

宿題 2

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 13 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 ~ D, Data)
```

	Estimate	Std. Error	t value	Pr(> t)	CI Lower	CI Upper	DF
(Intercept)	44.4	8.964374	4.952939	0.001116753	23.72812	65.07188	8
D	20.6	17.457377	1.180017	0.271889615	-19.65678	60.85678	8

Price	Size	D
100	75	1
98	55	1
50	35	1
55	45	1
22	30	1
53	40	0
27	25	0
49	50	0
71	75	0
22	30	0

```
lm_robust(Price ~ D + Size, Data)
```

	Estimate	Std. Error	t value	Pr(> t)	CI Lower	CI Upper	DF
(Intercept)	-11.75018	10.9743176	-1.070698	0.319827776	-37.700313	14.199962	7
D	15.49544	9.6212720	1.610539	0.151315392	-7.255255	38.246132	7
Size	1.27614	0.2741985	4.654075	0.002330679	0.627764	1.924517	7

1. 推定式に Size を加えると、2023 年と 2021 年の平均取引価格差が 20.6 から 15.5 まで縮小した。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. Size のバランスに、Double LASSO 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、 $Y = \text{Size}$, $D = \text{After}$ (または D) 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 Double Selection 両方を実行し、信頼区間も計算せよ