mwp-204

多階層混合効果モデルの概要　【 評価版 】

本 whitepaper では混合効果モデルについて、その機能概要と用例を紹介します。

1. はじめに

混合効果モデル (mixed-effects models) は固定効果 (fixed effects) と変量効果 (random effects) の双方を含むモデルとして特徴付けられます。固定効果は通常の回帰係数に対応するもので直接推定されます。これに対し変量効果は直接推定されるわけではなく、分散や共分散の推定値という要約された形でレポートされてきます（ただし個々の値は推定後機能を使って求めることはできます)。変量効果は切片 (random intercepts) のみならず勾配 (random coefficients) としてもモデル化できます。またデータのグループ構成は多階層のものが許容されます。従って混合効果モデルは多段モデル (multilevel models) とか階層モデル (hierarchical models) と呼ばれることもあります。混合効果をサポートしたコマンドは変量効果が正規分布に従うことを前提に、種々の分布に従う応答変数に対して混合効果モデルをフィットさせます。

2. コマンドの構文

英文マニュアルエントリ `[ME]` **me** のセクション“*Using mixed-effects commands*”をご参照ください。

3. 混合効果モデル

3.1 線形混合効果モデル

連続応答変数に対する混合効果モデル — 線形混合効果 (LME: linear mixed-effects) モデル — は、全般的な誤差項に加えてランダムな変動（効果）を許容できるように拡張された線形回帰モデルであり、そのモデル式は次のように記述されます。

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon} \tag{1}$$

ただし

- $\mathbf{y}$ は $n \times 1$ ベクトル
- $\mathbf{X}$ は固定効果 $\boldsymbol{\beta}$ に対する $n \times p$ のデザイン/共変量行列
- $\mathbf{Z}$ は変量効果 $\mathbf{u}$ に対する $n \times q$ のデザイン/共変量行列
- $n \times 1$ の誤差ベクトル $\boldsymbol{\epsilon}$ は平均値が 0 で分散行列が $\sigma_\epsilon^2\mathbf{R}$ であるような多変量正規分布に従う

ものとします。

(1) 式の固定効果部分 $\mathbf{X}\boldsymbol{\beta}$ は標準的な OLS 回帰における線形予測項に該当するものです。これに対し (1) 式の変量効果部分 $\mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}$ に対しては、$\mathbf{u}$ が分散共分散行列 $\mathbf{G}$ を有し、それは $\boldsymbol{\epsilon}$ と直交関係にあること、すなわち

$$\mathrm{Var}\begin{bmatrix}\mathbf{u} \\ \boldsymbol{\epsilon}\end{bmatrix} = \begin{bmatrix}\mathbf{G} & \mathbf{0} \\ \mathbf{0} & \sigma_\epsilon^2\mathbf{R}\end{bmatrix}$$

であることを仮定します。変量効果 $\mathbf{u}$ は予測はできても直接推定されるわけではありません。ただしそれは分散成分 (variance components) として知られる $\mathbf{G}$ の要素によって特徴付けられるわけですが、それらは全般的な残差分散 σ_ϵ^2 と $\mathbf{R}$ 内に含まれる残差分散パラメータと共に推定されます。

デザイン行列 $\mathbf{X}$ と $\mathbf{Z}$ の形式によってさまざまなデザインに対応した線形モデルが構成されます。それらはまたクラスタ内相関をモデリングする上での種々の手法を提供するものでもあります。すなわちランダム切片やランダム傾きを共用することによってクラスタ内相関を表現することができます。さらに $\mathbf{G}$ や $\mathbf{R}$ の規定のしかたによってもさまざまな柔軟性がもたらされます。

n 個の観測データが M 個の独立なグループ（クラスタ）から構成される場合には、(1) 式は

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j + \boldsymbol{\epsilon}_j, \quad j = 1, \ldots, M \tag{2}$$

のように表現できます。この場合、$\mathbf{y}_j$ はクラスタ j に対応する $\mathbf{y}$ の行からなるものとします。$\mathbf{X}_j, \boldsymbol{\epsilon}_j$ についても同様に定義されます。変量効果 $\mathbf{u}_j$ については平均が $\mathbf{0}$、分散行列が $\boldsymbol{\Sigma}$ の正規分布に従う $q \times 1$ ベクトルの M 個の実現値とみなすことができます。行列 $\mathbf{Z}_j$ はクラスタ j の変量効果に対応した $n_j \times q$ のデザイン行列です（クラスタ j に含まれる観測データの数を n_j とする)。これらを (1) 式と対応付けると次のようになります。

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z}_M \end{bmatrix}; \quad \mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_M \end{bmatrix}; \quad \mathbf{G} = \mathbf{I}_M \otimes \boldsymbol{\Sigma}; \quad \mathbf{R} = \mathbf{I}_M \otimes \boldsymbol{\Lambda} \tag{3}$$

ただし $\boldsymbol{\Lambda}$ はレベル 1 残差の分散行列であり、$\otimes$ は Kronecker 積を意味します。

モデル式 (2) は Laird and Ware (1982) に由来するものですが、それは次のような 2 つの利点を持ちます。第 1 に変量効果の規定をより容易なものとします。例えばクラスタが学校を意味するのであれば、学校のレベルにおける変量効果を考えれば良いことになるわけです。第 2 にそれは変量効果の多段化を可能にします。例えば学校内にネストされたクラスを考えた場合、学校レベルでの変量効果とクラスレベルでの変量効果を表現できるモデルへの拡張がより容易に行えるようになります。

線形混合モデルのフィットは `mixed` コマンドによって行えます。詳細については `[ME]` **mixed** (*mwp-018*) をご参照ください。`mixed` に対しては予測や診断に関する多様な postestimation 機能が用意されていますが、それらについては `[ME]` **mixed postestimation** (*mwp-169*) をご参照ください。

3.2 一般化線形混合効果モデル

評価版では割愛しています。

3.3 生存時間混合効果モデル

評価版では割愛しています。

3.4 非線形混合効果モデル

評価版では割愛しています。

3.5 その他のモデル記述様式

評価版では割愛しています。

3.6 尤度の計算

評価版では割愛しています。

3.7 計算時間とラプラス近似

評価版では割愛しています。

3.8 収束上の問題

評価版では割愛しています。

3.9 尤度比検定のための分布理論

評価版では割愛しています。

4. 用例

4.1 2 階層モデル

▷ Example 1: 成長曲線モデル（ランダム切片）

ここでは Example データセット pig.dta を用いた用例を紹介します。

```
. use https://www.stata-press.com/data/r19/pig.dta *1
(Longitudinal analysis of pig weights)
```

このデータセット中には 48 頭の豚の体重が 9 週間にわたって記録されています。参考までに先頭の豚についてのデータをリスト出力しておきます。

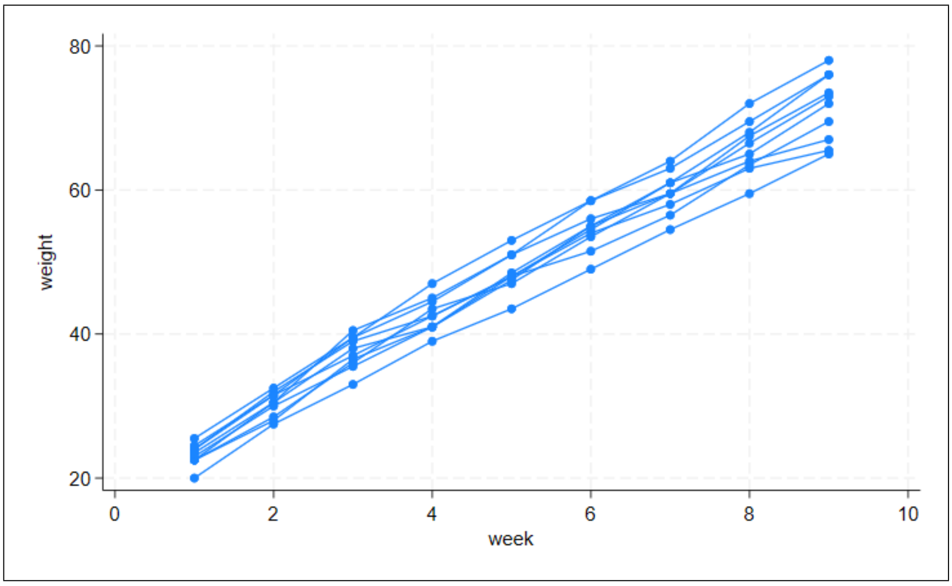
```
. list id week weight if id == 1 *2
```

	id	week	weight
1.	1	1	24
2.	1	2	32
3.	1	3	39
4.	1	4	42.5
5.	1	5	48
6.	1	6	54.5
7.	1	7	61
8.	1	8	65
9.	1	9	72

id = 1 から 10 までの範囲の豚についてその体重の変化をグラフ化してみると次のようになります。

[*1] メニュー操作：File ▷ Example Datasets ▷ Stata 19 manual datasets と操作、Multilevel Mixed-Effects Reference Manual [ME] の me の項よりダウンロードする。
[*2] メニュー操作： Data ▷ Describe data ▷ List data

```
. twoway connected weight week if id<=10, connect(L)
```
[*3]



それぞれの豚の体重の変化は線形であり、また個体差もそれなりにあることがわかります。ところで我々は pig.dta に含まれている特定の 48 頭の豚について関心があるわけではない点に注意してください。すなわちそれら 48 頭はより大きな母集団からランダムに抽出された標本に過ぎないととらえ、個体間の変動を変量効果としてモデル化することにします。ただしモデル式 (2) における変量効果項 $\mathbf{Z}_j\mathbf{u}_j$ としては定数項、すなわちランダム切片 (random-intercept) 項のみを想定することにします。このときのモデル式は次のようになります。

$$\texttt{weight}_{ij} = \beta_0 + \beta_1\texttt{week}_{ij} + u_j + \epsilon_{ij}, \quad i = 1, \ldots, 9;\ j = 1, \ldots, 48 \tag{4}$$

2 階層モデルの場合、第 2 レベルを表す j がクラスタに対応するため、豚を識別する変数 id に対応し、第 1 レベルを表す i は変数 week に対応する点に注意してください。

このモデル式における固定効果部 $\beta_0 + \beta_1\texttt{week}_{ij}$ は母集団平均に対応した 1 本の回帰直線を表します。これに対し変量効果項 u_j はその回帰直線を平行に上下させる個体ごとの効果を表すことになります。このようなモデルをフィットさせるには、クラスタを識別する j が変数 id に対応する点を考慮し、次のように操作します。

[*3] メニュー操作： Graphics ▷ Twoway graph (scatter, line, etc.)

- Statistics ▷ Multilevel mixed-effects models ▷ Linear regression と操作
- Model タブ: Estimation method: Maximum likelihood （デフォルト）
  Fixed-effects model: Dependent variable: `weight`
  Independent variables: `week`
  Random-effects equations: Create. . .

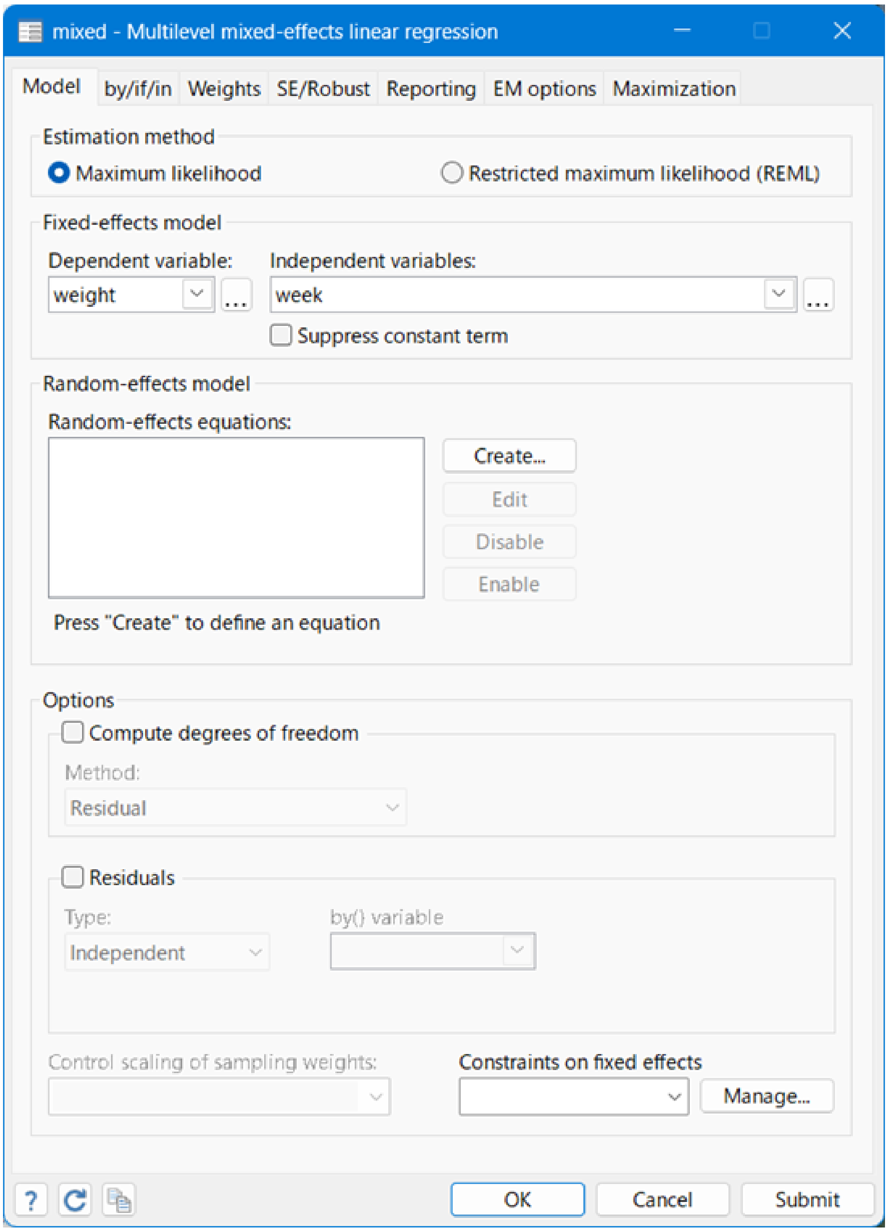


図 1 `mixed` ダイアログ- Model タブ

- Equation 1 ダイアログ: Level variable for equation: `id`

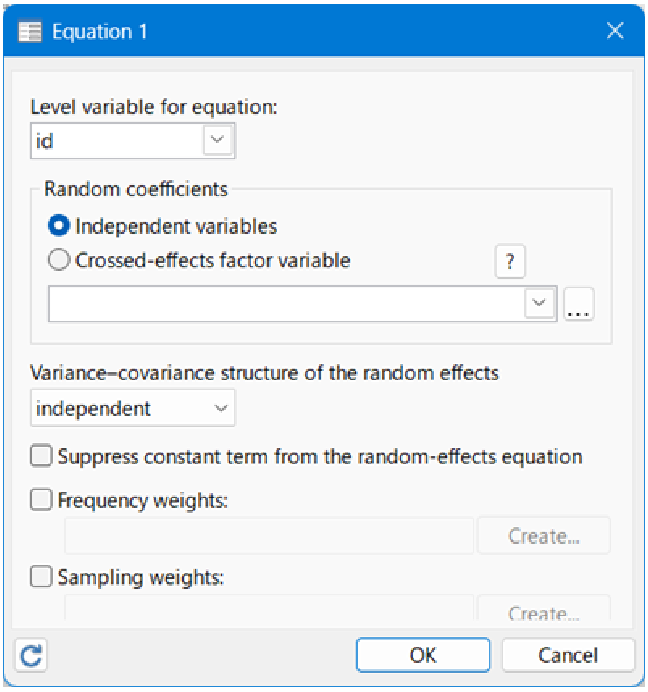


図 2 Equation 1 ダイアログ – ランダム切片

```
. mixed weight week || id:

Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -1014.9268
Iteration 1:  Log likelihood = -1014.9268

Computing standard errors ...

Mixed-effects ML regression                         Number of obs     =        432
Group variable: id                                  Number of groups  =         48
                                                    Obs per group:
                                                                  min =          9
                                                                  avg =        9.0
                                                                  max =          9
                                                    Wald chi2(1)      =   25337.49
Log likelihood = -1014.9268                         Prob > chi2       =     0.0000
```

weight	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
week	6.209896	.0390124	159.18	0.000	6.133433	6.286359
_cons	19.35561	.5974059	32.40	0.000	18.18472	20.52651

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
id: Identity				
var(_cons)	14.81751	3.124226	9.801716	22.40002
var(Residual)	4.383264	.3163348	3.805112	5.04926

LR test vs. linear model: chibar2(01) = 472.65 Prob >= chibar2 = 0.0000

なお、Equation 1 ダイアログ上で変量効果の分散共分散構造として independent がデフォルトとして選択されている点に注意してください。

評価版では割愛しています。

▷ Example 2: **分割区画デザイン**

評価版では割愛しています。

▷ Example 3: **二項カウント**

評価版では割愛しています。

4.2 共分散構造

評価版では割愛しています。

4.3 3 階層モデル

評価版では割愛しています。

4.4 Crossed-effects モデル

評価版では割愛しています。

4.5 非線形モデル

評価版では割愛しています。

■