

学力テストの下位領域に関する IRT を使った測定論的特徴の把握

○坂本佑太郎 (株式会社リクルートマネジメントソリューションズ)

柴山 直 (東北大学大学院)

キーワード：学力テスト, IRT, 双因子モデル

問題と目的

大学入学者選抜において従来の知識・技能の測定だけにとどまらず、「思考力・判断力・表現力」をも測定することが目指される(文部科学省, 2016)ことに象徴されるように、わが国においてもテストが測定すべき心理学的特性が多様化してきているのが現状である。その中で、IRT に代表される測定技術を単に分析の道具として用いるだけではなく、その前提として「テストがそもそも何を測っているのか」という構成概念についての精緻な検証が求められている(石井, 2014)。そこで本研究では、IRT における項目情報量的な観点から、学力テストの下位領域に関する計量的な特徴を明らかにすることを目的とする。

方 法

平成18年度新潟県全県学力調査における中学校2年生数学(N=9,102)を用いる。項目数は25であり、回答は正答/誤答の2値である。測定領域は、項目1から項目10は「数と式」、項目11から項目16は「図形」、そして項目17から項目25は「数量関係」であることが定められている(新潟県教育委員会, 2007; 文部科学省, 2006)。

本研究では補償型(compensatory)多次元2値2PLモデル

$$P(u_{ij} = 1 | \theta_i, \mathbf{a}_j, d_j) = \frac{\exp(\mathbf{a}_j \theta'_i + d_j)}{1 + \exp(\mathbf{a}_j \theta'_i + d_j)} \quad (1)$$

を用いる。このとき u_{ij} は受検者 i の項目 j に対する反応を示し、また次元数を m とすると \mathbf{a}_j は $1 \times m$ の項目 j の識別力パラメータベクトル、 θ_i は $1 \times m$ の受検者 i の潜在特性尺度値ベクトル、 d_j は困難度に関連するパラメータ(スカラー)を示している。今回はテスト全体が測定する構成概念(「数学力」と)、下位領域特有の情報量を計量的に比較するため、(1)を上記で示した測定領域別に確認的にモデリングした双因子モデル(bifactor model; Gibbons & Hedeker, 1992)を仮説として設定する。双因子モデル内で推定された「数学力」と下位領域ごとに推定された項目識別力パラメータの2乗値を比較することで、その相対的な項目情報量の差が求められる(Reckase, 2009; Reise, Morizot & Hays, 2007)、当該の項目が「数学力」あるいは下位領域特有の影響のどちらを強く受けているかを明らかにできる。

結 果

Table 1には双因子モデル内における項目識別力パラメータの2乗値を整理した。その比をとれば、item24, 25が測定する下位領域「数量関係」はテスト全体が測定する「数学力」のそれぞれ約1.333倍、約1.094倍の情報量を保有していることがわかる。

Table 1 双因子モデル内における項目識別力パラメータの2乗値の比較

数と式	下位領域の a^2	「数学力」の a^2	図形	数量関係				
item1	2.176	3.629	item11	1.309	1.496	item17	0.022	2.487
item2	0.746	2.190	item12	0.202	1.503	item18	0.091	2.280
item3	0.640	1.057	item13	1.175	3.312	item19	0.154	1.687
item4	0.658	1.323	item14	0.097	1.184	item20	0.208	3.021
item5	0.584	2.965	item15	5.359	7.177	item21	0.004	1.195
item6	1.138	3.244	item16	0.315	3.229	item22	0.001	0.714
item7	0.216	2.005				item23	0.112	2.550
item8	0.462	1.659				item24	3.463	2.599
item9	0.460	2.022				item25	2.796	2.554
item10	0.007	1.488						

また、各下位領域と「数学力」の潜在特性尺度値をEAP法によって推定し、それらをプロットしたものがFigure 1である。各下位領域と数学力との相関係数はそれぞれ0.366, 0.384, 0.279である。

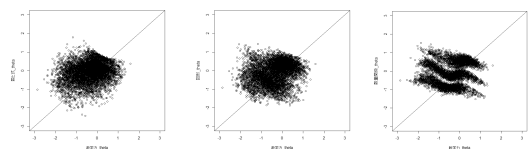


Figure 1 各下位領域と「数学力」における潜在特性尺度値の関係

これらの結果から、多次元IRTを基盤とした双因子モデルにより学力テストの測定領域についての測定論的な情報を項目レベルで把握できることが明らかになった。今後は領域別の潜在特性尺度値を従属変数にした教育社会学的分析や多値データに対する多次元段階反応モデル(Muraki & Carlson, 1995)への拡張が期待される。

【謝辞】本研究の遂行にあたってはJSPS 科研費25380867ならびに16H03731の助成を受けました。