パネルデータ分析と離散選択モデル:講義詳細サマリー

第1部:パネルデータの基本構造

1.1 表記法

パネルデータでは、変数に2つの添字を用いる:

- **i**:個人(クロスセクション方向)、i = 1, 2, ..., N
- **t**:時点(時系列方向)、t = 1, 2, ..., T

例:Y_{it} は「第i番目の個人の第t時点におけるデータ」を表す。

1.2 基本モデル(式1)

$$Y_{it} = X_{it}'\beta_1 + Z_i'\delta + U_{it}$$

各項の説明:

- X_{it}:時間とともに変化する説明変数(iとtの両方に依存)
- Z_i:時間不変の個人特性(iのみに依存、tには依存しない)
    - 例:生まれ持った能力、性別、出身地など
    - 時点が変わっても変わらないが、個人によって異なる値をとる
- U_{it}:誤差項

1.3 パネルデータ分析の核心的問題

Z_iが観測されない場合の推定問題:

- Z_i(例:個人の生まれ持った能力)は直接観測できないことが多い
- 代理変数を用意できる場合もあるが、本質的な変数そのものは測定困難
- Xのみで回帰すると、XとZの間に相関がある場合、欠落変数バイアスが発生

パネルデータの利点:欠落変数バイアスを解消する手法が使える

第2部:固定効果モデル

2.1 モデルの書き換え（式2）

観測されないZ_iを処理するため、以下のように書き換える：

$$\alpha_i = \beta_0 + \delta' Z_i$$

すると式1は：

$$Y_{it} = X_{it}'\beta_1 + \alpha_i + U_{it}$$

ポイント：α_iはtに依存しないため、添字にtがつかない。

2.2 ダミー変数表現（式3）

個人ダミー変数$D_{1i}, D_{2i}, ..., D_{Ni}$を定義：

- $D_{1it} = 1$（i = 1のとき）、それ以外は0
- $D_{2it} = 1$（i = 2のとき）、それ以外は0
- 以下同様

すると：

$$Y_{it} = X_{it}'\beta_1 + \alpha_1 D_{1it} + \alpha_2 D_{2it} + ... + \alpha_N D_{Nit} + U_{it}$$

式**2**と式**3**は数学的に完全に同値である。

第3部：固定効果推定の2つのアプローチ

3.1 アプローチ1：固定効果変換（Within変換）

手順：

Step 1：元の式の両辺について、t方向の平均をとる

$$\bar{Y}_i = \bar{X}_i'\beta_1 + \alpha_i + \bar{U}_i$$

ここで： $$\bar{Y}i = \frac{1}{T}\sum{t=1}^{T} Y_{it}, \quad \bar{X}i = \frac{1}{T}\sum{t=1}^{T} X_{it}, \quad \bar{U}i = \frac{1}{T}\sum{t=1}^{T} U_{it}$$

重要：α_iはtに依存しないため、平均をとってもα_iのまま。

Step 2：元の式から平均を引く（式5：固定効果変換）

$$(Y_{it} - \bar{Y}i) = (X{it} - \bar{X}i)'\beta_1 + (U{it} - \bar{U}_i)$$

結果：α_iが消去される！

Step 3：変換後のデータに対してOLSを適用

この方法では切片がi依存という複雑さを回避できる。

3.2 アプローチ2：LSDV推定（ダミー変数法）

手順：

1. Excelなどでダミー変数$D_{1}, D_{2}, ..., D_{N}$を作成
2. 式3の形で直接OLSを適用

データ構造の例（Excelイメージ）：

i	t	Y	X	D_1	D_2	...
1	1	...	...	1	0	...
1	2	...	...	1	0	...
2	1	...	...	0	1	...
2	2	...	...	0	1	...

3.3 両アプローチの同値性

定理：固定効果変換によるOLS推定とLSDV推定は、β_1について完全に同じ推定値を与える。

証明の概要（集中定理を使用）：

2段階最適化の考え方：

1. 第**1**段階：βを固定して、α_iについて最小化
    - 1階条件を解くと：$\hat{\alpha}_i(\beta) = \bar{Y}_i - \bar{X}_i'\beta$
2. 第**2**段階：$\hat{\alpha}_i(\beta)$を目的関数に代入し、βについて最小化

第2段階の目的関数は式5のOLS目的関数と完全に一致する。

3.4 推定量の呼称

同じ推定量に複数の名前がある：

- 固定効果推定量（Fixed Effect Estimator, FE）
- **LSDV**推定量（Least Squares Dummy Variable）

- 群内推定量（Within Estimator）

第4部：時間固定効果

4.1 時間方向の変動のみを持つ変数

$$Y_{it} = X_{it}'\beta_1 + Z_i'\delta + S_t'\gamma + U_{it}$$

S_t：iに依存せず、tのみに依存する変数

- 例：GDP、金利、歴史的イベントなど（マクロ変数）
- 全ての個人に共通だが、時点によって異なる

4.2 時間ダミー変数表現（式12）

$$Y_{it} = X_{it}'\beta_1 + \sum_{j=1}^{N}\alpha_j D_{jit} + \sum_{s=1}^{T}\lambda_s B_{sit} + U_{it}$$

ここで：

- $B_{1t} = 1$（t = 1のとき）、それ以外は0
- $B_{2t} = 1$（t = 2のとき）、それ以外は0

4.3 時間固定効果変換（式13）

i方向に平均をとって引く：

$$(Y_{it} - \bar{Y}{\cdot t}) = (X{it} - \bar{X}{\cdot t})'\beta_1 + (U{it} - \bar{U}_{\cdot t})$$

第5部：双方向固定効果モデル

5.1 個人効果と時間効果の両方を含むモデル

$$Y_{it} = X_{it}'\beta_1 + \alpha_i + \lambda_t + U_{it}$$

5.2 ダミー変数表現の注意点

$$Y_{it} = X_{it}'\beta_1 + \sum_{j=1}^{N}\alpha_j D_{jit} + \sum_{s=2}^{T}\lambda_s B_{sit} + U_{it}$$

重要：B_{1t}を除外している理由：

- $\sum_{j=1}^{N} D_{jit} = 1$（全ての観測で）
- $\sum_{s=1}^{T} B_{sit} = 1$（全ての観測で）
- 両方を全て含めると完全多重共線性が発生
- 解決策：どれか1つを除外（ここではB_1を除外）

5.3 双方向固定効果変換（式16）

$$(Y_{it} - \bar{Y}i - \bar{Y}{\cdot t} + \bar{Y}) = (X_{it} - \bar{X}i - \bar{X}{\cdot t} + \bar{X})'\beta_1 + (U_{it} - \bar{U}i - \bar{U}{\cdot t} + \bar{U})$$

ここで$\bar{Y}$はi, t両方向の全体平均。

ポイント：$\bar{Y}$を足すことで、左辺が平均ゼロに基準化される。

第6部：固定効果推定量の統計的性質

6.1 漸近理論の設定

パネルデータには2つのサンプルサイズ：N（個人数）とT（時点数）

典型的な設定：Tを固定し、N → ∞

- 理由：Tを増やすには時間が必要（年次データなら1年で1増加）
- 多くのパネルデータはN大、T小

例外：

- 株価データ：日次で取得可能 → Tが非常に大きい
- 東証：銘柄数（N）は1000数百、Tは数千～数万

その他の漸近設定：

- Nを固定、T → ∞
- N, T両方 → ∞（収束速度が異なる場合も含む）

6.2 仮定

仮定1：強外生性

$$E[U_{it} | X_{i1}, X_{i2}, ..., X_{iT}, \alpha_i] = 0$$

弱外生性との違い:

- 弱外生性:$E[U_{it} | X_{it}] = 0$(同時点のみ条件付け)
- 強外生性:全ての時点のXで条件付け

強外生性 ⇒ 弱外生性(逆は成り立たない)

なぜ強外生性が必要か:

- 固定効果変換後の説明変数$(X_{it} - \bar{X}_i)$には全時点のXが含まれる
- $\bar{X}i = \frac{1}{T}\sum{t=1}^{T} X_{it}$なので、$X_{i1}, ..., X_{iT}$全てが入る
- U_{it}とこれらが直交するためには強外生性が必要

仮定**2**:(Y_i, X_i)はiについてi.i.d.

仮定**3**:有限の4次モーメント

仮定**4**:多重共線性がない

6.3 推定量の導出

目的関数(式6): $$\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - X_{it}'\beta_1 - \alpha_i)^2$$

1階条件を解くと:

$$\hat{\beta}1 = \left(\sum{i=1}^{N}\sum_{t=1}^{T}\tilde{X}{it}\tilde{X}{it}'\right)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{X}{it}\tilde{Y}{it}$$

ここで$\tilde{X}{it} = X{it} - \bar{X}i$、$\tilde{Y}{it} = Y_{it} - \bar{Y}_i$

6.4 一致性と漸近正規性

展開: $$\hat{\beta}1 - \beta_1 = \left(\sum{i}\sum_{t}\tilde{X}{it}\tilde{X}{it}'\right)^{-1}\sum_{i}\sum_{t}\tilde{X}{it}\tilde{U}{it}$$

漸近分布(N → ∞、T固定): $$\sqrt{N}(\hat{\beta}_1 - \beta_1) \xrightarrow{d} N(0, V)$$

漸近分散の構造: $$V = Q^{-1} \Omega Q^{-1}$$

ここで:

- $Q = \text{plim} \frac{1}{N}\sum_{i}\sum_{t}\tilde{X}{it}\tilde{X}{it}'$
- Ω:中心極限定理による分散項

6.5 分散推定:クラスター頑健標準誤差

問題：U_{it}とU_{is}（同一個人、異なる時点）に自己相関がある可能性

- 同じ個人の誤差項は相関しやすい
- 分散が均一でない（不均一分散）可能性も

対処法：クラスター頑健標準誤差（Cluster-Robust Standard Errors）

- 自己相関があっても、不均一分散があっても妥当
- Stata：`vce(robust)`オプション

注意：U_{it}とU_{is}が無相関という仮定は不自然なことが多いため、クラスター頑健標準誤差を使うのが推奨される。

重要な注意点：単純な$\hat{U}_{it}^2$を使った分散推定は不整合であることが示されている（Stock and Watson, 2008）。

第7部：動学パネルデータモデル

7.1 問題設定

経済変数の時間を通じた相関（持続性）の源泉を識別したい。

例：所得の持続性

- 今年高所得の人は来年も高所得のことが多い
- 原因は？

仮説1：個人異質性（**Heterogeneity**）

- 能力が高い人は常に高所得
- 能力は時間不変 → α_iの影響

仮説2：状態依存性（**State Dependence**）

- たまたま高い職に就くと、労働市場の硬直性で翌年も高い
- 過去の状態が現在に影響 → ラグ変数の影響

7.2 パネルAR(1)モデル（式17）

$$Y_{it} = \rho Y_{i,t-1} + \alpha_i + U_{it}$$

- **AR(1)**：Auto-Regressive of order 1（1次自己回帰）
- 説明変数が被説明変数の1期ラグ

仮説の識別：

- 異質性仮説：$\rho = 0$、α_iが異なる
- 状態依存仮説：$\rho > 0$（正の自己相関）、$\alpha_i = 0$

7.3 固定効果変換の問題

試み：通常の固定効果変換を適用

$$(Y_{it} - \bar{Y}i) = \rho(Y{i,t-1} - \bar{Y}{i,-1}) + (U{it} - \bar{U}_i)$$

問題：説明変数と誤差項が相関する

$(Y_{i,t-1} - \bar{Y}{i,-1})$の中身を展開すると： $$\bar{Y}{i,-1} = \frac{1}{T}\sum_{s=0}^{T-1} Y_{is}$$

これには$Y_{i1}, Y_{i2}, ...$が全て含まれる。

一方、Y_{it}の構造から： $$Y_{it} = \rho Y_{i,t-1} + \alpha_i + U_{it}$$

よってY_{it}はU_{it}に依存し、$Y_{i,t+1}$はU_{it}に依存し...

結論：$\bar{Y}{i,-1}$の中に必ず$U{it}$と相関する項が入る → 強外生性が構造的に成り立たない → 固定効果推定量は不一致

7.4 1階差分変換

Anderson and Hsiao (1982) のアプローチ：

式17から1期前の式を引く： $$(Y_{it} - Y_{i,t-1}) = \rho(Y_{i,t-1} - Y_{i,t-2}) + (U_{it} - U_{i,t-1})$$

α_iは消去される。

新たな問題：$(Y_{i,t-1} - Y_{i,t-2})$と$(U_{it} - U_{i,t-1})$が相関

特に$Y_{i,t-1}$と$U_{i,t-1}$が相関（同時点の変数と誤差）

→ OLSは使えない → 操作変数法が必要

7.5 操作変数の選択

候補：$Y_{i,t-2}$（2期前のY）

操作変数の条件を確認：

1. 関連性：$Y_{i,t-2}$と$(Y_{i,t-1} - Y_{i,t-2})$は相関する

    - $Y_{i,t-1} = \rho Y_{i,t-2} + \alpha_i + U_{i,t-1}$より、$Y_{i,t-2}$が含まれる
2. 外生性：$Y_{i,t-2}$と$(U_{it} - U_{i,t-1})$は無相関
    - $Y_{i,t-2}$は$t-2$期以前の情報のみに依存
    - $U_{it}, U_{i,t-1}$はt期、$t-1$期の誤差

さらに遠い過去も使える：

- $Y_{i,t-3}, Y_{i,t-4}, ...$も操作変数として有効
- ただし、あまり遠い過去は関連性が弱くなる可能性

7.6 Arellano-Bond推定量

GMM（一般化モーメント法）の枠組みで、利用可能な全ての操作変数を効率的に使用。

参考文献：Arellano and Bond (1991) - 現在でも動学パネルで最も使われる手法の一つ

7.7 先決変数（Predetermined Variables）

より一般的な状況：

X_{it}がU_{it}と相関するが、$X_{i,t-s}$（$s \geq 1$）はU_{it}と無相関

このとき、$X_{i,t-s}$を操作変数として使用可能。

動学パネルは、この考え方の特殊ケース。

第8部：離散選択モデル入門

8.1 設定

被説明変数Yが二項変数（0または1のみ）の場合

例：

- 商品を買うかどうか（1 = 購入、0 = 非購入）
- 大学に進学するかどうか
- 投票するかどうか

注：説明変数Xが二項変数の場合は、通常のOLSで問題なく扱える（係数の解釈が「$X = 1$と$X = 0$の平均差」になるだけ）。

8.2 線形確率モデル（Linear Probability Model, LPM）

発想：Yが二項変数でも、とりあえず線形モデルを当てはめる

$$Y_i = X_i'\beta + U_i$$

期待値の解釈：

Yが二項変数のとき： $$E[Y_i | X_i] = 1 \cdot P(Y_i = 1 | X_i) + 0 \cdot P(Y_i = 0 | X_i) = P(Y_i = 1 | X_i)$$

線形モデルでは： $$E[Y_i | X_i] = X_i'\beta$$

したがって： $$P(Y_i = 1 | X_i) = X_i'\beta$$

解釈：「$Y = 1$となる確率」がXの線形関数

8.3 係数の解釈

β_k：X_kを1単位増加させたときの、$Y = 1$の確率の変化

8.4 線形確率モデルの利点

1. OLSがそのまま適用可能
2. 検定もこれまで通り使える
3. 解釈が直感的

8.5 線形確率モデルの問題点

問題**1**：予測確率が**[0, 1]**区間外になりうる

Xの定義域に制限がない場合：

- Xが大きい → $X'\beta > 1$（確率が1を超える）
- Xが小さい → $X'\beta < 0$（確率が負になる）

論理的矛盾：確率は定義上、0以上1以下

問題**2**：**R^2**の解釈が困難

通常の散布図（Y連続）：

```
Y |     * *
  |   *    *
  | *   *
  |____________ X
```

二項変数の散布図：

```
Y |
1 | * * * * * *
  |
0 | * * * * * *
  |____________ X
```

- 点は$Y = 0$の線と$Y = 1$の線にしか存在しない
- $R^2 = 1$になるのは、$Y = 1$の全員が同じX値、$Y = 0$の全員が別の同じX値を持つ極端なケースのみ
- R^2はむしろXの分散の小ささを反映してしまう

8.6 次への展開

線形確率モデルの問題を解決するため、以下のモデルが使われる:

- プロビットモデル（**Probit Model**）
- ロジットモデル（**Logit Model**）

これらは確率が必ず[0, 1]区間に収まる構造を持つ。

→ 次回の講義で詳細を扱う

まとめ表

トピック	手法	目的	注意点
個人固定効果	Within変換 / LSDV	時間不変の観測されない異質性を除去	強外生性が必要
時間固定効果	時間ダミー	マクロショックを制御	-
双方向固定効果	式16の変換	両方を制御	ダミー1つ除外が必要
動学パネル	1階差分 + IV	状態依存性の推定	通常のFEは不一致
離散選択	線形確率モデル	選択確率のモデル化	予測確率が区間外になりうる

何か特定のセクションについてさらに詳しく知りたい点があれば教えてください！