HMMA307 : Advanced Linear Modeling

Chapter 5 : Random Anova

Fanchon Herman Cassandre Lepercque Cherif Amghar

https://github.com/fanchonherman/HMMA307_CM_Random_Anova

Université de Montpellier

Table of Contents

Motivation

Mixed models can be used in practice to deal with disordered data and allow us to use all of our data. Indeed, we can have different grouping factors but also small sample sizes. So, mixed models can process the data even when we have small sample sizes, structured data, and many covariates to fit.

Statistical model

Model equation

$$y_{ij} = \mu^* + A_j + \varepsilon_{ij}$$

- $\mu^* \in \mathbb{R}$, fixed effect,
- $A_j \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_A^2),\ \sigma_A^2 > 0,\ \forall j \in [\![1, J]\!]$, random effect,
- $\varepsilon_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)\ \sigma_\varepsilon^2 > 0,\ \forall i \in [\![1, I]\!]\ , \forall j \in [\![1, J]\!]$ the noise,
- $A_j \perp\!\!\!\perp \varepsilon_{ij}\ , \forall i, j,$
- $n = \sum_{j=1}^{J} n_j$ and n_j the number of observations of the modality J.

Esperance and covariance

- $\mathbb{E}[y_{ij}] = \mu^*$,
- $\mathbb{V}(y_{ij}) = \sigma_A^2 + \sigma_\varepsilon^2$,
- $\mathrm{cov}(y_{ij}, y_{i'j'}) = \sigma_A^2 \delta_{jj'} + \sigma_\varepsilon^2 \delta_{ii'} \delta_{jj'}$,

Matrix Model

Model equation

$$y = \mu^* \mathbb{1}_n + ZA + \varepsilon$$

- $\mu^* \in \mathbb{R}$, fixed effect
- $Z = \begin{bmatrix} \mathbb{1}_{C_1} & \cdots & \mathbb{1}_{C_J} \end{bmatrix} \in \mathbb{R}^{n \times J}$, design matrix,
- $C_1 \sqcup \cdots \sqcup C_J = [\![1, n]\!]$, classes / modalities,
- $A = \begin{pmatrix} A_1 & \cdots & A_J \end{pmatrix}^T \in \mathbb{R}^J, \ A \sim \mathcal{N}(0, \sigma_A^2 \ I_{d_J}), \ \sigma_A^2 > 0$, random matrix,
- $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, $\sigma_\varepsilon^2 > 0$, the noise,
- $$ZA = \sum_{j=1}^{J} A_j \mathbb{1}_{C_j} \in \mathbb{R}^n.$$

Variance calculate

- We have that, $\mathbb{V}(ZA) = \underbrace{Z}_{n\times J}\underbrace{\mathbb{V}(A)}_{J\times J}\ \underbrace{Z^T}_{J\times n} = \sigma_A^2 ZZ^T \ \in \mathbb{R}^{n\times n}$,
- Where,
$$ZZ^T = \begin{bmatrix}\mathbb{1}_{C_1} & \cdots & \mathbb{1}_{C_J}\end{bmatrix}\begin{bmatrix}\mathbb{1}_{C_1}^T \\ \vdots \\ \mathbb{1}_{C_J}^T\end{bmatrix} = \sum_{j=1}^{J} \mathbb{1}_{C_j}\mathbb{1}_{C_j}^T,$$
- Then, $\mathbb{V}(y) = \sigma_A^2 ZZ^T + \sigma_\varepsilon^2 I_{d_n}$.

Reminder for the statistical model

- The model : $y_{ij} = \mu^* + A_j + \varepsilon_{ij}$,
- We have that
$$\overline{y_{\cdot j}} = \frac{1}{n_j} \sum_{i \in C_j} y_{ij},$$
- So, $\mathbb{V}(\overline{y_{\cdot j}}) = \sigma_A^2 + \frac{\sigma_\varepsilon^2}{n_j} := \tau_j^2$ and $\mathbb{E}(\overline{y_{\cdot j}}) = \mu^*$ without bias.

μ^*-estimator

Formula

$$\hat{\mu} = \sum_{j=1}^{J} \omega_j \overline{y_{:j}}, \quad \text{with } \omega_j \propto \frac{1}{\mathbb{V}(\overline{y_{:j}})} \text{the weighting.}$$

<u>**Remark :**</u>

In balance case, we have :

- $n_j = I$, as we already know that $n = IJ$,
- So we obtain,

$$\hat{\mu} = \frac{I}{J} \sum_{j=1}^{J} \overline{y_{:j}}.$$

Theorem

Let X_1,...,X_n independent random variables. We suppose that $E(X_1) = ... = E(X_n) = \mu^*$ Among the linear unbiased estimators in X_1, ...,X_n of μ^*, the minimal variance estimator is denoted by $\hat{\mu}$ and is worth:

$$\hat{\mu} = \sum_{i=1}^{n} \frac{X_i/var(X_i)}{\sum_{i'=1}^{n} 1/var(X_i)}$$

- **Demo**: We have

$$min_{(\alpha_1,...,\alpha_n)\in\mathbb{R}^n} Var(\sum_{i=1}^{n} \alpha_i X_i)$$

u.c: $\sum_{i=1}^{n} \alpha_i = 1$

$$E(\sum_{i=1}^{n} \alpha_i X_i) = \mu^*$$

$$Var(\sum_{i=1}^{n} \alpha_i X_i) = \sum_{i=1}^{n} \alpha_i^2 Var(X_i)$$

Minimize the last expression amounts to minimize this expression:

$$min_{(\alpha_1,\dots,\alpha_n)\in\mathbb{R}^n} \sum_{i=1}^{n} \alpha_i^2 Var(X_i) \quad u.c \quad \sum_{i=1}^{n} \alpha_i = 1$$

- **Lagrangian** :

$$\mathcal{L}(\alpha,\lambda) = \sum_{i=1}^{n} \alpha_i^2 Var(X_i) + \lambda(\sum_{i=1}^{n} \alpha_i - 1)$$

Resolution of the optimization system:

$$\nabla\mathcal{L}(\hat{\alpha},\hat{\lambda}) = 0$$

$$\begin{cases} \frac{\partial\mathcal{L}}{\partial\alpha} = 0 \\ \frac{\partial\mathcal{L}}{\partial\alpha_{i_0}} = 0 \ \forall i_0 \end{cases} \iff \begin{cases} \sum\limits_{i=1}^{n} \hat{\alpha}_i = 1 \\ 2\hat{\alpha}_{i_0} Var(x_{i_0} + \hat{\lambda} = 0, \forall i_0 \end{cases}$$

$$\iff \begin{cases} \hat{\alpha}_{i_0} = \frac{-\hat{\lambda}}{2Var(x_{i_0}} \\ \sum\limits_{i_0=1}^{n} \hat{\alpha}_{i0} = 1 = \frac{-\hat{\lambda}}{2}(\sum_{i_0=1}^{n} \frac{1}{Var(x_{i_0})}) \end{cases}$$

Finally,

$$\hat{\lambda} = -2(\frac{1}{\sum_{i=1}^{n} 1/var(X_i)})$$

$$\hat{\alpha}_{i_0} = \frac{\frac{1}{Var(X_{i0})}}{\sum_{i=1}^{n} 1/var(X_i)} \Longrightarrow \hat{\mu} = \sum_{j=1}^{J} \alpha_j \bar{y_{ij}}$$

- **Vraisemblance** :

$$\mathcal{L}(y, \mu, \hat{\sigma}^2_A, \hat{\sigma}^2_\epsilon) \ \propto |V|^{-1/2} \, exp(-\frac{1}{2} \left[y - \mu \mathbb{1}_n\right]^T \ V^{-1} \left[y - \mu \mathbb{1}_n\right])$$

$$-log(\mathcal{L}) = \frac{1}{2}(y - \mu\mathbb{1}_n)^T V^{-1}(y - \mu\mathbb{1}_n) + \frac{1}{2} log|V| + \frac{n}{2} log(2\pi)$$

On souhaite minimiser $-log(\mathcal{L})$

$$\implies \quad \underset{\mu, \sigma^2_A, \sigma^2_\epsilon}{min} \frac{1}{2}(y - \mu\mathbb{1}_n)^T V^{-1}(y - \mu\mathbb{1}_n) + \frac{1}{2} log|V|$$