

研究デザインを活用した因果効果推定について

慶應義塾大学 経済学部／大学院経済学研究科

星野崇宏

Rubin が 1974 年に提唱したルービン因果モデルに基づく因果効果・平均介入効果の推定はと統計学や計量経済学分野において種々の推定法の開発が盛んに行われただけではなく、応用分野においても 2000 年代以降、医学や社会科学全般で爆発的に普及している。特にこの研究動向において重要なのは、統計学としてはノンパラメトリック・セミパラメトリック推定であり、また応用分野においては差分の差や回帰分断デザインなどの研究デザインの活用、補助情報・交絡要因・共変量と言われる変数をなるべく多量に取得することを事前に意図するという発想の重要性である。

特に後者はここ数年社会的に重要な話題となっているビッグデータ解析との強い関連がある。近年では企業が実務において多変量多頻度かつ大量な対象についてのデータを取得しており、場合によっては国民の数割程度をカバーする購買・行動データを取得し蓄積している例もある。また医療分野では NBD や NCD など年間数百万から数千万件の患者の診療行為やバイタルデータが蓄積されている。また外国では ID で連結可能な形で消費・所得・投資等についての情報が長期間、数百万人単位で得られている。このような場合には膨大な数の共変量が利用可能であり、操作変数法などの強い仮定を置かずに推定を行うことがより現実的であり (Athey ら 2017)、また大規模データが存在することで周辺情報を積極的に利用することで通常求められる Ignorability の仮定を置かない解析を行う方法などの適用も可能になる (Hoshino and Igari, 2017)。

報告者はこれまで応用分野での利用を中心に因果効果推定についての研究に取り組んでいるが、本報告では今回の日本統計学会業績賞受賞対象となった研究や関連する一連の研究を紹介しながら、因果効果だけではなく近年関心の高まっている因果効果の異質性の理解を行うために必要な研究デザインやデータ取得と推定法の関連について議論を行う。

参考文献

- Athey, S., Imbens, G., Pham, T., and Wager, S. (2017) Estimating Average Treatment Effects: Supplementary Analyses and Remaining Challenges. *American Economic Review*, 107(5), 278-281.
- Hoshino, T. (2013). Semiparametric Bayesian Estimation for Marginal Parametric Potential Outcome Modeling: Application to Causal Inference. *Journal of the American Statistical Association*, 108, 1189-1204.
- Hoshino, T. and Igari, R. (2017), Quasi-Bayesian Inference for Latent Variable Models with External Information: Application to Generalized Linear Mixed Models for Biased Data, *Keio-IES- discussion paper series*.
- Imbens, G.W. and Rubin, D.B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.
- Rosenbaum, P.R. and Rubin, D.B. (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70, 41-55.
- Rubin, D.B. (1974) "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies". *Journal of Educational Psychology*, 66, 688-701.
- Wagner, S., Athery, S. (in press) Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*.
- 岩崎学(2015) 統計的因果推論 朝倉書店
- 星野崇宏(2009) 調査観察データの統計科学 岩波書店