

項目反応理論における欠測データに対する頑健性の比較検討 —未到達項目と省略項目の2つの欠測を考慮した数値的比較—

大杉 勇心

近年、テストを紙とペンではなくコンピュータで行うコンピュータベーステストが注目されている。コンピュータベーステストには様々な利点があり、その利点の1つにその項目を解こうとしたかどうかの意思の有無を把握することができる、というものがある。これにより従来の紙のテストでは不可能だった、欠測付きテストデータの欠測理由を2つに限定することが可能になった。1つが問題を解こうとしたうえでスキップした省略項目、もう1つが問題にたどり着かなかった未到達項目である。

テストの解析を行う際には項目反応理論と呼ばれる手法がよく用いられる。この手法ではまず、テストを受ける受験者の能力値とテストを構成する各項目に識別力や困難度などのパラメタを考える。前者は受験者パラメタ、後者は項目パラメタと呼ばれる。項目反応理論では、項目の正答確率を受験者パラメタと項目パラメタのロジスティック関数で表現することで、受験者やそのテスト自体の評価を行うことができる。

本論文では省略項目と未到達項目という2つの欠測に注目し、項目反応理論を用いて項目パラメタを推定したときの精度を欠測対処法ごとに比較した。欠測対処法には、欠測値に不正解を代入する方法、欠測に観測された項目平均を代入する方法、多重代入法、EM アルゴリズムを用いた代入法の4種類を採用した。項目反応理論のモデル式には2パラメタロジスティックモデルを採用した。このモデルでは項目パラメタに項目識別力と項目困難度の2種類が存在し、前者はその項目がどれだけ受験者の能力を区別するのに優れているかを示す変数で、後者はその項目がどれだけ正答するのが難しいかを示す変数である。

シミュレーションでは生成した受験者能力と項目パラメタから完全データを生成し、それに受験者の欠測傾向や受験者能力、項目パラメタから2種の欠測が含まれる欠測データを生成した。その欠測データを解析し、項目パラメタを推定し、真のパラメタとのバイアスを求めた。これを受験者数や省略項目の割合、未到達項目の割合という条件をそれぞれ変化させながら測定した。

結果としては、サンプルサイズや種々の欠測割合を問わず安定した精度を示したのは多重代入法を用いた推定であった。今回比較した他の3つの手法に比べて平均二乗誤差は低く、項目パラメタのバイアスは常に負の値をとることが特徴であった。また、欠測を不正解扱いする代入法もある程度の精度を保っており、こちらは逆に項目パラメタのバイアスを過大推定する特徴があった。

今後の課題としては、受験者数や欠測割合以外の条件を変化させたときの推定精度の比較や、提案されている他の欠測対処法を用いたときの推定精度についても確認する必要がある。(行動統計科学)