

補論: 不適切な統計的推論

川田恵介

Table of contents

1	多重検定問題	1
1.1	労働経済学における探索の例	2
1.2	多重検定問題	2
1.3	悪例:	2
1.4	数値例	2
1.5	Recap: 95 % 信頼区間	3
1.6	例. 多重検定	3
1.7	直感	4
1.8	Family-wise confidence interval	4
1.9	Bonferoni 法	4
1.10	例. Bonferoni 法	5
2	厳重注意: Y による分割	5
2.1	例. X による分割	6
2.2	例. Y による分割	6
2.3	“正しいモデル” の推定の前提	7
2.4	例. Y による分割	7
2.5	例. X による分割	8
	Reference	8

1 多重検定問題

- 多くの Y や X の中から、“顕著” な関係性を探したい
 - 一つ以上、明確なパラメータを発見したい
- 統計的推論の前提が崩れることに注意
 - Imbens (2021), List, Shaikh, and Xu (2019)

- Viviano, Wuthrich, and Niehaus (2021), Bowen (2022)

1.1 労働経済学における探索の例

- 複雑な現象/政策介入がなされており、大量の変数間の間での（因果的）関係性を理解したい
 - 労働条件のどのような項目が、労働供給を増加させるのか？
 - ハローワーク内でのさまざまな介入の中で、特に有効なものはどれか？
- ここまでの議論、特に信頼区間周り、は関心のある Y/D は一つであることが前提
 - X は無数にあっても良い

1.2 多重検定問題

- 複数の推定値から、大きな影響を持つものを発見しようとする場合、嚴重注意が必要
- Y と全く関係性がない/全ての負の関係性がある X から、正の関係性を持つものを誤って発見できてしまう
- ‘If you torture the data long enough, it will confess.’ (Ronald H. Coase)

1.3 悪例:

- 新聞を読む人/読まない人を比較し、どのような違いがあるのか明らかにしたい
 - “なぜか” 新聞を読む人の方が、良いことが起きると主張したい
- 「独自に収集したサーベイデータをもとに、新聞を読む人と読まない人を比較した。結果読む人の方が、“友達が多い” という結果が出た。解釈としては、、、」
 - という文言のみを発表
 - * どのくらいの特徴 ($= Y$) について推定したのか、一切記載がない
- 可能な限り事前に推定する estimand を絞り込む
 - 推定した結果はすべて公表すべき

1.4 数値例

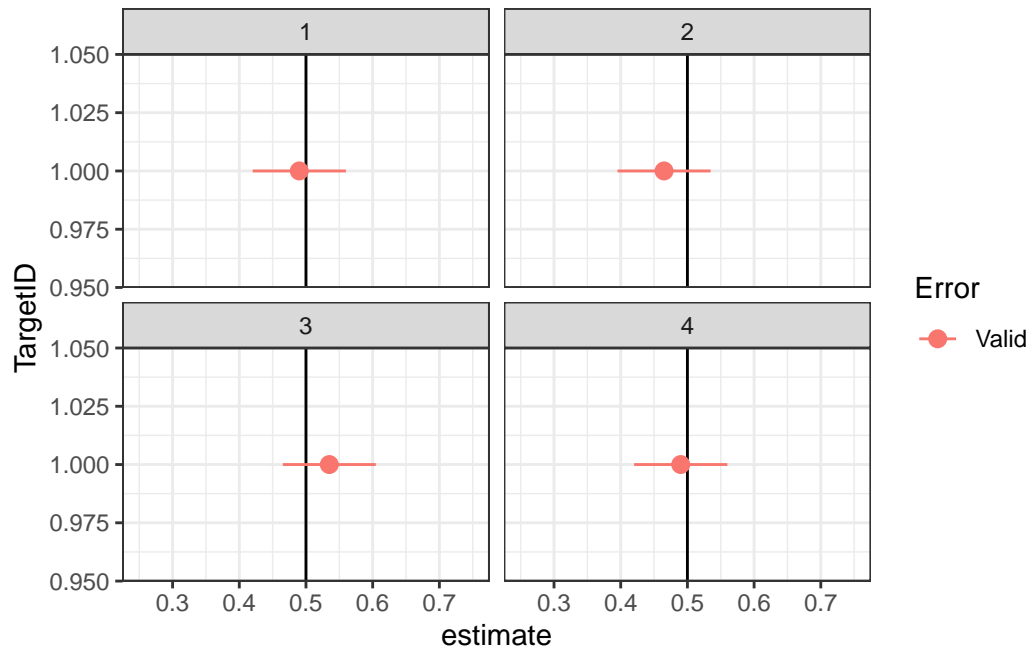
- あるプロスポーツ団体において、コイントスで先行後攻を決めている
 - 一部のチームが不正コインを使用しており、コイントスの結果が 50:50 になっていないのではないか？

* 実際には不正はない とする

- 各チームのコインを回収し、実際にコイントスを繰り返してテスト

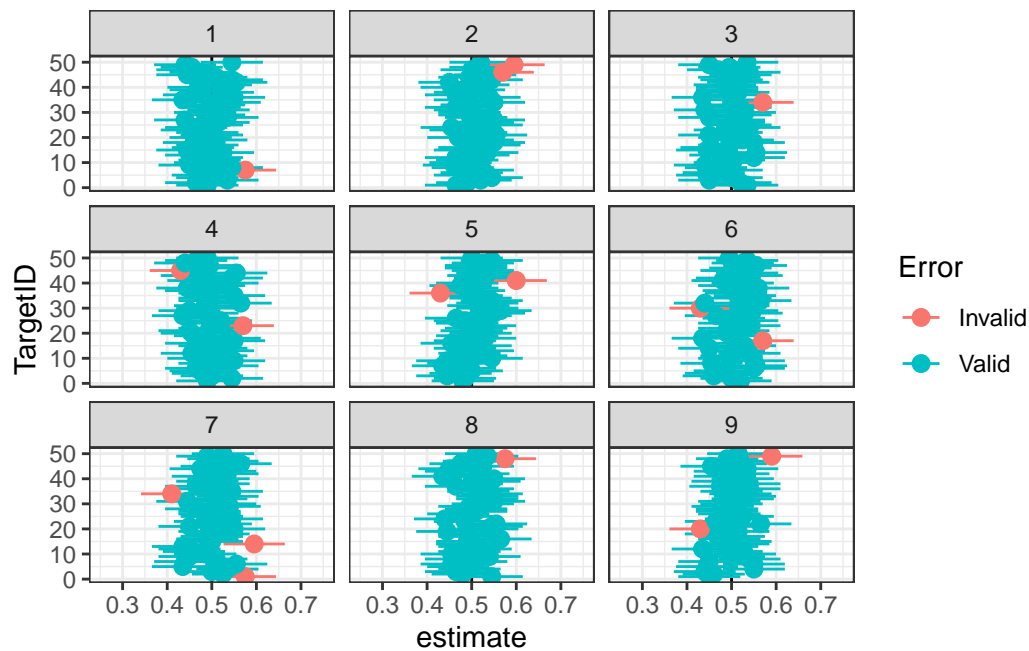
1.5 Recap: 95 % 信頼区間

- ある一つのコインについて、9 名の研究者がコイントスを 200 回行い推定した



1.6 例. 多重検定

- 50 個のコインについて、推定した



1.7 直感

- 間違っ出た明確な結果 = 目立ちたい研究者にとっての当たりくじ (5% の確率で引ける)
 - どんなに当たりにくいくじでも、無限に引けば当たる
- 何回引いた (パラメタを推定した) 結果、出たくじなのかを必ず明示する

1.8 Family-wise confidence interval

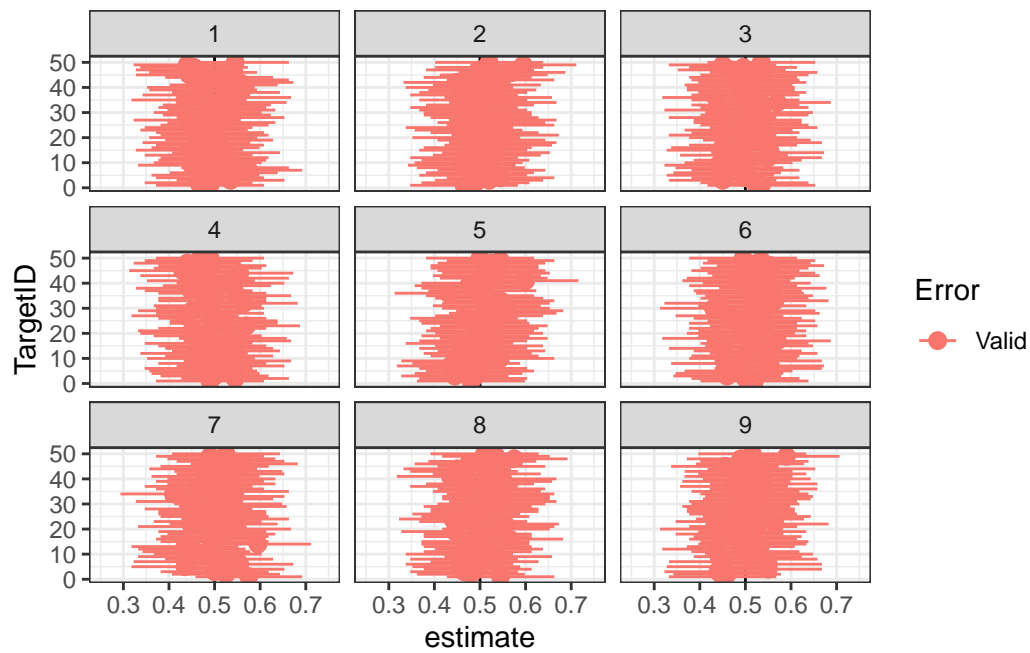
- 複数の信頼区間について、一つ以上ミスを犯す確率を一定以下 (5%) に抑える
 - \Leftrightarrow 通常の信頼区間 = 特定のパラメタについてミスを犯す確率を抑える
- Study-wise confidence interval と呼ばれる
- 信頼区間を適切に広げる形で修正
 - 広げ方について様々な提案

1.9 Bonferoni 法

- かなり保守的な修正法
- k 個のパラメタについて、一つ以上ミスを犯す確率を 5% 以下に抑えたいのであれば、 $5/k\%$ 信頼区間を計算

- パラメタが増えると無限に広がる
 - 他の方法については、[Introduction to statistical learning](#) 13 章 参照

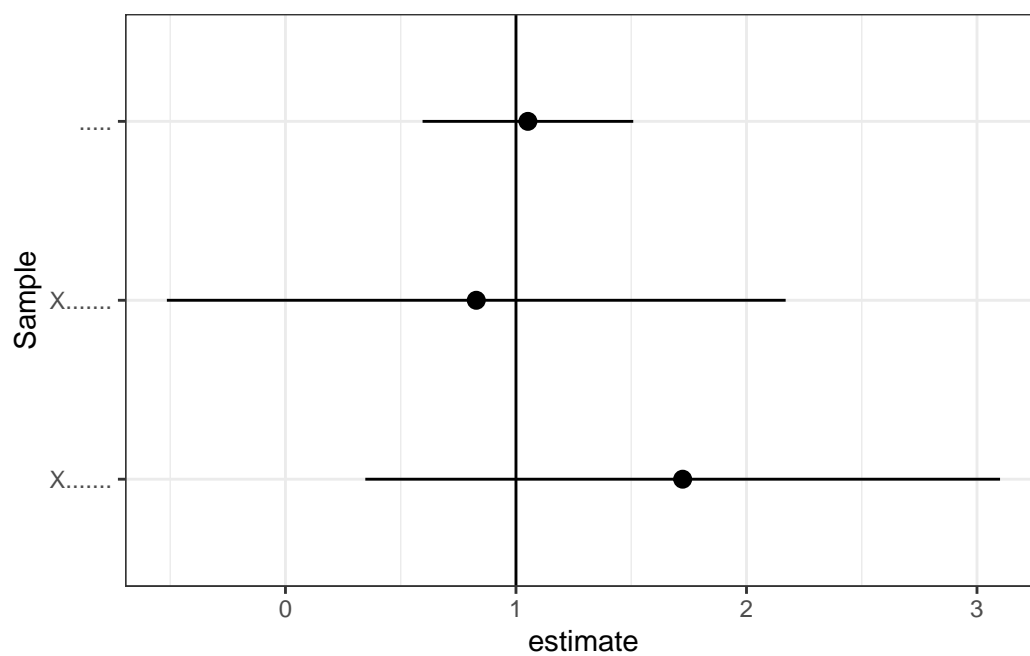
1.10 例. Bonferoni 法



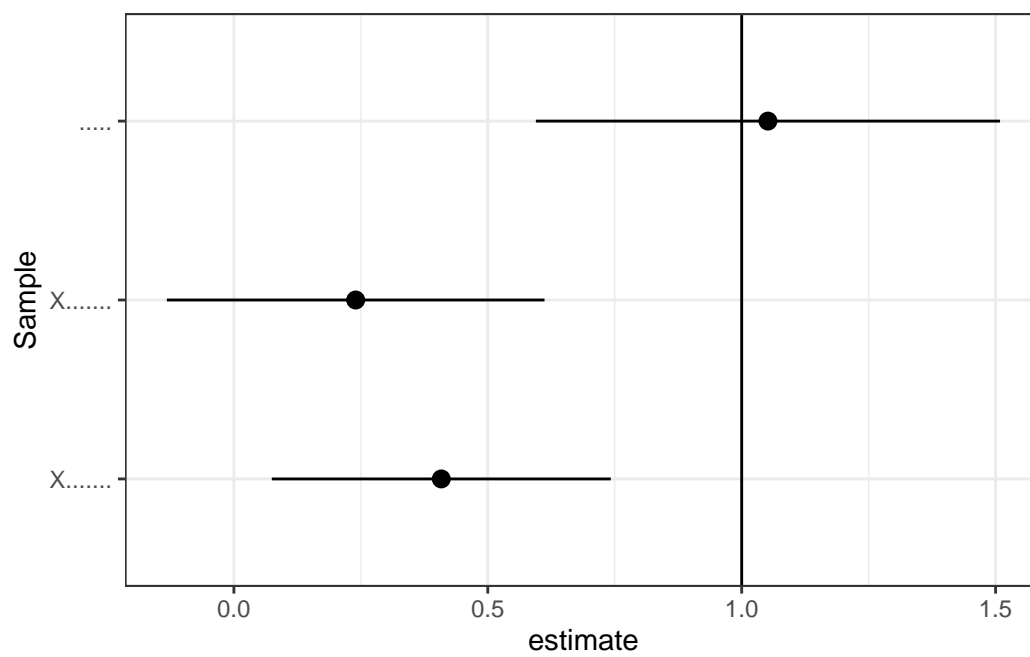
2 嚴重注意: Y による分割

- 労働経済学において、しばしば Y でサンプルを分割した、サブサンプル分析をしたくなる”誘惑”に駆られる
 - 通常の統計的推論は不可能になる
 - X による分解は問題ない
- Abadie, Chingos, and West (2018)

2.1 例. X による分割



2.2 例. Y による分割



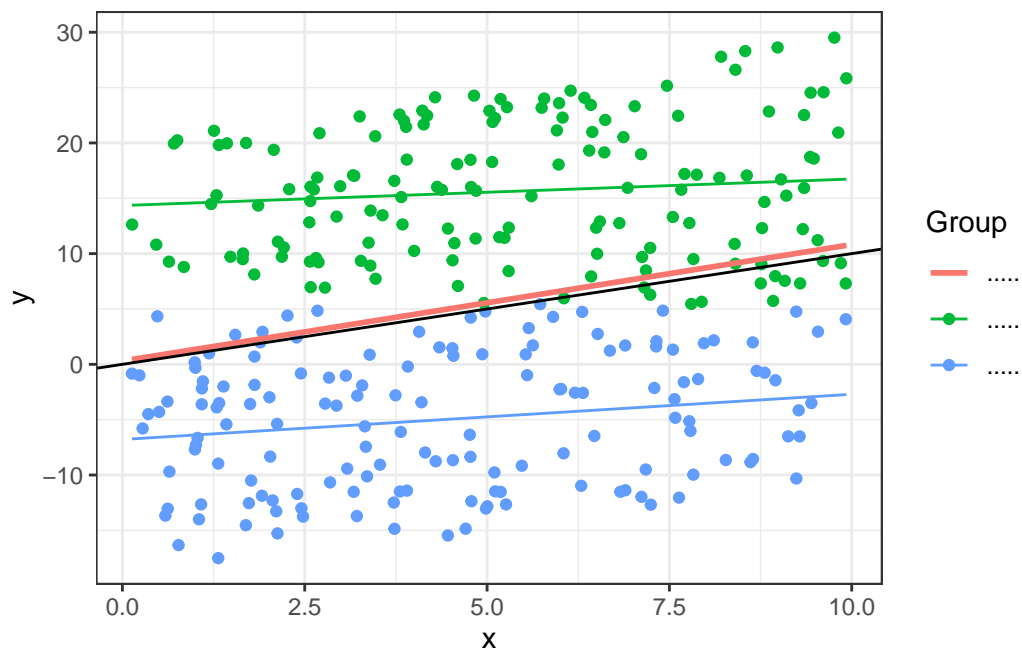
2.3 “正しいモデル” の推定的前提

- $E[Y|D] = \beta_0 + \beta_1 X$ ならば、

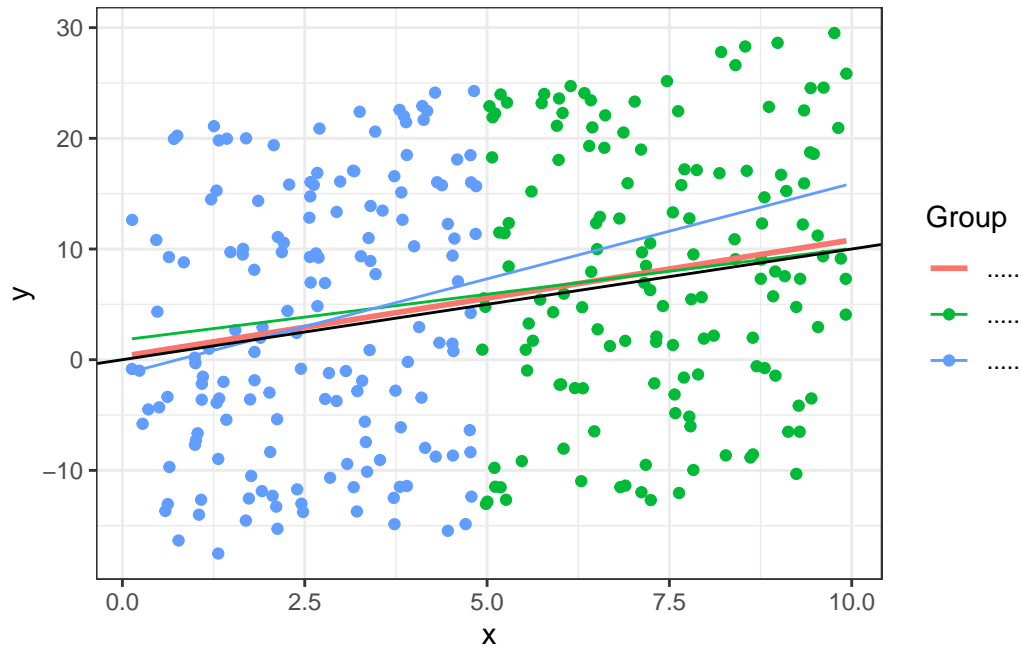
$$Y = \beta_0 + \beta_1 X + \underset{E[u|X]=0}{u}$$

- Y によって分割されると $E[u|X] \neq 0$
- X による分割では、 $E[u|X] = 0$

2.4 例. Y による分割



2.5 例. X による分割



Reference

- Abadie, Alberto, Matthew M Chingos, and Martin R West. 2018. “Endogenous Stratification in Randomized Experiments.” *Review of Economics and Statistics* 100 (4): 567–80.
- Bowen, Dillon. 2022. “A Social Scientist’s Guide to Multiple Inference.” *Available at SSRN 4182132*.
- Imbens, Guido. 2021. “Statistical Significance, p-Values, and the Reporting of Uncertainty.” *Journal of Economic Perspectives*.
- List, John A, Azeem M Shaikh, and Yang Xu. 2019. “Multiple Hypothesis Testing in Experimental Economics.” *Experimental Economics* 22: 773–93.
- Viviano, Davide, Kaspar Wuthrich, and Paul Niehaus. 2021. “(When) Should You Adjust Inferences for Multiple Hypothesis Testing?” In.