

マルチレベル混合効果モデル

はじめに

- 混合効果モデルは、固定効果と変量効果の両方を含みます。
- **mixed** コマンドは線形混合効果モデルをフィットします。これらのモデルはマルチレベルモデル、階層線形モデルとも呼ばれます。
- ☆ 本文中のコマンドをコピーし、Stata のコマンドウィンドウに貼り付けて実行できます。全ての操作のコマンドは、do ファイル `multilevel.do` にまとめられています。

2 レベルモデル

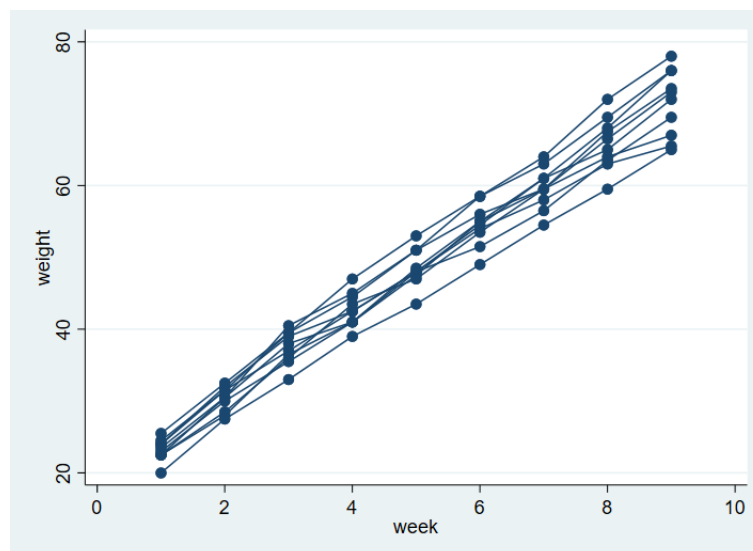
例題 1：ランダム切片の 2 レベルモデル

- Ruppert, Wand, and Carroll (2003) および Diggle et al. (2002) の 48 頭の豚のデータセットを使用します。体重 (**weight**) と 9 週間の飼育期間 (**week**) の関係をモデル化します。各個体は変数 **id** によって識別されます。

データセットをインポートし、はじめの 10 頭の成長曲線を作成します。

```
use https://www.stata-press.com/data/r16/pig
```

```
twoway connected weight week if id<=10, connect(L)
```



- 各個体の成長曲線は線形で、体重は飼育期間に比例して増加しています。10 頭の成長速度にある程度のばらつきが認められます。これらのサンプルを大きな母集団からのランダムサンプルとして扱います。
- 個体間の変動を変量効果としてモデル化します（個体レベルのランダム切片項）。モデルは次のようになります。

$$\text{weight}_{ij} = \beta_0 + \beta_1 \text{week}_{ij} + u_j + \epsilon_{ij} \quad (1)$$

i	飼育期間の週	$i = 1, \dots, 9$
j	豚の個体	$j = 1, \dots, 48$

モデルの固定効果部分 $\beta_0 + \beta_1 \text{week}_{ij}$ は、母集団の平均を表す 1 つの全体的な回帰直線を求めることを示しています。変量効果 u_j は、各個体に応じてこの回帰直線を上下にシフトします。

変量効果は各個体レベル（**id**）で発生するため、次のようにモデルをフィットします。

```
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

Group variable: id

Number of obs = 432

Number of groups = 48

Obs per group:

min = 9

avg = 9.0

max = 9

Log likelihood = -1014.9268

Wald chi2(1) = 25337.49

Prob > chi2 = 0.0000

weight	Coef.	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

`estimates store` コマンドで推定結果に名前を付けて保存します。

```
estimates store randint
```

例題 2：ランダムスロープの 2 レベルモデル

- モデル(1)を拡張し、成長速度のばらつきとして week_{ij} のランダムな傾きを可能にすると、次のようになります。

$$\text{weight}_{ij} = \beta_0 + \beta_1 \text{week}_{ij} + u_{0j} + u_{1j} \text{week}_{ij} + \epsilon_{ij} \quad (2)$$

このモデルでフィットを実行します。

```
mixed weight week || id: week
```

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log likelihood = -869.03825

Iteration 1: log likelihood = -869.03825

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
Log likelihood = -869.03825	Wald chi2(1)	=	4689.51
	Prob > chi2	=	0.0000

weight	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
week	6.209896	.0906819	68.48	0.000	6.032163	6.387629
_cons	19.35561	.3979159	48.64	0.000	18.57571	20.13551

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
id: Independent				
var(week)	.3680668	.0801181	.2402389	.5639103
var(_cons)	6.756364	1.543503	4.317721	10.57235
var(Residual)	1.598811	.1233988	1.374358	1.85992

LR test vs. linear model: chi2(2) = 764.42 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

推定結果に名前を付けて保存します。

```
estimates store randslope
```

- 変量効果 $(u_{0j}, u_{1j})'$ の共分散構造を指定しなかったため、**mixed** コマンドはデフォルトの独立 (**Independent**) 構造を使用しました。

$$\Sigma = \text{Var} \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} = \begin{bmatrix} \sigma_{u_0}^2 & 0 \\ 0 & \sigma_{u_1}^2 \end{bmatrix} \quad (3)$$

- 結果より、 $\hat{\sigma}_{u_0}^2 = 6.76$ 、 $\hat{\sigma}_{u_1}^2 = 0.37$ です。固定効果の点推定値は、モデル(1)の結果と同じになっていますが、これはレアケースです。
- $\hat{\sigma}_{u_1}^2$ の 95%信頼区間から、ランダムスロープが有意であると考えます。尤度比検定のコマンド **lrtest** と、これまでに保存した2つの推定結果を使用して、このことを検証します。

```
lrtest randslope randint
```

```
Likelihood-ratio test                    LR chi2(1) =    291.78
(Assumption: randint nested in randslope)  Prob > chi2 =    0.0000
```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

- ゼロに近い有意水準は、個体固有のシフトのみを許可するモデル(1)よりも、ランダムな個体固有の成長速度も許可するモデル(2)の方が当てはまりが良いことを示しています。

共分散構造

- 変量効果の共分散構造を指定しない場合、`mixed` コマンドはデフォルトの独立 (**Independent**) 構造を使用します。他の共分散構造を指定するときは `covariance()` オプションを使用します。

covariance()オプション一覧 (A)

independent (デフォルト)	$\begin{bmatrix} \sigma_{u0}^2 & 0 \\ 0 & \sigma_{u1}^2 \end{bmatrix}$	変量効果の間に相関なし。独立。
identity	$\begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}$	独立だが、分散はすべて同じ。
exchangeable	$\begin{bmatrix} \sigma_u^2 & \sigma_{01} \\ \sigma_{01} & \sigma_u^2 \end{bmatrix}$	分散は全て同じ。効果間の共分散もどの組み合わせも同じ。
unstructured	$\begin{bmatrix} \sigma_{u0}^2 & \sigma_{01} \\ \sigma_{01} & \sigma_{u1}^2 \end{bmatrix}$	全ての分散と共分散は異なる。

例題 3：相関のある変量効果の2レベルモデル

- 式 (3) を一般化して、 u_{0j} と u_{1j} を相関させることができます。`covariance(unstructured)` オプションで指定します。

$$\Sigma = \text{Var} \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} = \begin{bmatrix} \sigma_{u0}^2 & \sigma_{01} \\ \sigma_{01} & \sigma_{u1}^2 \end{bmatrix}$$

変量効果の共分散を仮定して、再度フィットを実行します。

```
mixed weight week || id: week, covariance(unstructured)
```


3 レベルモデル

- 2つのレベルのクラスタが入れ子状になった、3レベルモデルを考えます。

$$y_{jk} = X_{jk}\beta + Z_{jk}^{(3)}u_k^{(3)} + Z_{jk}^{(2)}u_{jk}^{(2)} + \epsilon_{jk} \quad (4)$$

$$u_k^{(3)} \sim N(0, \Sigma_3); \quad u_{jk}^{(2)} \sim N(0, \Sigma_2); \quad \epsilon_{jk} \sim N(0, \sigma_\epsilon^2 I)$$

$u_k^{(3)}$ 、 $u_{jk}^{(2)}$ 、 ϵ_{jk} は独立

i	レベル 1 (データの総数: jk)	$i = 1, \dots, n_{jk}$
j	レベル 2 (グループ数: k)	$j = 1, \dots, M_k$
k	レベル 3 (グループ数: M)	$k = 1, \dots, M$
$Z_{jk}^{(3)}$	レベル 3 変量効果 $u_k^{(3)}$ の $n_{jk} \times q_3$ のデザイン行列	
$Z_{jk}^{(2)}$	レベル 2 変量効果 $u_{jk}^{(2)}$ の $n_{jk} \times q_2$ のデザイン行列	

例題 4: ランダム切片の 3 レベルモデル

- Baltagi, Song, and Jung (2001)は、アメリカ合衆国各州の私的生産における公的資本の生産性を調べる Cobb–Douglas 型生産関数を推定しました。
- 48州を9つのグループに分け、1970~1986年にわたってデータが記録されました。

データセットをインポートして内容を確認します。

```
use https://www.stata-press.com/data/r16/productivity
```

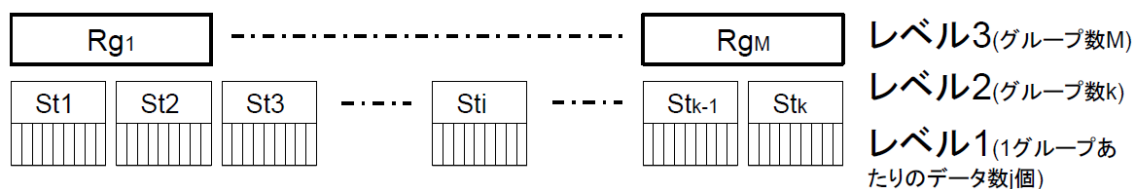
```
describe
```


Contains data from <https://www.stata-press.com/data/r16/productivity.dta>
 obs: 816 Public Capital Productivity
 vars: 11 29 Mar 2018 10:57
 (_dta has notes)

variable name	storage type	display format	value label	variable label
state	byte	%9.0g		states 1-48
region	byte	%9.0g		regions 1-9
year	int	%9.0g		years 1970-1986
public	float	%9.0g		public capital stock
hwy	float	%9.0g		log(highway component of public)
water	float	%9.0g		log(water component of public)
other	float	%9.0g		log(bldg/other component of public)
private	float	%9.0g		log(private capital stock)
gsp	float	%9.0g		log(gross state product)
emp	float	%9.0g		log(non-agriculture payrolls)
unemp	float	%9.0g		state unemployment rate

Sorted by:

データの構造は次のようになっています。



- `state` は `region` 内に含まれるため、`region` レベルと `region` 内 `state` レベルの両方で、ランダム切片を持つ 3 レベル混合モデルをフィットします。
- モデル(4)の $Z_{jk}^{(3)}$ および $Z_{jk}^{(2)}$ を $n_{jk} \times 1$ の行列に設定して使用します。 $\Sigma_3 = \sigma_3^2$ および $\Sigma_2 = \sigma_2^2$ はスカラーです。

フィットを実行します。

```
mixed gsp private emp hwy water other unemp || region: || state:
```

Performing gradient-based optimization:

Iteration 0: log likelihood = 1430.5017
 Iteration 1: log likelihood = 1430.5017

Computing standard errors:

Mixed-effects ML regression Number of obs = 816

Group Variable	No. of Groups	Observations per Group		
		Minimum	Average	Maximum
region	9	51	90.7	136
state	48	17	17.0	17

Log likelihood = 1430.5017 Wald chi2(6) = 18829.06
 Prob > chi2 = 0.0000

gsp	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
private	.2671484	.0212591	12.57	0.000	.2254814	.3088154
emp	.754072	.0261868	28.80	0.000	.7027468	.8053973
hwy	.0709767	.023041	3.08	0.002	.0258172	.1161363
water	.0761187	.0139248	5.47	0.000	.0488266	.1034109
other	-.0999955	.0169366	-5.90	0.000	-.1331906	-.0668004
unemp	-.0058983	.0009031	-6.53	0.000	-.0076684	-.0041282
_cons	2.128823	.1543854	13.79	0.000	1.826233	2.431413

Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]	
region: Identity					
	var(_cons)	.0014506	.0012995	.0002506	.0083957
state: Identity					
	var(_cons)	.0062757	.0014871	.0039442	.0099855
	var(Residual)	.0013461	.0000689	.0012176	.0014882

LR test vs. linear model: chi2(2) = 1154.73 Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

- 結果から、公的資本の高速道路と水道の要素は私的生産に有意にプラスの影響を与えたのに対し、公共施設建設その他の要素はマイナスの効果をもたらしたことがわかります。

ブロック型対角共分散構造

- モデル内の変量効果の共分散行列のオプションは表(A)の通りです。これらを組み合わせて、より複雑なブロック型対角共分散構造を生成し、分散成分に効果的に制約を課すことができます。

例題 5：繰り返しのあるレベルを使用してブロック型対角共分散構造を求める

- **region** レベルの **hwy** および **unemp** のランダム係数を考えます。これは固定効果の推定値にわずかな影響しか与えないので、分散成分に注目します。

フィットを実行します。

```
mixed gsp private emp hwy water other unemp || region: hwy unemp ||
state:, nolog nogroup nofetable
```

```
Mixed-effects ML regression          Number of obs   =          816
                                     Wald chi2(6)     =   17137.94
Log likelihood = 1447.6787           Prob > chi2     =    0.0000
```

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
region: Independent				
var(hwy)	.0000209	.0001103	6.71e-10	.6506971
var(unemp)	.0000238	.0000135	7.84e-06	.0000722
var(_cons)	.0030349	.0086684	.0000112	.8191302
state: Identity				
var(_cons)	.0063658	.0015611	.0039365	.0102943
var(Residual)	.0012469	.0000643	.001127	.0013795

```
LR test vs. linear model: chi2(4) = 1189.08          Prob > chi2 = 0.0000
```

Note: LR test is conservative and provided only for reference.

推定結果に名前を付けて保存します。

```
estimates store prodr
```

- このモデルは、モデル(4)の $Z_{jk}^{(3)}$ を $n_{jk} \times 3$ の行列に設定したものです(列は **hwy**、**unemp**、

切片項)。デフォルトの **Independent** 構造を使用するとき、 Σ_3 は次のようになります。

$$\begin{array}{c} \text{hwy} \quad \text{unemp} \quad \text{_cons} \\ \Sigma_3 = \begin{pmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_c^2 \end{pmatrix} \end{array}$$

- **state** レベルの変量効果仕様は変更ありません。 Σ_2 は **state** レベルのランダム切片のスカラ値です。
- このモデルと例題 4 のモデルの尤度比検定結果から、2つのランダム係数を含めることが支持されます。
- 推定された分散成分から、**hwy** および **unemp** のランダム係数の分散は等しいといえます。よって、 Σ_3 は次のようになります。

$$\begin{array}{c} \text{hwy} \quad \text{unemp} \quad \text{_cons} \\ \Sigma_3 = \begin{pmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_a^2 & 0 \\ 0 & 0 & \sigma_c^2 \end{pmatrix} \end{array}$$

- この構成の Σ_3 をブロック対角として扱うことにより、この等式制約を課すことができます。最初のブロックは、単位行列の 2×2 倍、つまり $\sigma_a^2 \mathbf{I}_2$ です。2つ目はスカラで、同じく単位行列の 1×1 倍です。

レベル指定を繰り返すことにより、ブロック型対角共分散を構築します。

```
mixed gsp private emp hwy water other unemp || region: hwy unemp,
cov(identity) || region: || state:, nolog nogroup nofetable
```

```

Mixed-effects ML regression          Number of obs   =       816
                                     Wald chi2(6)     =    17136.65
Log likelihood = 1447.6784           Prob > chi2     =       0.0000

```

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
region: Identity var(hwy unemp)	.0000238	.0000134	7.89e-06	.0000719
region: Identity var(_cons)	.0028191	.0030429	.0003399	.023383
state: Identity var(_cons)	.006358	.0015309	.0039661	.0101925
var(Residual)	.0012469	.0000643	.001127	.0013795

```
LR test vs. linear model: chi2(3) = 1189.08          Prob > chi2 = 0.0000
```

Note: LR test is conservative and provided only for reference.

- **region** レベルに 2 つの式を指定しました。
 1. **hwy** および **unemp** のランダム係数について、共分散を **Identity** に設定
 2. ランダム切片 **_cons** について、共分散は一次元であるため、デフォルトで **Identity** になる
- **mixed** は σ_a^2 の推定値を **var(hwy unemp)** としてラベル付けし、**hwy** と **unemp** の両方のランダム係数に共通であることを示しています。

対数尤度検定は、制約付きモデルが同等によくフィットすることを示しています。

```
lrtest . prodc
```

```

Likelihood-ratio test          LR chi2(1) =       0.00
(Assumption: _ nested in prodc) Prob > chi2 =     0.9784

```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

不等分散性変量効果

- ブロック型対角共分散構造と変量効果の反復レベル仕様を使用して、特定のレベルでの変量効果間の不等分散性をモデル化することが可能です。

例題 6：モデルの不等分散性に繰り返しのあるレベルを使用する

- 英国のコミュニティで最大 4 回、およそ 6 週間から 27 か月の年齢のアジア人の子供たちのデータを分析します。データセットは、Goldstein (1986) および Prosser, Rasbash, and Goldstein (1991)によって分析されたランダムサンプルデータです。

データセットをインポートして内容を確認します。

```
use https://www.stata-press.com/data/r16/childweight

describe
```

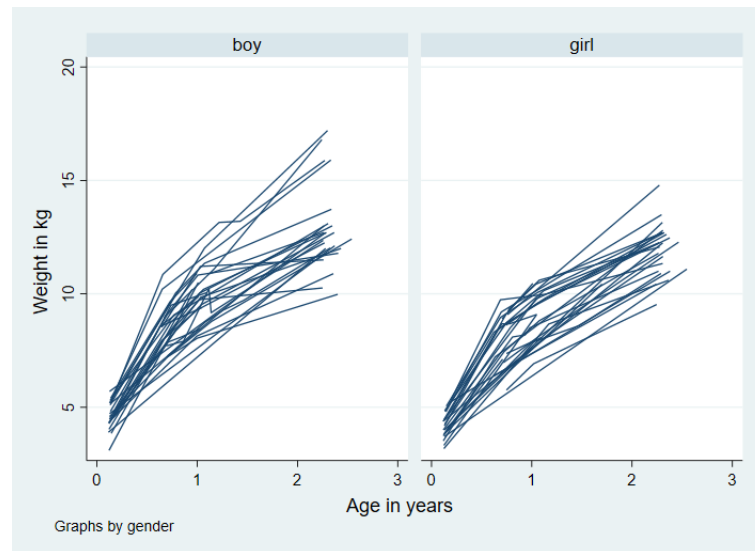
```
Contains data from https://www.stata-press.com/data/r16/childweight.dta
  obs:                198                Weight data on Asian children
  vars:                 5                23 May 2018 15:12
                                      (_dta has notes)
```

variable name	storage type	display format	value label	variable label
id	int	%8.0g		child identifier
age	float	%8.0g		age in years
weight	float	%8.0g		weight in Kg
brthwt	int	%8.0g		Birth weight in g
girl	float	%9.0g	bg	gender

Sorted by: id age

年齢と体重の関係を性別ごとにグラフ化します。

```
graph twoway (line weight age, connect(ascending)), by(girl) xtitle(Age
in years) ytitle(Weight in kg)
```



- はじめに、性別の影響を無視して、 j 番目の子供の i 番目の測定値について次のモデルを作成します。全体の平均成長を年齢の 2 次式としてモデル化し、子供固有の 2 つの変量効果を可能にします。

$$\text{weight}_{ij} = \beta_0 + \beta_1 \text{age}_{ij} + \beta_2 \text{age}_{ij}^2 + u_{j0} + u_{j1} \text{age}_{ij} + \epsilon_{ij}$$

u_{j0}	ランダム切片。全体の平均 (β_0) からの各子供の垂直シフトを表す。
u_{j1}	ランダムな年齢の傾き。全体の平均線形成長率 (β_1) からの線形成長率における各子供の偏差を表す。

フィットを実行します。

```
mixed weight age c.age#c.age || id: age, nolog
```

Mixed-effects ML regression
Group variable: id

Number of obs = 198
Number of groups = 68

Obs per group:

min = 1
avg = 2.9
max = 5

Log likelihood = -258.51915

Wald chi2(2) = 1863.46
Prob > chi2 = 0.0000

weight	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	7.693701	.2381076	32.31	0.000	7.227019	8.160384
c.age#c.age	-1.654542	.0874987	-18.91	0.000	-1.826037	-1.483048
_cons	3.497628	.1416914	24.68	0.000	3.219918	3.775338

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
id: Independent				
var(age)	.2987207	.0827569	.1735603	.5141388
var(_cons)	.5023857	.141263	.2895294	.8717298
var(Residual)	.3092897	.0474887	.2289133	.417888

LR test vs. linear model: chi2(2) = 114.70

Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

- 次に、全体的な平均成長の主要な性別効果、および性別と年齢の相互作用を含めることにより、モデルの固定部分に性別効果を導入します。
- `ibn.girl` と `noconstant` オプションを指定して、定数を省略し、男子と女子の別々の切片を推定します。
- `nofvlabel` オプションを指定して、値ラベルの代わりに変数 `girl` の値を結果に表示します。

フィットを実行します。

```
mixed weight ibn.girl i.girl#c.age c.age#c.age, noconstant nofvlabel ||
id: age, nolog
```


Mixed-effects ML regression
Group variable: id

Number of obs = 198
Number of groups = 68

Obs per group:

min = 1
avg = 2.9
max = 5

Log likelihood = -253.182

Wald chi2(5) = 6583.73
Prob > chi2 = 0.0000

weight	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
girl						
0	3.754275	.1726404	21.75	0.000	3.415906	4.092644
1	3.243808	.174255	18.62	0.000	2.902274	3.585341
girl#c.age						
0	7.806765	.2524583	30.92	0.000	7.311956	8.301574
1	7.577296	.2531318	29.93	0.000	7.081166	8.073425
c.age#c.age	-1.654323	.0871752	-18.98	0.000	-1.825183	-1.483463

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
id: Independent				
var(age)	.2772846	.0769233	.1609861	.4775987
var(_cons)	.4076892	.12386	.2247635	.7394906
var(Residual)	.3131704	.047684	.2323672	.422072

LR test vs. linear model: chi2(2) = 104.39

Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

主な性別の影響は 5%レベルで有意ですが、性別と年齢の相互作用は有意ではありません。

```
test 0.girl#c.age = 1.girl#c.age
```

(1) [weight]0bn.girl#c.age - [weight]1.girl#c.age = 0

chi2(1) = 1.66
Prob > chi2 = 0.1978

- 平均して、男子は女子より重いですが、平均線形成長率は有意に異なりません。
- 上記のモデルでは、平均成長率に性別の影響を導入しました。また、この平均からの子供固有の偏差の変動は、男子と女子で同じであると想定しました。この仮定を確認する

ために、モデルのランダム要素に性別を導入します。

フィットを実行します。

```
mixed weight ibn.girl i.girl#c.age c.age#c.age, noconstant nofvlabel ||
id: ibn.girl i.girl#c.age, noconstant nolog nofetable
```

```
Mixed-effects ML regression      Number of obs   =      198
Group variable: id              Number of groups =      68

Obs per group:
    min =      1
    avg =     2.9
    max =      5

Wald chi2(5) = 7319.20
Prob > chi2 = 0.0000

Log likelihood = -248.94752
```

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
id: Independent				
var(0.girl)	.3161091	.1557911	.1203181	.8305061
var(1.girl)	.5798676	.1959725	.2989896	1.124609
var(0.girl#age)	.4734482	.1574626	.2467028	.9085962
var(1.girl#age)	.0664634	.0553274	.0130017	.3397538
var(Residual)	.3078826	.046484	.2290188	.4139037

```
LR test vs. linear model: chi2(4) = 112.86      Prob > chi2 = 0.0000
```

Note: LR test is conservative and provided only for reference.

推定結果に名前を付けて保存します。

```
estimates store heteroskedastic
```

- 上記では、モデルの固定効果部分は、前回のモデルとあまり変わらないため、表示を抑制しています (`nofetable` オプション)。
- 前回のモデルの変量効果の仕様は次の通りです。


```
|| id: age
```

 今回のモデルでは次のような仕様になっています。


```
|| id: ibn.girl i.girl#c.age, noconstant
```
- 前者は年齢のランダム切片とランダムスロープをモデル化し、すべての子供を 1 つの

母集団からのランダムなサンプルとして扱います。後者も年齢に応じてランダムな切片とランダムな傾きを指定しますが、ランダムな切片と傾きの変動性は男子と女子の間で異なります（性別による変量効果の不等分散性）。

- **noconstant** オプションを使用すると、全体的なランダム切片を男子に固有のものと女子に固有のものに分離できます。
- これまでの結果から、線形成長率の変動には、性別が大きく影響していると考えられます。2つのモデルを尤度検定で比較します。

`lrtest homoskedastic heteroskedastic`

```
Likelihood-ratio test                    LR chi2(2) =      8.47
(Assumption: homoskedastic nested in heteroskedas~c) Prob > chi2 =    0.0145
```

Note: The reported degrees of freedom assumes the null hypothesis is not on the boundary of the parameter space. If this is not true, then the reported test is conservative.

- 結果から、不等分散性変量効果を持つ新しいモデルが支持されます。

不等分散性残差誤差

- これまでの例題では、レベル1の残差誤差（モデル内の ϵ ）は独立で、分散 σ_ϵ^2 の正規分布に従うと仮定してきました。 `mixed` コマンド実行後の変量効果表には、`var(Residual)`とラベル付けされた単一の残差誤差分散が示されています。
- 等分散性や残差誤差の独立性の仮定を緩和する際は、`residuals()`オプションを使用します。

例題7：独立残差分散構造

- West, Welch, and Gatacki (2015, chap. 7)は、歯科用セラミック被覆の配置が歯肉の健康に及ぼす影響の研究データを分析しています。データは12人の患者の55本の歯に関するものです。

データセットをインポートして内容を確認します。

```
use https://www.stata-press.com/data/r16/veneer, clear
```

```
describe
```

Contains data from <https://www.stata-press.com/data/r16/veneer.dta>

```
obs:          110          Dental veneer data
vars:          7           24 May 2018 12:11
                          (_dta has notes)
```

variable name	storage type	display format	value label	variable label
patient	byte	%8.0g		Patient ID
tooth	byte	%8.0g		Tooth number with patient
gcf	byte	%8.0g		Gingival crevicular fluid (GCF)
age	byte	%8.0g		Patient age
base_gcf	byte	%8.0g		Baseline GCF
cda	float	%9.0g		Average contour difference after veneer placement
followup	byte	%9.0g	t	Follow-up time: 3 or 6 months

Sorted by:

- 被覆は歯の元の輪郭にできるだけ一致するように配置されています。輪郭の違い（変数 `cda`）が歯肉の健康にどのように影響するか注目します。
- 歯肉の健康度の指標には各歯の歯肉溝滲出液（gingival crevicular fluid: GCF）の量を使用します。ベースライン（変数 `base_gcf`）および2回の治療後フォローアップ（3か

月後、6か月後)で測定します。

- 変数 **gcf** はフォローアップ時の GCF の記録、変数 **followup** はフォローアップタイムです。
- 患者一人につき複数の治療歯があり、また各歯について2回の測定が行われるため、次のような変量効果の3レベルモデルを作成します。
 1. 患者レベルでのフォローアップタイムのランダム切片とランダムスロープ
 2. 歯レベルでのランダム切片 k 番目の患者の j 番目の歯の i 番目の測定結果を次のようにモデル化します。

$$\text{gcf}_{ijk} = \beta_0 + \beta_1 \text{followup}_{ijk} + \beta_2 \text{base_gcf}_{ijk} + \beta_3 \text{cda}_{ijk} + \beta_4 \text{age}_{ijk} + u_{0k} + u_{1k} \text{followup}_{ijk} + v_{0k} + \epsilon_{ijk}$$

このモデルを **mixed** コマンドでフィットします。

```
mixed gcf followup base_gcf cda age || patient: followup, cov(un) ||  
tooth:, reml nolog
```