

交絡変数の数に比してイベントが少ない 二値アウトカムデータでのリスク差の推定

【背景】二値アウトカムデータの統計解析において、交絡変数の候補数に比してイベントが少ない状況に出くわすことはよくある。このような状況では、傾向スコアを用いた統計解析方針が推奨されることが多い。はたしてこの解析方針は、関心のある効果指標がリスク差の場合にも当てはまるのだろうか。

【目的】調整したい交絡変数の数に比してイベントが少ない二値アウトカムデータにおいて、アウトカムモデルを用いてリスク差を推定する方法を検討する。

【方法】リスク差の推定方法として、ロジスティック回帰を用いた回帰標準化、線型二項回帰、modified least squares 回帰について既知の事項をまとめた。modified least squares 回帰については、1 交絡変数当たりのイベント数とリスク差の推定性能との関係をシミュレーション実験で評価した。modified least squares 回帰、構造平均モデルの g 推定法、Mantel-Haenszel リスク差推定量の関係を整理した。実データ解析として、線型二項回帰と modified least squares 回帰を用いて、早産および出生後 1 年以内の乳幼児死亡に対する不適切な妊婦健診と適切な妊婦健診を比較したリスク差を推定した。

【結果】modified least squares 回帰は、交絡変数の数に比してイベントが少ない場合のリスク差の推定法として、魅力的な理論的性質を備えていた。シミュレーション実験の範囲内では、曝露群と非曝露群での期待イベント数がともに 5 イベントを上回っていれば、1 交絡変数当たりのイベント数と関係なく、modified least squares 回帰は妥当な点推定値と 95% 信頼区間を与えていた。実データ解析では、線型二項回帰が収束しなかったデータにおいても、modified least squares 回帰により、他の知見と整合するリスク差の推定値を得ることができた。

【結論】交絡変数の数に比してイベントが少ない二値アウトカムデータからリスク差を推定する手法として、modified least squares 回帰は有用である。

【主要文献】

1. Cheung YB. A modified least-squares regression approach to the estimation of risk difference. *Am J Epidemiol.* 2007;166:1337-44.
2. Greenland S. Introduction to regression modeling. In Rothman KJ, Greenland S, Lash TL. *Modern epidemiology.* 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008
3. Cepeda MS, Boston R, Farrar JT, et al. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol.* 2003;158:280-87.