

Summary: Balanced Comparison

労働経済学 2

川田恵介

1. Balanced Comparison のまとめ

- 因果推論/格差/比較研究における中心的な手法

1.1. 目標: Balanced Comparison

- X の分布についての差を補正した後の、平均値の比較
- $\theta(d = 1) - \theta(d = 0)$
- $\theta(d) = \underbrace{E[Y \mid d, X] \times Target(X)}_{\int E[Y \mid d, X] \times Target(X) dX}$ の X についての和
- $Target(X)$: 研究者が指定する集計用荷重

1.2. Key Quantity: Balancing Weight

- $\theta(d) = E[Y \mid d, X] \times \underbrace{f(X \mid d)}_{\text{実際の分布}} \times \omega(X, d)$ の総和

を達成する荷重 $\omega(X, d)$

- 定義より

$$f(X \mid d) \omega(X, d) = Target(X)$$

$$\omega(X, d) = \frac{Target(X)}{f(X \mid d)}$$

1.3. 課題

- $\omega(X, d)$ の直接的な推定は難しい
 - ▶ 事例数に比べて、 X の組み合わせが十分に少ないケースは例外

1.4. OLS

- X の”平均値”を、Target における”平均値”と一致させる Weight
 - ▶ 研究者がモデルを変えることで、分散や共分散も揃えることができる

- 問題点は、
 - Target を研究者が決められない
 - X の D 間での分断が大きければ、負の荷重を用いてしまう

1.5. Entropy weight

- “平均値”のみをバランスさせるが、
 - Target を研究者が決められる
 - 負の荷重を用いない

1.6. 代替アプローチ

- Modelling approach (Chattopadhyay et al., 2020)
 - OLS や Entropy Weight は、Balancing Approach と呼ばれる
- 代表例は、Inverse probability weight (IPW)
 - 傾向スコア (Propensity Score) の逆数
- 理論的性質がかなり解明されており、機械学習なども活用しやすい
 - 優れた大標本性質を保証しやすい

1.7. Balancing Weight = IWP

- 条件付き分布の定義を使用 (wiki)
- $$\underbrace{\frac{f(X)}{f(X | D = d)}}_{\text{Balancing Weight}} = \frac{f(X)}{\frac{f(X, D=d)}{f(D=d)}} = \underbrace{\frac{f(D = d)}{f(D = d | X)}}_{\text{Propensity Score}}$$

1.8. 推定問題

- 傾向スコア $f(D | X)$ を推定できれば良い
 - $D = 1, 0$ であれば、 $E[D | X]$ を推定できれば良い
 - データ分析の伝統的問題であり、多くの優れた選択肢がある

1.9. Double/Debiased Machine Learning

- 機械学習の応用が進む中で、非常に緩やかな仮定のもとで、信頼できる信頼区間を推定する方法が確立される (Chernozhukov et al., 2018; 2022; Van Der Laan & Rubin, 2006)
- 入門書
 - Applied Causal Inference Powered by ML and AI
 - Causal Inference: A Statistical Learning Approach

- Introduction to Modern Causal Inference

1.10. Modelling VS Balancing

- OLS や Entropy : どこまでバランスさせるのが、研究者の選択に委ねられる
- IPW (Double/Debiased Machine Learning) : 限られた事例数のもとの実際の性能に、強い理論的保証がない (Ben-Michael et al., 2021; Chernozhukov et al., 2023)
 - 特に Overlap が不十分なケース (D の分布が偏っている X の組み合わせが存在する) 場合、推定精度が大幅に悪化する。
- 私見では、まだ未決着

1.11. Reference

Bibliography

- Ben-Michael, E., Feller, A., Hirshberg, D. A., & Zubizarreta, J. R. (2021). The balancing act in causal inference. *Arxiv Preprint Arxiv:2110.14831*.
- Chattopadhyay, A., Hase, C. H., & Zubizarreta, J. R. (2020). Balancing vs modeling approaches to weighting in practice. *Statistics in Medicine*, 39(24), 3227–3254.
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning. *The Econometrics Journal*, 21(1).
- Chernozhukov, V., Escanciano, J. C., Ichimura, H., Newey, W. K., & Robins, J. M. (2022). Locally robust semiparametric estimation. *Econometrica*, 90(4), 1501–1535.
- Chernozhukov, V., Newey, W. K., & Singh, R. (2023). A simple and general debiased machine learning theorem with finite-sample guarantees. *Biometrika*, 110(1), 257–264.
- Van Der Laan, M. J., & Rubin, D. (2006). Targeted maximum likelihood learning. *The International Journal of Biostatistics*, 2(1).