

## 宿題 2

### 機械学習

川田恵介

2025-11-05

提出方法: Posit cloud からダウンロードした ScriptFile を、Musashi 3S をアップロードする

提出物: スクリプトファイルに分析コードと、記述問題の回答をコメントの形で書き込む

提出期限: 11 月 19 日の講義開始前まで

### 1 Q1

以下の 10 事例からなるデータについて、Price の予測モデルを作成したいとする

ただし  $D$  は取引年であり、 $D = 1$  は 2023、 $D = 0$  は 2021 年に取引されていることを表す。

OLS で推定を行い、信頼区間を計算した結果は以下であった

```
library(estimatr)
```

```
lm_robust(Price ~ Reform, Data)
```

	Estimate	Std. Error	t value	Pr(> t )	CI Lower	CI Upper	DF
Price	Size						Tenure
100	65						12
98	65						16
120	60						45
34	30						18
29	20						13
94	40						3
130	65						21
58	40						7
330	95						5
200	80						23

(Intercept)	162.4	48.02666	3.381455	0.009620895	51.65033	273.14967	8
Reform	-86.2	51.52669	-1.672919	0.132881417	-205.02077	32.62077	8

```
lm_robust(Price ~ Reform + Tenure + Size, Data)
```

	Estimate	Std. Error	t value	Pr(> t )	CI Lower	CI Upper
(Intercept)	-38.7200007	42.710813	-0.9065620	0.39958041	-143.2295963	65.789595
Reform	-27.9043758	29.760697	-0.9376251	0.38462744	-100.7261780	44.917426
Tenure	-0.6707635	1.551310	-0.4323852	0.68056965	-4.4666825	3.125155
Size	3.2661720	0.992716	3.2901375	0.01661096	0.8370835	5.695261

DF

(Intercept)	6
Reform	6
Tenure	6
Size	6

1. 推定式に Size と Tenure を加えると、改築済み/未改築間の平均取引価格差が縮小した。これはなぜか、バランスというキーワードを用いながら説明せよ
2. Tenure, Size のバランスに、二重選択を活用することの利点は何か?
3. 上記で推定された信頼区間は、推定に用いたデータの性質から、「あまり信頼できない」。その理由を述べよ

## 2 Q2

講義で用いたデータを使って、 $Y = \text{Size}$ ,  $D = \text{Reform}$  を、 $\text{Tenure}, \text{District}, \text{Price}$  をバランスさせて推定せよ

- 複雑な定式化を用いて、OLS と二重選択法両方を実行し、信頼区間も計算せよ