

生存時間解析における多重補完法のバイアス補正

【背景】 多重補完法において補完値のサンプリングに用いられる補完モデルは、曝露やイベント、共変量などの同時分布に基づく必要があり、その誤特定が生じた場合に補完後の推定値やその分散推定値にバイアスを生じることが知られている。しかし、一般的な多重補完法ではシンプルなパラメトリック補完モデルが用いられるため、程度の差はあれ、多くの場合で補完モデルは誤特定されていると考えられる。それにもかかわらず、欠測共変量を伴うデータを利用して生存時間解析を行う多くの疫学研究において、補完モデルの誤特定によるバイアスに特段の注意を払うことなく、多重補完法が漫然と使用されている。

【目的】 多重補完法を用いて生存時間解析における一致性のある推定量を得ることを目的として、ノンパラメトリック条件付確率推定法を利用した、補完モデルの明示的な特定を必要としない多重補完法の手順を提案する。

【方法】 共変量の一部が欠測した医療情報データを用いて、医薬品使用開始から有害事象発生までの生存時間解析を行うことを想定する。カテゴリ変数が欠測する場合、連続変数とカテゴリ変数の両者が欠測する場合のそれぞれについて、提案する多重補完法の理論的背景を整理し、その手順を示した。人工的に生成したデータおよび実際の薬剤疫学研究データに基づくシミュレーション実験を行い、提案法の性能評価を行った。

【結果】 提案法により、完全データ解析および補完モデルの誤特定に起因する推定値のバイアスが軽減された。

文献

Wang N, Robins JM. Large-sample theory for parametric multiple imputation procedures. *Biometrika*. 1998; 85(4), 935-948.

von Hippel PT. New confidence intervals and bias comparisons show that maximum likelihood can beat multiple imputation in small samples. *Struct Equ Modeling*. 2015; 23(3), 422-437.

Tomita H, Fujisawa H, Henmi M. A bias-corrected estimator in multiple imputation for missing data. *Stat Med*. 2018; 37(23), 3373-3386.