

数式自動採点システムの解答データを用いた能力推定法の調査と適用例

Survey and application example of ability estimation method using answer data of math e-learning system

伊藤 可子^{*1}, 中村 泰之^{*2}

Kako Ito^{*1}, Yasuyuki Nakamura^{*2}

^{*1}名古屋大学情報学部, ^{*2}名古屋大学大学院情報学研究所

^{*1} School of Informatics, Nagoya University, ^{*2} Graduate School of Informatics, Nagoya University

Email: ^{*1}itou.kako@b.mbox.nagoya-u.ac.jp

あらまし：数式自動採点システム STACK で行ったオンラインテストの解答データを用いて学生の能力値の推定法の調査と適用を行い、検討した。項目反応理論は一定の能力値の学生を選ぶ際の参考値を求める分析手法として、潜在ランク理論は学生をランク分けする際の分析手法として適切であると考えられる。
キーワード：STACK, 項目反応理論, 潜在ランク理論

1. はじめに

近年、教育の ICT 化が進み e ラーニングに注目が集まっている。数学のオンラインテスト教材の解答形式の多くは多肢選択式で、数式入力形式は少ない。しかし、多肢選択式や数値入力式の場合、選択肢から解答を推測できる場合があり、学生の実際の能力を測ることが難しい。一方、数式入力形式の場合、解答が推測できないため、学生の実際の能力を測ることができると考えられる。また、数式入力形式は誤答分析により、受験者の理解不足箇所の情報が得られることも期待できる。

本研究は STACK の解答データを用いて学生の能力の推定法を調査した。STACK は数式入力形式の問題の出題、正誤評価が可能である。そこでオンラインテストの受験者の能力値の推定法の調査、オンラインテストを受験した学生の実際の解答データへの適用を行い、能力推定法の調査と特徴の整理を行った。

2. STACK について

STACK⁽¹⁾は英国バーミンガム大学の Christopher Sangwin 氏らが開発した数学オンラインテストシステムである。学習管理システムの一つである Moodle の小テストで数式入力の解答が可能な問題タイプとして提供され、入力された数式を数式処理システムによって代数的に等価かを評価し、正誤評価、自動採点を行う。また、ポテンシャル・レスポンス・ツリーと呼ばれる機構を用いて学習者の解答に応じたフィードバックを返すことができる。

3. 解析手法

3.1 IRT(項目反応理論)

項目反応理論⁽²⁾(Item Response Theory 以下 IRT)は 1952 年米国の Lord が提出した現代テスト理論である。項目は試験を構成する各問題であり、反応は項目に対する正誤状況を示す。

IRT はある項目の困難度、識別力が判明している場合、項目に対する反応から測定できる受験者の能力を推定する。項目の識別力は受験者の能力の推定の正確性を示し、受験者の能力を正確に推定する項目ほど値が大きく、項目の困難度は、正答には受験者の能力値が高い必要があるほど値が大きくなる。

IRT は学生の能力値の推定が目標だが、直接能力値を求めることは不可能であるため、能力値の推定に必要な項目の識別力と困難度を先に推定する。この時、周辺最尤推定法を用い、項目の識別力と困難度を推定する。周辺最尤推定法は能力値パラメタを積分消去し、周辺尤度関数が最大となるように能力値に直接依存しない識別力と困難度を推定する手法である。

項目の識別力と困難度を推定した後、学生の能力値を推定する。学生の能力値の推定はベイズ推定法を用いる。能力値の事前確率分布を正規分布とし、尤度関数を用いて事後分布を求め、能力値に関する推定を行う。

また、IRT では横軸を能力値、縦軸を項目に対する正答確率とする項目特性曲線(Item Characteristic Curve 以下 ICC)を用い、項目の難易度と受験者の能力を分離して表す。ICC は正規累積モデルで表していたが、積分を含むため、Birnbaum がロジスティック分布を利用したロジスティックモデルを提案した。

2パラメタ・ロジスティックモデルは項目 j の識別力を a_j 、困難度を b_j とした時、能力値 θ の学生が項目 j に正答する確率 $p_j(\theta)$ を表す。

$$p_j(\theta) = \frac{1}{1 + \exp(-Da_j(\theta - b_j))}$$

D は Lord と Birnbaum が導入した $D = 1.7$ で全ての θ において誤差が 0.01 以下となる尺度因子を表す。項目特性曲線の変曲点での傾きが識別力によって表現される。

3.2 LRT(潜在ランク理論)

潜在ランク理論^(3,4)(Latent Rank Theory, 以下 LRT) は 2007 年に荘島が提唱したテスト理論である。

LRT は各学生の学力を測定するテストの解答から各学生を段階的なランクにわけ、各ランクに所属する確率を推定するため、テストの得点による連続的な能力評価ではなく、ランク分けという段階的な能力評価を行う。さらに LRT は潜在尺度を順序尺度としており、変数の値の大小に意味があるため、部分点のテストであってもランク分けすることが出来る。

また、荘島は各能力段階を示す学力進捗表を明らかにする必要があるとしている。学力進捗表は受験者をランク分けした上で、どのような学力を獲得しているかを判断するものである。テストの各問題に必要な要素を書き出し、これらの要素の各ランク帯での正答率を計算し、学力進捗表にしておくことで、あるランクに分類された学生が正答していない要素を明らかにできる。学力進捗表の作成により、テストが測定する具体的な内容を明確にし、テスト結果から学生の持つ具体的な能力を測ることが出来る。

4. 適用結果

7 月 2 日のオンラインテストに上記の理論を適用した。

最も困難度の高い問題は問題 3(図 1 中 X3, 以下同様)

$$\int_0^2 \cos \frac{2x}{3} dx$$

であり、三角関数の積分の要素を含む問題である。識別力の高い問題は問題 6

$$\int_0^{\frac{\pi}{2}} \frac{\cos x}{\cos x + 1} dx$$

問題 8

$$\int_0^3 \frac{x}{\sqrt{4-x}} dx$$

であり、問題 6 は分数関数の積分、三角関数の積分、置換積分の要素を含む問題であり、問題 8 は分数関数の積分と置換積分の要素を含む問題である。

また図 1 を見ると識別力の高い問題 6、問題 8 は変曲点付近での傾きが急である。困難度が最も高い問題 3 は単調増加の直線のグラフであり、変曲点が存在しないため、困難度が 2 番目に高い問題 6 が最も右側に変曲点があるグラフである。また、問題の困難度の低い問題 1、問題 7、問題 10 は変曲点がグラフの左側にあるため、変曲点前の部分が他のグラフよりも短くなっている。

困難度や識別力の高い問題の傾向について考察したが、困難度が高い問題には二つ以上の要素や三角関数や分数関数の要素が含まれていたが、三角関数や分数関数の要素を含んでも困難度が低い問題もあるため、困難度が高くなる要因に問題が含む要素が関係するかは判明しなかった。また、識別力の高い問題に関しては三角関数の積分、対数関数の積

分、置換積分の要素を含むことが多く、要素が複数含まれていることも多いことから、一つの問題に複数の要素を含むことで学生の能力値を正しく推定することのできる問題となると考えられる。

潜在ランク理論は今回の適用したデータではランク数が 2 の時のみ計算を行うことができた。ランク数が 3 以上の場合の計算が行えなかったのはデータ数が十分ではなかったことが原因の一つと考えられる。

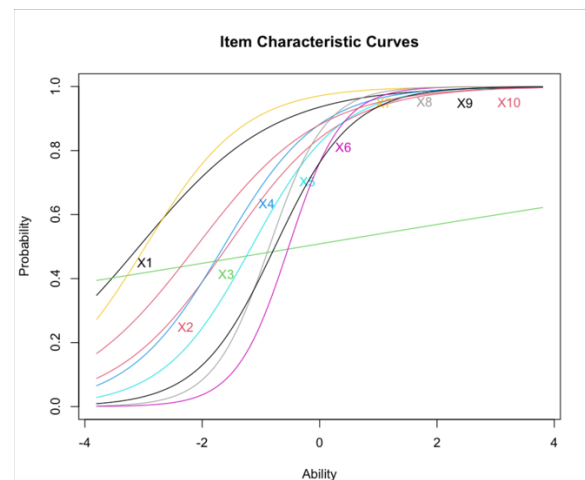


図 1 項目特性曲線

5. まとめ

項目反応理論は能力値などを参考にして問題を作成することで、一定の能力値を持つ学生を選抜できると考えられ、入試問題の作成時の問題分析の手法として適切であり、潜在ランク理論は学生をランク分けするため、成績処理などに用いる手法として適切であると考えられる。それぞれ求められるものが異なり、この手法を用いれば必ず学生の能力を推定できるという手法はなく、目的にあった手法を用いることが必要であると考えられる。また、項目パラメータ、能力値、ランク分けを計算することにより、個人の能力に応じた演習問題を提示することが可能となり、個別最適化教育に応用することも期待される。

6. 参考文献

参考文献

- (1) STACK, <https://www.ed.ac.uk/math/stack>, 参照日 2021 年 2 月 8 日
- (2) 豊田秀樹, 「項目反応理論[入門編] 第 2 版」, 朝倉書店, (2012)
- (3) 荘島宏二郎, ニューラルテスト理論—資格試験のためのテスト理論—, 平成 20 年度全国大学入学者選抜研究連絡協議会, pp.163-168, (2008)
- (4) 清水裕士, 「潜在ランク分析の R コードを作りました」, <https://norimune.net/2293>, 参照日 2021 年 2 月 8 日