

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

福島 紫織[†] 塩井 隆円^{††} 楠 和馬^{††} 波多野賢治^{†††}

[†] 同志社大学文化情報学部 〒610-0394 京都府京田辺市多々羅都谷 1-3

^{††} 同志社大学大学院文化情報学研究科 〒610-0394 京都府京田辺市多々羅都谷 1-3

^{†††} 同志社大学文化情報学部 〒610-0394 京都府京田辺市多々羅都谷 1-3

E-mail: [†]bin0022@mail4.doshisha.ac.jp, ^{††}{shioi,kusu}@ilab.doshisha.ac.jp, ^{†††}khatano@mail.doshisha.ac.jp

あらまし 現在の教育現場では各生徒に合わせた指導を行うことが重要視されており、個別の指導を実現するために e ラーニングを用いた個別指導や少人数指導などが行われている。個別指導を行うにあたってあらかじめ理解度を把握しておく必要があり、生徒の理解度把握の手段として S-P 表が利用されている。しかし従来の S-P 表は正解・不正解で採点する試験の回答データ（二値の回答データ）にのみ対応しているため、多肢選択式問題と記述式問題で構成された試験の回答データ（多値の回答データ）に対応していないという問題点がある。そこで本稿では教育現場で実施されている多値の回答データに対応する多値 S-P 表を提案し、二値 S-P 表と同じ性質を持たせ、二値 S-P 表よりも各生徒の理解度を正確に診断することを目指す。多値 S-P 表と二値 S-P 表で各生徒に対する理解度診断の精度を比較した結果、二値 S-P 表よりも多値 S-P 表の精度が上回る結果となった。これより提案した多値 S-P 表は二値 S-P 表よりも生徒の理解度を適切に把握できると結論付けることができ、教師の授業改善に貢献できると言える。また、多値 S-P 表で提案した各生徒の理解度をグルーピングする手法を提案することで試験の得点だけからでは知ることができない生徒の特徴把握も実現できた。

キーワード S-P 表, 個に応じた指導, 生徒の授業理解度診断

1. はじめに

文部科学省の中央教育審議会は平成 15 年中央審議会答申^(注1)において基礎・基本的な能力に加え、学ぶ意欲や自分で課題を発見し自ら学び、主体的に判断し、よりよく問題を解決する資質や能力を定着させることを目標としており、そのような能力を育むためには個に応じた指導を充実させることが重要だと提言している^(注2)。個に応じた指導を実現するために教育現場では e ラーニングを用いた個別指導 [1-3] や少人数での授業などさまざまな指導が行われている。しかしながら、生徒一人ひとりの授業理解度は異なるため、各生徒の理解度に合わせた指導を行うには、まず各生徒の授業理解度を正確に把握する必要がある。

理解度を把握するために従来から用いられている手段として S-P 表がある。S-P 表とは生徒の授業に対する理解度を教師が知るための理解度診断ツールであり、試験の得点一覧表から S-P 表を作成し、ケアレスミス回答とまぐれ正解回答の分布を考慮することで各生徒の学習課題を発見し、個別の理解度を把握することを可能としている [4]。S-P 表の利用事例としては小学校の図画工作の時間において描画技術の変化を検証するた

めに利用される場合や [5]、大学の講義で VR 教材を用いた指導では学習効果があるのかを検証するために利用されるという事例があり [6]、教育現場において S-P 表は生徒の理解度を把握するために広く利用されている。

現在、生徒一人ひとりの授業理解度の評価に利用されている S-P 表は、正解・不正解の二値で採点される試験の回答データのみを対象とするため、実際の教育現場で実施されている部分点がある記述式問題や正解・不正解の二値で採点できる多肢選択式問題が混同して構成されている試験に対応していないという問題点がある。

そこで本稿では、実際の教育現場で行われているテスト形式に対応可能な多肢選択式問題のように正解・不正解で判定できる問題に加え、記述式問題などの部分点がある問題で構成される試験に対応する多値 S-P 表を提案する。

また、教育現場で実施されている少人数指導による個別指導は教師の指導の負担がかかると言われているため [7]、ケアレスミス回答とまぐれ正解回答の分布の様子が似た生徒つまり重要視すべき学習課題の特徴が似た生徒をグルーピングすることで教師の指導の負担を軽減できると考えられる。よって、多値 S-P 表ではケアレスミス回答とまぐれ正解回答の分布の様子が似た生徒のグルーピングの手法についても提案する。

2. 従来の二値 S-P 表

S-P 表は生徒の試験の得点と得点分布を見ることで生徒の個別の理解度を把握する手段として利用されており、S-P 表の S

(注1)：文部科学省 中央教育審議会 平成 15 年 10 月 7 日 初等中等教育における当面の教育課程及び指導の充実・改善方策について（答申）、http://www.mext.go.jp/b_menu/shingi/toushin.htm (2017/3/20 閲覧)

(注2)：文部科学省 新学習指導要領のねらいの実現に向けて、http://www.mext.go.jp/a_menu/shotou/gakuryoku/point.htm (2017/3/20 閲覧)

は生徒 (Student) を表し, P は (Problem) を表している. 以下の節では二値 S-P 表における各キーワードに関する説明を行う.

2.1 回答データ

二値 S-P 表が対応可能な回答データは, 試験の採点を正解・不正解で行った回答データである. 例としては, e ラーニングの試験のように一問一答式の問題で構成された試験の回答データである. 紙媒体の学期末試験などは記述式問題のように部分点がある問題も存在するため, 正解・不正解で採点できる問題の部分抜き出ししたり, 部分点を付けずに先生の主観で正解・不正解に変換することで S-P 表は利用されてきた [11].

2.2 表の設計

表の列は問題 (Problem) であり, 行は生徒 (Student) である. 列の設計において問題の正解者数が多いほど難易度を低く定義して, 難易度が低い問題を表の左側から順に並べる. 行の設計は試験の得点が高ければ成績上位の生徒と定義して, 成績上位の生徒を表の上から順に並べる. このように表を設計することで表の左上には成績上位の生徒の難易度の低い問題に対する回答つまり正解が集まり, 表の右下には成績下位の生徒の難易度の高い問題に対する回答つまり不正解の回答が集まるという性質を持つ. 表の設計を行う際に, 列においては同難易度の問題が生じ, 行においては同得点の生徒が生じる場合がある. 以下では同難易度の問題の入替えと同得点の生徒の入替えについて説明を述べる.

(1) 同難易度の問題の入替え

0 の回答に対応する生徒の得点の総和をそれぞれ計算し, 同難易度の問題間で比較し, 得点の総和が小さいほうが表の左側に配列される. 例として図 1 の問題 2 と問題 3 について考える. 問題 2 において 0 となっている生徒は生徒 H と生徒 I であり, 得点の総和を計算すると 3 となる. 問題 3 において 0 となっている生徒は生徒 F と生徒 I であり, 得点の総和を計算すると 4 になる. 問題 2 と問題 3 の得点の総和を比較すると問題 2 の方が値が小さくなるため, 問題 2 が表の左側に配列される.

(2) 同得点の生徒の入替え

0 の回答に対応する問題の正解者数の総和をそれぞれ計算し, 同得点の生徒間で比較し, 正解者数の総和が小さいほうが表の上側に配列される. 例として図 1 の生徒 C と生徒 D について考える. 生徒 C において 0 となっている問題は問題 7 と問題 8 であり, 正解者数の総和を計算すると 5 になる. 生徒 D において 0 となっている問題は問題 5 と問題 7 であり, 正解者数の総和を計算すると 8 になる. 生徒 C と生徒 D の正解者数の総和を比較すると生徒 C の方が値が小さくなるため, 生徒 C が表の上側に配列される.

2.3 S 曲線と P 曲線

S 曲線は図 1 における実線であり, 試験に対して各生徒が取得可能な得点の水準を表している. S 曲線の引き方はまず始めに生徒の得点を見る. その次に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縦線を引く. 例として図 1 の生徒 C について考える. 生徒 C の得点は 6 点であるため, 生徒 C の S 曲線は表の左から数えて 6 マス目の右側に縦線が引かれる.

P 曲線とは図 1 における破線であり, 各問題において何人の正解者がいるかを表している. P 曲線の引き方はまず始めに問題の正解者数を見る. その次に表の上から問題の正解者数分だけマス目を数えてマス目の下側に横線を引く. 例として図 1 の問題 1 について考える. 問題 1 の正解者数は 8 人であるため, 問題 1 の P 曲線は表の上から 8 マス目の下側に横線が引かれる.

この S 曲線と P 曲線が離れていれば試験が生徒の理解度を図るために適切でなかったと考えることができ, S 曲線と P 曲線が近づいていれば質の良い試験であるということを知ることができる.

		問題番号								生徒の得点
		1	2	3	4	5	6	7	8	
生徒名	A	1	1	1	1	1	1	1	1	8
	B	1	1	1	1	1	1	1	0	7
	C	1	1	1	1	1	1	0	0	6
	D	1	1	1	1	0	1	0	1	6
	E	1	1	1	1	1	0	0	0	5
	F	1	1	0	1	1	0	0	0	4
	G	1	1	1	0	0	0	1	0	4
	H	1	0	1	0	0	1	0	0	3
	I	0	0	0	0	0	0	0	0	0
			8	7	7	6	5	5	3	2

図 1 S-P 表の S 曲線と P 曲線

2.4 ケアレスミス回答とまぐれ正解回答

S 曲線は, ある試験に対して各生徒が取得可能な得点の水準を表す問題の横に線が引かれるため, S 曲線よりも左側には各生徒が正解することが可能な回答が集まり, S 曲線よりも右側には各生徒が正解することが難しい回答が集まる.

ケアレスミスと判定される回答 (ケアレスミス回答) は図 2 における S 曲線よりも左にありグレーで色付けされた不正解の回答である. まぐれ正解と判定される回答 (まぐれ正解回答) は図 2 における S 曲線よりも右にありグレーで色付けされた正解の回答である. このようにケアレスミス回答とまぐれ正解回答の分布を見ることで試験の得点だけでは分からない生徒の理解度の特徴を知ることができる.

		問題番号								生徒の得点
		1	2	3	4	5	6	7	8	
生徒名	A	1	1	1	1	1	1	1	1	8
	B	1	1	1	1	1	1	1	0	7
	C	1	1	1	1	1	1	0	0	6
	D	1	1	1	1	0	1	0	1	6
	E	1	1	1	1	1	0	0	0	5
	F	1	1	0	1	1	0	0	0	4
	G	1	1	1	0	0	0	1	0	4
	H	1	0	1	0	0	1	0	0	3
	I	0	0	0	0	0	0	0	0	0
			8	7	7	6	5	5	3	2

図 2 ケアレスミス回答とまぐれ正解回答

3. 関連研究

二値 S-P 表をもとに各生徒の各問題に対する理解の定着度を定量化する研究が三輪らによって行われている [12]. 上位の生徒の難易度の低い問題に対する回答ほど理解の定着度が高い回

答であり、下位の生徒の難易度の高い問題に対する回答ほど理解の定着度が低い回答であると定義されており、この理解の定着度を表す指標を「理解度ポテンシャル」と呼んでいる。理解度ポテンシャルの算出方法は、各問題の正解者数から生徒が試験の取得可能な得点の水準の問題に対する正解者数を引き、その差を全生徒数で割ることで算出される。

		問題番号								生徒の得点
		1	2	3	4	5	6	7	8	
生徒名	A	0.67	0.56	0.56	0.44	0.33	0.33	0.11	0	8
	B	0.56	0.44	0.44	0.33	0.22	0.22	0	-0.1	7
	C	0.33	0.22	0.22	-0.6	0	0	-0.2	-0.3	6
	D	0.33	0.22	0.22	0.11	0	0	-0.2	-0.3	6
	E	0.33	0.22	0.22	0.11	0	0	-0.2	-0.3	5
	F	0.22	0.11	0.11	0	-0.1	-0.1	-0.3	-0.4	4
	G	0.22	0.11	0.11	0	-0.1	-0.1	-0.3	-0.4	4
	H	0.22	0	0	-0.1	-0.2	-0.2	-0.4	-0.6	3
	I	-0.1	-0.2	-0.2	-0.3	-0.4	-0.4	-0.7	-0.8	0
	J									
		8	7	7	6	5	5	3	2	
		各問題の正解者数								

図3 理解度ポテンシャル評価値を S-P 表に付与

以上のように各生徒の理解の定着度を定量化できた点がこの研究の優れている点であるが、二値 S-P 表をもとに研究を行っているため、多肢選択式の問題と記述式の問題で構成された試験の回答データに対応していないという点が課題である。

4. 提案手法

二値 S-P 表の多肢選択式の問題と部分点がある記述式の問題で構成された試験の回答データに対応していないという問題を解決するために、本稿では正解・不正解で採点される問題と部分点がある記述式問題で構成された試験の回答データに対応可能な多値 S-P 表を提案する。

また、本稿で提案する多値 S-P 表は二値 S-P 表と同じ性質を持たせるように設計することで関連研究の手法も適用可能になり、多値データならではの生徒の特徴を抽出することで二値 S-P 表では知ることができなかった生徒の理解度の特徴を明らかにすることを目的としている。さらに、少人数指導による個別指導の教師の負担を軽減するために理解度の似た生徒のグルーピングの手法についても提案する。

4.1 対象データと問題の基準値

以下の節でデータの扱い方と各問題の基準値について述べる。

4.1.1 対象データの扱い方

二値 S-P 表の回答データは各問題において得点者の階級数が 0 点と 1 点の二つに分布するという特徴がある。二値 S-P 表に対して提案する多値 S-P 表は各問題の配点をもとに採点するため、各問題によって得点者の階級数が異なるという点が特徴である。例としては図4の問題5と図5の問題8のように問題5であれば階級数が3であり、問題8であれば階級数が6であるというように問題によって階級数が異なる。二値 S-P 表であれば各問題が二値であるため、正解者を成績優秀者と定義することができていた。しかし、多値 S-P 表の回答データは得点者の階級数が異なるため、各問題において何点を超えていれば成績優秀であるという基準値を設定する必要がある。

4.1.2 各問題の基準値の決定手法

多値 S-P 表の回答データは得点者の階級数が異なるため、各問題において基準値を設定する必要がある。

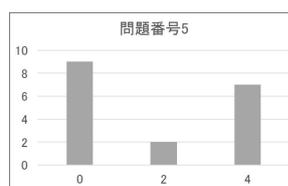


図4 問題5の得点分布

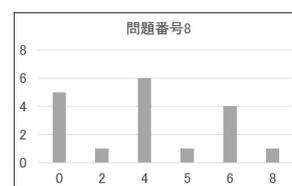


図5 問題8の得点分布

基準値とは、各問題において何点は超えておくべきだという境界線であり、本稿で提案する基準値は生徒の得点者数に対する人数の度数分布表の半分を示す値である平均値と中央値を提案する。この二つの基準値は、データの分布によって中央を指し示す数値が変化する。例として図6の問題5であれば平均値は1.77の値となり中央値は1の値となる。図7の問題6であれば平均値は2.22となり、中央値は2の値となる。このように、データの分布によってデータの中央を指し示す値が変化する [8]。

基準値を各問題ごとに選択する必要があり、正規性のある問題ほど平均値はデータの中央を示すといわれており、正規性のないデータには中央値を用いることが適切であるため、各問題のデータの正規性を調べる。提案する多値 S-P 表では各問題のデータの分布が正規性を持つかどうかを検定し、データに正規性がある問題には平均値、データに正規性が無い問題には中央値を基準値に採用する。本稿では、各問題つまり1標本のデータの正規性の有意を調査するために有用であるシャピロ・ウィルク検定を用いる [10]。シャピロ・ウィルク検定は帰無仮説を「問題のデータの分布が正規性に従う」とし、対立仮説を「問題のデータの分布が正規性に従わない」とし、p 値が 0.05 以下であれば帰無仮説は棄却されるため、その問題の基準値は中央値を採用する。p 値が 0.05 より大きければ帰無仮説は採択されその問題の基準値には平均値を採用する。

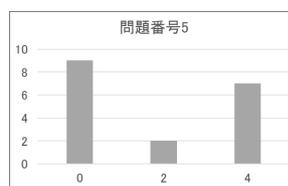


図6 問題5の得点分布

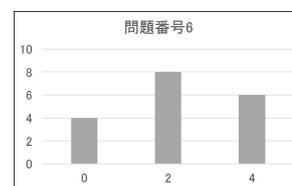


図7 問題6の得点分布

4.2 表の設計

二値 S-P 表と同じ性質を持たせるよう多値 S-P 表の設計を行う。二値 S-P 表は、表の左上には成績上位の生徒の難易度が低い問題に対する回答が集まるように設計され、表の右下には成績下位の生徒の難易度が低い問題に対する回答が集まるように設計されていた。以下、4.2.1 では表の設計を考えるうえで重要な問題の難易度の定義について述べ、4.2.2 では生徒の成績順位の定義について述べる。

4.2.1 問題の難易度定義

二値 S-P 表では正解者数が多ければ難易度の低い問題と定義されていた。そこで、同じ性質を持たせるために多値 S-P 表では得点率を用いる。得点率は、各問題の基準値を各問題の配点で割ることで算出される。二値 S-P 表の問題の難易度の定義と同じ性質を持たせるように多値 S-P 表を設計するために、得点率が低ければ難易度の高い問題と定義した。

同難易度の問題は、成績上位の生徒が基準値未満の得点を取った際に失点の重みが重くなるように同難易度の入替えを定義し、その重みが小さい問題が表の左側に配列されるように設計する。

4.2.2 生徒の成績順位の定義

二値 S-P 表であれば試験の得点が高ければ成績上位の生徒と定義している。そこで同じ性質を持たせるために多値 S-P 表でも同様に試験の得点を利用した。

同得点の生徒において、難易度の低い回答で基準値以下の得点を取った場合は、重みが重くなるように同得点の生徒の入替えを定義しており、その重みが小さいほうが表の上に配列されるように設計する。

4.3 S 曲線と P 曲線

二値 S-P 表と同じ性質を持つ S 曲線を定義するために、多値 S-P 表は列のマス目の見方に配点の累積を用いて S 曲線の引き方を定義する。本稿で提案する多値 S-P 表は各問題に配点が設けられており、隣合う問題の配点を足し合わせることで配点の累積を求めることができる。算出した配点の累積を利用して S 曲線を引く。S 曲線の引き方は生徒の得点が配点の累積が属する問題のマス目の右に縦線を引くという方法である。

二値 S-P 表と同じ性質を持つ P 曲線を設計するために、クラスにおいてどれだけの生徒が正解できているのかという性質を持たせるような P 曲線の引き方に得点者を提案する。二値 S-P 表では各問題について正解者数を求めその正解者数分だけ数えて線を引くという手法を用いていた。そこで提案する多値 S-P 表では各問題に対する正解者数を算出するために得点率を用いた。これは全生徒数に得点率をかけることで算出することができる。

		問題番号								生徒の得点
		1	2	3	4	5	6	7	8	
生徒名	A	4	4	4	4	6	8	6	0	36
	B	4	4	2	4	4	2	8	2	30
	C	4	4	2	4	4	10	0	0	28
	D	4	4	0	4	5	4	6	0	27
	E	4	4	4	2	6	2	2	0	24
	F	4	2	2	0	6	2	0	0	16
	G	0	4	0	0	5	5	0	0	14
	H	4	0	2	2	0	0	0	0	8
	I	0	0	2	0	0	0	0	0	2
配点		4	4	4	4	10	10	10	4	
配点累積		4	8	12	16	26	36	46	50	
基準値		4	4	2	2	5	2	0	0	
得点率		1	1	0.5	0.5	0.5	0.2	0	0	

図 8 S-P 表の設計

4.4 ケアレスミス判定とまぐれ正解判定

二値 S-P 表におけるケアレスミス回答は S 曲線よりも左に分布する回答かつ不正解の回答と定義されていたため、提案する多値 S-P 表では S 曲線よりも左に分布する回答かつ基準値

未満の回答と定義した。ケアレスミス回答は図 9 における S 曲線よりも左にあるグレーで色付けされた基準値未満の回答である。二値 S-P 表におけるまぐれ正解回答は S 曲線よりも右に分布する回答かつ正解の回答と定義されていたため、提案する多値 S-P 表では S 曲線よりも右に分布する回答かつ基準値以上の回答と定義した。まぐれ正解回答は図 9 における S 曲線よりも右にあるグレーで色付けされた基準値以上の回答である。

生徒のケアレスミス回答・まぐれ正解回答の分布を見ることで、試験の採点結果だけでは見ることができない生徒の理解度を診断することができるため、生徒の授業に対する理解度把握において重要である。

		問題番号								生徒の得点
		1	2	3	4	5	6	7	8	
生徒名	A	4	4	4	4	6	8	6	0	36
	B	4	4	2	4	4	2	8	2	30
	C	4	4	2	4	4	10	0	0	28
	D	4	4	0	4	5	4	6	0	27
	E	4	4	4	2	6	2	2	0	24
	F	4	2	2	0	6	2	0	0	16
	G	0	4	0	0	5	5	0	0	14
	H	4	0	2	2	0	0	0	0	8
	I	0	0	2	0	0	0	0	0	2
配点		4	4	4	4	10	10	10	4	
配点累積		4	8	12	16	26	36	46	50	
基準値		4	4	2	2	5	2	0	0	
得点率		1	1	0.5	0.5	0.5	0.2	0	0	

図 9 ケアレスミス回答とまぐれ正解回答

4.5 生徒の理解度別の分類

教育現場の少人数指導における教師の指導の負担を軽減を目指し、本節ではケアレスミス回答の分布の様子が似た生徒のグルーピングの手法について述べる。

4.5.1 各回答に対する重み付け

本稿では、生徒の各回答の特徴を考慮することで各生徒の理解度の特徴ごとにグルーピングを行う。まずはじめに重み付けについて説明する。重み付けをする理由としては生徒の順位と問題の難易度を考慮することで各生徒の重要な学習課題を見つけ、分布を考慮することができるため重みを付けする必要がある。多値 S-P 表は表の左上にいくほど理解の定着度が高くなり、表の右下にいくほど理解の定着度が低くなる [12] という性質を持っているため理解度ポテンシャルが高くなる範囲でケアレスミスをする重みが重くなりように重みづけを設計した。

		問題番号								生徒の得点
		1	2	3	4	5	6	7	8	
生徒名	A	0.0889	0.0889	0.0333	0.0333	0.0333	0	-0.022	-0.022	36
	B	0.0889	0.0889	0.0333	0.0333	0.0333	0	-0.022	-0.022	30
	C	0.0889	0.0889	0.0333	0.0333	0.0333	0	-0.022	-0.022	28
	D	0.0889	0.0889	0.0333	0.0333	0.0333	0	-0.022	-0.022	27
	E	0.0556	0.0556	0	0	0	-0.033	-0.056	-0.056	24
	F	0.0556	0.0556	0	0	0	-0.033	-0.056	-0.056	16
	G	0.0556	0.0556	0	0	0	-0.033	-0.056	-0.056	14
	H	0	0	-0.056	-0.056	-0.056	-0.089	-0.111	-0.111	8
	I	0	0	-0.056	-0.056	-0.056	-0.089	-0.111	-0.111	2
配点		4	4	4	4	10	10	10	4	
配点累積		4	8	12	16	26	36	46	50	
基準値		4	4	2	2	5	2	0	0	
得点率		1	1	0.5	0.5	0.5	0.2	0	0	

図 10 多値 S-P 表の各回答の重み付け

4.5.2 クラスタ数の決定

4.5.1 で述べた重み付けされた回答を用いて、生徒の理解度の特徴ごとにグルーピングの手法にはウォード法を用いたクラスタ分析を行った。また、そのデータをクラスタ分析した際に

クラスタ数を決定する必要があるのだが決定方法にはカリンスキーの方法という手法を用いてクラスタ数を決定することとする [9]。これは対象データに適切なクラスタ数を決定する手法であり、カリンスキーの基準値を最大化するクラスタ数をそのデータのクラスタ数に用いて分析すると良い。以下の数式 (1) にカリンスキーの基準値の計算方法について提示する。以下の n は全データ数、 g はクラスタ数、 k はクラスタ名、 $C(k)$ はカリンスキーの基準値、 $B(g)$ はクラスタ間の分散共分散行列、 $W(g)$ はすべてのクラスタ内の分散共分散行列の総和を示している。このように重み付けを設計しクラスタ分析を行うことで、各問題の特徴を考慮した生徒の理解度に対する特徴のグルーピングを行うことが可能となる。

$$C(k) = \frac{B(g) \cdot (n - g)}{W(g) \cdot (g - 1)} \quad (1)$$

5. 評価方法

二値 S-P 表は二値の回答データを対象としていた。そのため、多値の回答データを用いる際には教師の採点基準によって二値データに変換して二値 S-P 表の手法を適用させることで生徒の理解度診断が行われてきた。本来であれば記述式問題は得点が生徒によって異なるため正解・不正解で採点することは正確に生徒の個別の理解度を見ることができない。そのため、本稿で提案する多値 S-P 表は多値の回答データに対応可能であるため有用であると考えられ、二値 S-P 表よりも正解に理解度を診断できると考えられる。多値 S-P 表の有用性を確かめるために、生徒の理解度の特徴を知る上で重要なケアレスミス判定とまぐれ正解判定の精度を二値の S-P 表と比較する。また、多値 S-P 表で提案した生徒の理解度別のグルーピングの成果を提示する。

5.1 ケアレスミス回答とまぐれ正解回答の定義

表の左上にいくほど自信ありかつ正解の回答が集まり、S 曲線を境に表の右下にいくほど自信なしの回答が集まる [12] という S-P 表の性質を利用して、ケアレスミス回答とまぐれ正解回答を定義する。ケアレスミスの回答は自信ありかつ不正解・基準値未満の回答とし、まぐれ正解回答は自信なしかつ正解・基準値以上の回答と定義する。

5.2 試験の回答データ収集

本稿では各回答に対する生徒の自信あり・自信なしのデータを収集するために同志社大学文化情報学部の 2016 年秋学期に開講されたある講義の二回の試験の回答データを用意した。この受講者は 28 人であり、第一回の試験の受講者は 18 人で第二回の試験の受講者は 17 人である。第二回目の試験の受験者は第一回目の試験の共通の受講者であり、一人は試験を欠席していた。この試験は講義内容の基礎知識を問うものであり、多肢選択式の問題と記述式問題の全 12 問で構成されている。ケアレスミス判定とまぐれ正解判定に対する自信度の一致度合いを調べるために、各問題の回答欄に自信あり・やや自信あり・やや自信なし・自信なしの 4 段階の選択項目を設置し、被験者に解答時に各回答に対する自信度を選択してもらった。

5.3 二値 S-P 表の作成

多値 S-P 表と二値 S-P 表でケアレスミス判定とまぐれ正解判定の精度を評価するために、二値 S-P 表を用意する必要がある。二値 S-P 表は二値の回答データを対象としているため、多値の回答データを用いる際には教師の主観で正解・不正解の二値データに変換されてきた [11]。本稿では、従来のように教師の主観で 0/1 に変換することが困難であったため、4.1.2 節で提案した基準値を用いた変換手法を採用する。この基準値とは、データの半分の値を示す平均値または中央値であり各問題に与えられている。基準値以上の得点は 1 に変換し、基準値未満の得点は 0 に変換して二値の回答データを作成する。この二値の回答データから二値 S-P 表を作成し、多値 S-P 表との精度の比較評価に利用する。

5.4 評価方法

ケアレスミス判定とまぐれ正解判定に対する自信度の一致度合いの精度を二値 S-P 表と多値 S-P 表で比較するために判定の正確さを示す真陽性率と表全体の診断の正確さを示す正診率を評価値に用いる。

表 11 の a にはケアレスミス判定されている回答かつ生徒が解答時に自信ありと答えた回答の総数が当てはめられ、表 11 の b にはケアレスミス判定されている回答かつ生徒が解答時に自信なしと答えた回答の総数が当てはめられる。表 11 の c にはケアレスミス判定されていない回答かつ生徒が解答時に自信があった回答が当てはめられ、表 11 の d にはケアレスミス判定されていない回答かつ生徒が解答時に自信がないと答えた回答の総数が当てはめられる。これらの表 11 に当てはめられた a, b, c, d の値を用いて真陽性率は式 (2) に当てはめることで計算でき、正診率は式 (3) に当てはめて値を求めることができる。

図 12 の a にはまぐれ正解判定されている回答かつ生徒が解答時に自信なしと答えた回答の総数が当てはめられ、図 12 の b にはまぐれ正解判定されている回答かつ生徒が解答時に自信ありと答えた回答の総数が当てはめられる。図 12 の c にはまぐれ正解判定されていない回答かつ生徒が解答時に自信なしと答えた回答が当てはめられ、図 12 の d にはまぐれ正解判定されていない回答かつ生徒が解答時に自信ありと答えた回答の総数が当てはめられる。これらの図 12 に当てはめられた a, b, c, d の値を用いて真陽性率は (2) の式に当てはめることで計算でき、正診率は (3) の式に当てはめることで値を求めることができる。

図 11 ケアレスミス判定

		ケアレスミス	
		判定あり	判定なし
主観 評価	自信あり	a	c
	自信なし	b	d

図 12 まぐれ正解判定

		まぐれ正解	
		判定あり	判定なし
主観 評価	自信なし	a	c
	自信あり	b	d

$$\text{真陽性率} = \frac{a}{a + c} \quad (2)$$

$$\text{正診率} = \frac{a + d}{a + b + c + d} \quad (3)$$

5.5 結 果

第1回の試験の各問題のデータの正規性を検定した結果すべての問題において正規性がないという結果になったため、基準値には中央値を採用した。第2回の試験の各問題のデータの正規性の検定をした結果、問題12のみ正規性があるという結果になったため問題2の基準値には平均値を採用し、それ以外の問題の基準値には中央値を採用した。この基準値のもと多値S-P表を作成しケアレスミス判定とまぐれ正解判定の評価を行った。

結果は、図13の第1回試験の判定精度より第1回の試験において多値S-P表は二値S-P表よりもケアレスミス判定の真陽性率が3.15倍、ケアレスミス判定の正診率が1.11倍、まぐれ正解判定の真陽性率が2.78倍、まぐれ正解判定の正診率が1.26倍、精度が向上した。図14の第二回試験の判定精度より第2回の試験において多値S-P表は二値S-P表よりもケアレスミス判定の真陽性率が1.39倍、ケアレスミス判定の正診率が1.02倍、まぐれ正解判定の真陽性率が1.73倍、まぐれ正解判定の正診率が1.02倍、精度が向上した。よって、ケアレスミス判定とまぐれ正解判定の真陽性率と正診率は二値S-P表よりも多値S-P表の精度が上回る結果になり、多値S-P表の有用性を確認した。

図13 第1回試験の判定精度

	二値	多値
真陽性率(ケアレスミス)	0.067	0.211
正診率(ケアレスミス)	0.593	0.657
真陽性率(まぐれ正解)	0.135	0.375
正診率(まぐれ正解)	0.519	0.653

図14 第2回試験の判定精度

	二値	多値
真陽性率(ケアレスミス)	0.375	0.522
正診率(ケアレスミス)	0.563	0.575
真陽性率(まぐれ正解)	0.271	0.468
正診率(まぐれ正解)	0.649	0.662

各生徒の順位と各問題の難易度を考慮して、生徒の理解度の特徴ごとにグルーピングを行うためにクラスタ分析を行った。クラスタ分析を行った結果、第1回試験においては図13のように生徒番号1・2・3・4・5・6・7・10・11のグループ1と生徒番号8・9・12のグループ2と生徒番号13のグループ3と生徒番号14・15・16・17・18のグループ4に分けることができた。図14にはグループを反映させて色付けを行った多値S-P表である。第2回試験においては図15のように生徒番号1のグループ1と生徒番号2・3・5・6・7のグループ2と生徒番号4・8・9・10・12・14・16のグループ3と生徒番号13・15・17のグループ4に分けることができた。図16は図14でグループ分けされた生徒を反映させて色付けを行った多値S-P表である。

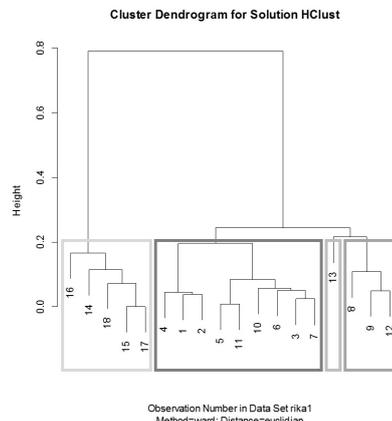


図15 第1回試験のクラスタ分析結果

	1	3	6	8	5	7	2	4	
1	4	4	4	6	4	8	6	0	36
2	4	4	4	4	4	2	8	2	30
3	4	4	4	4	4	8	0	0	28
4	4	4	0	4	4	4	6	0	26
5	4	4	2	4	4	10	0	0	24
6	4	4	4	8	0	0	2	2	24
7	4	4	4	6	2	2	1	0	23
8	4	0	4	4	4	2	0	2	20
9	4	2	2	6	0	4	2	0	20
10	4	4	4	2	4	0	0	0	18
11	4	4	2	4	0	4	0	0	18
12	4	2	2	6	0	2	0	0	16
13	0	4	0	5	0	5	0	0	13
14	4	0	2	0	2	0	0	0	8
15	4	2	2	0	0	0	0	0	8
16	4	2	0	0	0	0	0	2	8
17	0	0	2	0	0	0	0	0	4
18	0	0	0	0	0	0	0	0	0
配点	4	4	4	10	4	10	10	4	
配点累積	4	8	12	22	26	36	46	50	
中央値	4	4	2	4	1	2	0	0	
得点率	1	1	0.5	0.4	0.25	0.2	0	0	

図16 第1回試験の生徒分類

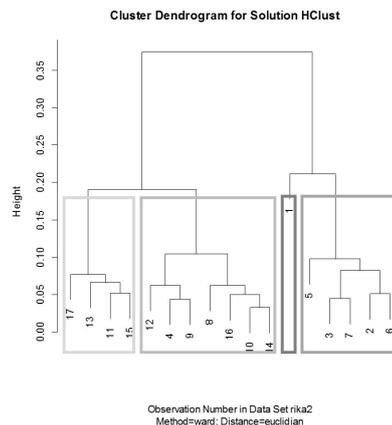


図17 第2回試験のクラスタ分析結果

	7	4	8	11	6	3	1	10	2	12	5	9	
1	2	2	2	10	8	4	4	2	2	7	10	2	55
2	2	2	2	10	4	2	3	2	2	8	6	1	44
3	2	1	2	8	8	4	4	1	2	1	6	0	39
4	2	2	2	8	4	2	4	1	2	3	0	1	36
5	2	2	1	4	8	2	2	1	2	4	6	2	36
6	2	2	2	8	4	2	2	0	1	6	6	0	35
7	2	0	2	8	4	2	4	1	1	2	6	0	32
8	2	1	1	8	6	4	2	1	1	2	2	1	31
9	2	2	1	8	8	0	2	0	1	6	0	0	30
10	2	2	1	6	5	2	2	1	2	4	2	0	29
11	1	2	0	4	8	4	4	0	2	3	0	0	28
12	2	2	2	4	4	2	0	1	2	4	0	2	25
13	0	0	0	4	4	4	4	1	0	4	0	1	22
14	2	2	2	4	4	0	0	0	0	4	2	0	20
15	0	1	1	4	3	4	2	1	2	2	0	0	17
16	1	2	2	4	2	0	0	0	1	2	2	0	16
17	0	0	0	4	2	0	0	0	0	2	0	0	8
配点	2	2	2	10	8	4	4	2	4	10	12	2	
配点率	2	4	6	16	24	28	32	34	38	48	60	62	
基準値	2	2	2	6	4	2	2	1	2	4	2	0	
得点率	1	1	1	0.6	0.5	0.5	0.5	0.5	0.4	0.17	0		

図 18 第 2 回試験の生徒分類

5.6 考 察

二値 S-P 表と比較して多値 S-P 表の方が精度高くなった要因は、正解・不正解では各回答に対する理解度を細かく診断できないが、多値 S-P 表各回答の得点のばらつきを利用して細かく理解度を見ることができているからだと考えられる。

クラスタ分析による第 1 回試験のグルーピング結果においては、グループ 1 は成績優秀でケアレスミスが少ない、グループ 2 は難易度の低い問題でケアレスミスをしている、グループ 3 はグループ 2 よりも難易度の低い問題でケアレスミスをしておりケアレスミスの回答が多い、グループ 4 はケアレスミス回答は少ないが得点が低いため全体的に失点の多い、と考えることができる。このケアレスミスに着目したグルーピングは二値の S-P 表ではどれだけの理解度があつてのケアレスミスなのかということを見ることができないためグルーピングが難しかったが、本稿で提案した多値 S-P 表であれば各回答の理解度を細かく見ることができると、生徒の試験の得点だけでは知ることができない生徒の特徴を把握することを可能とすることができたと考えられる。また、第 2 回試験のグルーピング結果においては、グループ 1 は成績優秀でケアレスミスがない、グループ 2 は難易度の低い問題でケアレスミスをしやすい、グループ 3 はグループ 2 よりもケアレスミスが多い、グループ 4 はケアレスミス回答が非常に多い、と考えることができる。このグルーピング結果を参考に少人数指導を行うことで、試験の得点だけでは分からない生徒の特徴で分類することができるため少人数指導の教師の負担軽減に貢献できると考えられる。

6. おわりに

本稿では、実際の教育現場で実施されている多値の回答データに対応する多値 S-P 表を提案した。また本稿で提案した多値 S-P 表は二値 S-P 表と同じ用途で利用できるようにすることを目的としているため、二値 S-P 表と同じ性質を持つように多値 S-P 表の設計を行った。また、多値 S-P 表が対応している試験の回答データは、各問題に与えられた配点をもとに採点を行うため各問題によって得点者の階級数が異なるという特徴があったため、各問題において基準値を提案した。基準値を用いて作成した多値 S-P 表と二値 S-P 表のケアレスミス判定およびまぐれ正解判定に対する精度の評価を行った結果、多値 S-P 表のケアレスミス判定とまぐれ正解判定の真陽性率および正診率の

全ての精度が向上したため、多値 S-P 表が有用であることを確認した。また、提案する多値 S-P 表は二値 S-P 表と同じ用途で使用できるよう設計してきたため、二値 S-P 表の研究で行われてきた各生徒の各問題に対する理解の定着度を定量化する手法 [12] やその他の二値 S-P 表に評価値を加えることを目的に行われてきた研究の手法を適用することが可能となる。

また本稿では、教師の少人数指導における指導の負担の軽減に貢献するために回答の分布の様子が似た生徒のグルーピングの手法を提案した。グルーピングには統計的手法であるクラスタ分析を用いた。多値 S-P 表を利用したグルーピングは各回答の得点のばらつきを考慮することで細かく理解度を見ることができていた。そのため基準値からどれだけ点数を落としているのか、生徒のクラスにおける位置づけと問題の難易度を考慮することでより生徒の理解度を細かく見たうえで分類することができた。

今後は、本稿で提案した基準値を用いた二値データへの変換手法が従来の教師の主観で 0/1 に変換するという手法とどれだけ近似しているのかを調べ、本稿で提案した多値 S-P 表の精度の評価に利用した二値 S-P 表の妥当性について検証することが課題である。

文 献

- [1] 宇佐美裕康, 尾崎正弘, “動的 Web 教材における学習上のトラブルを考慮した学習環境の提案”, 日本教育情報学会年會論文集. 日本教育情報学会, No. 25, pp. 288-289, Aug 2009.
- [2] 秋月治, “ネットワークを用いた講義時のコミュニケーション支援ツール”, 電子情報通信学会技術研究報告. ET, 教育工学. 一般社団法人電子情報通信学会, Vol.113, No. 482, pp. 183-186, Mar 2014.
- [3] 安達一寿, “ブレンディッドラーニングでの学習活動の類型化に関する分析”, 日本教育工学会論文誌. 日本教育工学会, Vol.31, No. 1, pp. 29-40, May 2007.
- [4] 佐藤 隆博, “S-P 表の作成と解釈”, 明治図書出版, 1975.
- [5] 糸満翔子, “体験を通して理解する伝統文化教材の研究”, 年會論文集. 日本教育情報学会, No. 29, pp. 464-465, November 2013.
- [6] 宮本文穂, 永田信人, 今野将顕, 中村秀明, “バーチャルリアリティを用いた橋梁維持管理技術者教育システムの提案”, 土木情報利用技術論文集. 日本教育情報学会, No. 14, pp. 109-118, 2005.
- [7] 宇佐美裕康, 杉村藍, ADHIKARI CHOLENDRA, 足達義則, 尾崎正弘, “対面授業と Web 習熟度別テストを併用した学習支援システムの提案 学習意欲の維持・向上を目的とした”, 情報教育シンポジウム 2013 論文集. 日本教育情報学会, No. 2, pp. 3-10, August 2013.
- [8] 白旗慎吾, “統計解析入門”, 共立出版, 1992.
- [9] Calinski.T and J.Harabasz, “A dendrite method for cluster analysis”, Communications in statistics, vol.3, No.1, pp.1-27, 1974.
- [10] S.S.SHAPIRO, M.B.WILK, “An analysis of variance test for normality. Biometrika”, Biometrika, vol.52, pp.591-611, 1965.
- [11] 川上省三, “サーバーデー英語試験結果の活用法の再考-多人数データの S-P 表分析-”, 日本工業大学, Vol. 45, No. 3, Sep 2015.
- [12] 三輪和久, 下村勉, “S-P 表に基づく理解度ポテンシャルの定義とその実証的考察”, 電子情報通信学会論文誌. A, 基礎・境界, 一般社団法人電子情報通信学会, Vol. 76, No. 9, pp. 1359-1367, Sep 1993.