M1 MIASHS

Données de panels

TP 1

---

1 Modèle de régression linéaire

On souhaite simuler des pseudo-données selon le modèle :

$$y_i = \beta_0 + x_i\beta_1 + \varepsilon_i \quad i = 1, ...100$$

- Simuler :
  - $\beta_0, \beta_1 \sim \mathcal{U}nif[-3;3]$
  - $x_i \sim \mathcal{U}nif[0;1]$
  - $\varepsilon_i \sim \mathcal{N}(0, 2)$

  pour obtenir les valeurs de y_i.
- Que vaut V (matrice de variance-covariance de Y dans ce modèle) ? V^{-1} ?
- Donner l'expression de $\hat{\beta}_{OLS}$ et $\hat{\beta}_{GLS}$. Calculer leurs valeurs numériquement dans R.
- Obtenir la valeur de la matrice de variance-covariance de l'estimateur $\hat{\beta}$.
- À l'aide de la fonction `lm` de R, retrouver ces informations.

2 Modèle d'ANCOVA

On souhaite maintenant simuler un modèle d'analyse en covariance (ANCOVA) :

$$y_{ij} = \beta_0 + x_{ij}\beta_1 + a_j + \varepsilon_{ij} \quad j = 1, ...10, i = 1, ...N_j$$

où a_j est l'effet du niveau j d'un facteur nommé fac.

- Simulations : on utilisera les valeurs de x et de $\beta = (\beta_0, \beta_1)'$ simulés dans l'exercice précédent.
  - Simuler l'appartenance de chaque donnée à l'un des 10 niveaux du facteur fac.
  - Simuler l'effet a_j de chacun des niveaux selon une loi gaussienne de variance 3.
  - Simuler les y_{ij} selon le modèle décrit.
- Construire la matrice X de design des effets fixes.
- Calculer les estimations de β, de $a = (a_1, ..., a_{10})'$ et de la variance résiduelle.
- Donner les matrice de variance-covariance des estimateurs associés.
- En utilisant la fonction `lm`, vérifier les résultats précédents.

3 Modèle d'ANOVA

On souhaite simuler :

$$y_{ik} = \beta + a_i + \varepsilon_{ik} \quad i = 1, ..., 10 \text{ et } k = 1, ..., 10$$

1) **Cas d'un plan équilibré**
   - Construire la matrice de design Z associée à l'effet aléatoire en utilisant la fonction `kronecker`
   - Simuler les a_i selon une loi $\mathcal{N}(0, \sigma_a^2)$ où $\sigma_a^2 = 9$
   - Simuler les y_{ik} avec une variance résiduelle $\sigma_0^2 = 1$
   - Calculer la matrice de variance-covariance marginale de Y selon le modèle simulé et comparer les estimations OLS et GLS de β
   - Comparer les résultats obtenus avec ceux de la fonction `lme` du package nlme et `lmer` du package lme4

2) **Cas d'un plan déséquilibré**
   - Supprimer aléatoirement des observations avec une probabilité de 0.3 et les lignes des matrices de design correspondantes
   - Comparer à nouveau les estimations OLS et GLS de β
   - Comparer les résultats obtenus avec ceux de la fonction `lme` du package nlme et `lmer` du package lme4