

アブストラクト：社会科学のデータでは、閾値の存在によって変数の値が操作される可能性がある。例えば、要介護認定区分を決定するための健康スコアのヒストグラムが、閾値付近で非連続となっていることが指摘されている。また、学術誌に掲載された論文の p 値のヒストグラムにおいて、 0.05 付近で非連続性が観察される場合、いわゆる **p-hacking** の存在が疑われる。このような操作を検出する方法として、密度関数が閾値において連続であるかどうかを検定するアプローチが提案されている。しかし、既存の密度推定手法は密度関数が滑らかであることを仮定しているため、操作が行われている場合には検出力が低下する可能性がある。本研究では、有限標本においても有効に機能する新たな検定手法を開発する。さらに、シミュレーションおよび実データ分析を通じて、提案手法が既存手法と比較して高い検出力をもつことを示す。