

外部周辺情報を利用した異質性のある因果効果の識別と推定

慶應義塾大学大学院経済学研究科・理化学研究所 AIP センター 高畑 圭佑
慶應義塾大学経済学部・理化学研究所 AIP センター 星野 崇宏

因果推論において通常推定の対象となるのは、Average Treatment Effects (ATE) や Average Treatment effects on the Treated (ATT) といった母集団やその部分集合上での期待値である。一方で、近年特に経済学、マーケティングや医療などの分野では、母集団で平均された値ではなく、ユニットの異質性を考慮した因果効果の推定が注目を集めている。そこで本研究では以下で定義される Heterogeneous Treatment Effects (HTE) を考える：

$$\text{HTE}(y_0) := E[y_1 - y_0 | y_0] = E[y_1 | y_0] - y_0.$$

$\text{HTE}(y_0)$ は「特殊な介入を受けなかった場合の結果変数が y_0 である人が、介入を受けた場合の因果効果の期待値」であり、 y_0 の関数となっている。

HTE は未介入時の結果変数に応じた因果効果の異質性を考慮できる一方で、観測されない結果変数への依存を考慮していることから欠測は NMAR となり、一般的にモデルの識別性は自明ではない。そこで本研究では、外部周辺情報を用いて HTE を識別するための十分条件を考える。具体的には、外部情報として未介入時の結果変数の周辺分布 $p(y_0)$ が既知であることを仮定する。これは、介入は通常関心のある母集団のごく一部に対して行われるため、介入が行われなかった場合の結果変数の分布は外部のデータベースなどから分かる場合を想定している。このとき、以下の3つの条件が識別のための十分条件となることを示す (Takahata and Hoshino, 2018)：未介入時の結果変数 (y_0) と共変量 (x) による条件付き割当て確率は、(1) y_0 と x に関する種の加法性が成立する、(2) ロジスティック回帰で特定化されている、(3) y_0 に関する項はパラメータに関して線形。

外部周辺情報を利用した NMAR モデルの識別性を扱った重要な研究として Hirano et al. (2001) があるが、彼らの結果を所与とすると、HTE の識別性はある積分方程式の解の一意性に帰着することが示される。そこで、同じく積分方程式の解の一意性が議論される Nonparametric Instrumental Variable models (NPIV) の識別性問題における完備性条件 (Newey and Powell, 2003) を援用することで、上で示した結果が得られることを示す。

さらに識別性の条件を提示するだけでなく、 $E[y_1 | y_0]$ の推定法も提案し、簡単なシミュレーションを通じてその挙動を検証する。

参考文献

1. Hirano, K., G.W. Imbens, G. Ridder, and D.B. Rubin (2001): “Combining Panel Data Sets with Attrition and Refreshment Samples,” *Econometrica*, 69, 1645-1659.
2. Takahata, K. and T. Hoshino (2018): “Identification of heterogeneous treatment effects as a function of potential untreated outcome under the nonignorable assignment condition,” *Keio-IES discussion paper series*, url: <https://ideas.repec.org/p/keo/dpaper/2018-005.html>.
3. Newey, W.K. and J.L. Powell (2003): “Instrumental Variable Estimation of Nonparametric Models,” *Econometrica*, 71, 1565-1578.