

カリキュラムに基づいたテストの計量的分析の試み

○ 坂本佑太郎（東北大学大学院）

柴山 直（東北大学大学院）

目的

近年、テストを支える技術的基盤としての項目反応理論（Item Response theory, IRT）はわが国でも注目されつつある。IRT では項目困難度をひとつのパラメータ b_j で表現し、項目固有の特徴を示す。しかしながら、項目困難度 b_j の推定値についてその要因がどこにあるのかに関してはテスト実施側や分析者側の推測に頼らざるを得なかった。そこで本研究では、ラッシュモデルを拡張した線形ロジスティックテストモデル（Linear Logistic Test Model, LLTM）（Fischer, 1973, 1983）をカリキュラムベースに作成されたテストに適用することによって、困難度の要因はどこにあるのかを明らかにする。

方法

データ 平成 18 年度新潟県全県学力調査の内、中学校 2 年生を対象に行われた数学に関する項目数 25 のテストにおける項目 1 から項目 10 の「式と計算」領域から、全問正答/誤答及び欠測値を含む受検者を除いた 9,102 名分のデータを使用した。
線形ロジスティックテストモデル ラッシュモデル；

$$P(u_{ij}|\theta_i, b_j) = \frac{\exp[u_{ij}(\theta_i - b_j)]}{1 + \exp(\theta_i - b_j)},$$

における項目困難度 b_j を、

$$b_j = \sum_{l=1}^p w_{jl}\alpha_l + c,$$

と分解する。このとき、 α_l は l 番目の測定領域の困難度を表す基本母数(basic parameters)、 w_{jl} は j 番目の項目の基本母数に対する重み(行列)、そして c は基準化のための定数である。当時の学習指導要領（文部科学省, 2006）を参考に Design Matrix W を作成した。なお分析には R3.0.2, パッケージ eRm を使用した。

結果と考察

Table1 は推定された基本母数 α_l を示しており、測定領域固有の困難度が同一尺度上で表現されていることがわかる。また、モデルの適合度の確認のために、Design Matrix W にしたがって該当する基本母数 α_l を合計した上でセンタリングした LLTM における項目困難度とラッシュモデルにおける項目困難度をプロットしたものが Figure1 である。相関係数は.997 であり、妥当な推定値が得られていることが確認できる。

Table1 基本母数 α_l の推定値

基本母数	推定値	SE	信頼下限	信頼上限
α_1	-1.41	0.06	-1.29	-1.53
α_2	3.08	0.05	3.17	2.99
α_3	1.65	0.05	1.74	1.55
α_4	1.37	0.05	1.47	1.27
α_5	0.10	0.04	0.17	0.03
α_6	1.66	0.04	1.74	1.57
α_7	1.42	0.04	1.51	1.34
α_8	3.88	0.05	3.97	3.79

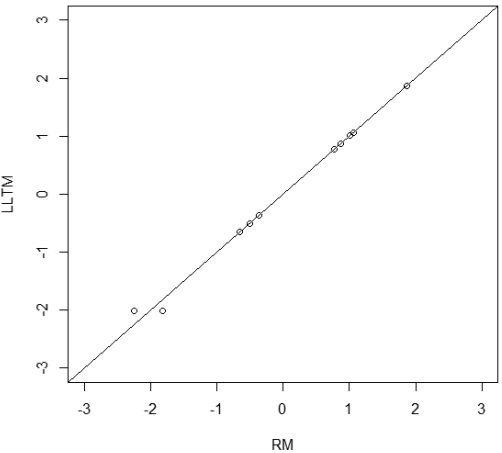


Figure1 ラッシュモデルと LLTM の推定値比較
【謝辞】：本研究は JSPS 科研費 25380867 の助成を受けました。