

数学オンラインテストの誤答傾向分析による評価指標の検討

Examination of evaluation indicator by incorrect answer tendency analysis of math online test

藤本 和侖^{*1}, 中村泰之^{*2}

Kazusa FUJIMOTO^{*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科

^{*1}School of Informatics, Nagoya University, ^{*2}Graduate School of Informatics, Nagoya University
Email:fujimoto.kazusa@f.mbox.nagoya-u.ac.jp

あらまし：数式自動採点システムの一つである STACK では、解答データの解析がまだ充分に行われていない。そこで本研究では、文系学部向けの数Ⅲを扱った授業で STACK を用いて得られた演習問題の解答データをもとに、誤答分類、S-P 表を用いた解答データの分析、評価指標の検討を行なった。その結果、問題ごとの誤答理由、学生の誤答理由が明らかになった。また、誤答傾向を読み取るために用いた新たな評価指標について検討した結果、学生の解答傾向を読み取ることができた。

キーワード：数式自動採点システム, STACK, 誤答分類, S-P 表

1. はじめに

2019年に文部科学省が GIGA スクール構想を打ち出す⁽¹⁾など、教育分野でも情報化が進んでいる。データ活用も進んでおり、その一つとしてオンラインテストデータの解析が進んでいる。オンラインテストの多くは解答選択型問題、数値入力型問題などの問題形式を用いている。そのため、バリエーションが少ない、当てずっぽうで正解できるといった問題点がある。この問題点を解決し、個々の問題に対する理解の度合いを確認するためには数式自動採点システムの利用が有効であると考えられる。

本研究では、誤答の分析により、学生の誤答傾向を明らかにすること、オンラインテストを用いて学生の能力指標を明らかにすることを目的として、数式自動採点システムの一つである STACK^(2,3,4)を用いて得られた解答データをもとに、誤答分類、S-P 表を用いた解答データの分析、評価指標の検討を行う。分析の対象とするデータは、文系学部向けに開講され、数Ⅲの内容を扱う 2020 年度「数学入門」の授業で、STACK を利用して出題された問題のうち、積分の問題である 15 問と、その問題を解き、提出した学生 63 名の解答データである。

2. STACK の概要紹介

STACK は、英国バーミンガム大学の Christopher Sangwin (現エディンバラ大学) が中心となり開発した数式自動採点システムである。日本では 2010 年に日本語化され、公開された。

STACK には、ポテンシャル・レスポンス・ツリーと呼ばれる機構を利用して、学生の解答に対応したフィードバックを与えたり、学生の解答に部分点を与えるなどの様々な機能がある⁽⁵⁾。また、中村、中原によって、解答だけではなく計算過程を記述したノートも提出でき、問題・解答と紐付けて管理することができる機能が開発された⁽⁶⁾。

3. S-P 表

3.1 S-P 表について

S-P 表は 1969 年に佐藤らが創案したテスト得点分析法であり、学習診断や評価を行う際の助けとなるものである。S-P 表の作成手順を示す。

- 1) n 個の問題 (小問) に対して正答には得点 1 を与え、誤答には得点 0 を与えて得た N 人の生徒の項目得点を集めて $N \times n$ の項目得点表を作成する。
- 2) 1) で作成した項目得点表において、生徒を合計得点の大きい順に上から、問題を正答者数の多い順に左から並びかえる。
- 3) 2) で並びかえた項目得点表中に、各生徒について、表の左から生徒の得点分だけマス目を数えてマス目の右側に線を引き、全生徒分の線を結んでできる S 曲線を書き入れる。S 曲線と同様にして、各問題について、P 曲線も書き入れる。

S 曲線の場合、ある生徒に対して S 曲線の左側には 1 だけ、右側には 0 だけというような 1 と 0 の入れかわりのない状態を完全反応パターンという。P 曲線の場合も S 曲線と同様である⁽⁷⁾。

正答と誤答の二値で構成される S-P 表をもとに、部分点に対応した多値で構成される S-P 表も福島らによって提案されている⁽⁸⁾。

3.2 注意係数

注意係数は学生の評価指標として用いられる。ある生徒の反応パターンの注意係数とは、区切線を境に 1 と 0 の入れかわりのない完全反応パターンを基準として、実際の反応パターンが完全な反応パターンからどの程度の差異があるかを示す値であり、一人ひとりの生徒に対して算出される⁽⁷⁾。

4. 誤答分析

STACK を用いて出題された問題の解答データの

うち、誤答データを抽出した。抽出した誤答データと解答ごとの誤答理由を問題ごとに表にまとめた。計算過程を記述したノートを提出している場合はノートから判明したものを、提出していない場合は予想できるものを、予想できないものは不明を誤答理由とした。

問題ごとに誤答理由を調べた結果、明らかになった典型的な誤答を示す。不定積分の問題では、最初の問題において積分定数の不足している解答が目立った。その後の問題では、最初の問題と比較すると、積分定数の不足している解答は少なくなっていた。置換積分法を用いる問題では、合成関数の微分の計算に関する間違いが多かった。また、 $\int \cos nx \, dx$ (n は定数)、 $\int e^{-x} \, dx$ の計算ミスも多かった。15問全体を通しては、ノートの解答は正解しているにもかかわらず、解答の一部を入力し忘れていた、事前に説明されている入力方法に従っていないといった、オンラインテスト特有の入力ミスによる間違いが多かった。

5. S-P 表分析

5.1 S-P 表の作成

63名の解答を、正答の場合は1、誤答の場合は0とする二値S-P表を作成した。また、正答の場合は10、誤答の場合は4章の誤答分析で明らかとなった誤答理由をもとに5段階の部分点を与えた多値S-P表も作成した。

5.2 注意係数の再検討

注意係数は完全反応パターンにおいては不正解であるべきとする問題を正解していると、正解するべきではないかのように扱われることになる。選択肢型問題であれば、たまたま正解している可能性があるため、注意係数を用いるので問題ないと考えられる。しかし、本研究のように、数式入力型問題の場合は、たまたま正解したのではなく、理解した上で正解していると考えられる。そこで、本研究では、正解するべき問題が不正解の場合はマイナス評価、不正解であるべき問題を正解した場合はプラス評価とする新しい係数を考えた。

この新たな係数の有用性を調べるために、学生それぞれの新たな係数と最終課題の得点を比較した。最終課題は微積分や極限の計算を必要とするペーパーテストである。

5.3 S-P 表を用いた結果

S-P表の上側に位置している、すなわち、比較的優秀な学生であっても、正答者数が多くため、難易度の低いとされる問題を間違えていることがわかった。また、S-P表の下側に位置している学生であっても、正答者数が少ないため、難易度が高いとされる問題を正解していることもわかった。多値S-P表は部分点を用いたため、誤答理由が何であったのかを学生の解答を見直すことなく知ることができた。

作成したS-P表をもとに求めた新たな係数と合計

得点を問題数で割った得点率を参考にすると、得点率が高く、新たな係数の値が小さい学生は難易度の低い問題を間違えていたといった学生の解答傾向を読み取ることができた。

5.4 最終課題との比較

新たな係数の有用性について調べるために、最終課題の得点と比較した結果、新たな係数と最終課題の得点に相関は見られなかった。そこで、注意係数と最終課題の得点も比較してみたが、相関は見られなかった。

6. まとめと考察

本研究では、STACKを用いて得られた解答データをもとに、誤答分類、S-P表を用いた解答データの分析、評価指標の検討を行なった。学生の解答データから抽出した誤答データをもとに、各問題について誤答理由をまとめた結果、問題に応じた種類の誤答やオンラインテスト特有の入力ミスによる誤答があることがわかった。オンラインテスト特有の誤答を減らすためには、数式入力方法の改善が望まれる。オンラインテスト特有の誤答理由や問題ごとの誤答理由をもとにして、想定される誤答に基づいた、自動部分点採点を含むオンラインテストの設計に生かすことができると考える。

S-P表を用いて解答データを分析する際には新たな係数の提案を行い、学生の誤答傾向を読み取ることができた。最終成績との相関は見られなかったが、これは、学生が最終課題までに勉強し、理解が進んだこと、ノートを提出していない学生の誤答理由が本来の誤答理由と異なって反映されていることなどが原因であると考えられる。

参考文献

- (1) 文部科学省:GIGA スクール構想の実現へ、https://www.mext.go.jp/content/20200625-mxt_syoto01-00003278_1.pdf, 参照 2021-1-20
- (2) Christopher Sangwin: “Computer Aided Assessment of Mathematics”, OUP Oxford, (2013)
- (3) STACK | The University of Edinburgh, <https://www.ed.ac.uk/math/stack>, 参照 2021-1-25
- (4) GitHub - maths/moodle-qtype_stack:Stack question type for Moodle, https://github.com/math/moodle-qtype_stack, 参照 2021-1-25
- (5) 中村泰之: “数学 e ラーニング 数式解答評価システム STACK と Moodle による理工系教育”, 東京電気大学出版局, (2010)
- (6) Yasuyuki Nakamura, Takahiro Nakahara: “NOTE-SUBMISSION FUNCTION FOR MOODLE QUIZ AND COLLECTING PEN-STROKE DATA”, Proc. of IADIS International Conference Mobile Learning, pp.163-164 (2019)
- (7) 佐藤隆博: “S-P 表の入門”, 明治図書, (1985)
- (8) 福島紫織, 塩井隆円, 楠和馬, 波多野賢治: “試験の得点とその分布を利用した生徒の授業に対する理解度把握手法の提案”, 第9回データ工学と情報マネジメントに関するフォーラム(発表資料 B8-5), (2017)