Modèles Linéaires Appliqués

Arthur Charpentier

Automne 2020

OLS #17 (ANOVA & modèles mixtes)

ANOVA (one-way)

One-way model, $y_{ij} = \mu_j + \varepsilon_{ij} = \mu + a_j + \varepsilon_{ij}$

- y_{ij} is the value for ith individual of group j,
- μ is the population (grand) mean,
- a_j is the random effect for the jth level of factor a,
- ε_{ij} is the random error effect.

Random effects, $a_j \sim \mathcal{N}(0, \sigma_a^2)$ while $\varepsilon \sim \mathcal{N}(0, \sigma^2)$
$\text{Var}[y] = \sigma_a^2 + \sigma^2$ if effects a_j and errors ε are independent

$\text{Cov}(Y_{ij}, Y_{ij'}) = \sigma_a^2$ so that $\text{Corr}(Y_{ij}, Y_{ij'}) = \dfrac{\sigma_a^2}{\sigma_a^2 + \sigma^2}$

(population intraclass correlation)

ANOVA (two-way)

Two-way model, $y_{ijk} = \mu_{jk} + \varepsilon_{ijk} = \mu + a_j + b_k + (ab)_{jk} + \varepsilon_{ijk}$
Random effects, $a_j \sim \mathcal{N}(0, \sigma_a^2)$, $b_k \sim \mathcal{N}(0, \sigma_b^2)$, $(ab)_{jk} \sim \mathcal{N}(0, \sigma_{ab}^2)$
while $\varepsilon \sim \mathcal{N}(0, \sigma^2)$
$\text{Var}[y] = \sigma_a^2 + \sigma^2$ if effects a_j and errors ε are independent
cross effect (ab)

$y_{ij} = \beta_0 + \beta_1 x_{ij} + a_j + \varepsilon_{ij} =$

- y_{ij} is the value of the response for ith individual of group j,
- β_0 is the (fixed) intercept
- β_1 is the (fixed) slope
- x_{ij} is the value of the predictor for ith individual of group j
- a_j is the random effect for the jth level of factor a,
- a_j is the random intercept, $a_j \sim \mathcal{N}(0, \sigma_a^2)$
- ε_{ij} is the random error effect, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$

Linear mixed-effect regression model

$$y_{ij} = \beta_0 + \mathbf{x}_{ij}^\top \boldsymbol{\beta} + a_{j0} + \mathbf{z}_{ij}^\top \mathbf{a}_j + \varepsilon_{ij}$$

where $\mathbf{a}_j = (a_{j1}, \cdots, a_{jq})$, $a_{jk} \sim \mathcal{N}(0, \sigma_k^2)$

If $\mathbf{a}_j \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$,

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{a}_j + \boldsymbol{\varepsilon}_j \sim \mathcal{N}(\mathbf{X}_j\boldsymbol{\beta}, \mathbf{\Sigma}_j)$$

where $\mathbf{\Sigma}_j = \mathbf{Z}_j\mathbf{\Sigma}\mathbf{Z}_j^\top + \sigma^2\mathbb{I}$, and

- $\mathbf{\Sigma}$ unstructured, $\mathbf{\Sigma} = (\Sigma_{ij})$
- $\mathbf{\Sigma}$ diagonal, $\text{diag}(\sigma_i^2)$
- $\mathbf{\Sigma}$ compound symmetric,
- $\mathbf{\Sigma}$ autoregressive
- $\mathbf{\Sigma}$ Toeplitz

$$\begin{pmatrix} \sigma_v^2+\sigma^2 & \sigma_v^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2+\sigma^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 & \sigma_v^2+\sigma^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 & \sigma_v^2 & \sigma_v^2+\sigma^2 \end{pmatrix}$$

If $\mathbf{a}_j \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$,

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{a}_j + \boldsymbol{\varepsilon}_j \sim \mathcal{N}(\mathbf{X}_j\boldsymbol{\beta}, \mathbf{\Sigma}_j)$$

where $\mathbf{\Sigma}_j = \mathbf{Z}_j\mathbf{\Sigma}\mathbf{Z}_j^\top + \sigma^2\mathbb{I}$, and

- $\mathbf{\Sigma}$ unstructured, $\mathbf{\Sigma} = (\Sigma_{ij})$
- $\mathbf{\Sigma}$ diagonal, $\text{diag}(\sigma_i^2)$
- $\mathbf{\Sigma}$ compound symmetric,
- $\mathbf{\Sigma}$ autoregressive
- $\mathbf{\Sigma}$ Toeplitz

$$\sigma^2\begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix} \text{ or } \sigma^2\begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{pmatrix}$$