

# HMMA 307 : Modèles linéaires avancés

Robustesse des modèles linéaires à effets mixtes face aux  
violations des hypothèses de distribution

**GAIZI IBRAHIM**

<https://github.com/ibrahimgaizi1>

Université de Montpellier



Simulation des données :

Les différents scénarios :

Violation des hypothèses de distribution :

Analyse des résultats :

# Introduction :

- ▶ Outils puissants pour analyser des ensembles de données complexes avec des observations répétées ou groupées.
- ▶ Impliquent des procédures d'ajustement sur la distribution des effets résiduels et aléatoires.
- ▶ les hypothèses de distribution des effets aléatoires ne peuvent pas être vérifiées aussi facilement que pour les effets fixes

# Les modèles à effets linéaires mixtes

## Modèle de base simulé

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \hat{\alpha} + \hat{\epsilon}$$

$$\hat{\alpha} \sim N(0, \hat{\sigma}_{\hat{\alpha}}^2)$$

$$\hat{\epsilon} \sim N(0, \hat{\sigma}_{\hat{\epsilon}}^2)$$

$\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\alpha}, \hat{\epsilon}$  sont respectivement les approximations de :  
 $\beta_0, \beta_1, \beta_2, \alpha, \epsilon$ .

- ▶ Simuler des données et à les ajuster dans un simple LMM
- ▶ Enfreindre les hypothèses sur les effets aléatoires et les distributions d'erreurs selon différents scénarios .

# Comparaison entre OLS RLM :

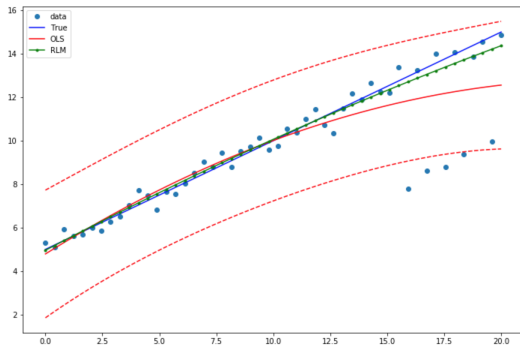


Figure: Comparaison entre les estimations avec les methodes RLM et OLS.

# Scénarios A et B

## ***Distributions biaisées (scénario A) :***

- ▶  $\alpha$  et  $\epsilon$  ont été tirés des distributions biaisées.

## ***Distributions bimodales (scénario B) :***

- ▶ Les distributions ont été tirées de deux distributions normales distinctes.
- ▶ Les distributions ont des moyennes déplacées de  $\pm 1,5$  unités
- ▶ La moitié des tirages est déplacée vers le bas et la moitié vers le haut.
- ▶ La variance a été réajustée

## Scénarios C et D et E

### ***Distributions hétéroscédastiques (scénario C) :***

- ▶  $\alpha$  et  $\epsilon$  sont tirées de distributions où la variance dépend d'une des covariables ( $x_1$ )
- ▶  $\lambda$  prend les valeurs 2, 4 ou 8.

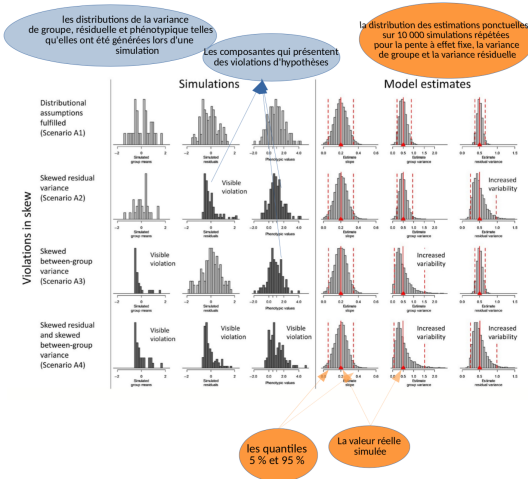
### ***Effets aléatoires manquants (scénario D) :***

- ▶ Effets aléatoires échantillonnés de manière déséquilibrée,
- ▶ Variances à effet aléatoire fixées.

### ***Prédicteurs corrélés (scénarios E) :***

- ▶ Prédicteurs corrélés avec des corrélations fixées à +0,2, +0,5 et +0,8

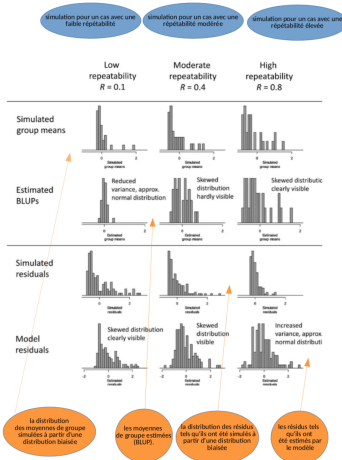
# Violation en biais



**Figure:** les effets des violations des hypothèses de distribution sur les estimations des paramètres d'intérêt majeur.

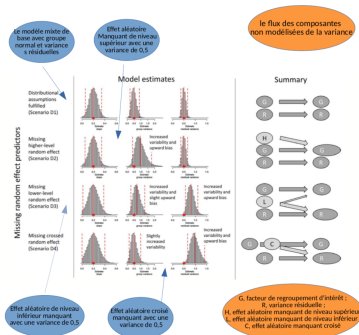


# Effets des distributions biaisées



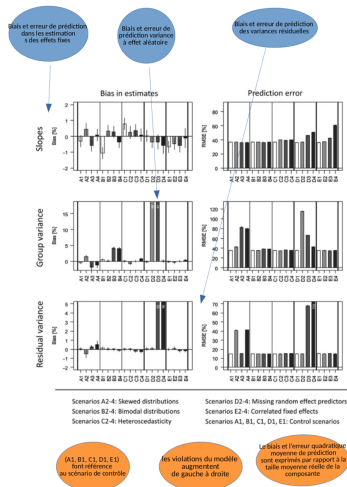
**Figure:** Effets des distributions biaisées et de la taille des composantes de la variance sur les moyennes et les résidus estimés des groupes.

# Impactes des effets aléatoires manquants



**Figure:** Effets des distributions biaisées et de la taille des composantes de la variance sur les moyennes et les résidus estimés des groupes.

# Biais et prédiction



**Figure:** Effets des distributions biaisées et de la taille des composantes de la variance sur les moyennes et les résidus estimés des groupes.

## Analyse :

- ▶ L'effet des violations des hypothèses de distribution des variances à effet aléatoire et des résidus est faible.
- ▶ Faible biais global sauf pour les distribution bimodale.
- ▶ Certaines violations simulées peuvent émerger de modèles incomplets .
- ▶ Les termes d'effet aléatoire manquants ont des effets systématiques sur les estimations d'autres composantes de la variance.
- ▶ Il existe une imprécision accrue des estimations d'effets fixes lorsque les covariables sont corrélées.
- ▶ Faible impact de l'hétéroscédasticité sur les estimations du modèle

## Conclusion :

- ▶ Modèles à effets mixtes sont largement robustes
- ▶ Ils ont un faible biais global sauf pour les distributions bimodales.
- ▶ Les violations hypothèses peuvent parfois entraîner une variabilité accrue des estimations.
- ▶ Outils puissants permettant de modéliser une grande variété d'ensembles de données.

# Bibliography

- ▶ Salmon, Joseph. *Modèle linéaire avancé : Anova*. 2019. URL: <http://josephsalmon.eu/enseignement/Montpellier/HMMA307/Anova.pdf>.
- ▶ Schielzeth, Holger. *Robustness of linear mixed-effects models to violations of distributional assumptions*. 2020. URL: <https://besjournals.onlinelibrary.wiley.com/doi/epdf/10.1111/2041-210X.13434>.