HMMA 307 : Modèle Linéaire Avancé

Modèle linéaire généralisé mixte

KHALIFI OUMAYMA

M2 MIND

Université de Montpellier

Sommaire

MLM

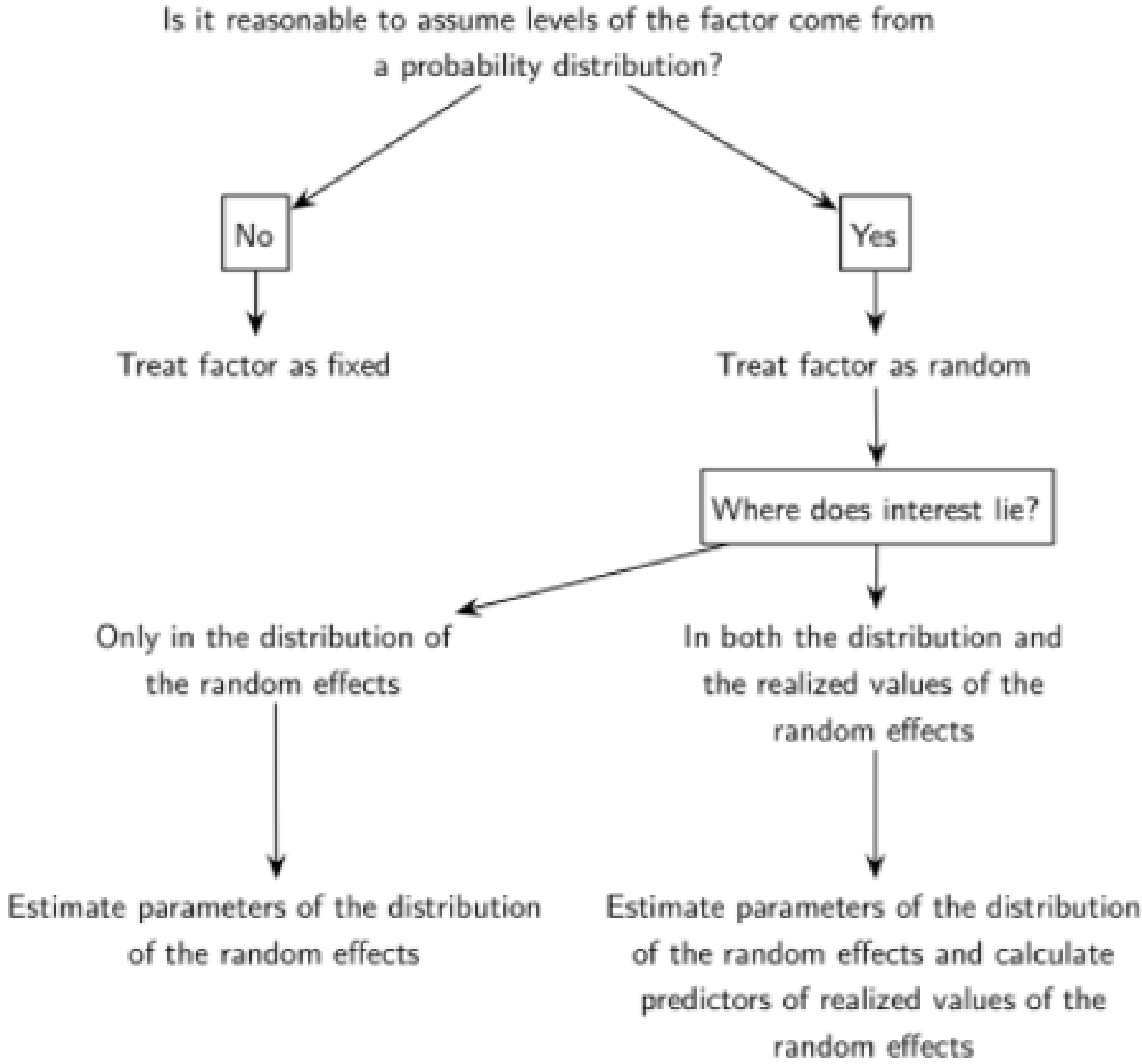
Définition:

En notation matricielle,un modèle mixte peut être représenté comme suit:

$$\boldsymbol{y} = X\boldsymbol{\beta} + Z\boldsymbol{u} + \boldsymbol{\epsilon}$$

avec:

- $\boldsymbol{y}$ un vecteur connu d'observations, de moyenne $E(\boldsymbol{y}) = X\boldsymbol{\beta}$.
- β un vecteur inconnu d'effets fixes.
- $\boldsymbol{u}$ un vecteur inconnu d'effets aléatoires, de moyenne $E(\boldsymbol{u}) = \mathbf{0}$ et ayant pour matrice de variance-covariance $\mathrm{var}(\boldsymbol{u}) = G$.
- ϵ un vecteur inconnu d'erreur aléatoires, de moyenne $E(\boldsymbol{\epsilon}) = \mathbf{0}$ et de variance $\mathrm{var}(\boldsymbol{\epsilon}) = R$.
- X et$Z$ des matrices liant les observations $\boldsymbol{y}$ à β et $\boldsymbol{u}$, respectivement.


Is it reasonable to assume levels of the factor come from a probability distribution?
No
Yes
Treat factor as fixed
Treat factor as random
Where does interest lie?
Only in the distribution of the random effects
In both the distribution and the realized values of the random effects
Estimate parameters of the distribution of the random effects
Estimate parameters of the distribution of the random effects and calculate predictors of realized values of the random effects

exemple d'un jeu de données simulé:

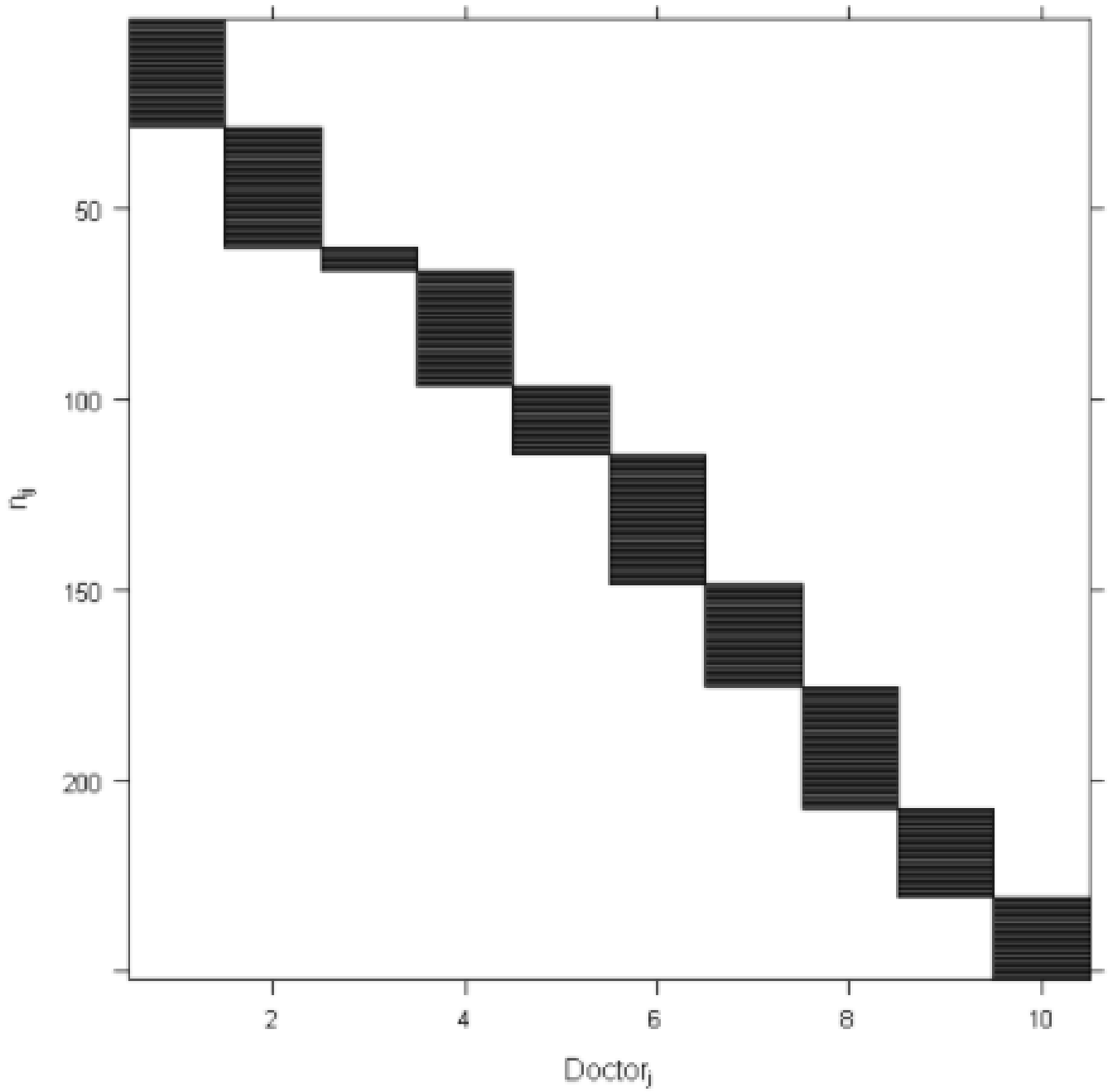
$$\overbrace{\mathbf{y}}^{8525\times 1} = \underbrace{\mathbf{X}}_{8525\times 6}\underbrace{\beta}_{6\times 1} + \underbrace{\mathbf{Z}}_{8525\times 407}\underbrace{\boldsymbol{u}}_{407\times 1} + \overbrace{\varepsilon}^{8525\times 1}$$

$$\mathbf{y} = \begin{bmatrix} \text{mobility} \\ 2 \\ 2 \\ \dots \\ 3 \end{bmatrix} \qquad \beta = \begin{bmatrix} 4.782 \\ .025 \\ .011 \\ .012 \\ 0 \\ -.009 \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} \text{Intercept} & \text{Age} & \text{Married} & \text{Sex} & \text{WBC} & \text{RBC} \\ 1 & 64.97 & 0 & 1 & 6087 & 4.87 \\ 1 & 53.92 & 0 & 0 & 6700 & 4.68 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 56.07 & 0 & 1 & 6430 & 4.73 \end{bmatrix}$$

	Intercept	Age	Married	sex	WBC	RBC	y
0	1	46.0	1	0	5991.528980	5.753987	2
1	1	62.0	1	0	5994.072082	3.175610	3
2	1	62.0	0	1	6018.001567	6.613776	2
3	1	53.0	0	0	5990.948873	4.675376	3
4	1	47.0	1	1	5985.482100	5.478123	2
...	...	...	...	...	...	...	...
8520	1	58.0	1	1	5988.672529	5.861227	3
8521	1	50.0	1	1	5993.302204	5.698575	2
8522	1	66.0	1	0	5979.884308	7.333334	3
8523	1	44.0	1	0	6006.765988	10.734307	3
8524	1	67.0	1	1	5999.967000	4.500191	2

8525 rows × 7 columns


First 10 Doctors Z: 252 x 10
50
100
150
200
n_{ij}
2
4
6
8
10
Doctor$_j$

Dep. Variable:	y	No. Observations:	8525
Model:	GLM	Df Residuals:	8519
Model Family:	Poisson	Df Model:	5
Link Function:	log	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-12373.
Date:	Sun, 08 Nov 2020	Deviance:	858.05
Time:	12:38:58	Pearson chi2:	853.
No. Iterations:	4		
Covariance Type:	nonrobust		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	0.5515	4.577	0.120	0.904	-8.420	9.523
sex	0.0029	0.014	0.208	0.835	-0.024	0.030
Age	0.0003	0.001	0.456	0.649	-0.001	0.002
Married	0.0048	0.014	0.348	0.728	-0.022	0.032
WBC	5.814e-05	0.001	0.076	0.939	-0.001	0.002
RBC	-0.0009	0.003	-0.253	0.800	-0.008	0.006

```
==========================================================
Model:                  MixedLM  Dependent Variable:  y
No. Observations:       8525     Method:              REML
No. Groups:             1        Scale:               0.2500
Min. group size:        8525     Log-Likelihood:      -6222.4240
Max. group size:        8525     Converged:           Yes
Mean group size:        8525.0
----------------------------------------------------------
                 Coef.    Std.Err.     z    P>|z| [0.025 0.975]
----------------------------------------------------------
Intercept         2.444        0.503  4.860 0.000  1.458  3.430
sex               0.059        0.077  0.772 0.440 -0.091  0.210
Married          -0.048        0.077 -0.622 0.534 -0.199  0.103
sex:Married       0.038        0.109  0.348 0.728 -0.175  0.251
Age               0.001        0.001  0.787 0.431 -0.001  0.003
sex:Age          -0.001        0.001 -0.599 0.549 -0.004  0.002
Married:Age       0.001        0.001  0.885 0.376 -0.001  0.004
sex:Married:Age  -0.001        0.002 -0.486 0.627 -0.005  0.003
Group Var         0.250 3587403.181
==========================================================
```

$$L1: \quad Y_{ij} = \beta_{0j} + \beta_{1j} Age_{ij} + \beta_{2j} Married_{ij} + \beta_{3j}\,\mathrm{Sex}_{ij} + \beta_{4j} WBC_{ij} + \beta_{5j} RBC_{ij} + e_{ij}$$

$$L2: \quad \beta_{0j} = \gamma_{00} + u_{0j}$$

$$L2: \quad \beta_{1j} = \gamma_{10}$$

$$L2: \quad \beta_{2j} = \gamma_{20}$$

$$L2: \quad \beta_{3j} = \gamma_{30}$$

$$L2: \quad \beta_{4j} = \gamma_{40}$$

$$L2: \quad \beta_{5j} = \gamma_{50}$$

$$Y_{ij} = (\gamma_{00} + u_{0j}) + \gamma_{10} Age_{ij} + \gamma_{20} \mathsf{Married}_{ij} + \gamma_{30} SEX_{ij} + \gamma_{40} WBC_{ij} + \gamma_{50} RBC_{ij} + e_{ij}$$