

**R 入門 (5) 操作変数法**

```

setwd("/dail")
install.packages("ivreg")                                ivreg を利用可能
library(ivreg)
install.packages("sem")                                  tsls を利用可能
library(sem)
install.packages("memisc")                               mtable を利用可能
library(memisc)
dat=read.table("20.csv",header = TRUE, sep =",")
weeks 女性の労働時間 age 母年齢(21-35) ダミー変数(morekids 子供 3 人以上 samesex
子供同性 black 黒人 5% hispan メキシコ 7% othrace その他非白人 5%) p=289
attach(dat)
reg1=lm(weeks~morekids)                                 OLS
reg0=lm(morekids~samesex)                             操作変数の check
summary(reg0)
reg2=ivreg(weeks~morekids | samesex)                  IV 法
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 21.421      0.487 43.988 < 2e-16 ***
morekids     -6.314      1.275 -4.953 7.3e-07 ***
Diagnostic tests:
             df1    df2 statistic p-value
Weak instruments      1 254652 1237.219 <2e-16 ***
Wu-Hausman           1 254651   0.531   0.466
Sargan                0     NA     NA     NA
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 21.71 on 254652 degrees of freedom
Multiple R-Squared:  0.01388,   Adjusted R-squared:  0.01388
Wald test: 24.54 on 1 and 254652 DF,  p-value: 7.296e-07
reg3=tsls(weeks~morekids, ~samesex)                   2SLS 法

```

```

xhat=predict(reg0)
reg4=lm(weeks~xhat)
mtable(reg1, reg2, reg4)
summary(reg3)

```

**演習**

- 白人ダミーwhite を操作変数と労働供給関数の説明変数に追加した分析をおこないなさい。
- (1) 白人ダミーwhite を作り、summary(white)で白人の割合を求めなさい  
Hint:nwhite=black+hispan+othrace で nwhite=0 なら white=1, それ以外は=0
  - (2) 操作変数(samesex,white)の関連性をチェックしなさい(妥当な操作変数か?)
  - (3) 労働供給関数 $week_i = \beta_0 + \beta_1 morekids_i + \beta_2 white_i + \varepsilon_i$ を OLS 推定しなさい
  - (4) 操作変数(samesex と white)を用いて、労働供給関数を IV 推定しなさい
  - (5) tsls コマンドを用いて、労働供給関数を 2SLS 推定しなさい
  - (6) 1 段階目と 2 段階目の回帰をおこなって、その結果が(5)の結果と一致することを確認しなさい
  - (7) OLS 推定の結果と IV 推定の結果を比較しなさい。有意な差はみられるか?

OLS  
操作変数の check

IV 法

Week instrument
有意であれば無問題
Wu-Hausman 検定
OLS と IV に有意な差があるか?
Sargan 検定
過剰識別検定

2SLS 法

1 段階目(X の予測値)
2 段階目
結果表を整理して表示

## (1)白人の割合

```
Min. 1st Qu. Median Mean 3rd Qu. Max.
0.0000 1.0000 1.0000 0.8483 1.0000 1.0000
```

## (2)操作変数の関連性

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.437823 0.002638 165.97 <2e-16 ***
samesex 0.067752 0.001914 35.41 <2e-16 ***
white -0.107873 0.002667 -40.44 <2e-16 ***
---
Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4828 on 254651 degrees of freedom
Multiple R-squared: 0.01119, Adjusted R-squared: 0.01118
F-statistic: 1440 on 2 and 254651 DF, p-value: < 2.2e-16
```

## (3)労働供給関数の OLS 推定

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 25.11279 0.11782 213.14 <2e-16 ***
morekids -5.65977 0.08863 -63.86 <2e-16 ***
white -4.64502 0.11997 -38.72 <2e-16 ***
---
Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1

Residual standard error: 21.65 on 254651 degrees of freedom
Multiple R-squared: 0.02008, Adjusted R-squared: 0.02007
F-statistic: 2608 on 2 and 254651 DF, p-value: < 2.2e-16
```

## (4) IV 推定

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 25.3524 0.6076 41.72 <2e-16 ***
morekids -6.1677 1.2664 -4.87 1.12e-06 ***
white -4.6997 0.1813 -25.92 <2e-16 ***

Diagnostic tests:
df1 df2 statistic p-value
Weak instruments 1 254651 1253.511 <2e-16 ***
Wu-Hausman 1 254650 0.162 0.688
Sargan 0 NA NA NA
---
Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1

Residual standard error: 21.65 on 254651 degrees of freedom
Multiple R-Squared: 0.01995, Adjusted R-squared: 0.01994
Wald test: 581.3 on 2 and 254651 DF, p-value: < 2.2e-16
```

## (5)2段階最小2乗法

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 25.3524324 0.6076281 41.72360 < 2.22e-16 ***
morekids -6.1676614 1.2664408 -4.87007 1.1162e-06 ***
white -4.6996659 0.1813056 -25.92123 < 2.22e-16 ***
---
Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 21.6481518 on 254651 degrees of freedom  
(6)1段階目と2段階目の回帰の実施

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 25.3524 0.6124 41.398 < 2e-16 ***
xhat -6.1677 1.2764 -4.832 1.35e-06 ***
white -4.6997 0.1827 -25.719 < 2e-16 ***
---
Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 21.82 on 254651 degrees of freedom
Multiple R-squared: 0.004474, Adjusted R-squared: 0.004466
F-statistic: 572.2 on 2 and 254651 DF, p-value: < 2.2e-16

## (7)OLS 推定の結果と IV 推定の結果を比較

白人の割合は 84.8%

samesex, white ともに有意水準 1%で有意  
※内生性のある morekids を操作変数 samesex と white で回帰

morekids, white ともに有意に負の効果

(4)(5)(6)は(3)の労働供給関数を推計しますが、white を説明変数に加えないレポートが多い。

有意な負の効果は変わらないが、OLS にはややバイアスがある。  
Hausman 検定からも OLS と IV の差は有意ではない

2SLS と IV の結果は一致している

2段階の回帰を自分でやっても(5)と一致する

ややバイアスがあるが、有意な差はない