

# 因果推論におけるセミパラメトリックアプローチのための $C_p$ 基準

二宮 嘉行

統計数理研究所，数理・推論研究系

## 概要

因果推論で重用される周辺構造モデルに対し，逆確率重み付け推定や二重頑健推定を用いるセミパラメトリックアプローチをとったときのモデル選択問題を考える．このアプローチにおいてモデリングされているのは潜在結果変数であり，それは欠測変数ともみなされるため，例えば M 推定を用いるパラメトリックアプローチに対する AIC あるいは拡張型 AIC を適用することはできない．つまり，古典的な情報量規準元来の定義に基づいて解析的に導かれた基準は存在しない．そこで，まずは潜在結果変数に対応できる平均二乗誤差として自然に考えられるものを与える．そして，セミパラメトリック推定量の漸近的性質と無視できる割り当て条件を用いてその漸近不偏推定量を導き，それを  $C_p$  基準として提案する．数値実験では，導いた  $C_p$  基準を形式的な議論で導かれた既存の情報量規準と比較し，性能が上回ることを確認する．具体的には，平均二乗誤差がいずれのシミュレーション設定においてもはっきりと小さくなること，および真の構造を含むモデルあるいはそれに近いモデルを選択する確率がはっきりと高くなることを確認する．なお，本結果は馬場崇充氏（塩野義製薬），金盛誠之氏（トーマ再保険）との共同研究によるものである．

キーワード：逆確率重み付け推定，周辺構造モデル，統計的漸近理論，二重頑健推定，モデル選択