## TD N° 4 : Modèles linéaires mixtes (mais pas trop)

**Exercice 1.** Pour $j = 1, \dots, J$ et pour $i = 1, \dots, n_J$ on définit des variables *i.i.d.* $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, avec $\sum_{j=1}^J n_j = n$. On note $\bar{\varepsilon}_{:,j} = \frac{1}{n_j} \sum_{i=1}^{n_j} \varepsilon_{i,j}$ et $\bar{\varepsilon}_n = \frac{1}{n} \sum_{j=1}^J \sum_{i=1}^{n_j} \varepsilon_{i,j}$. On suppose que l'on définit de manière indépendante pour $j = 1, \dots, J$ des variables *i.i.d.* $A_j \sim \mathcal{N}(0, \sigma_A^2)$. Enfin le modèle d'observation est :

$$y_{i,j} = \mu + A_j + \varepsilon_{i,j} \quad . \tag{1}$$

1) Montrer que

$$\mathbb{E}\left(\frac{1}{J-1} \sum_{j=1}^J \sum_{i=1}^{n_j} (\bar{\varepsilon}_{:,j} - \bar{\varepsilon}_n)^2\right) = \sigma_\varepsilon^2 \quad . \tag{2}$$

2) Montrer que

$$\mathbb{E}\left(\sum_{j=1}^J \sum_{i=1}^{n_j} \left(A_j - \frac{1}{n} \sum_{j'=1}^J \sum_{i'=1}^{n_j} A_{j'}\right)^2\right) = \frac{n^2 - \sum_{j=1}^J n_j^2}{n} \sigma_A^2 \quad . \tag{3}$$

3) En déduire $\mathbb{E}\left(\sum_{j=1}^J \sum_{i=1}^{n_j} (\bar{y}_{:,j} - \bar{y}_n)^2\right)$ (avec les notations similaires).
4) Montrer que $\mathbb{E}\left(\sum_{j=1}^J \sum_{i=1}^{n_j} (y_{i,j} - \bar{y}_{:,j})^2\right) = (n - J)\sigma_\varepsilon^2$
5) En déduire des estimateurs $\hat{\sigma}_A^2$ et $\hat{\sigma}_\varepsilon^2$ de $\sigma_A^2$ et $\sigma_\varepsilon^2$.

**Exercice 2.** Prenons un modèle à effet aléatoire pour une catégorie à $J$ modalités $C_1, \dots, C_J$ et $n$ observations (avec $n = n_1 + \dots + n_J$) :

$$y = \mu \mathbb{1}_n + \sum_{j=1}^J \mathbb{1}_{C_j} A_j + \varepsilon \quad . \tag{4}$$

avec $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2 \operatorname{Id}_n)$ et $A \sim \mathcal{N}(0, \sigma_A^2 \operatorname{Id}_J)$.

1) Reprendre la feuille de TD précédente pour obtenir $(\operatorname{var}(y))^{-1}$.
2) Donner les conditions du premier ordre pour l'estimateur du maximum de vraisemblance de $\mu$ (en supposant connu $\sigma_\varepsilon^2$ et $\sigma_A^2$)
3) Retrouver la valeur de l'estimateur vu en cours

$$\hat{\mu} = \sum_{j=1}^J w_j \left(\frac{1}{n_j} \sum_{i \in C_j} y_i\right) \quad , \tag{5}$$

où $w_j \propto \frac{1}{\sigma_A^2 + \frac{\sigma_\varepsilon^2}{n_j}}$