

ミクロデータサイエンス

Problemset3

2125178
廣江友哉

July 25, 2024

1 Step 2 回帰分析

A write_regression_models の修正

```
1 write_regression_models <- function() {  
2   regression_models <- list(  
3     # correct model  
4     "(1)" = log10(income_child) ~ effort + log10(income_parent),  
5  
6     # omitted variables models  
7     # model 2  
8     "(2)" = log10(income_child) ~ log10(income_parent),  
9     # model 3  
10    "(3)" = log10(income_child) ~ effort,  
11  
12    # measurement error models  
13    # model 4  
14    "(4)" = log10(income_child_noisy) ~ effort + log10(income_parent),  
15    # model 5  
16    "(5)" = log10(income_child) ~ effort_noisy + log10(income_parent),  
17    # model 6  
18    "(6)" = log10(income_child) ~ effort + log10(income_parent_noisy)  
19  )  
20  
21  return(regression_models)  
22 }
```

B 回帰分析表のアウトプット

Regression Table

	(1)	(2)	(3)	(4)	(5)	(6)
Own effort	1.09		1.27	1.05		1.12
	(0.17)		(0.16)	(0.17)		(0.16)
Log Parents Income	0.39	0.63		0.39	0.52	
	(0.15)	(0.14)		(0.15)	(0.15)	
Own effort with error					0.47	
					(0.12)	
Log Parents Income with error						0.35
						(0.13)
Constant	-0.68	0.27	0.92	-0.59	-0.08	-0.49
	(0.78)	(0.77)	(0.40)	(0.78)	(0.77)	(0.70)
Num.Obs.	1000	1000	1000	1000	1000	1000
Std.Errors	by: household_id by: household_id by: household_id by: household_id by: household_id by: household_id					

Heteroskedasticity robust standard errors clustered at household level

C モデル 1 の推定値の 95%信頼区間

$$1.09 - 1.96 \times SE(\hat{\beta}_1) \leq \hat{\beta}_1 \leq 1.09 + 1.96 \times SE(\hat{\beta}_1) \quad (1)$$

$$1.09 - 1.96 \times 0.17 \leq \hat{\beta}_1 \leq 1.09 + 1.96 \times 0.17 \quad (2)$$

$$0.7568 \leq \hat{\beta}_1 \leq 1.4232 \quad (3)$$

$$0.39 - 1.96 \times SE(\hat{\beta}_2) \leq \hat{\beta}_2 \leq 0.39 + 1.96 \times SE(\hat{\beta}_2) \quad (4)$$

$$0.39 - 1.96 \times 0.15 \leq \hat{\beta}_2 \leq 0.39 + 1.96 \times 0.15 \quad (5)$$

$$0.096 \leq \hat{\beta}_2 \leq 0.684 \quad (6)$$

D 帰無仮説 $\beta_1 = 0$ と $\beta_2 = 0$ をそれぞれ棄却できるか

有意水準を 5% で検定をすると、 $\beta_1 = 0$ は 95% の信頼区間から外れるため、帰無仮説は棄却される。同様に、 $\beta_2 = 0$ は 95% の信頼区間から外れるため、帰無仮説は棄却される。

E 帰無仮説 $\beta_1 = 1$ を棄却できるか

F モデル 2, モデル 5 の解釈

モデル 2 は“努力”が欠落した分、親の所得の log の係数 γ_2 は 0.63 と真のモデルの係数 β_2 の 0.39 よりも大きい値となっている。つまり、このモデルには欠落変数バイアスが存在する。

モデル 5 は“努力”に測定誤差があることで、真のモデルと比較して係数が 0.47 と低い値となっている。これは測定誤差が説明変数にある時に発生する減衰バイアスである。

G モデル 2 は必ずしも問題ではない場合

分析目的が推論ではなく、予測でモデル 2 でも十分にフィットする問題はない。

また、欠落変数がモデル 2 の誤差項とほとんど相関せず、真のモデルの β_1 の値がとても小さい場合は、そこまで問題は生じないと考えられる。