

第 13 章: 単純なパネルデータ

Jeffrey Wooldridge (2016).

Introductory Econometrics: A Modern Approach
Seventh Edition. Cengage Learning.

2026-03-06

準備

必要なパッケージの読み込み

- ▶ `wooldridge` パッケージの読み込み

```
library(wooldridge)
```

- ▶ `plm` パッケージの読み込み (パネルデータ分析に使用)

```
library(plm)
```

13-1 独立した複数の横断面データの結合

独立した横断面データの結合 (Pooled Cross Sections)

- ▶ 異なる時点 ($t = 1, 2, \dots$) において、それぞれ独立に無作為抽出されたデータを結合したもの。
- ▶ サンプルサイズが大きくなるため、推定の精度（標準誤差）が向上する。
- ▶ 重要なのは、「時間とともに母集団の構造が変化しているか」を考慮すること。
- ▶ 通常、各年ダミー変数を含めて推定する。

例：賃金関数の年次変化 (cps78_85)

- ▶ 1978 年と 1985 年のデータを結合。
- ▶ モデル: $\log(wage) = \beta_0 + \delta_0 y85 + \beta_1 educ + \dots + u$
- ▶ δ_0 は、他の条件を一定としたとき、1978 年から 1985 年の間に実質賃金がどれだけ変化したかを表す。

13-2 差の差法 (Difference-in-Differences)

政策評価と差の差法 (DiD)

- ▶ 自然実験 (Natural Experiments) の評価に用いられる手法。
- ▶ 処置群 (Treatment group) と対照群 (Control group) の2つのグループを、政策の前後で比較する。
- ▶ **モデル:** $y = \beta_0 + \beta_1 \text{treated} + \delta_0 \text{after} + \delta_1 (\text{treated} \cdot \text{after}) + u$
- ▶ **DiD 推定量:** $\hat{\delta}_1 = (\bar{y}_{T,A} - \bar{y}_{T,B}) - (\bar{y}_{C,A} - \bar{y}_{C,B})$
- ▶ 平行トレンド仮定 (介入がなければ処置群・対照群の平均変化が同じ) の下で、交差項の係数 $\hat{\delta}_1$ が政策の因果的効果を表す。

DiD の例 (kielmc)

- ▶ 1978 年と 1981 年の住宅価格。
- ▶ 1981 年にゴミ焼却炉が建設される予定の地域 (nearinc) の住宅価格がどう変化したか。

```
data(kielmc)
# y81: 1981 年ダミー, nearinc: 焼却炉近隣ダミー
res_did <- lm(rprice ~ nearinc + y81 + I(nearinc * y81), data = kielmc)
summary(res_did)$coefficients[1:4, ]
```

##	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	82517.23	2726.910	30.260341	1.709246e-95
## nearinc	-18824.37	4875.322	-3.861154	1.368017e-04
## y81	18790.29	4050.065	4.639502	5.116892e-06
## I(nearinc * y81)	-11863.90	7456.646	-1.591051	1.125948e-01

13-3 2 期間のパネルデータ

パネルデータ (Panel Data) とは

- ▶ 同じ個体 (i) を、異なる時点 (t) にわたって追跡したデータ。
- ▶ 独立した横断面データの結合とは異なり、個体の固有の特性 (能力、性格、歴史など) をコントロールできる。

個体固定効果モデル

- ▶ $y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 x_{it} + a_i + u_{it}$
- ▶ a_i : **個体固有の効果** (Unobserved heterogeneity)。時間を通じて一定。
- ▶ u_{it} : 攪乱項。時間とともに変化する。
- ▶ もし a_i が x_{it} と相関している場合、通常の OLS (Pooled OLS) はバイアスが生じる。

13-4 1 階差分法 (First-Differencing)

1 階差分法 (First-Differencing, FD)

- ▶ 2 期間のパネルデータにおいて、個体固有の効果 a_i を消去する手法。
- ▶ $t = 2$ の式から $t = 1$ の式を引く：
- ▶ $(y_{i2} - y_{i1}) = \delta_0 + \beta_1(x_{i2} - x_{i1}) + (u_{i2} - u_{i1})$
- ▶ $\Delta y_i = \delta_0 + \beta_1 \Delta x_i + \Delta u_i$
- ▶ これにより時間不変の未観測効果 a_i は除去できるが、 Δx_i と Δu_i が相関しない（差分後の外生性）ことが一致性の条件である。

FD の例 (crime4)

- ▶ 郡 (county) ごとの犯罪率に対する逮捕率の影響。
- ▶ a_i は郡固有の地理的、文化的要因 (時間不変)。

```
data(crime4)
# 1981年と1982年に絞る
d8182 <- subset(crime4, year %in% c(81, 82))
# crmrte: 犯罪率 (log), prbarr: 逮捕率 (log)
# plmパッケージを使ってデータをパネル形式に変換
pdata <- pdata.frame(d8182, index = c("county", "year"))
# 1階差分モデルの推定
res_fd <- plm(crmrte ~ prbarr, data = pdata, model
  = "fd")
summary(res_fd)$coefficients
```

##	Estimate	Std. Error	t-value	Pr(> t)
## (Intercept)	-3.112476e-05	0.0004228138	-0.07361341	0.9414851
## prbarr	-5.663059e-03	0.0039700288	-1.42645298	0.1572758

まとめ

- ▶ **結合した横断面データ:** サンプルサイズの拡大と時間的变化の分析。
- ▶ **DiD:** 政策介入の効果を、処置群と対照群の「変化の差」で測定。
- ▶ **パネルデータ:** 個体固有の観測不能な要因 (a_i) を考慮できる。
- ▶ **1 階差分法:** a_i を引き算で消去し、時間不変の未観測要因によるバイアスを軽減する (因果解釈には差分後の外生性など追加条件が必要)。