

欠測データ解析と因果推論

観察研究において治療効果を評価する際、治療変数や交絡因子を含むベースライン共変量に欠測が生じることは少なくない。共変量欠測が存在する場合に広く用いられている方法は完全データ (complete case; CC) 解析である。CC 解析は欠測割合が高い場合に推定精度が低下してしまう一方、欠測が共変量に依存する場合には、多重代入法などの MAR (missing at random) 仮定に基づく手法とは異なり、MNAR (missing not at random) 下でも妥当な推定値を与え得ることが知られている。しかしながら、CC 解析によりどのような因果効果が、いかなる仮定の下で推定されているのかについては十分に整理されていない。例えば、膀胱癌術後化学療法の効果推定した研究では、傾向スコア重み付けにより ATT (average treatment effect in the treated) が推定されているが、これは一定の条件下では CC 集団における ATT と解釈されるにとどまり、CC と非 CC 集団で効果が同様である等の強い仮定が成立しなければ、全体集団に対する ATT とは解釈できない。また近年では、効果修飾を評価するため条件付き平均処置効果 (conditional average treatment effect; CATE) を推定する研究も増加している。本研究では、条件付き効果推定を目標とする場合に、(i) CC 集団における CATE および (ii) 非 CC 集団を含む全体集団における CATE がそれぞれ識別可能となる条件を整理し、因果グラフに基づく判定の仕方について述べる。さらに周辺効果についても同様に検討し、特に欠測が共変量自体に依存する MNAR 状況下では全体集団の ATE 推定が困難であることを示した上で、いくつかの付加的な仮定のもとでの ATE の bounds を提案する。

参考文献

- Blake HA, Leyrat C, Mansfield KE, Seaman S, Tomlinson LA, Carpenter J, Williamson EJ. Propensity scores using missingness pattern information: a practical guide. *Stat Med.* 2020;39(11):1641–1657.
- Choi BY, Gelfond J. The validity of propensity score analysis using complete cases with partially observed covariates. *Eur J Epidemiol.* 2020;35(1):87–88.
- Choi J, Dekkers OM, le Cessie S. A comparison of different methods to handle missing data in the context of propensity score analysis. *Eur J Epidemiol.* 2019;34(1):23–36.
- Leyrat C, Seaman SR, White IR, et al. Propensity score analysis with partially observed covariates: How should multiple imputation be used? *Stat Methods Med Res.* 2019;28(1):3–19.
- Ross RK, Breskin A, Westreich D. When Is a Complete-case approach to missing data valid? the importance of effect-measure modification. *Am J Epidemiol.* 2020;189(12):1583–1589.