

## xtgee - 一般化推定方程式 【 評価版 】

xtgee はパネルデータを対象に一般化線形モデルのフィットを行います。

- |              |                        |
|--------------|------------------------|
| 1. 一般化線形モデル  |                        |
| 2. xtgee の用例 | Example 1<br>Example 2 |

## 1. 一般化線形モデル

xtgee はパネルデータを対象にして次のような一般化線形モデル (GLM: generalized linear models) のフィットを行います。

$$g\{E(y_{it})\} = \mathbf{x}_{it}\boldsymbol{\beta}, \quad y \sim F \text{ (パラメータ}\theta_{it}\text{を含む)} \quad (1)$$

ただし  $i = 1, \dots, m, t = 1, \dots, n_i$  とします。  $g()$  はリンク関数 (link function) を、  $F$  は分布族 (distributional family) を表します。種々の定義を  $g()$  と  $F$  に当てはめることにより多様なモデルが構成されることとなります。例えば  $y_{it}$  が正規分布に従い、  $g()$  を恒等関数 (identity function) とした場合にはモデル式は

$$E(y_{it}) = \mathbf{x}_{it}\boldsymbol{\beta}, \quad y \sim N() \quad (2)$$

となり、どのような相関構造を前提にするかによって線形回帰モデル、変量効果回帰モデル等が表現されることとなります。

$y_{it}$  がベルヌーイ分布 (二項分布) に従い、  $g()$  をロジット関数とした場合には

$$\text{logit}\{E(y_{it})\} = \mathbf{x}_{it}\boldsymbol{\beta}, \quad y \sim \text{Bernoulli} \quad (3)$$

すなわちロジスティック回帰が誘導されます。また  $y_{it}$  の分布をポアソン分布とし、  $g()$  を自然対数とした場合には

$$\ln\{E(y_{it})\} = \mathbf{x}_{it}\boldsymbol{\beta}, \quad y \sim \text{Poisson} \quad (4)$$

という形でポアソン回帰が導かれます。もちろんこれら以外の組合せも想定可能です。



xtgee は一般化推定方程式 (GEE: generalized estimating equations) という手法を用いて GLM のフィットを行います。

評価版では割愛しています。

## 2. xtgee の用例

以下においては xtgee と他の推定コマンドとの対応関係を具体例によって示します。確かに xtgee は他のコマンドを一般化するものと言えますが、推定手法には違いがあります。従って得られる結果は必ずしも一致するとは限らない点に注意してください。

### ▷ Example 1

評価版では割愛しています。

### ▷ Example 2

Example 1 では正規分布と恒等リンクの組合せで xtgee を実行したわけですが、この場合には線形回帰型のモデルが生成されます。特に独立型の相関構造を指定した場合、それは通常の OLS (ordinary least squares) 回帰と等価なものとなります。これに対し exchangeable な相関構造を選択した場合には等相関線形回帰推定法が構成されることとなります。

xtgee, family(gaussian) link(identity) corr(exchangeable) は xtreg, re によって提供される重み付き GLS 推定法、及び xtreg, mle によって提供される最尤推定法と漸近的に等価なものとなります。特にバランスのとれたデータの場合、xtgee, family(gaussian) link(identity) corr(exchangeable) と xtreg, mle とは同一の結果をもたらします。しかしバランスが取れていないデータの場合には両方でデータの扱いが異なるため、結果は近いものではあっても同一とはなりません。バランスが取れているか否かにかかわらず、xtgee, family(gaussian) link(identity) corr(exchangeable) と xtreg, mle の結果は xtreg, re とは一致しません。Example データセット nlswork2.dta のパネルデータ構成は unbalanced なものである点に注意した上で、これら 3 つの推定法を対比させてみます。結果は次の順番で掲載することになります。

- (1) xtgee
- (2) xtreg, mle
- (3) xtreg, re
- (4) xtgee, vce(robust)

### (1) xtgee

最初に xtgee をデフォルトの設定で実行します。

- Statistics ▷ Longitudinal/panel data ▷ Generalized estimating equations (GEE) ▷ Generalized estimating equations (GEE) と操作
- Model タブ: Dependent variable: ln\_wage  
Independent variables: grade age c.age#c.age  
Family and link choices: (Gaussian, Identity) (デフォルト)

- Correlation タブ: Exchangeable (デフォルト)
- Optimization タブ: Iteration log: Suppress iteration log

```
. xtgee ln_wage grade age c.age#c.age, family(gaussian) link(identity) corr(exc
> hangeable) nolog
```

GEE population-averaged model		Number of obs	=	16,085
Group variable:	idcode	Number of groups	=	3,913
Link:	identity	Obs per group:		
Family:	Gaussian	min =		1
Correlation:	exchangeable	avg =		4.1
		max =		9
		Wald chi2(3)	=	2918.26
Scale parameter:	.1416586	Prob > chi2	=	0.0000

  

ln_wage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
grade	.0717731	.00211	34.02	0.000	.0676377	.0759086
age	.1077645	.006885	15.65	0.000	.0942701	.1212589
c.age#c.age	-.0016381	.0001362	-12.03	0.000	-.001905	-.0013712
_cons	-.9480449	.0869277	-10.91	0.000	-1.11842	-.7776698

この場合にはグループ内相関は等相関、すなわち (6) 式のような相関構造を仮定してフィットが行われたわけです。なお、推定された  $\rho$  の値は `estat wcorrelation` コマンドによって確認できるわけですが、これについては [XT] `xtgee postestimation` (*mwp-166*) をご参照ください。

(2) `xtreg, mle`

評価版では割愛しています。

(3) `xtreg, re`

評価版では割愛しています。

(4) `xtgee, vce(robust)`

評価版では割愛しています。

