- 6 Limites des résultats et Conclusion**

# Limites des résultats

- Le recrutement des participants parmi ceux ayant donné leur autorisation et exprimé un intérêt pourrait introduire un biais si les élèves plus motivés ou mieux équipés sont surreprésentés.
- Les plans non équilibrés compliquent le calcul et l'interprétation des tailles d'effet, car une partie de la variabilité est perdue, rendant les tailles d'effet moins précises mais toujours indicatives de l'importance des facteurs.
- Les résultats de l'étude sont basés sur des données recueillies en Alsace et en Île-de-France la généralisation de ces résultats à d'autres régions de France doit être effectuée avec prudence.
- L'examen du niveau d'engagement physique en EPS basé sur l'écart de MVPA par rapport à la moyenne de MVPA de chaque classe ne permet pas de prendre en compte les effets aléatoires liés aux collègues et aux classes

# Conclusion et perspective

- Les garçons sont généralement plus impliqués que les filles.
- Une différence significative est observée uniquement dans le champ 4 (activités d'opposition).
- Les autres champs (sports de performance, sports de plein air, activités artistiques) ne montrent pas de différences significatives entre les sexes.
- Analyse du niveau d'engagement en EPS basé sur le MVPA des participants avec prise en compte des effets aléatoires des collèges et des classes appartenant à ces collèges via un modèle mixte.