

広いデータセットに応用可能な統計モデルの研究として、セミパラメトリックな操作変数法に関する研究を行った。

近年、因果効果の推定のために、完全に外生的な情報として与えられる遺伝的情報を操作変数として用いるメンデルアン・ランダムマイゼーションが経済学や医学分野で用いられている。主に計量経済学で研究が蓄積されている既存の操作

変数法との違いは、ここでは遺伝的情報が膨大であるために多次元の操作変数が利用可能であることである。そのため、操作変数と内生変数の関係を記述する誘導方程式を正しく特定することは非常に困難であると言える。

そこで本研究では、構造方程式にはパラメトリックモデルを仮定する一方、誘導方程式にはノンパラメトリックモデルを仮定した、セミパラメトリック操作変数法を提案した。より具体的には、ノンパラメトリックモデルの表現として無限混合回帰モデルとしてディリクレ過程混合(DPM)モデルを用いる。特にここでは、DPM を柔軟な形に発展させた **probit stick-breaking process mixture (PSBPM)** モデルを用いた。このモデルでは、既存モデルのように線形回帰モデルを仮定することによる誘導方程式の誤設定を回避できるため、誘導方程式の母数に関して効率性の良い推定値を得ることが可能である。

乱数を用いたシミュレーションの結果、操作変数の次元が大きく、かつ内生変数と操作変数の関係が非線形という状況において、提案手法は既存の GMM や制限情報最尤法などと比べて、推定値の分散が小さく、効率性の高い推定量が得られることが示された。本研究の成果は、査読付き海外学術雑誌 *Stat* にて公開されている。

また、社会科学への応用研究として、位置情報に紐づく効率的な広告配信の研究を行った。先行研究ではクーポンを受け取った地点から店舗までの距離が近いと、クーポンが使用される傾向があることが明らかにされてきた。

一方本研究では、新しい観点としてクーポンを受け取った地点への親近性(Area familiarity)と、当該店舗への知識(Store knowledge)を考慮した。RCT を基に得られたデータから、これら 2つの要素はクーポンの効果に大きな影響を与えることが明らかになった。本研究の成果はディスカッションペーパーとして公開されている。

