

項目間に依存性があるときの ニューラルテスト理論の適用可能性について

安井彩乃 (名古屋大学)

問題と目的

ニューラルテスト理論 (Neural Test Theory: NTT) (荘島, 2007) は, 受検者を順序尺度である潜在ランクに割り当てるという特徴をもつテスト理論である。項目応答理論 (Item Response Theory: IRT) との主な違いは, 潜在特性が順序尺度か間隔尺度か, また, モデルがノンパラメトリックモデルであるかパラメトリックモデルであるかだという (荘島, 2009)。

本研究では, 項目間に依存性をもつテストと独立なテストの推定について, IRT と NTT を用いて比較を行う。これにより, 項目間に依存性をもつテストへの NTT の適用可能性を検討する。

方法

項目反応パタンの発生モデル

誤答の後の正答を許さないという項目間に依存性をもつテストと, 項目間で独立なテストという 2 つのテストを仮定した。項目間に依存性をもつ条件では, 小問 3 項目からなる大問 10 題, 計 30 項目のテストを, 項目間で独立な条件では, 30 項目のテストを仮定した。個人の潜在特性真値に応じて項目反応パタンの確率を決定する。

項目反応データは以下に述べる 2PLM を用いて発生させた。

$$P_j(\theta_i) = \frac{1}{1 + \exp[-1.7a_j(\theta_i - b_j)]} \quad (1)$$

$$Q_j(\theta_i) = 1 - P_j(\theta_i) \quad (2)$$

ここで, θ_i は個人の潜在特性値を, a_j は項目 j の識別力母数を, b_j は項目 j の困難度母数を, それぞれ表している。

識別力母数と困難度母数

項目間に依存性をもつテストと独立なテストで識別力母数と困難度母数は同一のものとした。

識別力母数は, 全ての項目で一様乱数 $U(0.3, 0.9)$ から設定した。

困難度母数は, 小問 3 項目を一つのセットとし, b_1 を $U(-1.5, -0.5)$, b_2 を $U(-0.5, 0.5)$, b_3 を $U(0.5, 1.5)$ として, 3 項目間で困難度が徐々に上がる範囲で一様乱数を用いて設定した。

シミュレーションの手続き

Step1, 受検者の潜在特性値 θ_i を標準正規分布 $N(0, 1)$ から 1000 名分発生させ, θ_i に応じて (1), (2) に基いて項目反応パタンの確率を求めた。

Step2, 解答パターン作成のため, 一様分布 $U(0, 1)$ を発生させ, Step1 で求めた確率に応じて 1000 名分の項目反応パターンを作成し, 1000×3 の回答パターン行列を作成した。また, それを 10 回繰り返した。
Step3, こうして発生させたデータにもとづいて, 潜在特性値と潜在ランクの推定を行った。
Step4, Step1 ~ Step3 について 100 回繰り返した。

用いる指標

潜在特性の真値と潜在特性推定値, 潜在ランク推定値を比較するための指標として, 宇佐美 (2009) を参考に, 相関係数の平均を用いた。

結果

結果を Table1 に示した。

潜在ランク, 潜在特性値のどちらも, 依存性をもつテストで独立なテストよりも真値との相関係数の平均値が低くなった。

Table1 真値との相関係数の平均値

潜在ランク		潜在特性値	
依存	独立	依存	独立
.749	.890	.890	.921

考察

今回, 独立なテストと依存性をもつテストの間で, 依存性をもつテストの方が相関係数の平均値が小さくなった。また, NTT では, IRT よりもその傾向が顕著となった。そのため, 項目間の依存性を無視して NTT を適用することは難しいと考えられる。しかし, 項目を積み上げるテストレットモデル (Wainer & Kiely, 1987) を適用することで, 依存性をもつ項目であっても IRT での推定精度を高められることが知られている。そのため, NTT でもテストレットモデルを導入するなど検討を重ね, 依存性をもつテストへの適用可能性の有無を吟味する必要があるだろう。