

項目反応理論に基づいた学習評価支援システムにおける 能力可視化機能の検証

Verification of the Ability Visualization Function in an Assessment Support System Based on Item Response Theory

蛸名 哲也^{*1}, 宮澤 芳光^{*1}, 森本 康彦^{*1}

Tetsuya EBINA^{*1}, Yoshimitsu MIYAZAWA^{*1}, Yasuhiko MORIMOTO^{*1}

^{*1}東京学芸大学

^{*1}Tokyo Gakuji University

あらまし：本研究では、あらゆる学びの記録である学習記録データ(e ポートフォリオ)を用いた、継続的な学習評価の実現を目的として、学習評価支援システムを開発してきた。具体的には、タブレット端末等の ICT を活用して学習記録データを蓄積し、この学習記録データを用いて、項目反応理論に基づき学習者の資質・能力を測定・可視化する。本論文では、能力値推定プログラムについて、シミュレーションの結果及び考察について述べる。

キーワード：項目反応理論, 学習評価, 学習記録データ, e ポートフォリオ

1. はじめに

近年、学習者の継続的な学習プロセスを通して、21 世紀型スキル、学士力、社会人基礎力に代表される「テストだけでは測ることのできない資質・能力」についての学習評価が強く求められるようになった。この学習評価においては、学習の結果だけではなく、学習過程が重視される。学習過程を評価するためには、学習者が生成するあらゆる学びの記録である学習記録データ(e ポートフォリオ)を用いて、学習者のパフォーマンスを評価する必要がある。しかし、現状では、学習記録データを収集・蓄積し、学習評価に利活用する具体的な方法は確立されていない。また、学習者のパフォーマンス中に発揮された資質・能力を評価するためには、学習過程全体を通したパフォーマンスの評価が求められる。そのため、学習者によって生成される異なる学習での学習記録データから、継続的に資質・能力を測定することが求められる。

そこで、本研究では、学習記録データを用いた、継続的な学習評価の実現を目的とする。具体的には、ICT を活用して学習記録データを蓄積し、その学習記録データを用いて、項目反応理論に基づき学習者の資質・能力を同一の尺度で測定し、学習状況を可視化する。筆者らはこれまで、学習記録データを用いた、学習評価支援システムの開発を行ってきた⁽¹⁾。本論文では、開発したシステムにおいて、能力可視化機能の検証をするために、シミュレーションを行った。以下、このシミュレーションの結果及び考察について述べる。

2. 項目反応理論に基づいた資質・能力の測定

本研究では、学習記録データから読み取れる学習者のパフォーマンスを、テスト、チェックリスト、ルーブリックに基づいて学習評価を行う。特に、本研究では、ルーブリックを用いた学習評価に焦点を

当てる。ここでは、ルーブリックの評価基準に応じた点数を反応データとする。学習者の資質・能力を測定するため、テスト理論である項目反応理論を用いる。項目反応理論では、異なる項目から構成されたテストの結果を比較できる⁽²⁾。本研究では、学習者によって生成される異なる学習での学習記録データから、学習者の能力値を推定することが可能になるため、学習者の資質・能力を継続的に能力値として見る事が可能になる。

本研究では、多値の反応データを扱うことを想定しているため、項目反応理論における段階反応モデルを適用する⁽³⁾。段階反応モデルは、学習者 j が項目 i に対してカテゴリ $k(k = 1, \dots, K)$ と反応する確率 P_{ijk} を次式で与える。

$$P_{ijk} = P_{ijk-1}^* - P_{ijk}^* \quad (1)$$

$$\begin{cases} P_{ijk}^* = \frac{1}{1 + \exp(-a_i(\theta_j - b_{ik}))} & k = 1, \dots, K-1 \\ P_{ij0}^* = 1 \\ P_{ijK}^* = 0 \end{cases}$$

ここで、 a_i は項目 i の識別力パラメータ、 b_{ik} は項目 i のカテゴリ k の難易度パラメータ、 θ_j は学習者 j の能力パラメータを表す。ただし、 $b_{i1} < b_{i2} < b_{i3} < \dots < b_{iK} < \dots < b_{iK-1}$ と制約される。

開発したシステムの能力値推定プログラムは、オープンソースである R のパッケージ ltm⁽⁴⁾を適用する。

3. 実データを用いた能力値推定のシミュレーション

3.1 概要

本章では、開発したシステムの能力値推定プログラムの動作検証のために実データを用いたシミュレ

ーションについて述べる．シミュレーションに用いるデータは，小学校で実践したループリックを用いた学習評価の結果である．ループリックは評価基準が 21 項目あり，評価基準がそれぞれ 3 段階ある．また，シミュレーションのプログラムは Java を用いて作成した．ここでは，評価基準のそれぞれを項目，評価基準の段階をカテゴリとして段階反応モデルを用いる．

3.2 手順

シミュレーションの流れは以下のとおりである．

- (1) 小学校で実践しているループリックを用いた学習評価のデータの結果から IRTPRO⁽⁵⁾を用いてベイズ推定で項目の識別力パラメータと困難度パラメータを推定する．推定の結果，識別力パラメータの値は 0.92~3.04 であり，困難度パラメータはカテゴリ 1~3 でそれぞれ $k=1$ が -3.03~-0.47, $k=2$ が -1.81~0.80, $k=3$ が 0.01~1.53 であった．
- (2) 真の能力値 θ が -2.0, -1.0, 0.0, 1.0, 2.0 となる受検者を各 100 人分用意する．
- (3) 学習者の能力推定値の初期値を 0.0 とする．
- (4) 学習者の各項目への反応データを，真の能力値を所与した 2 章の(1)式から求める．
- (5) 反応データから学習者の能力値を推定する．
- (6) 上記の(4)と(5)を 21 項目，繰り返す．

以上の手順でシミュレーションを行うことで，真の能力値 θ を持った学習者 100 名が，項目 1~21 で，どの評価基準に解答したか，また，解答したときの能力推定値を出力する．

3.3 結果及び考察

ここでは，シミュレーションの結果及び考察について述べる．結果は，真の能力値ごとに分け，100 名の学習者について課題ごとに能力推定値の平均を算出した．図 1 にシミュレーションの結果をグラフに表したものを示す．

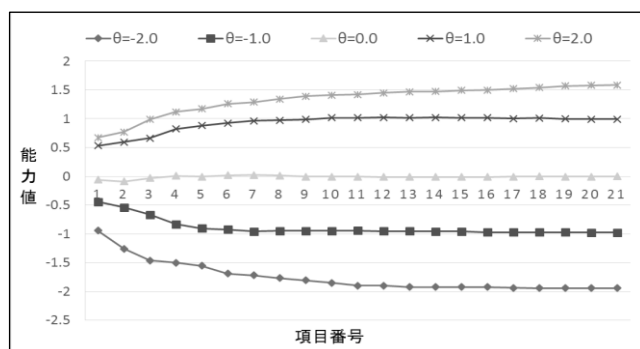


図 1 推定した能力値の結果

真の能力値が -2.0, -1.0 と 1.0, 0.0 の学習者についてはそれぞれおよそ項目が 13, 8, 9 で真の能力値に収束している．

一方，真の能力値が 2.0 の学習者については，能力推定値が 1.583213 で収束し，真の能力値と乖離している．これは， $k=3$ の困難度パラメータが 1.53 ま

でであり，能力値推定に適切な項目が出題されなかったことが原因と考えられる．ここから，困難度パラメータが高い評価基準をより多く作成することで，推定される能力値はより真の能力値に近づくことが予想される．図 2 に真の能力値 θ が -2.0 である学習者を 1 人抜粋して，学習者の能力推定値を，開発したシステムの能力可視化機能で表示する画面を示す．



図 2 能力値表示の画面

これらのシミュレーションの結果及び考察から，開発したシステムの能力値推定プログラムは正常に動作していることが確認できた．

4. おわりに

本論文では，開発したシステムの能力値推定プログラムの動作検証をするために，シミュレーションを行った．シミュレーションの結果より，開発したシステムの能力値推定プログラム及び，能力可視化機能の有効性が確認できた．

今後は，開発したシステムを実際に授業等の実践の場で継続的に利用し，本システムの信頼性と妥当性，また，利用することでの教育効果について評価・検証していく予定である．

謝辞

本研究は，東京学芸大学「日本における次世代対応型教育モデルの研究開発」〔文部科学省特別経費(プロジェクト分)〕の研究成果の一部である．

参考文献

- (1) 蛭名哲也，宮澤芳光，榎本命，森本康彦：“アクティブ・ラーニングにおける学習評価のための項目反応理論に基づく CBT の開発”，日本教育工学会研究報告集, 17(1), pp.483-488 (2017)
- (2) 豊田秀樹：“項目反応理論 [入門編]【第 2 版】”，朝倉書店 (2012)
- (3) Samejima, F.: “Estimation of latent ability using a response pattern of graded scores”, Psychometrika Monography, no.17, pp.1-100 (1969)
- (4) Rizopoulos, D. I.: “An R package for latent variable modelling and item response theory analyses”, Journal of Statistical Software 17(5), 1-25 (2006)
- (5) 「SSI-Scientific Software」：
<http://www.ssi-central.com/irt/index.html>
(2017 年 5 月 24 日)